Developments in Mathematics

Russell Johnson
Rafael Obaya
Sylvia Novo
Carmen Núñez
Roberta Fabbri

Nonautonomous Linear Hamiltonian Systems: Oscillation, Spectral Theory and Control

Springer

Developments in Mathematics

VOLUME 36

Series editors
Krishnaswami Alladi, *University of Florida, Gainesville, FL, USA*
Hershel M. Farkas, *Hebrew University of Jerusalem, Jerusalem, Israel*

More information about this series at http://www.springer.com/series/5834

Russell Johnson • Rafael Obaya • Sylvia Novo •
Carmen Núñez • Roberta Fabbri

Nonautonomous Linear Hamiltonian Systems: Oscillation, Spectral Theory and Control

Springer

Russell Johnson
Università degli Studi di Firenze
Firenze, Italy

Rafael Obaya
Universidad de Valladolid
Valladolid, Spain

Sylvia Novo
Universidad de Valladolid
Valladolid, Spain

Carmen Núñez
Universidad de Valladolid
Valladolid, Spain

Roberta Fabbri
Università degli Studi di Firenze
Firenze, Italy

ISSN 1389-2177 ISSN 2197-795X (electronic)
Developments in Mathematics
ISBN 978-3-319-29023-2 ISBN 978-3-319-29025-6 (eBook)
DOI 10.1007/978-3-319-29025-6

Library of Congress Control Number: 2016935697

Mathematics Subject Classification: 37B55, 34C10, 54H20, 34D08, 34D09, 34H05, 49N05, 37D25, 53D12, 37C55

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

This book is devoted to a study of the oscillation theory of nonautonomous linear Hamiltonian differential systems and that of a spectral theory which is adapted to such systems. Systematic use will be made of basic facts concerning Lagrange subspaces of $\mathbb{R}^{2n}$ and argument functions on the set of symplectic matrices. We will also consistently apply some fundamental methods of topological dynamics and of ergodic theory, including Lyapunov exponents, exponential dichotomies, and rotation numbers. Further, we will show that our results concerning oscillation theory can be fruitfully applied to several basic issues in the theory of linear-quadratic control systems with time-varying coefficients.

Nonautonomous Oscillation Theory

In due course, we will give an outline of the specific problems, methods, and results to be discussed in the body of the book. Before doing that, it seems appropriate to collocate them in a priori way in the vast and nonhomogeneous area called oscillation theory of ordinary differential equations. In fact, the word "oscillation" has various meanings in this context. For example, it can refer to the study of the zeroes contained in some interval $\mathcal{I} \subseteq \mathbb{R}$ of a solution of an ordinary differential equation (ODE). In the case of a two-dimensional ODE, it can refer to the variation of the polar angle along a solution, i.e., to the "rotation" associated to that solution. Still again, it may indicate one of the many themes encountered in the study of the periodic solutions of an ordinary differential equation.

This book is about "rotation." Let us try to be a bit more precise. We will focus attention on various issues concerning the solutions of a linear Hamiltonian differential system

$$\mathbf{z}' = H(t)\,\mathbf{z}\,, \tag{1}$$

where $\mathbf{z} \in \mathbb{R}^{2n}$ and $t \in \mathbb{R}$. The coefficient $H(\cdot)$ is a bounded measurable real $2n \times 2n$ matrix-valued function satisfying the symplectic condition $(JH)^T(t) = JH(t)$ for all $t \in \mathbb{R}$, where the "T" indicates the transpose and $J = \left[\begin{smallmatrix} 0_n & -I_n \\ I_n & 0_n \end{smallmatrix}\right]$ is the usual $2n \times 2n$ antisymmetric matrix: I_n is the $n \times n$ identity matrix and 0_n the $n \times n$ zero matrix. Generally speaking, we will be interested in the "rotation" of the solutions of (1). Of course, this notion is initially problematic because it is not immediately clear how to define it precisely, especially if $n \geq 2$. One of our main goals will be to do this. It will turn out that our concept of rotation is closely related to a more or less standard notion of a "point of verticality" of a solution of (1), namely, a focal point. It will also turn out that the concept of rotation considered here can be used to study some basic questions in spectral theory, which are formulated in terms of equation (1) and which will be discussed shortly.

