Chapter 1
Linear differential systems with parameter excitation

This chapter intends to introduce to the general setting of linear stochastic systems

$$\dot{Z}_t \;=\; \mathbf{A}(t)\, Z_t,$$

where $(\mathbf{A}(t))_{t\geq 0}$ is a matrix valued stochastic process and $\dot{Z}_t \equiv \frac{d}{dt} Z_t$ denotes the derivative with respect to the "time" variable t. Such differential systems are called *parametrically excited (perturbed)* or *real noise* linear systems; in the engineering literature the terminology *rheo linear* system is also used. The system matrix is assumed to be a continuous mapping defined on the state space of a Markov process which serves as stochastic input for the system. Hence, the above matrix process is of the form

$$\mathbf{A}(t) \;=\; \mathbf{A}(X_t),$$

where $(X_t)_{t\geq 0}$ is the input process. Note for preciseness that some authors further differentiate the nature of the noise by distinguishing "real noise" on the one hand which is defined on the two-sided time set $\mathbb{R}$ and *Markovian noise* on the other hand which is a Markov process with time set $\mathbb{R}_+$; see Arnold and Kliemann [Ar-Kl 83, p.4]. In this book both terms will be used interchangeably in the latter sense. Linear real noise systems with state-dependent coefficient matrix $\mathbf{A}(X_t)$ and Markovian input noise X_t as above have for example been investigated by Frisch [Fs 66]. In this reference it is argued heuristically how a Fokker-Planck equation might be obtained for $(X_t, Z_t)_t$, if X_t is stationary. Real noise systems with Markovian input noise are also subject to the considerations of Kats and Krasovskii [Ka-Kv 60]; however, these authors consider the case that $(X_t)_{t\geq 0}$ is a Markov chain with finite state space. Now the setting of our work also assumes that the input process has finitely many "states of preference", but is a continuous process defined by a stochastic differential equation with respect to Brownian motion. These preferential ("metastable") states correspond to certain time scales which will be made precise in the next chapter.

W. Siegert, *Local Lyapunov Exponents.*
Lecture Notes in Mathematics 1963.

Since the real noise process $(X_t)_{t\geq 0}$ will be defined by some SDE, the coupled system (X, Z) is also given by the resulting SDE. Note that this differential equation is degenerate by definition in the sense that no noise term, i.e. no summand of the type $g(X_t, Z_t)\, dB_t$ containing the differential of a Brownian motion (or some more general stochastic process) B, is visible in the Z-direction, but solely in the X-variable; see e.g. Bunke [Bu 72, Ch.6] for standard results on such equations. A different type of randomly perturbed linear SDE is the so called *white noise case*; see e.g. Khasminskii [Kh 67], [Kh 80] and [Kh 60]. Here, there is no real noise input but stochastic integration with respect to Brownian motion in the SDE for Z itself. The latter type of systems is not subject to our investigations, but also will be commented on in this chapter.

From an applications' point of view real noise stochastic systems are often considered as more realistic for describing "real"-world problems. The reason is that most processes to be modeled are concerned with variables of a specified magnitude or even restricted to a bounded interval such as concentrations (e.g. in chemistry), population fluctuations (in life sciences) or strictly positive parameters in technical systems; see e.g. Kliemann [Kl 80, App.IV: p.3] and [Kl 83b], Arnold et al. [Ar-Hh-Lf 78], Ahmadi and Morshedi [Ah-Mr 78] and Griesbaum [Gb 99] for further examples and discussions. In contrast using white noise instead is mostly considered as too drastic an idealization; see Kliemann [Kl 80, App.IV: p.3] and Wihstutz [Wh 75, p.3].

1.1 The model

The system which lies at the heart of these investigations is the real-noise driven system

$$\begin{aligned} dZ_t^\varepsilon &= \mathbf{A}\left(X_t^\varepsilon\right)\, Z_t^\varepsilon\, dt \\ dX_t^\varepsilon &= b\left(X_t^\varepsilon\right)\, dt + \sqrt{\varepsilon}\, \sigma\left(X_t^\varepsilon\right)\, dW_t \end{aligned} \tag{1}$$

where $\mathbf{A} \in C(\mathbb{R}^d, \mathbb{K}^{n\times n})$ is a continuous matrix function ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$); $d, n \in \mathbb{N}$ denote the dimensions of the state spaces of X^ε and Z^ε, respectively; $\varepsilon \geq 0$ parametrizes the intensity of $(W_t)_{t\geq 0}$ which denotes a Brownian motion in $\mathbb{R}^d$, defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and X^ε is a diffusion, defined by the above SDE with coefficient functions $b \in C^\infty(\mathbb{R}^d, \mathbb{R}^d)$ and $\sigma \in C^\infty(\mathbb{R}^d, \mathbb{R}^{d\times d})$.

A detailed study of the SDE for X^ε will follow in the next chapter specifying the assumptions; see (2.1) and the corresponding set of assumptions 2.1.1. For the moment, just assume that there exists a unique non-exploding solution[1] $X^{\varepsilon,x}$, where the superscript $x \in \mathbb{R}^d$ denotes its initial value.

[1] For standard terminology concerning SDEs, such as solution, uniqueness and non-explosiveness, we refer to Hackenbroch and Thalmaier [Hb-Th 94, Ch.6], Khasminskii [Kh 80, Ch.I-III] and Freidlin and Wentzell [Fr-We 98, Ch.1] among many others.

Since the stochastic process $X^{\varepsilon,x}$ is non-explosive and has ($\mathbb{P}$-almost surely) continuous paths, the SDE (1) defines a continuous stochastic process $(X^\varepsilon, Z^\varepsilon)$ in $\mathbb{R}^d \times \mathbb{K}^n$; more precisely, the differential equation for Z^ε,

$$dZ_t^\varepsilon \;=\; \mathbf{A}\big(X_t^{\varepsilon,x}(\omega)\big)\, Z_t^\varepsilon \;\, dt\,, \qquad Z_0^\varepsilon \;=\; z \in \mathbb{K}^n\,, \tag{1.1}$$

is a random differential equation (RDE) which ω-wise, i.e. for any realization $X_t^{\varepsilon,x}(\omega)$, is solved as an ODE; hence, the continuity of $t \mapsto \mathbf{A}(X_t^{\varepsilon,x}(\omega))$ guarantees existence and uniqueness of the solution path $Z_t^\varepsilon(\omega,x,z)$ in $\mathbb{K}^n$ (see e.g. Coppel [Cp 65, p.42]) which together with $X^{\varepsilon,x}$ forms the solution $(X^\varepsilon, Z^\varepsilon)$ of (1) in $\mathbb{R}^d \times \mathbb{K}^n$, starting in (x,z). Collecting these solution paths (possibly defined as 0 on an exceptional subset of Ω of zero $\mathbb{P}$-measure) yields a well-defined mapping

$$\begin{aligned} Z^\varepsilon : \mathbb{R}_+ \times \Omega \times \mathbb{R}^d \times \mathbb{K}^n &\longrightarrow \mathbb{K}^n\,, \\ (t,\omega,x,z) &\;\mapsto\; Z^\varepsilon(t,\omega,x,z) \;:=\; Z^\varepsilon(t,\omega,x)\, z \\ &\qquad := Z_t^\varepsilon(\omega,x)\, z \;:=\; Z_t^\varepsilon(\omega,x,z) \end{aligned}$$

and these notations will be used equivalently. Being linear in the z-variable the mappings $Z^\varepsilon(t,\omega,x)$ form a matrix process which (as the corresponding fundamental matrix) solves the random matrix differential equation

$$dZ_t^\varepsilon \;=\; \mathbf{A}\big(X_t^{\varepsilon,x}(\omega)\big)\, Z_t^\varepsilon \;\, dt\,, \qquad Z_0^\varepsilon \;=\; \mathrm{id}_{\mathbb{K}^n}.$$

Furthermore, as the fundamental solution of a linear equation, $Z^\varepsilon(t,\omega,x)$ takes its values in the set of invertible matrices, its inverse being governed by the matrix differential equation

$$d(Z^{-1})_t \;=\; -\,(Z^{-1})_t\, \mathbf{A}(X_t^{\varepsilon,x})\, dt\,.$$

In particular, the *Wronski-determinant* $\det(Z_t^\varepsilon(\omega,x))$ does not vanish and differentiating pathwise with respect to t yields the random differential equation

$$d\,\det(Z_t^\varepsilon(\omega,x)) \;=\; \mathrm{trace}\Big(\mathbf{A}\big(X_t^{\varepsilon,x}(\omega)\big)\Big)\, \det(Z_t^\varepsilon(\omega,x))\;\, dt, \quad \det(Z_0^\varepsilon(\omega,x)) = 1,$$

and therefore one obtains the *Jacobi equation (Liouville equation)*

$$\begin{aligned} \det(Z_t^\varepsilon(\omega,x)) &= \det(Z_0^\varepsilon(\omega,x))\;\; \exp\left(\int_0^t \mathrm{trace}\left(\mathbf{A}\big(X_u^{\varepsilon,x}(\omega)\big)\, du\right)\right) \\ &= \exp\left(\int_0^t \mathrm{trace}\left(\mathbf{A}\big(X_u^{\varepsilon,x}(\omega)\big)\, du\right)\right) \;\; ; \end{aligned} \tag{1.2}$$

see e.g. Coppel [Cp 65, p.44] for the pathwise calculations. Note that for $\varepsilon = 0$ all objects remain well-defined as the solutions of the resulting ODE.

Under appropriate conditions it can be proved that X^ε and hence also Z^ε depend continuously on the parameter $\varepsilon > 0$; see Blagovescenskii and Freidlin [Bla-Fr 61]. However, this will not be used in the sequel.

In this work we are interested in *Lyapunov exponents*, i.e. exponential growth rates

$$\frac{1}{t} \log |\, Z_t^\varepsilon(\omega, x, z) \,| \ ,$$

which are realized for large times $t = T(\varepsilon)$; such growth rates will then be called *local* Lyapunov exponents. In the following section spherical coordinates will be introduced in order to get an integral decomposition for the growth rate mentioned previously. This representation which is also an ingredient of the Furstenberg-Khasminskii formula (see p.27) provides quantitative information concerning $|Z^\varepsilon|$ and will play a fundamental role in subsection 4.4.2.

1.2 Spherical coordinates for linear systems

In the following the system Z^ε will be decomposed by means of spherical coordinates; for this section we take $\mathbb{K} = \mathbb{R}$, i.e. $Z_t^\varepsilon(\omega, x, z)$ is an element of $\mathbb{R}^n$ where $n \in \mathbb{N}$. In the literature on stochastic systems the use of such coordinates (also called the *projection method*) is accredited to Khasminskii [Kh 67] and Infante [In 68]; see e.g. Kliemann [Kl 79, p.465] and [Kl 80, p.144] and Crauel [Cra 84, p.13]. It is, however, interesting to note that this use of spherical coordinates is by no means restricted to the stochastic case. In fact, this method had been a well established tool in investigating deterministic linear differential systems before; see e.g. Levi-Civita [LC 11] and Wintner [Wi 50] and [Wi 57].

Since the matrix $Z^\varepsilon(\omega, x)$ is invertible, the process $Z_t^\varepsilon(\omega, x, z)$ is non-zero for $z \neq 0$ and can hence be characterized by its radial and spherical components via

$$\varrho_t^\varepsilon(\omega, x, z) \ := \ |\, Z_t^\varepsilon(\omega, x, z) \,| \ \in (0, \infty)$$

and

$$\psi_t^\varepsilon(\omega, x, z) \ := \ \frac{Z_t^\varepsilon(\omega, x, z)}{\varrho_t^\varepsilon(\omega, x, z)} \ \in S^{n-1} \ ,$$

respectively, if $z \neq 0$; here, $S^{n-1} := \{y \in \mathbb{R}^n : |y| = 1\}$ denotes the unit sphere in $\mathbb{R}^n$.

For the system Z^ε as given by (1), the defining pathwise differential equation (1.1) is equivalent to the system of the two RDEs

$$d\varrho_t^\varepsilon = Q(X_t^\varepsilon, \psi_t^\varepsilon)\, \varrho_t^\varepsilon \, dt \ , \tag{1.3}$$

$$d\psi_t^\varepsilon = h(X_t^\varepsilon, \psi_t^\varepsilon)\, dt \ \ , \tag{1.4}$$

where

$$Q(x,\psi) \;:=\; Q\big(\mathbf{A}(x),\psi\big) \;:=\; \big\langle\, \mathbf{A}(x)\psi\,,\,\psi\,\big\rangle$$

and

$$h(x,\psi) \;:=\; h\big(\mathbf{A}(x),\psi\big) \;:=\; \mathbf{A}(x)\psi \;-\; \big\langle\, \mathbf{A}(x)\psi\,,\,\psi\,\big\rangle\,\psi\;;$$

here, $\langle\,.\,,\,.\,\rangle$ denotes the standard scalar product of $\mathbb{R}^n$. This follows in a straightforward manner by calculating the pathwise derivative $\frac{d}{dt}$ for fixed (ω, x, z) where $z \neq 0$, since

$$\begin{aligned}
\frac{d}{dt}\,\varrho_t^\varepsilon \equiv \frac{d}{dt}\,\langle Z_t^\varepsilon, Z_t^\varepsilon\rangle^{1/2}
&= \frac{1}{2}\,\frac{1}{\varrho_t^\varepsilon}\,\frac{d}{dt}\,\langle Z_t^\varepsilon, Z_t^\varepsilon\rangle \\
&= \frac{1}{2}\,\frac{1}{\varrho_t^\varepsilon}\,2\,\left\langle \frac{d}{dt}\,Z_t^\varepsilon, Z_t^\varepsilon\right\rangle \\
&= \frac{1}{\varrho_t^\varepsilon}\,\big\langle \mathbf{A}(X_t^\varepsilon)\,Z_t^\varepsilon, Z_t^\varepsilon\big\rangle \;\equiv\; \big\langle \mathbf{A}(X_t^\varepsilon)\,\psi_t^\varepsilon, \psi_t^\varepsilon\big\rangle\,\varrho_t^\varepsilon
\end{aligned}$$

and hence

$$\begin{aligned}
\frac{d}{dt}\,\psi_t^\varepsilon \;\equiv\; \frac{d}{dt}\,\frac{Z_t^\varepsilon}{\varrho_t^\varepsilon}
&= \frac{\varrho_t^\varepsilon\,\frac{d}{dt}\,Z_t^\varepsilon \;-\; Z_t^\varepsilon\,\frac{d}{dt}\,\varrho_t^\varepsilon}{(\varrho_t^\varepsilon)^2} \\
&= \frac{\varrho_t^\varepsilon\,\mathbf{A}(X_t^\varepsilon)\,Z_t^\varepsilon \;-\; Z_t^\varepsilon\,\big\langle \mathbf{A}(X_t^\varepsilon)\,\psi_t^\varepsilon, \psi_t^\varepsilon\big\rangle\,\varrho_t^\varepsilon}{(\varrho_t^\varepsilon)^2} \\
&\equiv \mathbf{A}(X_t^\varepsilon)\,\psi_t^\varepsilon \;-\; \big\langle \mathbf{A}(X_t^\varepsilon)\,\psi_t^\varepsilon, \psi_t^\varepsilon\big\rangle\,\psi_t^\varepsilon
\end{aligned}$$

due to (1.1); conversely, these differential equations for $(\varrho^\varepsilon, \psi^\varepsilon)$ imply that

$$\begin{aligned}
\frac{d}{dt}\,Z_t^\varepsilon \equiv \frac{d}{dt}\,(\varrho_t^\varepsilon\,\psi_t^\varepsilon)
&= \psi_t^\varepsilon\,\frac{d}{dt}\,\varrho_t^\varepsilon \;+\; \varrho_t^\varepsilon\,\frac{d}{dt}\,\psi_t^\varepsilon \\
&= \psi_t^\varepsilon\,\big\langle \mathbf{A}(X_t^\varepsilon)\,\psi_t^\varepsilon, \psi_t^\varepsilon\big\rangle\,\varrho_t^\varepsilon \;+\; \varrho_t^\varepsilon\,\mathbf{A}(X_t^\varepsilon)\,\psi_t^\varepsilon \;-\; \varrho_t^\varepsilon\,\big\langle \mathbf{A}(X_t^\varepsilon)\,\psi_t^\varepsilon, \psi_t^\varepsilon\big\rangle\,\psi_t^\varepsilon \\
&\equiv \mathbf{A}(X_t^\varepsilon)\,Z_t^\varepsilon \quad .
\end{aligned}$$

Note that

$$h(\mathbf{A},\psi) \equiv \mathbf{A}\psi - \langle \mathbf{A}\psi, \psi\rangle\psi$$

vanishes at an element $\psi \in S^{n-1}$, if and only if ψ is an eigenvector of the matrix $\mathbf{A}$. The above vector field $h(\mathbf{A},\,.\,)$ is considered as the *projection* of the linear vector field $\psi \mapsto \mathbf{A}\psi$ onto S^{n-1} for a fixed $\mathbf{A} \in \mathbb{R}^{n\times n}$.

Due to the multiplicative structure of the RDE for ϱ^ε, it can be integrated pathwise by separation of variables, thus giving

$$|\,Z_t^\varepsilon(\omega,x,z)\,| \;\equiv\; \varrho_t^\varepsilon(\omega,x,z) \;=\; |\,z\,|\,\exp\left(\int_0^t Q\Big(X_u^{\varepsilon,x}(\omega), \psi_u^\varepsilon(\omega,x,z)\Big)\,du\right)$$

and therefore

$$\frac{1}{t}\log|Z_t^\varepsilon(\omega,x,z)| = \frac{\log|z|}{t} + \frac{1}{t}\int_0^t Q\Big(X_u^{\varepsilon,x}(\omega),\psi_u^\varepsilon(\omega,x,z)\Big)\,du. \tag{1.5}$$

Now since $h(x,-\psi) = -h(x,\psi)$ for all x and ψ, the SDE for ψ_t^ε is symmetric with respect to the origin; more precisely, $h(x,\,.\,)$ can be viewed as a vector field on the projective space P^{n-1}, i.e. on S^{n-1} where opposite points are identified; therefore, ψ^ε is also considered as a P^{n-1}-valued process. The SDE for ϱ^ε and formula (1.5) remain unaffected, since $Q(x,-\psi) = Q(x,\psi)$ for all x and ψ. Furthermore, the projective process ψ^ε is decoupled from the radial process ϱ^ε; therefore, $(X^\varepsilon,\psi^\varepsilon)$ is a Markov process in $\mathbb{R}^d \times S^{n-1}$ or $\mathbb{R}^d \times P^{n-1}$, respectively, whose generating partial differential operator is given by

$$\mathcal{L}^\varepsilon \;:=\; \mathcal{G}^\varepsilon \,+\, h\,\frac{\partial}{\partial\psi}\;,$$

where

$$\mathcal{G}^\varepsilon \;:=\; \sum_{i=1}^d b_i\,\frac{\partial}{\partial x_i} \,+\, \frac{\varepsilon}{2}\sum_{i,j=1}^d a_{ij}\,\frac{\partial^2}{\partial x_i\,\partial x_j}$$

is the generator of X^ε to be discussed later; see section 2.2.

Remark 1.2.1 ($n = 2$). Consider the two-dimensional case, i.e. $\mathbb{K}^n = \mathbb{R}^2$. Here, we can canonically identify $\psi \in S^2$ with its angle $\alpha \in [0,2\pi)$ via

$$\begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix} \;=\; \psi\;,$$

for which we will use the abbreviation

$$\psi \;\doteq\; \alpha$$

in the following. In particular, the spherical process ψ_t^ε on S^1 defines the *angle process* $\alpha_t^\varepsilon \in [0,2\pi)$ by

$$\begin{pmatrix}\cos\alpha_t^\varepsilon\\ \sin\alpha_t^\varepsilon\end{pmatrix} \;:=\; \psi_t^\varepsilon\;.$$

The previous SDE (1.4), $d\psi_t^\varepsilon \;=\; h(X_t^\varepsilon,\psi_t^\varepsilon)\,dt$, implies the following equation for α_t^ε:

$$d\alpha_t^\varepsilon \;=\; \bar h(X_t^\varepsilon,\alpha_t^\varepsilon)\,dt\;, \tag{1.6}$$

for the drift $\bar h$ being defined by

$$\bar h(x,\alpha) := \bar h\big(\mathbf{A}(x),\alpha\big)\;,$$

where

$$\bar{h}\big(\mathbf{A}, \alpha\big) := -\, a_{12} \sin^2 \alpha \,+\, a_{21} \cos^2 \alpha \,+\, \big(a_{22} - a_{11}\big) \sin \alpha \cos \alpha$$

for matrices

$$\mathbf{A} \;=\; \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \;\in\; \mathbb{R}^{2\times 2}\;;$$

also see Arnold and Kliemann [Ar-Kl 83, p.59]. More precisely, let $\alpha \doteq \binom{\cos\alpha}{\sin\alpha} =: \psi(\alpha)$ and define

$$\psi'(\alpha) \;:=\; \frac{d}{d\alpha}\,\psi(\alpha) \;=\; \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}\;;$$

then one directly calculates that

$$\bar{h}(\mathbf{A}, \alpha) \;=\; \Big\langle\, \mathbf{A}\psi(\alpha)\,,\, \psi'(\alpha)\,\Big\rangle$$

for any $\mathbf{A} \in \mathbb{R}^{2\times 2}$; furthermore, $(\psi(\alpha), \psi'(\alpha))$ is an orthonormal basis of $\mathbb{R}^2$ for any α and one obtains altogether for the pathwise derivatives that

$$\begin{aligned}
\frac{d}{dt}\,\psi^\varepsilon_t &\equiv \frac{d}{dt}\,\psi(\alpha^\varepsilon_t) \\
&= \psi'(\alpha^\varepsilon_t)\;\frac{d}{dt}\,\alpha^\varepsilon_t \\
&= \psi'(\alpha^\varepsilon_t)\;\;\bar{h}(\mathbf{A}(X^\varepsilon_t), \alpha^\varepsilon_t) \\
&= \big\langle\, \mathbf{A}(X^\varepsilon_t)\psi(\alpha^\varepsilon_t)\,,\, \psi'(\alpha^\varepsilon_t)\,\big\rangle\;\; \psi'(\alpha^\varepsilon_t) \\
&= \mathbf{A}(X^\varepsilon_t)\psi(\alpha^\varepsilon_t) \;-\; \big\langle\, \mathbf{A}(X^\varepsilon_t)\psi(\alpha^\varepsilon_t)\,,\, \psi(\alpha^\varepsilon_t)\,\big\rangle\;\; \psi(\alpha^\varepsilon_t) \\
&\equiv h(\mathbf{A}(X^\varepsilon_t), \psi^\varepsilon_t)
\end{aligned}$$

which proves that the RDEs for ψ^ε and α^ε are equivalent, since this calculation can be read in both directions.

Note that it follows from the above definition of $\bar{h}$ that $\bar{h}(\mathbf{A} + c\, I_2\,,\,.\,) = \bar{h}(\mathbf{A},\,.\,)$, for all $\mathbf{A} \in \mathbb{R}^{2\times 2}$ and $c \in \mathbb{R}$, where I_2 denotes the two-dimensional unit matrix; an example for such a system is given by Ahmadi and Morshedi [Ah-Mr 78, Sec.IV]; also see Kliemann and Rümelin [Kl-Rm 81, p.17].

Also note that if $\mathbf{A} \in \mathbb{R}^{2\times 2}$ is symmetric, then $\bar{h}$ reads

$$\begin{aligned}
\bar{h}(\mathbf{A}, \alpha) &\equiv -\, a_{12} \sin^2 \alpha \,+\, a_{21} \cos^2 \alpha \,+\, (a_{22} - a_{11}) \sin \alpha \cos \alpha \\
&= a_{12} \cos 2\alpha \,+\, \frac{1}{2}\,(a_{22} - a_{11}) \sin 2\alpha\;.
\end{aligned}$$

In any case the RDE (1.3) for ϱ^ε now reads

$$d\varrho^\varepsilon_t = \bar{Q}(X^\varepsilon_t, \alpha^\varepsilon_t)\; \varrho^\varepsilon_t\; dt\;,$$

for $\bar{Q}$ being defined by

$$\bar{Q}(x,\alpha) := \bar{Q}(\mathbf{A}(x),\alpha) \ ,$$

where

$$\bar{Q}(\mathbf{A},\alpha) := a_{11}\,\cos^2\alpha \,+\, a_{22}\,\sin^2\alpha \,+\, (a_{12}+a_{21})\,\sin\alpha\,\cos\alpha \tag{1.7}$$

is just designed such that

$$\begin{aligned} Q(\mathbf{A},\psi) &\equiv \langle \mathbf{A}\psi, \psi\rangle \\ &\equiv \left\langle \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}, \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} \right\rangle \\ &= a_{11}\,\cos^2\alpha \,+\, a_{22}\,\sin^2\alpha \,+\, (a_{12}+a_{21})\,\sin\alpha\,\cos\alpha \\ &\equiv \bar{Q}(\mathbf{A},\alpha) \ , \end{aligned}$$

if $\alpha \doteq \psi$. Formula (1.5) hence becomes

$$\frac{1}{t}\log|Z_t^\varepsilon(\omega,x,z)| = \frac{\log|z|}{t} + \frac{1}{t}\int_0^t \bar{Q}\Big(X_u^{\varepsilon,x}(\omega), \alpha_u^\varepsilon(\omega,x,z)\Big)\,du. \tag{1.8}$$

Furthermore, the drift function $\bar{h}(x,\,.\,)$ is π-periodic which corresponds to the fact that $h(x,-\psi) = -h(x,\psi)$ for $\psi \in S^1$; α^ε is the solution of the above real noise SDE modulo π. Thus $(X_t^\varepsilon, \alpha_t^\varepsilon)$ is a Markov process in $\mathbb{R}^2 \times [0,\pi)$ with generator

$$\bar{\mathcal{L}}^\varepsilon \ := \ \mathcal{G}^\varepsilon \,+\, \bar{h}\,\frac{\partial}{\partial\alpha} \ .$$

Another way of describing the motion of the angle is by means of

$$\xi_t^\varepsilon \ := \ \tan\alpha_t^\varepsilon$$

whenever $\alpha_t^\varepsilon \notin \{(2k+1)\pi/2 : k \in \mathbb{Z}\}$; defining

$$F(\mathbf{A},\xi) := a_{21} \,-\, a_{12}\,\xi^2 \,+\, [\,a_{22} - a_{11}\,]\;\xi$$

for $\mathbf{A} \in \mathbb{R}^{2\times 2}$ and $\xi \in \mathbb{R}$, the dynamics of ξ^ε is described by the differential equation

$$\begin{aligned} d\xi_t^\varepsilon &= F(\mathbf{A}(X_t^\varepsilon), \xi_t^\varepsilon) \\ &\equiv \Big\{ a_{21}(X_t^\varepsilon) \,-\, a_{12}(X_t^\varepsilon)\,(\xi_t^\varepsilon)^2 \,+\, \big[\,a_{22}(X_t^\varepsilon) - a_{11}(X_t^\varepsilon)\,\big]\;\xi_t^\varepsilon \Big\}\,dt\,, \end{aligned} \tag{1.9}$$

since

$$\begin{aligned}\frac{d}{dt}\,\xi_t^\varepsilon &= \frac{1}{\cos^2\alpha_t^\varepsilon}\;\bar h(X_t^\varepsilon,\alpha_t^\varepsilon)\\ &\equiv \frac{1}{\cos^2\alpha_t^\varepsilon}\Big\{-a_{12}(X_t^\varepsilon)\,\sin^2\alpha_t^\varepsilon\;+\;a_{21}(X_t^\varepsilon)\,\cos^2\alpha_t^\varepsilon\\ &\qquad\qquad\qquad+\big[a_{22}(X_t^\varepsilon)-a_{11}(X_t^\varepsilon)\big]\,\sin\alpha_t^\varepsilon\,\cos\alpha_t^\varepsilon\Big\}\\ &= a_{21}(X_t^\varepsilon)\;-\;a_{12}(X_t^\varepsilon)\,(\xi_t^\varepsilon)^2\;+\;\big[a_{22}(X_t^\varepsilon)-a_{11}(X_t^\varepsilon)\big]\,\xi_t^\varepsilon.\end{aligned}$$

□

Remark 1.2.2. The above pathwise considerations can be rewritten for any linear differential equation with time dependent system matrix,

$$\dot Z_t \;=\; \mathbf{A}(t)\,Z_t \qquad\qquad (t\ge 0)$$

in $\mathbb{R}^n$, as

$$\begin{aligned}\dot\varrho_t &= Q(t,\psi_t)\;\varrho_t\;,\\ \dot\psi_t &= h(t,\psi_t)\;\;,\end{aligned}$$

where $\varrho_t := |\,Z_t\,| \in (0,\infty)$ and $\psi_t := \frac{Z_t}{\varrho_t} \in S^{n-1}$ whenever $Z_0\neq 0$. Here, the notation is adapted as

$$Q(t,\psi)\;:=\;Q\big(\mathbf{A}(t),\psi\big)\;\equiv\;\big\langle\,\mathbf{A}(t)\psi\,,\,\psi\,\big\rangle$$

and

$$h(t,\psi)\;:=\;h\big(\mathbf{A}(t),\psi\big)\;:=\;\mathbf{A}(t)\psi\;-\;\big\langle\,\mathbf{A}(t)\psi\,,\,\psi\,\big\rangle\,\psi\;.$$

In particular,

$$\frac{1}{t}\,\log|\,Z_t\,|\;=\;\frac{\log|\,Z_0\,|}{t}\;+\;\frac{1}{t}\int_0^t Q(u,\psi_u)\,du\,. \tag{1.10}$$

For the two-dimensional case $n=2$,

$$\mathbf{A}(t) := \begin{pmatrix} a_{11}(t) & a_{12}(t)\\ a_{21}(t) & a_{22}(t)\end{pmatrix}\;\;,$$

and the then well defined angle process $\alpha_t \doteq \psi_t$, i.e.

$$\begin{pmatrix}\cos\alpha_t\\ \sin\alpha_t\end{pmatrix}\;:=\;\psi_t\;\;,$$

it follows that

$$\dot\alpha_t\;=\;\bar h(t,\alpha_t)\;,$$

where

$$\begin{aligned}\bar{h}(t,\alpha) &:= \bar{h}(\mathbf{A}(t),\alpha)\\ &\equiv -a_{12}(t)\,\sin^2\alpha + a_{21}(t)\,\cos^2\alpha + (a_{22}(t)-a_{11}(t))\,\sin\alpha\,\cos\alpha\\ &\equiv \left\langle\, \mathbf{A}(t)\psi(\alpha)\,,\,\psi'(\alpha)\,\right\rangle .\end{aligned}$$

The formula (1.10) becomes

$$\frac{1}{t}\,\log|\,Z_t\,| \;=\; \frac{\log|\,Z_0\,|}{t} + \frac{1}{t}\int_0^t \bar{Q}(u,\alpha_u)\,du \quad , \tag{1.11}$$

where $\bar{Q}(t,\alpha) := \bar{Q}(\mathbf{A}(t),\alpha)$. Furthermore,

$$\xi_t \;:=\; \tan\alpha_t$$

is driven by the differential equation

$$\begin{aligned}\dot{\xi}_t &= F(\mathbf{A}(t),\xi_t)\\ &\equiv a_{21}(t) \,-\, a_{12}(t)\,\xi_t^2 + [\,a_{22}(t)-a_{11}(t)\,]\xi_t \;.\end{aligned}$$

□

The following definition is taken from Arnold and Kliemann [Ar-Kl 83, p.13] and from Kliemann [Kl 79, p.464].

Definition 1.2.3 (Switching surfaces of a drift function). Let $F : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ be a continuous function. The points $(x,y) \in \mathbb{R}^d \times \mathbb{R}$ with $F(x,y) = 0$ define connected surfaces in $\mathbb{R}^d \times \mathbb{R}$. These are called the *switching surfaces (of F)*. If $d = 1$, these sets will also be called the *switching curves (of F)*.

This terminology will now be used for investigating the behavior of ξ^ε and α^ε, respectively. Here, we follow Arnold and Kliemann [Ar-Kl 83, p.59f.]; also see Kliemann [Kl 80, p.148f.] and [Kl 79, p.465f.].

Definition-Remark 1.2.4 (Switching surfaces corresponding to (1.1)). Consider the RDE (1.1) which is the pathwise differential system defined by the SDE (1) and suppose that $\mathbb{K}^n = \mathbb{R}^2$, i.e. the system matrix is given by a mapping $\mathbf{A} \in C(\mathbb{R}^d, \mathbb{R}^{2\times 2})$. By (1.9) $\mathbf{A}(\,.\,)$ induces the drift vector field for $\xi^\varepsilon \equiv \tan\alpha^\varepsilon$,

$$F(x,\xi) \;:=\; F(\mathbf{A}(x),\xi) \equiv a_{21}(x) \,-\, a_{12}(x)\,\xi^2 \,+\, [\,a_{22}(x)-a_{11}(x)\,]\;\xi \;\;,$$

where $x \in \mathbb{R}^d$ and $\xi \in \mathbb{R}$. According to the previous definition 1.2.3 we look for the zeros,

$$F(x,\xi) \;\;=\;\; F(\mathbf{A}(x),\xi) = 0 \;\;;$$

from the solution formula for quadratic equations it follows that one gets

no switching curve, if $a_{12} \neq 0$ and $(a_{22} - a_{11})^2 < -4a_{12}a_{21}$;

one switching curve

$$\xi_1(\mathbf{A}(x)) = \frac{a_{22} - a_{11}}{2a_{12}}, \text{ if } a_{12} \neq 0 \text{ and } (a_{22} - a_{11})^2 = -4a_{12}a_{21},$$

and

$$\xi_1(\mathbf{A}(x)) = \infty, \text{ if } a_{12} = 0 \text{ and } a_{11} = a_{22} \text{ and } a_{21} \neq 0,$$

respectively;

two switching curves

$$\xi_{1,2}(\mathbf{A}(x)) = \frac{(a_{22} - a_{11}) \pm \sqrt{(a_{22} - a_{11})^2 + 4a_{12}a_{21}}}{2a_{12}}, \quad \text{if } a_{12} \neq 0 \text{ and } (a_{22} - a_{11})^2 > -4a_{12}a_{21},$$

and

$$\xi_1(\mathbf{A}(x)) = \infty, \ \xi_2(\mathbf{A}(x)) = \frac{a_{21}}{a_{11} - a_{22}}, \text{ if } a_{12} = 0 \text{ and } a_{11} \neq a_{22},$$

respectively;

infinitely many switching curves

$$\xi \in \mathbb{R} \cup \{\infty\}, \text{ if } a_{12} = 0 \text{ and } a_{11} = a_{22} \text{ and } a_{21} = 0.$$

The above switching curves for F, i.e. for the motion of $\xi^\varepsilon \equiv \tan \alpha^\varepsilon$, directly translate into switching curves of the drift

$$\bar{h}(\mathbf{A}(x), \alpha) \equiv -a_{12}(x) \sin^2 \alpha + a_{21}(x) \cos^2 \alpha + (a_{22}(x) - a_{11}(x)) \sin \alpha \cos \alpha$$

of the angle motion; since $\bar{h}(\mathbf{A}(x), .)$ is π-periodic, we can restrict our investigations on an interval of that length, e.g. $(-\frac{\pi}{2}, \frac{\pi}{2}]$; then we get the following switching curves of $\bar{h}$:

$$\mathcal{A}_i(x) := \mathcal{A}_i(\mathbf{A}(x)) := \arctan \xi_i(\mathbf{A}(x)),$$

if there are $i \in \{1, 2\}$ many zeros $\xi_i(\mathbf{A}(x))$ of $F(\mathbf{A}(x), \xi)$.