Equation (1) is of course very significant. As a special case, one can set $\mathbf{z} = \left[\begin{smallmatrix} \mathbf{x} \\ \mathbf{y} \end{smallmatrix}\right]$ for $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, and

$$H(t) = \begin{bmatrix} 0_n & I_n \\ G(t) & 0_n \end{bmatrix},$$

where $G^T = G$ is a real symmetric $n \times n$ matrix-valued function. Then (1) is equivalent to the second-order system

$$\mathbf{x}'' = G(t)\,\mathbf{x}\,, \tag{2}$$

which is often encountered in the study of mechanical systems near an equilibrium. Another special case is obtained by setting $n = 1$ and

$$H(t) = \begin{bmatrix} 0 & 1/p(t) \\ g(t) - \lambda\, d(t) & 0 \end{bmatrix}$$

for a real parameter λ; in this case (1) is equivalent to the classical Sturm–Liouville problem

$$-\,(p\,x')' + g(t)\,x = \lambda\, d(t)\,x\,. \tag{3}$$

Problem (3) has been studied with success from various points of view for over 150 years. The number and the location of the zeroes of a solution $x(\cdot)$ are a recurring theme. Information concerning these zeroes has implications for the spectral problem obtained by varying λ and by imposing boundary conditions, for example, of Dirichlet type: $x(a) = x(b) = 0$ where $a < b \in \mathbb{R}$. Then, as is well known, if p, g, and d satisfy certain general hypotheses, then the nth eigenfunction of (3) has $n - 1$ zeroes in (a, b), for $n = 1, 2, \ldots$

A more general spectral problem is obtained by using (1) as a point of departure. One introduces a parameter $\lambda \in \mathbb{R}$ and a positive semidefinite real weight function $\Gamma(t)$ in (1), so as to obtain

$$\mathbf{z}' = \big(H(t) + \lambda J^{-1} \Gamma(t)\big)\, \mathbf{z}\,. \tag{4}$$

This problem was studied systematically by Atkinson in [5]. It is noteworthy that if Γ is semidefinite but not everywhere definite, then the study of the boundary-value problem associated to (4) cannot be naturally carried out using standard functional-analytic techniques (due to the fact that one cannot multiply (4) by Γ^{-1}). However, in [5], one finds an "Atkinson condition" which, when imposed on (4), allows the development of a satisfactory spectral theory for (4).

Another of our goals is to show that our oscillation theory of (1) can be fruitfully applied to the spectral problem (4) especially when "the boundary conditions are imposed at $t = \pm\infty$," i.e., when (4) is considered on the whole line. Let us explain some of the issues involved in relating oscillation theory and spectral theory in the context of problem (4). Consider for a moment the version of (3) obtained by setting $p = d \equiv 1$:

$$-\,x'' + g(t)\,x = \lambda\,x\,. \tag{5}$$

This is the Schrödinger equation with potential $g(t)$ (a most important ordinary differential equation, due to its basic role in one-dimensional quantum mechanics). Fix $\lambda \in \mathbb{R}$, and consider a solution $x(t)$ of (5), say, that defined by the initial conditions $x(a) = 0$ and $x'(a) = 1$. This solution is called nonoscillatory in the interval (a, b) if it has no zeroes there; otherwise, it oscillates. There is a simple and fruitful way to study the presence/absence of zeroes of $x(\cdot)$ on (a, b), which is at the heart of the classical Sturm–Liouville theory. Namely, one introduces the polar angle $\theta(t)$ of the vector $\left[\begin{smallmatrix} x(t) \\ x'(t) \end{smallmatrix}\right]$ in the two-dimensional phase plane $\mathbb{R}^2$. It is clear that if $a < t < b$, then $x(t) = 0$ if and only if $\theta(t) = \pi/2 \bmod \pi$. Moreover, $\theta'(t) < 0$ at each zero t of $x(t)$, so we can determine the number of zeroes of $x(\cdot)$ in (a, b) by studying the evolution of $\theta(\cdot)$ there, that is, the "rotation" of $x(\cdot)$.