In the last case of infinitely many switching curves any α is a zero of $\bar{h}(\mathbf{A}(x), .)$.

It follows from the above that $a_{12}(x) = 0$ implies that $\frac{\pi}{2}$ is a switching curve; furthermore, if $a_{21}(x) = 0$, then 0 is a switching curve.

In case that there are exactly two switching curves $\mathcal{A}_{1,2}(\mathbf{A}(x))$, the previous notation is redefined such that $\mathcal{A}_1(\mathbf{A}(x))$ is *attracting* and $\mathcal{A}_2(\mathbf{A}(x))$ is *repelling*, in the sense that

$$\alpha \mapsto \bar{h}(\mathbf{A}(x), \alpha)$$

changes its sign from + to − at $\mathcal{A}_1(\mathbf{A}(x))$ and changes from − to + at $\mathcal{A}_2(\mathbf{A}(x))$. □

Remark 1.2.5. The above considerations have been accomplished in the real noise situation of the differential equation (1). Actually, this technique of projecting the linear system onto the sphere using spherical and angular coordinates can also be transferred to the white noise case; the Itô-formula then provides the SDEs for the radial, spherical and angular component, respectively. We do not give details here, since this goes beyond the framework of this book, but refer to the literature instead: See Khasminskii [Kh 67] and [Kh 80, Sec.VI.7-VI.9], Nishioka [Nk 76], Böhme [Bm 80], Auslender and Milshtein [Al-Mi 82], Arnold et al. [Ar-Oe-Pd 86], Arnold and Kliemann [Ar-Kl 87a], Pinsky and Wihstutz [Pi-Wh 88], Pardoux and Wihstutz [Pd-Wh 88] and [Pd-Wh 92] and Imkeller and Lederer [Im-Ld 99] and [Im-Ld 01].

1.3 The Multiplicative Ergodic Theorem: Lyapunov exponents

In this section the "classical" Lyapunov exponents are to be discussed as the result of Oseledets' [Os 68] Multiplicative Ergodic Theorem. Since this result and its arguments are not to be used in the sequel, we refer to Arnold [Ar 98] for details.

Here, the solution of the RDE (1.1) is modeled as random dynamical system (RDS) over the canonical metric dynamical system. More precisely, let

$$\Omega := C(\mathbb{R}_+, \mathbb{R}^d)$$

denote the path space of X^ε,

$$dX_t^\varepsilon = b\,(X_t^\varepsilon)\;dt + \sqrt{\varepsilon}\;\sigma\,(X_t^\varepsilon)\;dW_t \qquad (t \geq 0)\,,$$

where $\varepsilon > 0$. Ω is fixed as canonical probability space in this section. If endowed with the topology of uniform convergence on compacts, Ω is a Polish space which can then be equipped with its Borel-σ-algebra; the latter is the trace σ-algebra of $\mathcal{B}(\mathbb{R}^d)^{\mathbb{R}_+}$ in Ω, i.e.

$$\mathcal{F} := \mathcal{B}(\Omega) = \Omega \cap \mathcal{B}(\mathbb{R}^d)^{\mathbb{R}_+} \;;$$

see Arnold [Ar 98, p.544f.] and Hackenbroch and Thalmaier [Hb-Th 94, 1.24]. For all $t \in \mathbb{R}_+$, there is the canonical *shift transformation* on Ω defined by

$$\begin{aligned} \theta(t) \,:\, \Omega \to \Omega\,, \\ \omega \mapsto \theta(t)\omega \,:=\, \omega(t+\bullet\,) \,\equiv\, \big(s \mapsto \omega(t+s)\big)\,. \end{aligned}$$

Here, the mapping $(t,\omega) \mapsto \theta(t)\omega$ is continuous, hence measurable with respect to the underlying Borel-σ-algebras.

Let $(P^\varepsilon_t)_t$ denote the Markov transition probabilities of the Markov process X^ε. Furthermore, suppose that there exists a unique *stationary probability distribution for* X^ε, i.e. a unique probability measure[2] ρ^ε on $\mathbb{R}^d$ which is *invariant* with respect to the Markov transition probabilities $(P^\varepsilon_t)_t$ in the sense that

$$\rho^\varepsilon(\,.\,) \,=\, \int_{\mathbb{R}^d} P^\varepsilon_t(x,\,.\,)\,\rho^\varepsilon(dx) \qquad (t \geq 0)\,;$$

conditions assuring existence and uniqueness of such a measure ρ^ε will be given in section 2.2 where the SDE of X^ε is investigated in detail; see p.62f. Then there is a unique probability $\mathbb{P}_{\rho^\varepsilon}$ on $\big((\mathbb{R}^d)^{\mathbb{R}_+}, \mathcal{B}(\mathbb{R}^d)^{\mathbb{R}_+}\big)$ such that the coordinate process is a (time-homogeneous) Markov process with transition semigroup $(P^\varepsilon_t)_t$ and initial distribution ρ^ε; see e.g. Arnold [Ar 98, p.548] and Hackenbroch and Thalmaier [Hb-Th 94, 2.5]. Since the outer measure of Ω is full, $\mathbb{P}^*_{\rho^\varepsilon}(\Omega) = 1$, due to the continuity of the paths of X^ε, $\mathbb{P}_{\rho^\varepsilon}$ also induces a probability distribution on $\mathcal{F}$; see e.g. Hackenbroch and Thalmaier [Hb-Th 94, p.43f.]. It will be denoted by the same symbol $\mathbb{P}_{\rho^\varepsilon}$. Let $X^{\varepsilon,\rho^\varepsilon}$ denote the process with distribution $\mathbb{P}_{\rho^\varepsilon}$ on the path space Ω, i.e. the system X^ε with initial distribution ρ^ε.

Note that the above stationarity of the measure ρ^ε (stationarity of the canonical Markov process), i.e. the $(P^\varepsilon_t)_t$-invariance of ρ^ε, is equivalent to $\mathbb{P}_{\rho^\varepsilon}$ being $(\theta(t))_t$*-invariant* in the sense that

$$\mathbb{P}_{\rho^\varepsilon} \circ \theta(t)^{-1} \,=\, \mathbb{P}_{\rho^\varepsilon} \qquad \text{for all } t \in \mathbb{R}_+\,;$$

this fact is also expressed by stating that all shifts $\theta(t)$ *preserve* the measure $\mathbb{P}_{\rho^\varepsilon}$; see Arnold [Ar 98, p.545,549].

Altogether, the above system $(\Omega, \mathcal{F}, \mathbb{P}_{\rho^\varepsilon}, (\theta(t)_{t\in\mathbb{R}_+})$ satisfies

1) $(\omega,t) \mapsto \theta(t)\omega$ is $\mathcal{F}\otimes\mathcal{B}(\mathbb{R}_+)/\mathcal{F}$-measurable,

2) $\theta(0) = \mathrm{id}_\Omega$ and $\theta(s+t) = \theta(s)\circ\theta(t)$ for all $s,t \in \mathbb{R}_+$,

3) $\mathbb{P}_{\rho^\varepsilon} \circ \theta(t)^{-1} = \mathbb{P}_{\rho^\varepsilon}$ for all $t \in \mathbb{R}_+$,

which are the characterizing properties for calling $\big(\Omega, \mathcal{F}, \mathbb{P}_{\rho^\varepsilon}, (\theta(t)_{t\in\mathbb{R}_+}\big)$ a *metric (measure preserving) dynamical system*; see Arnold [Ar 98, p.536f.].

[2] This measure ρ^ε is not to be confused with the radial component process ϱ^ε_t of the previous section.

Furthermore, this metric dynamical system is *ergodic* meaning that

$$\mathbb{P}_{\rho^\varepsilon}(A) \;\in\; \{0,1\}$$

on all sets $A \in \mathcal{F}$ which are $(\theta(t))_{t\in\mathbb{R}_+}$*-invariant*, i.e.

$$\theta(t)^{-1}\, A \;=\; A \quad \text{for all } t \in \mathbb{R}_+ \,.$$

This is due to the fact that the measures which are ergodic in this sense are the extreme points of the convex set of $(\theta(t))_t$-invariant measures; however, due to the postulated uniqueness of ρ^ε, there is only one such measure which is $\mathbb{P}_{\rho^\varepsilon}$; see Arnold [Ar 98, p.539].

Let $X^{\varepsilon,\rho^\varepsilon}$ denote the stationary Markov process which realizes $\mathbb{P}_{\rho^\varepsilon}$ as defined above. This process is now specified as stochastic input for the differential system (1); more precisely, consider the RDE

$$dZ_t^\varepsilon \;=\; \mathbf{A}\big(\, X_t^{\varepsilon,\rho^\varepsilon}(\omega)\big)\; Z_t^\varepsilon \;\; dt \,, \qquad Z_0^\varepsilon \;=\; z \;\in\; \mathbb{K}^n$$

and let

$$\begin{aligned}
\Phi^\varepsilon : \mathbb{R}_+ \times \Omega \times \mathbb{K}^n &\longrightarrow \mathbb{K}^n \,, \\
(t,\omega,z) &\longmapsto \Phi^\varepsilon(t,\omega,z) \;\equiv\; \Phi^\varepsilon(t,\omega)z \\
&\quad = z \;+\; \int_0^t \mathbf{A}\big(\, X_u^{\varepsilon,\rho^\varepsilon}(\omega)\big)\; \Phi^\varepsilon(u,\omega,z)\, du
\end{aligned}$$

denote its unique (up to indistinguishability) solution. Then the following holds true:

1) Φ^ε is $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F} \otimes \mathcal{B}(\mathbb{K}^n)/\mathcal{B}(\mathbb{K}^n)$-measurable,
2) Φ^ε is a *cocycle over* $(\theta(t))_{t\in\mathbb{R}_+}$ in the sense that for all $\omega \in \Omega$ and $s,t \in \mathbb{R}_+$:

$$\Phi^\varepsilon(0,\omega) = \mathrm{id}_{\mathbb{K}^n} \qquad \text{and} \qquad \Phi^\varepsilon(t+s,\omega) = \Phi^\varepsilon(t,\theta(s)\omega) \circ \Phi^\varepsilon(s,\omega)\,,$$

3) $\Phi^\varepsilon(\,.\,,\omega,\,.\,)$ is continuous for any $\omega \in \Omega$,
4) $\Phi^\varepsilon(t,\omega) \equiv \Phi^\varepsilon(t,\omega,\,.\,)$ is linear for any $t \in \mathbb{R}_+$ and $\omega \in \Omega$;

see Arnold [Ar 98, 2.2.12]. These properties are summarized by calling Φ^ε a *linear random dynamical system (RDS)* on $\mathbb{K}^n$ over the metric dynamical system $\big(\Omega, \mathcal{F}, \mathbb{P}_{\rho^\varepsilon}, (\theta(t)_{t\in\mathbb{R}_+}\big)$ with time $\mathbb{R}_+$; see Arnold [Ar 98, p.5f.]. In particular, since Φ^ε is defined by the characterizing RDE (1), it will be called the RDS *generated* by (1).

Being the solution of a linear differential equation the linear operator $\Phi^\varepsilon(t,\omega)$ is even invertible, i.e. $\Phi^\varepsilon(t,\omega) \in GL(n,\mathbb{K})$ for all $t \in \mathbb{R}_+$ and $\omega \in \Omega$.

The pathwise Jacobi (Liouville) equation (1.2) now reads

$$\det(\Phi^\varepsilon(t,\omega)) = \exp\left(\int_0^t \mathrm{trace}\left(\mathbf{A}\big(X_u^{\varepsilon,\rho^\varepsilon}(\omega)\big)\, du\right)\right) \quad . \tag{1.12}$$

The adjoint matrix of $\Phi^\varepsilon(t,\omega)$ will be denoted by $\Phi^\varepsilon(t,\omega)^*$ in the sequel. Then the Multiplicative Ergodic Theorem states the following in the above setting:

Theorem 1.3.1 (Multiplicative Ergodic Theorem). *Suppose that X^ε has a unique stationary distribution ρ^ε and let Φ^ε be the linear, invertible RDS generated by (1), where $\varepsilon > 0$. Assume that $\beta^+ \in L^1(\mathbb{P}_{\rho^\varepsilon})$ and $\beta^- \in L^1(\mathbb{P}_{\rho^\varepsilon})$ which are defined by*

$$\beta^\pm(\omega) := \sup_{0\le t\le 1} \log^+ \left\| \Phi^\varepsilon(t,\omega)^{\pm 1} \right\| ,$$

where $\log^+ a := \max(0,\log a)$ and $\|\,.\,\|$ denotes the operator norm. Then there exists a $(\theta(t))_{t\in\mathbb{R}_+}$-invariant set $\widetilde{\Omega}^\varepsilon \in \mathcal{F}$ of full measure, $\mathbb{P}_{\rho^\varepsilon}(\widetilde{\Omega}^\varepsilon) = 1$, such that the following holds:

$$\lim_{t\to\infty} \left(\Phi^\varepsilon(t,\omega)^*\, \Phi^\varepsilon(t,\omega)\right)^{1/2t} =: \Psi^\varepsilon(\omega)$$

exists for any $\omega \in \widetilde{\Omega}^\varepsilon$ and is non-negative definite, its eigenvalues being given by n values

$$e^{\Lambda_1^\varepsilon} \ge e^{\Lambda_2^\varepsilon} \ge \cdots \ge e^{\Lambda_n^\varepsilon} > 0\,,$$

where $\Lambda_1^\varepsilon \ge \Lambda_2^\varepsilon \ge \cdots \ge \Lambda_n^\varepsilon > -\infty$ do not depend on $\omega \in \widetilde{\Omega}^\varepsilon$; writing the distinct numbers in this list of eigenvalues of $\Psi^\varepsilon(\omega)$ as

$$e^{\lambda_1^\varepsilon} > e^{\lambda_2^\varepsilon} > \cdots > e^{\lambda_{p^\varepsilon}^\varepsilon} > 0$$

and defining d_i^ε as the counting multiplicity of λ_i^ε in the list $\{\Lambda_1^\varepsilon,\ldots,\Lambda_n^\varepsilon\}$, it follows that

$$d_i^\varepsilon = \dim U_i^\varepsilon(\omega) \qquad \textit{for all} \quad \omega\in\widetilde{\Omega}^\varepsilon \quad \textit{and} \quad i\in\{1,\ldots,p^\varepsilon\}\,,$$

where $U_i^\varepsilon(\omega)$ denotes the eigenspace of $\Psi^\varepsilon(\omega)$ corresponding to $e^{\lambda_i^\varepsilon}$. Further define

$$V_i^\varepsilon(\omega) := \begin{cases} U_p^\varepsilon(\omega)\oplus\cdots\oplus U_i^\varepsilon(\omega) & ,i\in\{1,\ldots,p^\varepsilon\} \\ \{0\} & ,i = p^\varepsilon+1 \end{cases}$$

for any $\omega \in \widetilde{\Omega}^\varepsilon$, entailing the flag

$$\{0\} \equiv V_{p+1}^\varepsilon(\omega) \subset V_p^\varepsilon(\omega) \subset \cdots \subset V_1^\varepsilon(\omega) = \mathbb{K}^n \quad .$$

Then the Lyapunov exponent

$$\lambda^\varepsilon(\omega,z) := \lim_{t\to\infty} \frac{1}{t} \log|\,\Phi^\varepsilon(t,\omega) z\,|$$

exists for all $\omega \in \widetilde{\Omega}^\varepsilon$ *and* $z \in \mathbb{K}^n \setminus \{0\}$ *and it holds that*

$$z \in V_i^\varepsilon(\omega) \setminus V_{i+1}^\varepsilon(\omega) \iff \lambda^\varepsilon(\omega, z) = \lambda_i^\varepsilon \ .$$

The latter fact is equivalent to the following characterization,

$$V_i^\varepsilon(\omega) \ = \ \left\{ z \in \mathbb{K}^n \ : \ \lambda^\varepsilon(\omega, z) \leq \lambda_i^\varepsilon \right\} \ .$$

For all $\omega \in \tilde{\Omega}^\varepsilon$ *and* $t \in \mathbb{R}_+$ *it holds that*

$$\lambda^\varepsilon(\theta(t)\omega, \Phi^\varepsilon(t,\omega)z) = \lambda^\varepsilon(\omega, z) \qquad\qquad (z \in \mathbb{K}^n \setminus \{0\})$$

and hence

$$\Phi^\varepsilon(t,\omega)\, V_i^\varepsilon(\omega) = V_i^\varepsilon(\theta(t)\omega) \qquad\qquad (i \in \{1, \ldots, p^\varepsilon\}) \ .$$

The Multiplicative Ergodic Theorem has initially been proved by Oseledets [Os 68]. The one-sided version considered above is taken from Arnold [Ar 98, 3.4.1]. The fact that the proof there is presented for the case $\mathbb{K} = \mathbb{R}$ is no restriction: According to Arnold [Ar 98, 3.4.10.(ii)], all arguments hold true for the complex case, $\mathbb{K} = \mathbb{C}$, as can be also read off from Ruelle's [Ru 82] generalization to real or complex Hilbert spaces.

The numbers $\Lambda_1^\varepsilon, \Lambda_2^\varepsilon, \ldots, \Lambda_n^\varepsilon$ are the possible exponential growth rates of Φ^ε due to the above theorem and are called the *Lyapunov exponents of* Φ^ε *(under* ρ^ε*).*

Furthermore the proof of the Multiplicative Ergodic Theorem (via the Furstenberg-Kesten Theorem) and the above version (1.12) of the pathwise Jacobi (Liouville) equation yield together with the ergodic theorem that

$$\begin{aligned}\sum_{i=1}^{n} \Lambda_i^\varepsilon \equiv \sum_{i=1}^{p^\varepsilon} d_i^\varepsilon \, \lambda_i^\varepsilon & \\ &= \lim_{t\to\infty} \frac{1}{t} \log |\det \Phi^\varepsilon(t, \,.\,)| \qquad\qquad \mathbb{P}_{\rho^\varepsilon}\text{-a.s.} \\ &= \lim_{t\to\infty} \frac{1}{t} \left(\int_0^t \operatorname{trace}\Big(\mathbf{A}\big(X_u^{\varepsilon,\rho^\varepsilon}(\,.\,)\big)\, du\Big)\right) \\ &\equiv \int_{\mathbb{R}^d} \operatorname{trace}\big(\mathbf{A}(x)\big)\, \rho^\varepsilon(dx) \qquad\qquad \mathbb{P}_{\rho^\varepsilon}\text{-a.s.} \ ,\end{aligned} \tag{1.13}$$

provided that $\|\mathbf{A}\| \in L^1(\rho^\varepsilon)$. This formula is called the *trace formula* for the sum of the Lyapunov exponents; see Arnold [Ar 98, 3.4.15, 3.3.11 & 3.3.4] and Oseledets [Os 68, p.203].

Further note that the previous condition

$$\|\mathbf{A}\| \in L^1(\rho^\varepsilon) \tag{1.14}$$

is sufficient for the integrability assumption $\beta^{\pm} \in L^1(\mathbb{P}_{\rho^\varepsilon})$ of the Multiplicative Ergodic Theorem, since by the defining RDEs for Φ^ε and $(\Phi^\varepsilon)^{-1}$,

$$\| \Phi^\varepsilon(t,\omega)^{\pm 1} \| \le 1 + \int_0^t \left\| \mathbf{A}\left(X_u^{\varepsilon,\rho^\varepsilon}(\omega) \right) \right\| \, \| \Phi^\varepsilon(u,\omega)^{\pm 1} \| \, du$$

and hence by the general version of the Gronwall lemma (see e.g. Arnold [Ar 98, p.557]),

$$\| \Phi^\varepsilon(t,\omega)^{\pm 1} \| \le \exp\left(\int_0^t \left\| \mathbf{A}\left(X_u^{\varepsilon,\rho^\varepsilon}(\omega) \right) \right\| du \right)$$

which implies that

$$\begin{aligned}
\int_\Omega \left| \beta^{\pm}(\omega) \right| \mathbb{P}_{\rho^\varepsilon}(d\omega) &\equiv \int_\Omega \sup_{0\le t\le 1} \log^+ \| \Phi^\varepsilon(t,\omega)^{\pm 1} \| \;\; \mathbb{P}_{\rho^\varepsilon}(d\omega) \\
&\le \int_\Omega \sup_{0\le t\le 1} \log^+ \exp\left(\int_0^t \left\| \mathbf{A}\left(X_u^{\varepsilon,\rho^\varepsilon}(\omega)\right) \right\| du \right) \mathbb{P}_{\rho^\varepsilon}(d\omega) \\
&= \int_\Omega \int_0^1 \left\| \mathbf{A}\left(X_u^{\varepsilon,\rho^\varepsilon}(\omega) \right) \right\| du \;\; \mathbb{P}_{\rho^\varepsilon}(d\omega) \\
&= \int_0^1 \int_\Omega \left\| \mathbf{A}\left(X_u^{\varepsilon,\rho^\varepsilon}(\omega) \right) \right\| \mathbb{P}_{\rho^\varepsilon}(d\omega) \;\; du \\
&= \int_0^1 \int_{\mathbb{R}^d} \| \mathbf{A}(x) \| \left(\mathbb{P}_{\rho^\varepsilon} \circ \left(X_u^{\varepsilon,\rho^\varepsilon} \right)^{-1} \right)(dx) \;\; du \\
&= \int_{\mathbb{R}^d} \| \mathbf{A}(x) \| \;\; \rho^\varepsilon(dx) \\
&< \infty \;\; ;
\end{aligned}$$

also see Arnold [Ar 98, 3.4.15 & 4.2.10].

At the end of this section we would like to comment on further research concerning Lyapunov exponents, notably the so-called Furstenberg-Khasminskii formula and the associated law of large numbers which are due to Arnold et al. [Ar-Kl-Oe 86, Th.4.1]. This remark is in particular intended to underline the basic difference between their setting and the framework of our work: While Arnold et al. [Ar-Kl-Oe 86] make sure that the generator $\bar{\mathcal{L}}^\varepsilon \equiv \mathcal{G}^\varepsilon + \bar{h}\frac{\partial}{\partial\alpha}$ of $(X^\varepsilon, \alpha^\varepsilon)$ is hypoelliptic (where the notation is as in the previous section), it will be proposed here that the operator $\mathcal{L}^\varepsilon + \frac{\partial}{\partial t}$ is hypoelliptic: By assuming that $\bar{h}$ is "strongly hypoelliptic" on the sets of interest as will be made precise in definition 4.4.4, it will be made sure that $(X^\varepsilon, \alpha^\varepsilon)$ can "essentially" be replaced by a process called $(\hat{X}^\varepsilon, \hat{\alpha}^\varepsilon)$ such that its corresponding operator $\hat{\mathcal{L}}^\varepsilon + \frac{\partial}{\partial t}$ is hypoelliptic according to 3.1.3 (c); see the proofs of theorems 4.4.6 and 4.4.7. However, note that the hypoellipticity of $\mathcal{L}^\varepsilon + \frac{\partial}{\partial t}$ is stronger than hypoellipticity of $\mathcal{L}^\varepsilon$; see Arnold et al. [Ar-Kl-Oe 86, p.104]. The reason is basically that the hypoellipticity of $\mathcal{L}^\varepsilon$ is related to the existence of a C^∞ density of the invariant measure (*limiting distribution*), while the hypoellipticity of $\mathcal{L}^\varepsilon + \frac{\partial}{\partial t}$, meaning "ellipticity" in the sense of Ichihara and Kunita [Ic-Ku 74], assured that already the finite time transition probabilities have C^∞ densities; also see Arnold [Ar 84, p.794].

Remark 1.3.2. Again assume that (1.14) is satisfied and suppose that $\mathbb{K} = \mathbb{R}$, $\mathbf{A}$ is analytic and that $\sigma(\,.\,) = \mathrm{id}_{\mathbb{R}^d}$ for simplicity in addition to the previous assumptions. Using the spherical and angular coordinates as defined in the previous section for Φ^ε, one gets the processes

$$\psi^\varepsilon(t,\omega,\psi) \;\equiv\; \frac{\Phi^\varepsilon(t,\omega,z)}{|\,\Phi^\varepsilon(t,\omega,z)\,|} \qquad \text{starting in} \qquad \psi^\varepsilon(0,\omega,\psi) \;=\; \psi \;:=\; \frac{z}{|\,z\,|}$$

and in dimension $n = 2$,

$$\alpha^\varepsilon(t,\omega,\alpha) \;\doteq\; \psi^\varepsilon(t,\omega,\psi) \qquad \text{starting in} \qquad \alpha \;\doteq\; \psi\;,$$

respectively. These systems can be regarded as random dynamical systems over $(\theta(t))_{t\in\mathbb{R}_+}$ in their own right; see Arnold [Ar 98, 6.2.1].

Now suppose that the operator $\mathcal{L}^\varepsilon$ is *hypoelliptic*, i.e. that for any distribution $v \equiv v(x)$ on $\mathbb{R}^d$, it follows that v is a C^∞ function in every open subset of $\mathbb{R}^d$, where $\mathcal{L}^\varepsilon v$ is a C^∞ function. Hörmander [Hö 67, p.149ff.] has shown that this is equivalent to proposing that

$$\dim\;\mathrm{LA}\left\{ \begin{pmatrix} b(x) \\ h(x,\psi) \end{pmatrix}, \sqrt{\varepsilon}\begin{pmatrix} e_1 \\ 0 \end{pmatrix}, \ldots, \sqrt{\varepsilon}\begin{pmatrix} e_d \\ 0 \end{pmatrix} \right\}(x,\psi) \;=\; d+n-1$$

for all $(x,\psi) \in \mathbb{R}^d \times P^{n-1}$, where $\mathrm{LA}\{\mathcal{V}\}$ denotes the Lie-algebra generated by the set $\{\mathcal{V}\}$ of vector fields, e_i denotes the i-th canonical basis vector (column vector) of $\mathbb{R}^d$ and where the simplifying assumption that $\sigma = \mathrm{id}_{\mathbb{R}^d}$ such that

$$\mathcal{L}^\varepsilon \;=\; \sum_{i=1}^{d} b_i(x)\,\frac{\partial}{\partial x_i} \;+\; h(x,\psi)\,\frac{\partial}{\partial \psi} \;+\; \frac{\varepsilon}{2}\,\Delta$$

has also been used; further see Arnold [Ar 84, p.794]. Due to Arnold et al. [Ar-Kl-Oe 86, Prp.2.2 & Rem.2.2], this is equivalent to demanding that

$$\dim\;\mathrm{LA}\left\{\, h(x,\,.\,)\;:\; x\in\mathbb{R}^d \,\right\}(\psi) \;=\; n-1 \qquad \text{for all } \psi \in P^{n-1}\;. \tag{1.15}$$

If $n = 2$, this hypoellipticity condition (1.15) is equivalent to assuming that for any $\psi \in P^1$, there is $x \in \mathbb{R}^d$ such that $h(x,\psi)$ does not vanish; see Arnold et al. [Ar-Kl-Oe 86, Prp.2.1] and Pardoux and Wihstutz [Pd-Wh 88, p.444] and [Pd-Wh 92, p.291].

As Arnold et al. [Ar-Kl-Oe 86, Cor.3.1] further show under the above hypoellipticity assumption (1.15), the Markov process $(X^\varepsilon, \psi^\varepsilon)$ has a unique stationary (invariant with respect to Markov transition probabilities) probability distribution μ^ε on $\mathbb{R}^d \times P^{n-1}$ whose $\mathbb{R}^d$-marginal is ρ^ε; μ^ε has support $\mathbb{R}^d \times C$, where C is an "invariant control set" of a certain associated control problem; we do not go further into details but refer to Arnold et al. [Ar-Kl-Oe 86, p.88-95] instead. The stationary measure μ^ε has a density m^ε which satisfies the forward equation (Fokker-Planck equation) $(\mathcal{L}^\varepsilon)^* \, m^\varepsilon = 0$, where $(\mathcal{L}^\varepsilon)^*$ denotes the formal adjoint operator of $\mathcal{L}^\varepsilon$. Using the formula (1.5) and the ergodic theorem one gets the following *Furstenberg-Khasminskii formula*

$$\lambda^\varepsilon \; := \; \int_{\mathbb{R}^d \times P^{n-1}} Q(x,\psi) \quad \mu^\varepsilon(dx, d\psi)$$

for the Lyapunov exponent (exponential growth rate) of the unique stationary process $(X^\varepsilon, \psi^\varepsilon)$. Furthermore, λ^ε coincides with the top Lyapunov exponent from the Multiplicative Ergodic Theorem,

$$\lambda^\varepsilon \; = \; \lambda_1^\varepsilon \; \equiv \; \Lambda_1^\varepsilon \; ,$$

and the following *law of large numbers* holds true: For all $z \in \mathbb{R}^n \setminus \{0\}$,

$$\lambda^\varepsilon(\omega, z) \; \equiv \; \lim_{t\to\infty} \frac{1}{t} \log |\, \Phi^\varepsilon(t,\omega) z \,| \; = \; \lambda^\varepsilon \qquad \text{for } \mathbb{P}_{\rho^\varepsilon}\text{-almost all } \omega \, . \tag{1.16}$$

For the proofs of these results see Arnold et al. [Ar-Kl-Oe 86, Th.4.1].

For $n = 2$, the above assertions can also be stated in terms of the angle process $\alpha^\varepsilon(t,\omega,\alpha)$: The hypoellipticity condition (1.15) is met if and only if for any $\alpha \in [0,\pi)$ at least one vector $\bar{h}(x,\alpha)$ does not vanish. In this case the above reasoning implies that $(X^\varepsilon, \alpha^\varepsilon)$ has a unique stationary measure $\bar{\mu}^\varepsilon$ with a C^∞ density $\bar{m}^\varepsilon$ satisfying the Fokker-Planck-equation $(\bar{\mathcal{L}}^\varepsilon)^* \, \bar{m}^\varepsilon \; = \; 0$, as well as the periodicity constraint

$$\bar{m}^\varepsilon(x,0) \; = \; \bar{m}^\varepsilon(x,\pi) \qquad\qquad \text{for all } x \in \mathbb{R}^d \, .$$

Unfortunately no explicit formula for $\bar{m}^\varepsilon$ is known to us. Its $\mathbb{R}^d$-marginal is ρ^ε and the *Furstenberg-Khasminskii formula* for the almost surely observed top Lyapunov exponent now reads

$$\lambda^\varepsilon = \lim_{t\to\infty} \frac{1}{t} \int_0^t \bar{Q}(X_u^\varepsilon, \alpha_u^\varepsilon) du \; = \int_{\mathbb{R}^d \times [0,\pi)} \bar{Q}(x,\alpha) \;\; \bar{m}^\varepsilon(x,\alpha) \;\; dx d\alpha \; , \tag{1.17}$$

where $\bar{Q}$ had been defined in the previous section. □

Remark 1.3.3. Our goal in this paper is to obtain local Lyapunov characteristic numbers, i.e. growth numbers for Z_t^ε for certain time scales $t \le T(\varepsilon)$. For this aim neither the MET nor the Furstenberg-Khasminskii formula are applicable, since these theorems cover the case that the time t tends to ∞ for fixed ε and hence the respective invariant measures ρ^ε and μ^ε arise which precisely describe this asymptotic behavior. However, our rationale here is *not* to observe the asymptotic behavior as $t \to \infty$, but "before", i.e. on shorter time scales $T(\varepsilon)$. Therefore we need to use an argument which reproduces the "limit-matrices" $\mathbf{A}(K_i)$ on the time scales on which the "metastable" state K_i is observed; see definition 2.5.4 for the precise explanation of a metastable state corresponding to a time scale.

For that purpose we next recall, what happens in the deterministic case ($\varepsilon = 0$)

$$\frac{dz_t}{dt} = \mathbf{A}(t)\, z_t \ ,$$

if $\mathbf{A}(t)$ has a limit matrix $\mathbf{A}$, i.e. $\mathbf{A}(t) \xrightarrow{t\to\infty} \mathbf{A}$. The easiest way would be to obtain a closed form for the dynamical system z_t and to manipulate this explicit expression; e.g. $z_t = \mathrm{e}^{\mathbf{A}t} z_0$ for the constant case $\mathbf{A}(t) \equiv \mathbf{A}$. However, this does not work, since even in the general deterministic case, if $\mathbf{A}(\,.\,)$ is non-constant, the formula for the propagator, involving the time-order-operator, seems too bulky for further manipulations. Hence we investigate the dynamics of the absolute value $\varrho_t \equiv |z_t|$ and relate it to the Jordan decomposition of the limit matrix.

1.4 The deterministic case: Lyapunov exponents for asymptotically constant linear systems

This section is dedicated to discussing the deterministic case $\varepsilon = 0$ in equation (1),

$$\begin{aligned} dZ_t^0 &= \mathbf{A}\left(X_t^0\right)\, Z_t^0 \, dt \\ dX_t^0 &= b\left(X_t^0\right)\, dt \ . \end{aligned}$$

In the situations we are interested in, the drift vector field b confines the deterministic system $X^{0,x}$ to converge to a certain attracting point K_i,

$$X_t^{0,x} \xrightarrow{t\to\infty} K_i \ ,$$

where $i \in \{1,\dots,l\}$ denotes an index determined by the initial value x; see assumption 2.1.1**(K)**. Since $\mathbf{A}(\,.\,)$ is assumed to be continuous in (1), it follows that

$$\mathbf{A}(t) \ := \ \mathbf{A}(X_t^{0,x}) \xrightarrow{t\to\infty} \mathbf{A}(K_i) \ ;$$

in other words, the (deterministic) coefficient matrix $\mathbf{A}(t)$ is *asymptotically constant.*

In the sequel, we will consider general asymptotically constant, linear, deterministic differential systems $\dot{z}_t = \mathbf{A}(t)\, z_t$ which we write as

$$dz_t \ = \ \mathbf{A}(t)\, z_t \, dt \tag{1.18}$$

again, where $t \mapsto \mathbf{A}(t)$ is a continuous function taking its values in $\mathbb{K}^{n\times n}$, the $n \times n$ matrices with real or complex entries and $z_t \in \mathbb{K}^n$. Assuming that the system (1.18) is *asymptotically constant* as defined above, i.e. that there is a fixed matrix $\mathbf{A}$ such that

$$\mathbf{A}(t) \xrightarrow{t\to\infty} \mathbf{A}$$

(where $\mathbb{K}^{n\times n}$ is equipped with the operator norm $\| \, . \, \|$), one can rewrite (1.18) as

$$dz_t \; = \; [\, \mathbf{A} + \mathbf{G}(t) \,] \, z_t \, dt \, , \tag{1.19}$$

where of course

$$\mathbf{G}(t) \; := \; \mathbf{A}(t) - \mathbf{A} \xrightarrow{t\to\infty} 0 \, .$$

In this deterministic case (1.19), the statements to be discussed in the sequel are due to Hartman and Wintner [Ha-Wi 55] and Perron [Pe 29]; the reference underlying the exposition here is Coppel [Cp 65, Ch.IV].