This simple observation does not generalize easily to the Hamiltonian system (1). It is rather straightforward to generalize the concept of zero of $x(\cdot)$: one sets $\mathbf{z} = \left[\begin{smallmatrix} \mathbf{x} \\ \mathbf{y} \end{smallmatrix}\right]$, requires that $\mathbf{x}(t) = \mathbf{0}$, and arrives at the concept of focal point, alias point of verticality. But it is not easy to extend the concept of polar angle in an appropriate way; in fact, it seems that this was only done in the 1950s and 1960s. One way is to introduce argument functions in the symplectic group, as done by Gel'fand, Lidskii, and Yakubovich. Another is to introduce the Maslov cycle and the corresponding Maslov index in the manifold of Lagrange subspaces of $\mathbb{R}^{2n}$. There is a corresponding angle, as was pointed out by Arnol'd (and by Conley in a little-known paper), which can be used to develop a Sturm–Liouville-type theory for (4). Still another method to generalize the Sturm–Liouville theory to Hamiltonian systems can be based on the polar coordinates of Barret and Reid.

A point which we will emphasize in this book is that one can study the argument functions, the index, and the polar coordinates from a dynamical point of view, more precisely, by using basic tools from topological dynamics and ergodic theory. One point of arrival in our theory is a quantity called the rotation number and its "complexification," the Floquet exponent for system (1). Using these quantities, we will connect the oscillation theory of (1) with the spectral theory of the Atkinson problem (4), much as the Sturm–Liouville theory connects the oscillation of solutions of (3) for each fixed λ to the spectral theory of (3).

Let us explain this matter in more detail. Let $\Gamma \geq 0$ be a real symmetric matrix-valued function. Consider the boundary-value problem

$$\begin{aligned} &\mathbf{z}' = \big(H(t) + \lambda\, J^{-1}\Gamma(t)\big)\,\mathbf{z}\,, \qquad \mathbf{z} = \begin{bmatrix}\mathbf{x}\\ \mathbf{y}\end{bmatrix} \in \mathbb{R}^{2n}\,, \\ &\mathbf{x}(a) = \mathbf{x}(b) = \mathbf{0}\,, \end{aligned} \tag{6}$$

where $a < b \in \mathbb{R}$. In [5] an analytic theory of the eigenvalues and eigenfunctions of (6) is worked out. Let us first try to extend that theory to the entire real axis: thus set $a = -\infty$ and $b = \infty$. One can expect that this will involve some analogue of the classical Weyl m-functions $m_{\pm}(\lambda)$ for (3), and in fact there is a rich literature concerning the "Weyl–Titchmarsh M-matrices" for (6). We will assume that $H(\cdot)$ and $\Gamma(\cdot)$ are uniformly bounded and will impose a natural "Atkinson condition" on the solutions of (5). It will then turn out that the dynamical concept of exponential dichotomy together with the above-mentioned notion of rotation number permits one to develop a satisfactory spectral theory for (6) with $a = -\infty$ and $b = \infty$. In particular, the introduction of the exponential dichotomy concept permits one to clarify the dynamical significance of the M-matrices.

To summarize what has been said so far, we will supplement the analytic methods which have been previously used to study the oscillation theory of (1) and the spectral theory of (4) with certain geometrical and dynamical techniques. The geometrical methods derive from the structure of the group of symplectic matrices and from that of the manifold of Lagrangian subspaces of $\mathbb{R}^{2n}$. Using dynamical methods, we define the rotation number and the Floquet exponent, which permit one to count the focal points of (1) and to develop the spectral theory of (4) using the exponential dichotomy concept.