As one expects from the well-known case of linear deterministic systems with constant coefficients,

$$\dot{z}_t = \mathbf{A} \, z_t \quad ,$$

the Lyapunov exponents will turn out to be the real parts of the eigenvalues of $\mathbf{A}$ and the proof uses the Jordan decomposition of $\mathbf{A}$. However, a closed form of $\mathrm{e}^{\mathbf{A}t}$ only exists in the case of constant coefficients, $\mathbf{G}(t) \equiv 0$. The theory of *asymptotic integration* by Hartman and Wintner then treats the general situation.

This will then motivate our general rationale in the stochastic case for the system (1): The "local" Lyapunov exponents (if existing as stochastic limits) should be obtained as the real parts of the eigenvalues of a certain "sublimit" matrix $\mathbf{A}_{\mu(x,\zeta)}$, where the deviation $\mathbf{G}(\,.\,)$ from $\mathbf{A}_{\mu(x,\zeta)}$ gets small in a stochastic sense; see proposition 4.1.1.

The deterministic considerations here shall be started by recalling some facts concerning the simplest possible case of a linear system with constant coefficients, $\dot{z}_t = \mathbf{A} \, z_t$.

Remark 1.4.1 (Decomposition of the state space). For the matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ (or $\mathbb{C}^{n\times n}$) and a given $\lambda \in \mathbb{R}$, the state space $\mathbb{K}^n$ ($\mathbb{R}^n$ or $\mathbb{C}^n$), on which the linear transformation $\mathbf{A}$ acts, can be uniquely decomposed into a direct sum

$$\mathbb{K}^n \; = \; E_{<\lambda} \oplus E_{\lambda} \oplus E_{>\lambda}$$

such that the three subspaces $E_{<\lambda}$, E_{λ} and $E_{>\lambda}$ are invariant under $\mathbf{A}$ and the eigenvalues (i.e. the complex roots of the respective characteristic polynomial) of $\mathbf{A}$ restricted to these spaces have real parts less than, equal to or greater than λ, respectively. Namely, let $D_{<\lambda}$, D_{λ} and $D_{>\lambda}$ denote open domains in $\mathbb{C}$ including precisely the respective eigenvalues; if the respective domain is nonempty, then we assume that its boundary is rectifiable; the subspaces

$E_{<\lambda}$, E_λ and $E_{>\lambda}$ are then given as the images of the corresponding *Riesz projections* $P_{<\lambda}$, P_λ and $P_{>\lambda}$ defined by

$$P_\nu \;:=\; \frac{1}{2\pi \mathrm{i}} \int_{\partial D_\nu} \left(\chi\, I_n - \mathbf{A} \right)^{-1} d\chi\,, \qquad\qquad \nu \in \left\{ <\!\lambda\,;\, \lambda\,;\, >\!\lambda \right\},$$

where I_n denotes the $n \times n$ unit matrix; P_ν is defined to be the zero projection, if D_ν is empty.

In particular, for the linear ordinary differential equation with constant coefficients $\dot{z}_t = \mathbf{A}\, z_t$, or

$$dz_t \;=\; \mathbf{A}\, z_t\, dt\,,$$

one recovers the (Oseledets-)splitting

$$\mathbb{K}^n \;=\; E_1 \oplus \cdots \oplus E_p$$

and the Lyapunov spectrum

$$\lambda_1 \;>\; \lambda_2 \;>\; \cdots \;>\; \lambda_p$$

consisting of the distinct real parts of eigenvalues of $\mathbf{A}$, where E_i is the sum of the generalized eigenspaces to eigenvalues with real part equal to λ_i and is dynamically characterized as

$$\lim_{t\to\pm\infty} \frac{1}{t} \log |\, z_t\,| \;=\; \lambda_i \quad\Longleftrightarrow\quad z_0 \;\in\; E_i \setminus \{0\}\,.$$

In this terminology, the above spaces $E_{<\lambda}$, E_λ and $E_{>\lambda}$ are given as the direct sums of the corresponding spaces $E_1, \ldots, E_p$.

The solution flow can also be described in the Riesz calculus as

$$\mathrm{e}^{t\mathbf{A}} \;=\; \frac{1}{2\pi \mathrm{i}} \int_{\partial D} \left(\chi\, I_n - \mathbf{A} \right)^{-1} \mathrm{e}^{t\chi}\; d\chi\,,$$

where the interior of the domain D now contains all eigenvalues of $\mathbf{A}$ and ∂D is supposed to be rectifiable.

For the above statements see e.g. Coppel [Cp 65, Sec. II.1 & III.2], Riesz and Szökefalvi-Nagy [Ri-Na 55, Ch.XI], Kato [Kt 80, §I.5] and Arnold [Ar 98, 3.2.3]. □

Now we discuss the result by Hartman and Wintner [Ha-Wi 55] and Perron [Pe 29]. As the above cited Oseledets-splitting and the corresponding Lyapunov spectrum for the constant coefficient differential equation $\frac{dz_t}{dt} = \mathbf{A}\, z_t$ is derived from the Jordan canonical form of the matrix $\mathbf{A}$, it does not come as a surprise that the same strategy is used for the perturbed equation (1.19). The following lemma will turn out to be crucial in this argumentation.

In the sequel let t_0 be an arbitrary initial time of the system and let $\|\,.\,\|$ denote the operator-norm. Furthermore, d/dt and the dot "$\cdot$" will be used interchangeably for the derivative with respect to the "time" variable t.

Lemma 1.4.2 (Riccati-type differential inequality). *Let* $v \in C^1([t_0,\infty),\mathbb{R})$ *satisfy the differential inequality*

$$\dot v(t) \;\geq\; \beta\big(v(t)\big) \;-\; \gamma(t) \qquad\quad (t > t_0)\,,$$

where $\gamma \in C^0([t_0,\infty),\mathbb{R}_+)$ *fulfills*

$$\int_t^{t+1} \gamma(u)\,du \;\xrightarrow{t\to\infty}\; 0$$

and β *is a real-valued function; let* $v^* \in \mathbb{R}$ *be a subsequential limit point of* v *and assume that* β *is continuous at* v^**. Then*

$$\beta(v^*) \;\leq\; 0\,.$$

In particular, if v *satisfies the* Riccati-type differential inequality

$$\dot v(t) \;\geq\; b\,v(t)\,\big(1 - v(t)\big) \;-\; \gamma(t) \qquad\quad (t > t_0)\,,$$

where $b > 0$ *is constant and* γ *is as above, then either*

$$\limsup_{t\to\infty}\, v(t) \;\leq\; 0 \qquad \textit{or} \qquad \liminf_{t\to\infty}\, v(t) \;\geq\; 1\,.$$

Proof. 1) In order to obtain a contradiction we assume that $\beta(v^*) > 0$. As v^* is a continuity point of β, this implies that one can choose constants $\eta_1, \eta_2 > 0$ such that

$$\beta(v) \;\geq\; \eta_1 \;>\; 0 \qquad \text{for all} \qquad v \in B_{\eta_2}(v^*)\,,$$

where the latter set $B_{\eta_2}(v^*)$ denotes the closed ball with center v^* and radius η_2.

Due to the convergence assumption on $\int_t^{t+1} \gamma(u)du$, one can fix a positive constant $\eta_3 < \min(\eta_1,\eta_2)$ and a time $t_1 \geq t_0$ such that

$$\int_t^{t+1} \gamma(u)\,du \;\leq\; \eta_3 \qquad\quad (t \geq t_1)\,;$$

in particular, this implies that

$$\int_t^{t+T} \gamma(u)\,du \;\leq\; (T+1)\,\eta_3 \qquad\quad (t \geq t_1, T > 0)\,.$$

Let $[\tau, \tau + T]$ denote some time interval such that $\tau \geq t_1$, $T > 0$ and

$$v(t) \in B_{\eta_2}(v^*) \qquad \text{for all} \qquad t \in [\tau, \tau + T]\,.$$

Applying the proposed differential inequality $\dot{v}(t) \geq \beta(v(t)) - \gamma(t)$ over this time interval $[\tau, \tau + T]$ and using the previous estimate of $\int_\tau^{\tau+\bullet} \gamma(u)\,du$ now yields that

$$
\begin{aligned}
v(\tau + s) - v(\tau) \;=\; \int_\tau^{\tau+s} \dot{v}(u)\,du &\geq \int_\tau^{\tau+s} \beta\big(v(u)\big) \;-\; \gamma(u)\,du \\
&\geq \eta_1\, s \;-\; (s+1)\,\eta_3 \\
&> \eta_3\, s \;-\; (s+1)\,\eta_3 \\
&= -\,\eta_3
\end{aligned}
\tag{1.20}
$$

for all $s \in [0, T]$.

Next we use this inequality (1.20) to prove that t_1 can be enlarged such that also

$$|v(t) - v^*| \;<\; \eta_2 \qquad \text{for all} \qquad t \geq t_1 \;;$$

for this purpose let $[\,\widetilde{\tau}, \widetilde{\tau} + \widetilde{T}\,]$ denote some interval such that $\widetilde{\tau} \geq t_1$, $\widetilde{T} > 0$ and

$$|v(t) - v^*| \;<\; \eta_2 - \eta_3 \qquad\qquad (\widetilde{\tau} \leq t \leq \widetilde{\tau} + \widetilde{T})\,;$$

as v^* is a subsequential limit of $t \mapsto v(t)$, such time intervals exist for any previous choice of t_1 ; due to the above reasoning (1.20) and since $v(\widetilde{\tau}) \in [v^* - (\eta_2 - \eta_3), v^* + (\eta_2 - \eta_3)]$,

$$v\,(\,\widetilde{\tau} + s\,) \;\overset{(1.20)}{>}\; v(\widetilde{\tau}) - \eta_3 \;\geq\; v^* - \eta_2 \qquad\qquad (0 \leq s \leq \widetilde{T})$$

so that $t \mapsto v(t)$ cannot exit the interval $[v^* - \eta_2, v^* + \eta_2]$ via the lower boundary; on the other hand, the same reasoning, since also $v(\widetilde{\tau} + \widetilde{T}) \in [v^* - (\eta_2 - \eta_3), v^* + (\eta_2 - \eta_3)]$, yields that

$$v\,(\widetilde{\tau}) \;\overset{(1.20)}{<}\; v\left(\,\widetilde{\tau} + \widetilde{T}\,\right) + \eta_3 \;\leq\; v^* + \eta_2$$

so that $t \mapsto v(t)$ cannot enter the interval $[v^* - \eta_2, v^* + \eta_2]$ via the upper boundary; this proves altogether that t_1 can be chosen large enough such that

$$|v(t) - v^*| \;<\; \eta_2 \qquad \text{for all} \qquad t \geq t_1 \;,$$

as had been claimed above.

In particular, (1.20) can be applied for any $\tau \geq t_1$ and $s := T > 0$, hence implying that

$$v(\tau + T) \;\geq\; [v(\tau) - \eta_3] \;+\; (\eta_1 - \eta_3)\,T \;\xrightarrow{T \to \infty}\; \infty$$

in contradiction to the previously proven boundedness $|v(t) - v^*| < \eta_2$ for large t.

2) Now we consider the Riccati-part of the statement, i.e. the case that

$$\beta(v) := b\,v\,(1-v)\,.$$

Since $b > 0$, β is a downward sloped quadratic function which is strictly positive on $(0,1)$, strictly negative on $\mathbb{R}\setminus[0,1]$ and vanishes at 0 and 1.

Now let v^* be some subsequential limit point of v. Then the first part of the lemma, shown above, yields that $\beta(v^*) \le 0$ which — according to the special choice of β — is only possible for $v^* \notin (0,1)$. But this is just a reformulation of the above claim. □

Theorem 1.4.3 (Hartman-Wintner-Perron). *Let $\mathbb{K}$ denote the real or complex number field, $\mathbb{K}=\mathbb{R}$ or $\mathbb{C}$, and consider the linear ODE*

$$dz_t \;=\; [\,\mathbf{A}+\mathbf{G}(t)\,]\,z_t\,dt \tag{1.19}$$

in $\mathbb{K}^n$, where $\mathbf{A}\in\mathbb{K}^{n\times n}$ is constant and the continuous map $\mathbf{G} : [t_0,\infty) \to \mathbb{K}^{n\times n}$ satisfies

$$\int_t^{t+1} \|\,\mathbf{G}(u)\,\|\,du \xrightarrow{t\to\infty} 0\,. \tag{1.21}$$

Then either $z_t = 0$ for all large t or the Lyapunov exponent of z_t exists and is equal to the real part of one of the eigenvalues Λ_j of $\mathbf{A}$, i.e.[3]*:*

$$\lambda(z_t) \;:=\; \lim_{t\to\infty}\frac{1}{t}\log|\,z_t\,| \;\in\; \{-\infty\,,\,\mathrm{Re}(\Lambda_1)\,,\,\ldots\,,\,\mathrm{Re}(\Lambda_n)\,\}\;.$$

Proof. We first choose an enumeration of the eigenvalues such that

$$\mathrm{Re}(\Lambda_1) \;\ge\; \mathrm{Re}(\Lambda_2) \;\ge\; \cdots \;\ge\; \mathrm{Re}(\Lambda_n)\;.$$

After a constant, invertible coordinate transformation we can assume that $\mathbf{A}$ is given in (complex) Jordan canonical form

$$\mathbf{A} \;=\; \begin{pmatrix} \boxed{\begin{matrix}\Lambda_1 & 0 & 0\\ \mathfrak{a} & \Lambda_2 & 0 \\ 0 & \mathfrak{a} & \Lambda_3\end{matrix}} & & & \\ & \boxed{\Lambda_4} & & \\ & & \ddots & \\ & & & \boxed{\begin{matrix}\Lambda_{n-1} & 0 \\ \mathfrak{a} & \Lambda_n\end{matrix}} \end{pmatrix}$$

---

[3] Here we adopt to the common convention that the Lyapunov exponent of the trivial solution $z_t = 0$ is defined as $\lambda(0) := -\infty$. Furthermore, $\mathrm{Re}(\,.\,)$ denotes the real part operation as usual.

where in this example $\Lambda_1 = \Lambda_2 = \Lambda_3$ and $\Lambda_{n-1} = \Lambda_n$, so that the boxes depict lower Jordan blocks of the following qualitative shape:

$$J^{\mathfrak{a}} := \begin{pmatrix} \Lambda & 0 & & & & & \\ \mathfrak{a} & \Lambda & 0 & & & & \\ & \mathfrak{a} & \Lambda & 0 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \mathfrak{a} & \Lambda & 0 & \\ & & & & \mathfrak{a} & \Lambda & 0 \\ & & & & & \mathfrak{a} & \Lambda \end{pmatrix} ;$$

here, $\mathfrak{a} > 0$ is some positive parameter and all matrix entries not mentioned are zero. Note that we have taken the "ordinary" lower Jordan blocks J^1 and carried out another transformation

$$\tilde{z}_t := C z_t ,$$

where

$$C := \operatorname{diag}(1, \mathfrak{a}, \ldots, \mathfrak{a}^{n-1}) ;$$

then

$$\frac{d}{dt}\,\tilde{z}_t = C\,\frac{d}{dt}\,z_t = C\mathbf{A}z_t + C\mathbf{G}(t)z_t \equiv C\mathbf{A}C^{-1}\tilde{z}_t + C\mathbf{G}(t)C^{-1}\tilde{z}_t ;$$

hence, we can redefine $\mathbf{A}$ as $C\mathbf{A}C^{-1}$ and $\mathbf{G}(t)$ as $C\mathbf{G}(t)C^{-1}$; after this transformation the Jordan-blocks of $\mathbf{A}$ are of the form $J_j^{\mathfrak{a}}$ and $\mathbf{G}(t)$ still enjoys the prescribed assumption. The possible limits of $\frac{1}{t}\log|z_t|$ are invariant under such invertible transformation of variables (or more generally under so-called Lyapunov-transformations). More precisely, the norm $|C\,.\,|$ on $\mathbb{K}^n$ is equivalent to the "old" norm $|\,.\,|$ as $\mathbb{K}^n$ is finite dimensional; since $t \to \infty$, factors from converting these norms into each other, end up as vanishing summands; therefore, we can redefine z_t as $\tilde{z}_t$ for the following considerations without loss of generality.