The use of dynamical methods is made possible by carrying out a construction named after Bebutov, which we now explain. Begin with linear Hamiltonian differential system (1): we first view the coefficient function $H(\cdot)$ as an element of an appropriate functional space. This will often be the space of bounded continuous functions $\widetilde{H}$ from $\mathbb{R}$ to the Lie algebra of real infinitesimally symplectic matrices $\mathfrak{sp}(n,\mathbb{R}) = \{\widetilde{H} \in \mathbb{M}_{2n\times 2n}(\mathbb{R}) \mid \widetilde{H}^T J + J\widetilde{H} = 0_{2n}\}$. Next introduce the translation flow σ_t by setting $\sigma_t(\widetilde{H})(\cdot) = \widetilde{H}(\cdot + t)$ for all $t \in \mathbb{R}$. If the coefficient $H(\cdot)$ of (1) is uniformly continuous, then the closure $\mathrm{cls}\{\sigma_t(H) \mid t \in \mathbb{R}\}$ is compact (in the compact-open topology). Call the closure Ω: it is clearly invariant with respect to the translation flow. The idea now is to let H vary over Ω; to emphasize that we do not deal only with the "original" function $H(\cdot)$, we write ω to indicate a generic point of Ω. Note that each $\omega \in \Omega$ gives rise to a linear differential system of the form (1); call this system $(1)_\omega$.

At this point, one introduces the so-called cocycle obtained by considering the fundamental matrix solution of $(1)_\omega$ and letting ω run over Ω. One can now apply the Oseledets theory of the Lyapunov indices of solutions of $(1)_\omega$ $(\omega \in \Omega)$. One can also apply the Sacker–Sell–Selgrade approach to the theory of exponential dichotomies. In addition, one can define the rotation number of the family of equations $(1)_\omega$. We will see that all these dynamical methods permit one to gain important insight into the oscillation theory of (1) and the spectral theory of (4).

In fact the main tool in the analysis consists in the systematic use of the rotation number, the Lyapunov index, the exponential dichotomy concept, and the Weyl matrices. These objects are also important in the discussion of two more notions which are of fundamental significance in the context of the linear Hamiltonian system (1): the property of disconjugacy, which is of basic significance in the calculus of variations, and the related property of existence of principal solutions, which in many interesting cases can be understood as a generalization to the nonuniformly hyperbolic case of the bundles provided by the existence of exponential dichotomy.

Applications to Control Theory

There are numerous applications of the oscillation theory of equation (1) to the theory of mechanical systems, to the calculus of variations, to control theory, and to other areas. We will not give an exhaustive account of these applications. But we will apply our results concerning equations (1) and (4) to certain problems in linear-quadratic (LQ) control theory. Among these are the linear-quadratic regulator problem, the Kalman–Bucy filter, the Yakubovich frequency theorem, and the question of Willems-type dissipativity in (linear) control systems. We now discuss in a bit more detail these applications to control theory.

First we recall the formulation of the LQ regulator problem. The point of departure consists of a linear control problem

$$\begin{aligned} &\mathbf{x}' = A(t)\,\mathbf{x} + B(t)\,\mathbf{u}\,, \quad \mathbf{x} \in \mathbb{R}^n,\ \mathbf{u} \in \mathbb{R}^m\,, \\ &\mathbf{x}(0) = \mathbf{x}\,. \end{aligned} \tag{7}$$