Hence, the underlying ODE (1.19), $\frac{d}{dt}\,z_t = [\,\mathbf{A} + \mathbf{G}(t)\,]\,z_t$, is of the form

$$\begin{aligned} \frac{d}{dt}\,z^1 &= \left[\Lambda_1\,z^1 + (\mathbf{G}z)^1\right] \\ \frac{d}{dt}\,z^i &= \left[\mathfrak{a}_{i-1}\,z^{i-1} + \Lambda_i\,z^i + (\mathbf{G}z)^i\right] \qquad (i = 2, \ldots, n) , \end{aligned}$$

where the superscript i indicates the i-th coordinate and

$$\mathfrak{a}_i := \mathbf{A}_{i+1,i} \in \{0, \mathfrak{a}\} \qquad (i = 1, \ldots, n-1)$$

denotes the $n-1$ entries of $\mathbf{A}$ below the diagonal. Putting $\mathfrak{a}_0 := 0$ as well as $z^0 := 0$ for definiteness and defining

$$\varrho_t^i \;:=\; \left| \, z_t^i \, \right| \;,$$

this implies that

$$\begin{aligned}
\frac{d}{dt}\,\left(\varrho^i\right)^2 &= \frac{d}{dt}\,\left(\bar{z}^i z^i\right) \;=\; \overline{\left(\frac{d}{dt}\,z^i\right)}\,z^i \,+\, \bar{z}^i\,\left(\frac{d}{dt}\,z^i\right) \\
&= 2\,\mathrm{Re}\left\{\,\bar{z}^i\left(\frac{d}{dt}\,z^i\right)\right\} \\
&= 2\,\mathrm{Re}\left\{\,\bar{z}^i\,\left[\,\mathfrak{a}_{i-1}z^{i-1} \,+\, \Lambda_i z^i \,+\, (\mathbf{G}z)^i\,\right]\right\} \\
&= 2\mathfrak{a}_{i-1}\;\mathrm{Re}(\bar{z}^i\,z^{i-1}) \,+\, 2\,\mathrm{Re}(\Lambda_i)\,(\varrho^i)^2 \,+\, 2\,\mathrm{Re}(\bar{z}^i\,(\mathbf{G}z)^i)
\end{aligned} \tag{1.22}$$

for all $i = 1, \ldots, n$, where "$^{-}$" denotes complex conjugation; hence, we can estimate

$$\begin{aligned}
\left|\frac{d}{dt}\,\left(\varrho^i\right)^2 - 2\,\mathrm{Re}(\Lambda_i)\,\left(\varrho^i\right)^2\right| &= \left|2\mathfrak{a}_{i-1}\,\mathrm{Re}(\bar{z}^i z^{i-1}) + 2\,\mathrm{Re}(\bar{z}^i(\mathbf{G}z)^i)\right| \\
&\le 2\mathfrak{a}_{i-1}\,\varrho^i\,\varrho^{i-1} \,+\, 2\varrho^i\,\left|(\mathbf{G}z)^i\right| \;.
\end{aligned} \tag{1.23}$$

Now define

$$\lambda_1 \;>\; \lambda_2 \;>\; \cdots \;>\; \lambda_p$$

as the *distinct* numbers in the list

$$\mathrm{Re}(\Lambda_1) \;\ge\; \mathrm{Re}(\Lambda_2) \;\ge\; \cdots \;\ge\; \mathrm{Re}(\Lambda_n)\;.$$

Note in particular that by definition, for all $k = 1, \ldots, p$

$$\underline{i} \;:=\; \min\{\,i \;:\; \mathrm{Re}(\Lambda_i) = \lambda_k\} \qquad \text{implies that} \qquad \mathfrak{a}_{\underline{i}-1} \;=\; 0\;. \tag{1.24}$$

We now define auxiliary processes describing the norms of the projection of z_t onto the subspaces belonging to the different real parts of eigenvalues; more precisely:

$$\begin{aligned}
L_t^k &:= \sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)=\lambda_k\}} (\varrho_t^i)^2 \qquad , k = 1, \ldots, p \\
M_t^k &:= \begin{cases} 0 & , k = 1 \\ \displaystyle\sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)>\lambda_k\}} (\varrho_t^i)^2 & , k = 2, \ldots, p \\ \displaystyle\sum_{i=1}^{n} (\varrho_t^i)^2 \;\equiv\; |\,z_t\,|^2 & , k = p+1 \end{cases} \\
N_t^k &:= \begin{cases} \displaystyle\sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)\le\lambda_k\}} (\varrho_t^i)^2 & , k = 1, \ldots, p \\ 0 & , k = p+1 \quad . \end{cases}
\end{aligned}$$

Then

$$| z_t |^2 \;=\; M_t^k + N_t^k \;=\; L_t^1 + \cdots + L_t^p \;\; . \tag{1.25}$$

Now it follows that

$$\begin{aligned}
&\left| \frac{dL^k}{dt} \,-\, 2\lambda_k \, L^k \right| \\
&\quad = \left| \sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)=\lambda_k\}} \left[\frac{d(\varrho^i)^2}{dt} \,-\, 2\lambda_k \, (\varrho^i)^2 \right] \right| \\
&\quad \le 2 \sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)=\lambda_k\}} \mathfrak{a}_{i-1} \, \varrho^i \, \varrho^{i-1} \,+\, 2 \sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)=\lambda_k\}} \varrho^i \, |(\mathbf{G}z)^i| \\
&\quad \le 2 \left(\sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)=\lambda_k\}} (\varrho^i)^2 \right)^{1/2} \left(\sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)=\lambda_k\}} \mathfrak{a}_{i-1}^2 \, (\varrho^{i-1})^2 \right)^{1/2} \\
&\qquad\quad + \, 2 \left(\sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)=\lambda_k\}} (\varrho^i)^2 \right)^{1/2} \left(\sum_{i=1}^{n} |(\mathbf{G}z)^i|^2 \right)^{1/2} \\
&\quad = 2 \, (L^k)^{1/2} \left(\sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)=\lambda_k \,,\, i-1 \ge \underline{i}\}} \mathfrak{a}_{i-1}^2 \, (\varrho^{i-1})^2 \right)^{1/2} \,+\, 2 \, (L^k)^{1/2} \;\; |(\mathbf{G}z)| \\
&\quad \le 2 \, \mathfrak{a} \, (L^k)^{1/2} \left(\sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)=\lambda_k\,\}} (\varrho^i)^2 \right)^{1/2} \,+\, 2 \, (L^k)^{1/2} \;\; \|\mathbf{G}\| \, |z| \\
&\quad = 2 \, \mathfrak{a} \, L^k \,+\, 2 \, \|\mathbf{G}\| \;\, (L^k)^{1/2} \;\, \big(M^k + N^k \big)^{1/2} \\
&\quad = L^k \left(2 \, \mathfrak{a} \,+\, 2 \, \|\mathbf{G}\| \sqrt{\frac{M^k + N^k}{L^k}} \right)
\end{aligned} \tag{1.26}$$

for all $k = 1, \ldots, p$, where we used in consecutive order the definition of L_t^k, the calculation (1.23), the Cauchy-Schwarz inequality, (1.24), the fact that $\mathfrak{a}_i \in \{0, \mathfrak{a}\}$ and (1.25).

Similarly, it follows from the definition of M_t^k, (1.22), the estimate that $\mathrm{Re}(a\, b) \ge -|a|\,|b|$ for complex numbers a and b, (1.24) and (1.25) that

$$\begin{aligned}
\frac{d\, M^k}{dt} &= \sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)>\lambda_k\}} \frac{d}{dt} \, (\varrho^i)^2 \\
&= \sum_{\{i:\mathrm{Re}(\Lambda_i)>\lambda_k\}} \Big[2\mathfrak{a}_{i-1} \; \mathrm{Re}(\bar{z}^i \, z^{i-1}) + 2 \, \mathrm{Re}(\Lambda_i) \, (\varrho^i)^2 + 2 \, \mathrm{Re} \left(\bar{z}^i \, (\mathbf{G}z)^i \right) \Big]
\end{aligned}$$

$$
\begin{aligned}
&\geq -2 \sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)>\lambda_k\}} \mathfrak{a}_{i-1}\,\varrho^i\,\varrho^{i-1} \;+\; 2\lambda_{k-1} \sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)>\lambda_k\}} (\varrho^i)^2 \\
&\qquad -\; 2 \sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)>\lambda_k\}} \varrho^i\,|\,(\mathbf{G}z)^i\,| \\
&\geq -2\,\mathfrak{a}\,M^k \;+\; 2\,\lambda_{k-1}\,M^k \;-\; 2\,\|\mathbf{G}\|\,(M^k)^{1/2}\;|z| \\
&= 2\,(\lambda_{k-1}-\mathfrak{a})\,M^k -\; 2\,\|\mathbf{G}\|\,\sqrt{M^k}\,\sqrt{M^k+N^k}
\end{aligned}
\tag{1.27}
$$

for all $k = 2, \ldots, p+1$.

In the same manner it follows from the definition of N_t^k, (1.22), the estimate that $\mathrm{Re}(a\,b) \leq +|a|\,|b|$ for complex numbers a and b, (1.24) and (1.25) that

$$
\begin{aligned}
\frac{d\,N^k}{dt} &= \sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)\leq\lambda_k\}} \frac{d}{dt}\,(\varrho^i)^2 \\
&= \sum_{\{i:\mathrm{Re}(\Lambda_i)\leq\lambda_k\}} \Big[2\mathfrak{a}_{i-1}\;\mathrm{Re}(\bar{z}^i\,z^{i-1}) + 2\,\mathrm{Re}(\Lambda_i)\,(\varrho^i)^2 + 2\,\mathrm{Re}(\bar{z}^i\,(\mathbf{G}z)^i)\Big] \\
&\leq 2 \sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)\leq\lambda_k\}} \mathfrak{a}_{i-1}\,\varrho^i\,\varrho^{i-1} \;+\; 2\lambda_k \sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)\leq\lambda_k\}} (\varrho^i)^2 \\
&\qquad +\; 2 \sum_{\{i\,:\,\mathrm{Re}(\Lambda_i)\leq\lambda_k\}} \varrho^i\,|\,(\mathbf{G}z)^i\,| \\
&\leq 2\,\mathfrak{a}\,N^k \;+\; 2\,\lambda_k\,N^k \;+\; 2\,\|\mathbf{G}\|\,(N^k)^{1/2}\,|z| \\
&= 2\,(\lambda_k+\mathfrak{a})\,N^k +\; 2\,\|\mathbf{G}\|\,\sqrt{N^k}\,\sqrt{M^k+N^k}\;\;.
\end{aligned}
\tag{1.28}
$$

for all $k = 1, \ldots, p$.

From (1.27) and (1.28) one gets for $|z|^2 \equiv M^{p+1} \equiv N^1$ that

$$
2\left(\lambda_p - \mathfrak{a} - \|\mathbf{G}(t)\|\right)|z_t|^2 \;\leq\; \frac{d}{dt}\,|z_t|^2 \;\leq\; 2\left(\lambda_1 + \mathfrak{a} + \|\mathbf{G}(t)\|\right)|z_t|^2\;,
$$

so that integration over $t \geq t_0$ yields that

$$
\begin{aligned}
\exp\left\{(\lambda_p - \mathfrak{a})(t-t_0) - \int_{t_0}^{t} \|\mathbf{G}(s)\|\,ds\right\}\,|z_{t_0}| \;\leq\; |z_t| \\
\leq \exp\left\{(\lambda_1 + \mathfrak{a})(t-t_0) + \int_{t_0}^{t} \|\mathbf{G}(s)\|\,ds\right\}\,|z_{t_0}|\;\;.
\end{aligned}
$$

This proves again (besides the invertibility of the fundamental matrix solution) that if $z_{t_1} = 0$ for some time t_1, then also $z_t = 0$ for all times $t \geq t_1$. Furthermore, due to the standing assumption (1.21) it follows that

$$
\frac{1}{t}\int_{t_0}^{t} \|\mathbf{G}(s)\|\,ds \;\xrightarrow{t\to\infty}\; 0
$$

and hence the previous string of inequalities implies the statement about the Lyapunov exponent for a non-vanishing solution in case $p = 1$:

$$\frac{1}{t} \log |\, z_t \,| \xrightarrow{t\to\infty} \lambda_1$$

since $\mathfrak{a}$ can be chosen arbitrarily small.

Thus it is left to show the claim concerning the Lyapunov exponents in case $p \geq 2$; again $\mathfrak{a} > 0$ serves as the small parameter and since $p \geq 2$, we can choose it such that

$$2\,\mathfrak{a} < \min_{k=1,\dots,p-1} \lambda_k - \lambda_{k+1} \ . \tag{1.29}$$

One needs to find for each non-vanishing solution z_t an index $\mathcal{J} \in \{1,\dots,p\}$ such that

$$\lim_{t\to\infty} \frac{1}{t} \log |\, z_t| = \lambda_{\mathcal{J}} \ . \tag{1.30}$$

Having excluded the case that z vanishes after some time, we can consider the auxiliary processes

$$v_t^k := \frac{M_t^k}{|\, z_t|^2} \equiv \frac{M_t^k}{M_t^k + N_t^k} \in [0,1]$$

for $k = 1,\dots,p+1$; in particular $v^1 \equiv 0$ and $v^{p+1} \equiv 1$. These processes satisfy

$$\dot{v}^k = \frac{N^k\,\dot{M}^k - M^k\,\dot{N}^k}{(M^k + N^k)^2} \ ,$$

where $\dot{}$ denotes the derivative with respect to time, $\frac{d}{dt}$, as usual; furthermore,

$$v^k\,(\,1 - v^k\,) = \frac{M^k\,N^k}{(M^k + N^k)^2} \ ;$$

since for the positive quantities M^k, N^k also the trivial manipulation

$$\begin{aligned}(N^k\sqrt{M^k} + M^k\sqrt{N^k})^2 &\leq 2(N^k)^2 M^k + 2(M^k)^2 N^k \\ &\leq 3(N^k)^2 M^k + 3(M^k)^2 N^k + (M^k)^3 + (N^k)^3 \\ &= (M^k + N^k)^3\end{aligned}$$

holds true, we get together with (1.27) and (1.28) that

$$\dot{v}^k = \frac{N^k\,\dot{M}^k - M^k\,\dot{N}^k}{(M^k + N^k)^2}$$

$$
\begin{aligned}
&\geq \frac{1}{(M^k+N^k)^2}\ \Big\{ N^k \Big[2(\lambda_{k-1}-\mathfrak{a})M^k - 2\,\|\mathbf{G}\|\,\sqrt{M^k}\,\sqrt{M^k+N^k} \Big] \\
&\qquad\qquad - \ M^k \Big[2\,(\,\lambda_k\ +\ \mathfrak{a}\,)\,N^k\ +\ 2\,\|\mathbf{G}\|\,\sqrt{N^k}\,\sqrt{M^k+N^k} \Big] \Big\} \\
&= \frac{1}{(M^k+N^k)^2}\ \Big\{ 2\,M^k\,N^k\,(\lambda_{k-1}-\lambda_k-2\mathfrak{a}) \\
&\qquad\qquad -2\,\|\mathbf{G}\|\,\sqrt{M^k+N^k}\,\Big(N^k\sqrt{M^k} + M^k\sqrt{N^k} \Big) \Big\} \\
&\geq \frac{1}{(M^k+N^k)^2}\ \big\{ 2\,M^k\,N^k\,(\lambda_{k-1}-\lambda_k-2\mathfrak{a})\ -\ 2\,\|\mathbf{G}\|\,(M^k+N^k)^2 \big\} \\
&= 2\,(\lambda_{k-1}-\lambda_k-2\mathfrak{a})\,v^k\,(\,1-v^k\,)\ -\ 2\,\|\mathbf{G}\|\,.
\end{aligned}
$$

Hence, lemma 1.4.2 applies to the Riccati-type differential inequality

$$\dot{v}\ \geq\ b\,v\,\big(1-v\big)\ -\ \gamma\ ,$$

with $v := v^k$, $b := 2\,(\lambda_{k-1}-\lambda_k-2\mathfrak{a}) > 0$ due to (1.29) and $\gamma := 2\,\|\mathbf{G}\|$ due to (1.21). Therefore, as $v \equiv v^k$ takes its values in $[0,1]$,

$$\lim_{t\to\infty} v_t^k\ \in\ \{0,1\} \qquad\qquad (k=1,\dots,p+1)\ .$$

Since $v^1 \equiv 0$ and $v^{p+1} \equiv 1$, the following quantity is thus well-defined:

$$\mathcal{J}\ :=\ \max\Big\{\,k=1,\dots,p\ :\ \lim_{t\to\infty} v_t^k\ =\ 0\,\Big\}\ .$$

This implies, as $M^{k+1} = L^1+\dots+L^k$, that for $k=1,\dots,p$

$$\frac{L_t^k}{|z_t|^2}\ =\ \frac{M_t^{k+1}-M_t^k}{|z_t|^2}\ =\ v_t^{k+1}-v_t^k\ \xrightarrow{t\to\infty}\ \begin{cases} 0\,,\ k\neq\mathcal{J}\,; \\ 1\,,\ k=\mathcal{J}\,. \end{cases} \tag{1.31}$$

In particular one also gets

$$\frac{L_t^k}{L_t^{\mathcal{J}}}\ =\ \frac{L_t^k}{|\,z_t|^2}\,\frac{|\,z_t|^2}{L_t^{\mathcal{J}}}\ \xrightarrow{t\to\infty}\ 0 \qquad\qquad (k\neq\mathcal{J}). \tag{1.32}$$

The importance about (1.31) is that it reduces the study of $|z_t|$ to $\sqrt{L_t^{\mathcal{J}}}$ for large t, namely

$$\log\sqrt{L_t^{\mathcal{J}}}\ -\ \log|z_t|\ =\ \frac{1}{2}\,\log\frac{L_t^{\mathcal{J}}}{|z_t|^2}\ \xrightarrow{t\to\infty}\ 0$$

and hence the claim (1.30) concerning the Lyapunov exponent $\lambda(z_t)$ reduces to proving that

$$\lim_{t\to\infty}\frac{1}{t}\log L_t^{\mathcal{J}}\ =\ 2\,\lambda_{\mathcal{J}}\ .$$

Considering (1.26) for $k = \mathcal{J}$ after dividing by $L_t^{\mathcal{J}}$ ($\neq 0$) now yields:

$$\begin{aligned}
\limsup_{t\to\infty} \left| \frac{1}{t} \log L_t^{\mathcal{J}} - 2\lambda_{\mathcal{J}} \right| &= \limsup_{t\to\infty} \frac{1}{t} \left| \int_{t_0}^{t} \left(\frac{d}{du} \log L_u^{\mathcal{J}} - 2\,\lambda_{\mathcal{J}} \right) du \right| \\
&\leq \limsup_{t\to\infty} \frac{1}{t} \int_{t_0}^{t} \left| \frac{d}{du} \log L_u^{\mathcal{J}} - 2\lambda_{\mathcal{J}} \right| du \\
&\overset{(1.26)}{\leq} 2\mathfrak{a} + 2 \limsup_{t\to\infty} \frac{1}{t} \int_{t_0}^{t} \|\mathbf{G}(u)\| \sqrt{\frac{M_u^{\mathcal{J}} + N_u^{\mathcal{J}}}{L_u^{\mathcal{J}}}}\, du \\
&\overset{(1.31)}{\leq} 2\,\mathfrak{a} + 4 \limsup_{t\to\infty} \frac{1}{t} \int_{t_0}^{t} \|\,\mathbf{G}(u)\,\|\, du \\
&\overset{(1.21)}{=} 2\,\mathfrak{a}\,.
\end{aligned}$$

Since $\mathfrak{a}$ can be chosen arbitrarily small, this shows that the Lyapunov exponent $\lambda(z_t)$ exists and is equal to $\lambda_{\mathcal{J}}$. □

Remark 1.4.4 (Asymptotic growth of the projections). Due to the structure of the differential equation in Jordan canonical form in the proof of theorem 1.4.3, it follows in particular from remark 1.4.1 that for any $k = 1,\ldots,p$

$$L_t^k = \left| P_{\lambda_k} z_t \right|^2 .$$

Thus (1.32) implies that by choosing $\lambda := \lambda_{\mathcal{J}} \equiv \lambda(z_t)$, the Lyapunov exponent of one particular solution z_t, one gets that

$$\frac{\left| P_{\lneqq\lambda} z_t \right|}{|\,P_{\lambda} z_t\,|} \xrightarrow{t\to\infty} 0\,.$$

□

Remark 1.4.5. The proof of the above theorem 1.4.3 shows that the linearity of the perturbation $\mathbf{G}(t)z$ is not crucial. The argument also goes through unchanged for the nonlinearly perturbed equation

$$dz_t = [\,\mathbf{A}\,z_t + g(t,z_t)\,]\,dt\,,$$

where the continuous map $g : [t_0,\infty) \times \mathbb{R}^n \to \mathbb{R}^n$ satisfies

$$|\,g(t,z)\,| \leq \gamma(t)\,|z|$$

together with

$$\int_t^{t+1} \gamma(u)\,du \xrightarrow{t\to\infty} 0\,.$$

□

Remark 1.4.6 ($\mathbb{K}^n = \mathbb{R}^2$). As in theorem 1.4.3 again consider the linear ODE

$$dz_t \;=\; [\,\mathbf{A} + \mathbf{G}(t)\,]\, z_t\, dt \tag{1.19}$$

under the same assumptions as imposed there; let $\mathbb{K}^n = \mathbb{R}^2$ for simplicity. The exponential growth rate $\lim_{t\to\infty} \frac{1}{t} \log|\, z_t\,|$ has been calculated as $\mathrm{Re}(\Lambda_1)$ or $\mathrm{Re}(\Lambda_2)$ in the non-trivial cases $z_t \not\equiv 0$. Despite these numerically exact results, the formulas (1.10) or (1.11),

$$\frac{1}{t}\log|\, z_t\,| \;=\; \frac{\log|\, z_0\,|}{t} + \frac{1}{t}\int_0^t \bar{Q}(u,\alpha_u)\, du \quad ,$$

which evaluate the growth rate of z_t by a functional of its angle α_t, have not been used in the course of the proof. It is the purpose of this remark to bridge this gap in the illustrative two-dimensional situation.

If $p = 1$, i.e. if $\mathrm{Re}(\Lambda_1) = \mathrm{Re}(\Lambda_2)$, we get in the notion of the previous proof (see p.35) that the norms of the relevant projections are $L_t^1 = M_t^2 = N_t^1 = |z_t|^2$ and $M_t^1 = N_t^2 = 0$; the proof then estimates $|z_t|^2$ and does not contain a conclusion concerning the angle.

If $p = 2$, i.e. if $\mathrm{Re}(\Lambda_1) > \mathrm{Re}(\Lambda_2)$, the situation changes. Both eigenvalues are real,

$$\lambda_1 \;:=\; \Lambda_1 \;>\; \Lambda_2 \;=:\; \lambda_2 \;\; ,$$

each of which has an eigenvector. After the transformation mentioned at the beginning of the previous proof,

$$\mathbf{A} \;=\; \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

and the canonical unit vector e_i is the eigenvector corresponding to λ_i , $i = 1,2$. Then it follows from the notion of p.35 that

$$\begin{aligned}
L_t^k &= (\varrho_t^k)^2 \qquad\qquad , k \in \{1,2\}\ , \\
M_t^k &= \begin{cases} 0 & , k=1 \\ (\varrho_t^1)^2 & , k=2 \\ \sum_{i=1}^2 (\varrho_t^i)^2 \;\equiv\; |\, z_t\,|^2 & , k=3\ , \end{cases} \\
N_t^k &= \begin{cases} \sum_{i=1}^2 (\varrho_t^i)^2 \;\equiv\; |\, z_t\,|^2 & , k=1 \\ (\varrho_t^2)^2 & , k=2 \\ 0 & , k=3\ , \end{cases}
\end{aligned}$$

where as before $\varrho_t^i \;\equiv\; |\, z_t^i\,|$. The auxiliary processes v^1, v^2 and v^3 as defined on p.38 read in this case as $v_t^1 = \frac{M_t^1}{|\,z_t|^2} \equiv 0\,,\; v_t^3 = \frac{M_t^3}{|\,z_t|^2} \equiv 1$ and

$$v_t \;:=\; v_t^2 \;:=\; \frac{M_t^2}{|\, z_t|^2} \;\equiv\; \frac{(\varrho_t^1)^2}{(\varrho_t^1)^2 + (\varrho_t^2)^2} \;=\; (\cos\alpha_t)^2 \quad ,$$

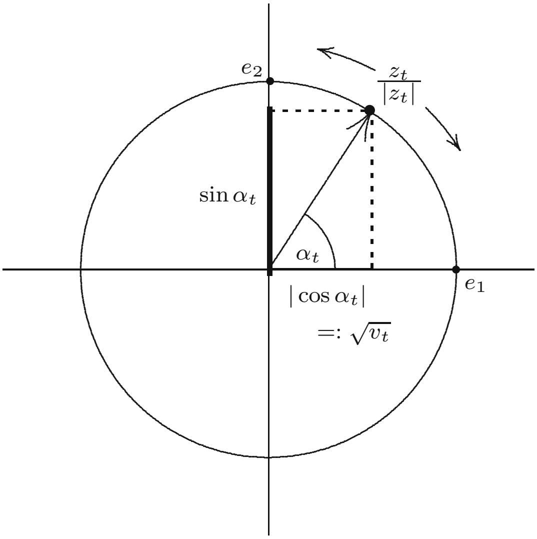

Fig. 1.1 Geometrical interpretation of the quantity $v_t := (\cos\alpha_t)^2$

where α_t denotes the angle of z_t as measured canonically with respect to the e_1-coordinate axis. As figure 1.1 illustrates, $v_t = \cos^2\alpha_t$ quantifies the distance between $\frac{z_t}{|z_t|}$ and e_2 and hence induces a metric on the projective space S^1. Another popular choice to obtain a metric on S^1 is to work with $|\sin\alpha_t|$ which measures the distance between the projective lines $\frac{z_t}{|z_t|}$ and e_1; the latter metric is commonly used in the proof of the Multiplicative Ergodic Theorem; see Arnold [Ar 98, Prop. 3.2.8 & Lem. 3.4.6].

The system v_t then satisfies the ODE

$$\dot{v}_t = \frac{(\varrho_t^2)^2\,\left[\frac{d}{dt}(\varrho_t^1)^2\right] \;-\; (\varrho_t^1)^2\,\left[\frac{d}{dt}(\varrho_t^2)^2\right]}{|z_t|^4}$$

from which the following Riccati-type differential inequality follows (see p.38f.),

$$\dot{v}_t \geq 2\,(\lambda_1 - \lambda_2)\,v_t\,(\,1 - v_t\,) \;-\; 2\,\|\,\mathbf{G}(t)\,\|\,.$$

Lemma 1.4.2 applies with $b := 2\,(\lambda_1 - \lambda_2) \;>\; 0$ and $\gamma := 2\,\|\mathbf{G}\|$ due to (1.21). Therefore,

$$\lim_{t\to\infty} v_t \;\in\; \{0,1\}$$

and the following quantity is thus well-defined,

$$\mathcal{J} \;:=\; \begin{cases} 1 & , \;\text{ if } \lim_{t\to\infty} v_t \;=\; 1\,, \\ 2 & , \;\text{ if } \lim_{t\to\infty} v_t \;=\; 0\,. \end{cases}$$

It displays the direction $e_{\mathcal{J}}$ to which the projective line $\psi_t := \frac{z_t}{|z_t|} \doteq \alpha_t$ converges; see figure 1.1; due to continuity the angle also converges, $\alpha_\infty := \lim_{t\to\infty} \alpha_t \in \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$. Therefore, it follows altogether that

$$\bar{Q}\big(\mathbf{A} + \mathbf{G}(t), \alpha_t\big) \;\equiv\; Q\big(\mathbf{A} + \mathbf{G}(t), \psi_t\big) \;\xrightarrow{t\to\infty}\; Q\big(\mathbf{A}, e_{\mathcal{J}}\big) \;\equiv\; \bar{Q}(\mathbf{A}, \alpha_\infty)$$

and hence the formulas (1.10) and (1.11) imply that

$$\frac{1}{t}\log|z_t| = \frac{\log|z_0|}{t} + \frac{1}{t}\int_0^t \bar{Q}\big(\mathbf{A} + \mathbf{G}(u), \alpha_u\big)\, du \xrightarrow{t\to\infty} \bar{Q}(\mathbf{A}, \alpha_\infty) \in \{\lambda_1, \lambda_2\},$$

where in the last step the definition of $\bar{Q}$ (see (1.7)) as well as the diagonal form of $\mathbf{A}$ have been used.

Thus the gap between theorem 1.4.3 and the formulas (1.10) and (1.11) is closed. $\square$

The following theorem shows that every real part of an eigenvalue can indeed be seen as a Lyapunov exponent. Since it will not be used in the sequel, the arguments will not be provided here. Instead, see Coppel [Cp 65, p.100–102] for the proof which consists of using remark 1.4.1, defining an appropriate integral transformation and applying the fixed point theorem by Schauder and Tychonoff then.

Theorem 1.4.7. *Suppose that the assumptions of the Hartman-Wintner-Perron theorem 1.4.3 hold, i.e. that we consider*

$$dz_t \;=\; \big[\,\mathbf{A} + \mathbf{G}(t)\,\big]\, z_t\, dt\,, \tag{1.19}$$

where $\mathbf{A}$ *is a constant matrix and* $\mathbf{G}$ *is a continuous matrix function again satisfying*

$$\int_t^{t+1} \|\,\mathbf{G}(u)\,\|\, du \;\xrightarrow{t\to\infty}\; 0\,. \tag{1.21}$$

Then for any real part λ *of an eigenvalue* Λ_i *of* $\mathbf{A}$*,*

$$\lambda \;\in\; \{\,\mathrm{Re}(\Lambda_1)\,,\,\ldots\,,\,\mathrm{Re}(\Lambda_n)\,\}\;,$$

there exists an $\eta \equiv \eta(\mathbf{A}) > 0$ *and a time* $T \equiv T(\mathbf{A}, \mathbf{G}) > t_0$ *such that for all* $t_1 \geq T$ *and for all* $\chi_{<\lambda} \in E_{<\lambda}$ *and* $\chi_\lambda \in E_\lambda$ *satisfying*

$$|\,\chi_{<\lambda}\,| \;<\; \eta\,|\,\chi_\lambda\,|\,,$$

the equation (1.19) has a unique solution $(z_t)_{t\geq t_1}$ *such that*

$$P_{<\lambda}\, z_{t_1} \;=\; \chi_{<\lambda}\;, \qquad P_\lambda\, z_{t_1} \;=\; \chi_\lambda$$

and

$$\lambda(z_t) \;\equiv\; \lim_{t\to\infty} \frac{1}{t}\, \log|\,z_t\,| \;=\; \lambda\,.$$

Remark 1.4.8. As has been already mentioned, the results of this section are due to Hartman and Wintner [Ha-Wi 55]. There are many more papers on asymptotically constant linear ODEs of which we only mention Levinson [Lv 48] and Harris and Lutz [Harr-Lutz 77]; also see Eastham [Ea 89] and the references therein for an overview.

There also exists a version of the Hartman-Wintner theorem for functional differential equations which is due to Pituk [Pu 99].

A vast amount of the literature is furthermore dedicated to the non-linear case, that is to considering ODEs

$$\dot{x}_t \;=\; f(x,t)\,,$$

which are *asymptotically autonomous*; this means that the time-dependent vector field $f(\,.\,,t)$ approaches an autonomous (time-independent) vector field $f(\,.\,)$ as $t \to \infty$ in a certain sense; see e.g. Markus [Mar 56] and Strauss and Yorke [Stra-Yor 67].

1.5 Sample systems

In this section several examples are presented which shall illustrate where linear, real noise driven stochastic systems (1) appear; these sample systems can be regarded as toy models for the different situations described by the definitions and assumptions in chapter 4.

Example 1.5.1 (Linearized SDEs with constant noise coefficient $\boldsymbol{\sigma}$). Let X^ε be the diffusion given by the SDE

$$dX_t^\varepsilon \;=\; b\,(X_t^\varepsilon)\;dt \;+\; \sqrt{\varepsilon}\;\sigma\,dW_t\;\;,$$

where $b \in C^\infty(\mathbb{R}^d,\mathbb{R}^d)$, $\varepsilon \geq 0$, and the Brownian motion W are as before and where $\sigma \in GL(d,\mathbb{R})$ is now supposed to be a constant (invertible) matrix. Defining the matrix-valued mapping $\mathbf{A} : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ as the Jacobian of b,

$$\mathbf{A}(x) \;:=\; Db(x) \;:=\; \left(\frac{\partial\, b_i(x)}{\partial x_j}\right)_{i,j=1,\ldots,d}\qquad,$$

yields the *linearized (variational) equation*

$$\begin{aligned} dZ_t^\varepsilon &= \mathbf{A}\,(X_t^\varepsilon)\;\,Z_t^\varepsilon\;dt \\ dX_t^\varepsilon &= b\,(X_t^\varepsilon)\;dt \;+\; \sqrt{\varepsilon}\;\sigma\,dW_t \end{aligned} \tag{1.33}$$

in the sense that the coefficient functions of the SDE for X^ε are linearized, i.e. differentiated with respect to the space variable x; as the noise coefficient $\sqrt{\varepsilon}\,\sigma$ does not depend on x by assumption, its derivative vanishes and thus there is no noise component of Z^ε meaning that the resulting stochastic system is indeed real noise driven; also see Arnold [Ar 88, p.194f.]. By taking derivatives it follows, in particular, that the dimensions of the state spaces of Z^ε and X^ε are equal, $n = d$.

For further assertions (and the corresponding assumptions), concerning how Z^ε can then be considered as the linearization of X^ε itself, we would like to refer to Blagovescenskii and Freidlin [Bla-Fr 61], Khasminskii [Kh 80, Sec.V.6], Bismut [Bis 81] and Arnold [Ar 98, Sec.2.3] among others.

By considering the linearized equation as above, it follows in other words that any system X^ε to be investigated later (satisfying the general assumptions 2.1.1, more precisely) for which σ is constant serves as an example of (1) by taking the Jacobian $\mathbf{A} := Db$. Since in most cases $\sigma \equiv \mathrm{id}_{\mathbb{R}^d}$ for simplicity anyway, the sample models as will be discussed in section 2.6 already provide a whole class of sample systems for (1). To be specific two cases of linearized SDEs will be discussed in the following two examples.

Example 1.5.2 (Multi-dimensional Ornstein-Uhlenbeck process). Consider the SDE

$$dX_t^\varepsilon \;=\; \mathbf{A}X_t^\varepsilon\, dt \;+\; \sqrt{\varepsilon}\;\sigma\, dW_t \;\;,$$

where $\mathbf{A}$ and σ are constant elements of $\mathbb{R}^{d\times d}$, the latter being invertible in addition. This SDE is understood as random perturbation of the linear, deterministic dynamical system $\dot{X}_t^0 = \mathbf{A}X_t^0$. The linearization (1.33) of this SDE is

$$dZ_t^\varepsilon \;=\; \mathbf{A}Z_t^\varepsilon\, dt \;\;, \qquad\qquad Z_0^\varepsilon \;=\; z \;\in \mathbb{R}^d \;\;;$$

it does not depend on X^ε and is therefore independent of ε, $Z_t^\varepsilon(\omega, x, z) = Z_t^\varepsilon(z) = X_t^{0,z}$; also see Arnold [Ar 88, p.194]. In this case the Lyapunov exponents of the system (1.33) from the Multiplicative Ergodic Theorem 1.3.1 coincide with the Lyapunov exponents from 1.4.1 and are given by the real part of the eigenvalues of $\mathbf{A}$. Since the stochasticity is not present after the linearization, it also amounts to call the real parts of the eigenvalues of $\mathbf{A}$ the *local Lyapunov exponents of the system* X^ε.