The matrices $A(\cdot)$, $B(\cdot)$ are taken to be bounded continuous functions; the time dependence is otherwise arbitrary. Let $\tau \in (0,\infty]$ be an extended positive real number. Introduce a quadratic functional

$$\mathcal{I}_{\mathbf{x}}(\mathbf{x},\mathbf{u}) = \langle \mathbf{x}(\tau), S\mathbf{x}(\tau)\rangle + \int_0^\tau \big(\langle \mathbf{x}(t), G(t)\,\mathbf{x}(t)\rangle + \langle \mathbf{u}(t), R(t)\,\mathbf{u}(t)\rangle\big)\,dt\,.$$

where S is a symmetric positive semidefinite matrix and $G(\cdot)$, $R(\cdot)$ are bounded continuous functions such that $G^T(t) = G(t) \geq 0$ and $R^T(t) = R(t) > 0$ for all $t \in \mathbb{R}$. If the upper limit τ is finite, one speaks of a finite-horizon problem, otherwise one has an infinite-horizon problem. If $\tau = \infty$ one sets $S = 0_n$. For each fixed initial condition $\mathbf{x} \in \mathbb{R}$, one seeks a control $\mathbf{u}\colon [0, \tau] \to \mathbb{R}^m$ which, when taken together with the corresponding solution of (7), minimizes $\mathcal{I}_{\mathbf{x}}(\mathbf{x}, \mathbf{u})$.

This basic problem has been studied in detail and has been solved both when $\tau < \infty$ and when $\tau = \infty$. Our contribution is to give a solution in the infinite-horizon case $\tau = \infty$ which uses the theory of exponential dichotomies and the rotation number as applied to an appropriate linear Hamiltonian system of the form (1). In this way one obtains, among other things, detailed information concerning the regular dependence of the optimal control on parameters.

The appropriate system (1) is obtained via a formal application of the Pontryagin maximum principle. According to this principle, a minimizing control $\mathbf{u}$ must maximize the Hamiltonian

$$\mathcal{H}(t, \mathbf{x}, \mathbf{y}, \mathbf{u}) = \langle \mathbf{y}, A(t)\, \mathbf{x} + B(t)\, \mathbf{u} \rangle - \frac{1}{2}\left(\langle \mathbf{x}, G(t)\, \mathbf{x} \rangle + \langle \mathbf{u}, R(t)\, \mathbf{u} \rangle \right),$$

for each $t \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$, and an appropriate $\mathbf{y} \in \mathbb{R}^n$. Here $\mathbf{y}$ is interpreted as a variable dual to $\mathbf{x}$. This leads immediately to the "feedback rule"

$$\mathbf{u} = R^{-1}(t)\, B^T(t)\, \mathbf{y}\,.$$

Substituting for $\mathbf{u}$ in the Hamiltonian equations $\mathbf{x}' = \partial\mathcal{H}/\partial\mathbf{y}$, $\mathbf{y}' = -\partial\mathcal{H}/\partial\mathbf{x}$ leads to the differential system

$$\mathbf{z}' = \begin{bmatrix} A(t) & B(t)\, R^{-1}(t)\, B^T(t) \\ G(t) & -A^T(t) \end{bmatrix} \mathbf{z}\,. \tag{8}$$

Of course, (8) is a special case of (1).

We now arrive at the main point, which is that (under standard controllability and observability conditions on (7)) the system (8) admits exponential dichotomy. This is easily proved when one has available the basic facts concerning the rotation number of (8) and its relation to the existence of exponential dichotomy. Now, the existence of exponential dichotomy for (8) means that there is a linear projection $P = P^2\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ such that if $\mathbf{z} = \left[\begin{smallmatrix} \mathbf{x} \\ \mathbf{y} \end{smallmatrix}\right]$ is in the image of P, then the solution $\mathbf{z}(t)$ of (8) satisfying $\mathbf{z}(0) = \mathbf{z}$ decays exponentially as $t \to \infty$. It further turns out that $\mathbf{z}(t) = \left[\begin{smallmatrix} \mathbf{x}(t) \\ \mathbf{y}(t) \end{smallmatrix}\right] = \left[\begin{smallmatrix} \mathbf{x}(t) \\ M(t)\, \mathbf{x}(t) \end{smallmatrix}\right]$ where $\mathbf{x}(0) = \mathbf{x}$ and $M(t)$ is a function taking values in the set of negative definite symmetric $n \times n$ matrices. Set $\mathbf{u}(t) = R^{-1}(t)\, B^T(t)\, M(t)\, \mathbf{x}(t)$ and note that $\mathbf{u}(t) \to \mathbf{0}$ exponentially as $t \to \infty$. So it is not so surprising that this $\mathbf{u}$ is in fact the unique control which minimizes $\mathcal{I}_{\mathbf{x}}(\mathbf{x}, \mathbf{u})$. If one varies $\mathbf{x}$, the dichotomy projection P and the symmetric matrix-valued function $M(t)$ do not change, so in fact we have solved the LQ regulator problem.