This system X^ε will be further discussed in 2.6.1 under the assumption that $\mathbf{A}$ be normal and with the specification $\sigma = \mathrm{id}_{\mathbb{R}^d}$; these two conditions then allow to calculate a certain "cost" function (quasipotential). All in all this system X^ε together with its linearization serves as a bridge between the stochastic case (i.e. the Multiplicative Ergodic Theorem 1.3.1) and the familiar deterministic fact 1.4.1, since the resulting Lyapunov exponents coincide and since it additionally incorporates a simple example for the Freidlin-Wentzell theory (with only one metastable state) which will come into play later; see 2.6.1. □

Example 1.5.3 (Linearized gradient SDE with constant $\boldsymbol{\sigma}$. Potential $\boldsymbol{U_1}$). Now consider the linearized SDE (1.33) in case that the drift is derived as the gradient of a function which is then called the *potential (function)* of the drift. More precisely, consider

$$dX_t^\varepsilon \;=\; -\,\nabla U\,(X_t^\varepsilon)\;dt \;+\; \sqrt{\varepsilon}\;\sigma\,dW_t\;\;,$$

where the drift $b := -\nabla U$ is given by $U \in C^\infty(\mathbb{R}^d, \mathbb{R})$, as for example in figure 1, and where again $\sigma \in GL(d, \mathbb{R})$. An important special case is given by $\sigma = \mathrm{id}_{\mathbb{R}^d}$ for which this SDE will be considered later as equation (2.2). By differentiation it follows that the linear component of the variational system (1.33) is given by

$$dZ_t^\varepsilon \;=\; -\,H_U\left(X_t^\varepsilon\right)\,Z_t^\varepsilon\;\;dt\;,$$

where the coefficient matrix $\mathbf{A}(x) := -H_U(x)$ is defined via the Hesse matrix

$$H_U(x) \;:=\; \left(\frac{\partial^2 U}{\partial x_i\,\partial x_j}\,(x)\right)_{i,j=1,\ldots,d}\;\;,$$

the symmetric matrix of second derivatives of U at x. Again note that $n = d$ due to the linearization.

For $n = d = 2$ the formulas as obtained in remark 1.2.1 read as follows: The drift $\bar{h}$ of the nonlinear real noise SDE (1.6), $d\alpha_t^\varepsilon = \bar{h}(X_t^\varepsilon, \alpha_t^\varepsilon)\;dt$, for the angle α^ε of Z^ε is

$$\begin{aligned}
\bar{h}(x,\alpha) &\equiv \bar{h}\big(-H_U(x),\alpha\big)\\
&= \frac{\partial^2 U}{\partial x_1\,\partial x_2}\,(\,\sin^2\alpha\;-\;\cos^2\alpha\,) + \left(\frac{\partial^2 U}{\partial x_1^2}\;-\;\frac{\partial^2 U}{\partial x_2^2}\right)\sin\alpha\,\cos\alpha\\
&= -\,\frac{\partial^2 U}{\partial x_1\,\partial x_2}\,\cos 2\alpha\;+\;\frac{1}{2}\left(\frac{\partial^2 U}{\partial x_1^2}\;-\;\frac{\partial^2 U}{\partial x_2^2}\right)\sin 2\alpha
\end{aligned}$$

for the drift of the angle process. According to (1.7) the depiction (1.8) of the growth rate,

$$\frac{1}{t}\,\log|\,Z_t^\varepsilon(\omega,x,z)\,| \;=\; \frac{\log|\,z\,|}{t}\;+\;\frac{1}{t}\int_0^t \bar{Q}\Big(X_u^{\varepsilon,x}(\omega),\alpha_u^\varepsilon(\omega,x,z)\Big)\,du\;\;,$$

is determined by the kernel function

$$\begin{aligned}
\bar{Q}(x,\alpha) &\equiv \bar{Q}\big(-H_U(x),\alpha\big)\\
&= -\,\frac{\partial^2 U}{\partial x_1^2}\,\cos^2\alpha\;-\;\frac{\partial^2 U}{\partial x_2^2}\,\sin^2\alpha\;-\;2\,\frac{\partial^2 U}{\partial x_1\,\partial x_2}\,\sin\alpha\,\cos\alpha\;.
\end{aligned}$$

The symmetry of the Hesse matrix implies that there are either two or infinitely many switching curves of the angle drift $\bar{h}$; see 1.2.4. Note that the latter assertion is also a consequence of the spectral decomposition theorem for symmetric matrices.

In order to examine a specific numerical example consider the potential function

$$U_1(x) \;:=\; \frac{3}{2}\,x_1^4 \,-\, \frac{2}{3}\,x_1^3 \,-\, 3\,x_1^2 \,+\, c\,x_1\,x_2 \,+\, \frac{3}{2}\,x_2^4\,, \qquad (1.34)$$

where $c \in \mathbb{R}$ is a constant. This function U_1 is plotted in figure 2.1 for $c = 1$. Here, the coefficient matrix of the linearized system is given by

$$\mathbf{A}_1(x) \;:=\; -\,H_{U_1}(x) \;=\; \begin{pmatrix} -\,18x_1^2 \,+\, 4\,x_1 \,+\, 6 & -\,c \\ -\,c & -\,18\,x_2^2 \end{pmatrix} \;;$$

furthermore, one gets

$$\begin{aligned} \bar{h}(x,\alpha) &= c\,(\,\sin^2\alpha \,-\, \cos^2\alpha\,) \,+\, (\,18x_1^2 \,-\, 18\,x_2^2 \,-\, 4\,x_1 \,-\, 6\,)\,\sin\alpha\,\cos\alpha \\ &= -\,c\,\cos 2\alpha \,+\, (\,9x_1^2 \,-\, 9\,x_2^2 \,-\, 2\,x_1 \,-\, 3\,)\,\sin 2\alpha\,. \end{aligned}$$

In particular, $\left\{\, k\,\frac{\pi}{2} : k \in \mathbb{Z} \,\right\}$ are zeros of $\bar{h}(x,\,.\,)$ for any x, if $c = 0$. If $c \neq 0$, one can find for any α an $x \in \mathbb{R}^2$ such that $\bar{h}(x,\alpha) \neq 0$; in particular, $\bar{h}$ is hypoelliptic in the sense of (1.15) then; also see definition 4.4.4. However, $\bar{h}$ is not "strongly hypoelliptic" in the sense of definition 4.4.4 as will be discussed in remark 4.4.11. The integral kernel for the exponential growth rate in (1.8) is calculated as

$$\bar{Q}(x,\alpha) \;=\; (\,-\,18x_1^2 \,+\, 4\,x_1 \,+\, 6\,)\,\cos^2\alpha \,-\, 18\,x_2^2\,\sin^2\alpha \,-\, 2\,c\,\sin\alpha\,\cos\alpha\,.$$

$\square$

Example 1.5.4 (Diagonal matrix plus small skew-symmetric perturbation). Consider the coefficient matrix

$$\mathbf{A} \,:\, \mathbb{R} \to \mathbb{R}^{2\times 2}\,, \qquad \mathbf{A}(x) \;:=\; \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \,+\, x\,\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \;=\; \begin{pmatrix} \lambda_1 & x \\ -x & \lambda_2 \end{pmatrix}$$

for the system (1), where

$$\lambda_1 \;>\; \lambda_2\,.$$

In particular, the dimensions are $n = 2$ and $d = 1$ in (1), the latter fact meaning that X^ε enters as one-dimensional diffusion

$$dX_t^\varepsilon \;=\; -\,U'\,(X_t^\varepsilon)\;dt \,+\, \sqrt{\varepsilon}\;dW_t\;\;,$$

where we used the fact that in dimension $d = 1$ any drift can be written as the gradient $b = -U'$ of a potential function by direct integration $U(x) := -\int_0^x b(y)dy$ and where for simplicity $\sigma := \mathrm{id}_{\mathbb{R}} \equiv 1$. For the system Z^ε as defined by (1) with the above coefficient matrix function $\mathbf{A}$ it follows from remark 1.2.1 that its angle α_t^ε follows the RDE (1.6), $d\alpha_t^\varepsilon = \bar{h}(X_t^\varepsilon,\alpha_t^\varepsilon)\;dt$, with velocity

$$\begin{aligned}\bar{h}(x,\alpha) &\equiv \bar{h}\big(\mathbf{A}(x),\alpha\big)\\ &= -\,x\,\sin^2\alpha \;-\; x\,\cos^2\alpha \;+\; \big(\lambda_2-\lambda_1\big)\,\sin\alpha\,\cos\alpha\\ &= -\,x \;+\; \frac{\lambda_2-\lambda_1}{2}\;\sin 2\alpha\;;\end{aligned}$$

hence, the switching curve can be written in the simple form

$$x = \frac{\lambda_2-\lambda_1}{2}\;\sin 2\alpha\;;$$

this switching curve together with the tendencies of the drift $\bar{h}$ is sketched in figure 1.2.
The formula for the exponential growth rate (1.8) is determined by the integrand function (1.7) which reads

$$\begin{aligned}\bar{Q}(x,\alpha) &\equiv \bar{Q}(\mathbf{A}(x),\alpha)\\ &= \lambda_1\,\cos^2\alpha \;+\; \lambda_2\,\sin^2\alpha \;+\; (x-x)\,\sin\alpha\,\cos\alpha\\ &= \lambda_1 \;+\; (\lambda_2-\lambda_1)\,\sin^2\alpha\end{aligned}$$

here. For these facts also see Arnold [Ar 79, p.136f.].

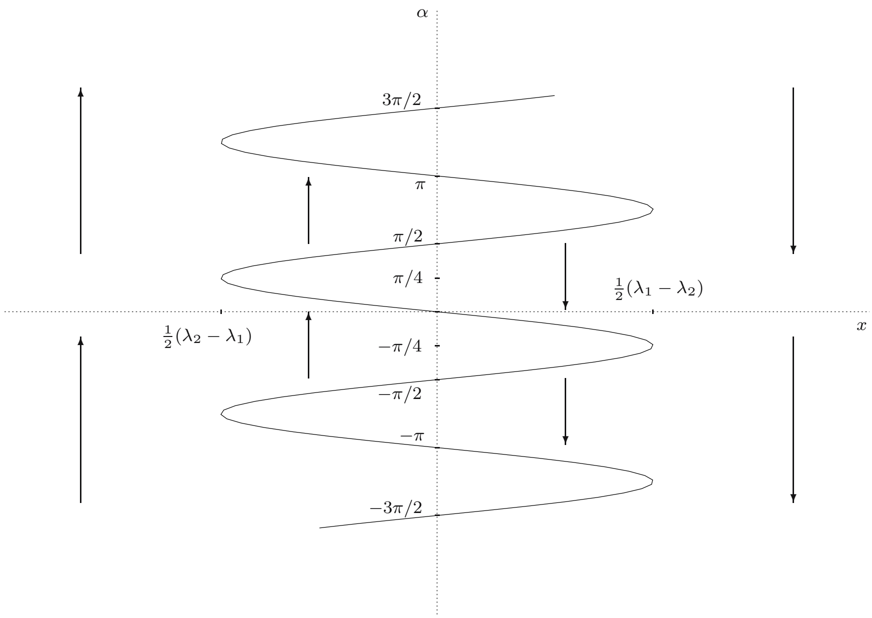

Fig. 1.2 The switching curve and tendencies of $\bar{h}(x,\alpha) = -\,x \,+\, \frac{1}{2}(\lambda_2-\lambda_1)\,\sin 2\alpha$

The above matrix function

$$\mathbf{A}(x) \;\equiv\; \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} + x \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \;=\; \begin{pmatrix} \lambda_1 & x \\ -x & \lambda_2 \end{pmatrix}$$

has been investigated by Arnold [Ar 79] for Ornstein-Uhlenbeck noise as an example of an unstable (if $\lambda_1 > 0$) system being stabilized by random parameter noise; also see Arnold and Kliemann [Ar-Kl 83, p.67f.]. However, here we are not interested in stabilization (as $t \to \infty$), but in the behavior on time scales; more precisely, the parameter noise X^ε will be assumed as small on the time scales under consideration which embodies that 0 is a metastable point of X^ε; in other words, 0 will be supposed to be a local minimum of U entailing that the skew-symmetric "perturbation" $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ of $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ is small on the time scales on which X^ε sojourns near 0. This terminology of metastability will be made precise in the subsequent chapter. The matrix function $\mathbf{A}$ of this example provides a toy model for the following investigations; more precisely, our findings of subsection 4.4.2 will turn out to be applicable to this model; see example 4.4.14.

Furthermore, note that "superposing" the matrices $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is also a popular sample model for investigating the parameter dependence of Lyapunov exponents in the case of white noise; see Pardoux and Wihstutz [Pd-Wh 88, p.455f.] and [Pd-Wh 92, p.293]. □

After having discussed our main examples in 1.5.3 and 1.5.4 we would like to close this section by giving references for further sample systems without going into details any further.

Example 1.5.5 (nonlinear, stochastic systems driven by real noise). Here, one considers nonlinear systems

$$\frac{d}{dt}\, y_t \;=\; F(X_t, y_t) \ ,$$

where the motion of the vector $y_t \in \mathbb{R}^n$ describes the evolution of the system under consideration and the other influences are modeled as perturbing real noise process $X \equiv X^\varepsilon$, its state space being $\mathbb{R}^d$ for example; F is a drift function on the joint state space which is assumed to be differentiable. If $\mathcal{O}$ is an equilibrium of $F(x, \,.\,)$, i.e. if $F(x, \mathcal{O}) = 0$ for all x, then linearization at $\mathcal{O}$ leads to the real noise system

$$\frac{d}{dt}\, Z_t \;=\; \mathbf{A}(X_t)\, Z_t \ ,$$

where $\mathbf{A}(x) := D_y\, F(x, y)|_{y=\mathcal{O}}$, where D_y denotes the differentiation operator with respect to the y-variable; see e.g. Arnold and Kliemann [Ar-Kl 83, p.68]; such a system with dichotomous ("telegraphic") noise, i.e. X_t being a Markov process with two states, has been considered by Arnold and Kloeden [Ar-Kd 89, p.1242f.]; the latter authors also point towards a real-world system in electrohydrodynamics: See Behn et al. [Bh-Lg-Jh 98] and Müller and Behn [M-Bh 87] for details.

Another physical example for such a linearized system with telegraphic noise is the LRC circuit described by Kats and Martynyuk [Ka-My 02, p.44 & p.20f.]; note that their linearized system only takes its values in the diagonal matrices. Such a situation will constitute the subject of our investigations in section 4.3.

For applications e.g. to economics, we refer to Ruelle [Ru 88] and Weintraub [Wt 70] among others. □

Example 1.5.6 (Harmonic oscillator with real-noise input). Consider the random differential equation

$$\ddot{z} + \left(2\beta + a_2 X_t^2\right)\dot{z} + \left(\varpi^2 + a_1 X_t^1\right) z = 0\,,$$

where $\varpi, \beta \in \mathbb{R}$ and $a_1, a_2 \geq 0$ are constants and X_t^1, X_t^2 are the components of a stochastic process $X \equiv X^\varepsilon \in \mathbb{R}^2$. Defining

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} := \begin{pmatrix} z \\ \dot{z} \end{pmatrix},$$

this equation can be rewritten as usual as

$$\frac{d}{dt}\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \left[\begin{pmatrix} 0 & 1 \\ -\varpi^2 & -2\beta \end{pmatrix} + a_1 X_t^1 \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} + a_2 X_t^2 \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}\right]\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}.$$

In this example $a_1 X^1$ is considered as random restoring force, whereas $a_2 X^2$ is thought of as random perturbation of the damping constant β. In case that X is a fixed stochastic process, the related control theoretic analysis is undertaken for different specifications of the parameters by Benderskii and Pastur [Be-Pt 75], Kliemann and Rümelin [Kl-Rm 81, p.18-24], Arnold and Kliemann [Ar-Kl 83, p.12 & p.63ff.], Kliemann [Kl 88, p.91ff.], Kliemann [Kl 80, p.120,158], Kliemann [Kl 79, p.468], Wihstutz [Wh 75], Arnold and Wihstutz [Ar-Wh 78] and Rümelin [Rm 78] and [Rm 79]; also see Pinsky and Wihstutz [Pi-Wh 91], Arnold et al. [Ar-Pp-Wh 86], Pinsky [Pi 86], Arnold [Ar 98, p.160] and Hernández-Lerma [HL 79, p.39f.]. Applications to engineering systems are given by Griesbaum [Gb 99, p.22]; for the related modeling of the ship roll motion in particular of the capsizing of vessels see Colonius and Kliemann [Cu-Kl 00, p.497] and [Cu-Kl 97, p.137f.].

Benderskii [Be 69] considers the following oscillator with telegraphic noise,

$$\ddot{z} + p\dot{z} + (q - \gamma\, r_t)\, z = 0\,,$$

where $p, q, \gamma > 0$ are constants and $(r_t)_t$ is a random process assuming the values ± 1. The times at which $(r_t)_t$ changes its sign form a Poisson process with a certain intensity.

The case that X is a fixed telegraphic noise process (i.e. a stationary, ergodic two-state Markov process) is treated by Arnold and Kloeden [Ar-Kd 89, p.1269f.]; more precisely, the latter authors investigate asymptotic formulas for $a_1 \to 0$ and $a_1 \to \infty$ for the case that $\beta = a_2 = 0$; this case has also been considered by Leizarowitz [Lz 89] who assumes the real noise to be a finite-state Markov process.

For white noise perturbations of the harmonic oscillator see Pardoux and Wihstutz [Pd-Wh 88, p.450] and Pinsky and Wihstutz [Pi-Wh 88] among others. $\square$

Triangular matrices have been of interest, too: For explicit calculations on upper triangular matrices see Arnold [Ar 98, 3.3.13 & 3.4.16] and the references therein. Lower triangular matrices are e.g. considered by Kliemann [Kl 80, p.157] and Kliemann and Rümelin [Kl-Rm 81, p.15].

Brockett and Willems [Bro-Wil 72, p.253] propose a model for a closed loop dynamics with feedback interaction which leads to a linear real noise differential system.