Let us note in passing that we have also solved the feedback stabilization problem for the control system (7). In fact, set $\mathbf{u}(t) = R^{-1}(t)\,B^T(t)\,M(t)\,\mathbf{x}(t)$ as above. Note that if $\mathbf{z}(t) = \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{y}(t) \end{bmatrix}$ is the solution of (8) mentioned above, then $\mathbf{x}(t)$ solves (7) with precisely this control $\mathbf{u}(t)$. Since $\mathbf{u}$ has the "feedback form" $\mathbf{u}(t) = K(t)\,\mathbf{x}(t)$ with $K(t) = R^{-1}(t)\,B^T(t)\,M(t)$, and since the linear system $\mathbf{x}' = \big(A(t) + B(t)\,R^{-1}(t)\,B^T(t)\,M(t)\big)\,\mathbf{x}$ is exponentially stable, we have "feedback stabilized" the system (7).

We can also study certain important properties of the Kalman–Bucy filter by applying our methods to an appropriate Hamiltonian system of the form (1). This is because, as Kalman and Bucy observed, the construction of their filter is closely tied to a "time-reversed" LQ regulator problem. We briefly describe the filter and the relevance of the theory of linear Hamiltonian systems in this context.

Let $\boldsymbol{\xi}(t) \in \mathbb{R}^n$ $(t \geq 0)$ denote the state of a linear system which is disturbed by a d-dimensional white noise process: thus

$$d\boldsymbol{\xi}(t) = A(t)\,\boldsymbol{\xi}(t)\,dt + S(t)\,d\mathbf{w}(t)\,. \tag{9}$$

Here $\mathbf{w}(t)$ is a d-dimensional standard Brownian motion, and equation (9) is understood to be of Itô type. The state $\boldsymbol{\xi}(t)$ can only be partially observed; it is assumed that the observation process $\boldsymbol{\eta}(t)$ satisfies the Itô equation:

$$d\boldsymbol{\eta}(t) = B(t)\,\boldsymbol{\xi}(t)\,dt + S_1(t)\,d\mathbf{w}_1(t)\,.$$

where $\mathbf{w}_1(t)$ is a second, m-dimensional Brownian motion which is independent of $\mathbf{w}(t)$. The functions A, B, S, S_1 are assumed to be continuous and bounded and to have the appropriate dimensions. It is assumed that $\boldsymbol{\eta}(0) = \mathbf{0}$ and that $\boldsymbol{\xi}(0)$ is Gaussian, which implies that $\boldsymbol{\xi}(t)$ is Gaussian for all $t \geq 0$.

Let Σ_t be the σ-algebra generated by the set $\{\boldsymbol{\eta}(r) \mid 0 \leq r \leq t\}$ of measurements up to time t. The goal is to describe an estimate $\boldsymbol{\gamma}(t)$ for $\boldsymbol{\xi}(t)$, which minimizes the mean-square error $E\{(\mathbf{x}^T(\boldsymbol{\xi}(t) - \boldsymbol{\gamma}(t)))^2\}$ for all vectors $\mathbf{x} \in \mathbb{R}^n$; here the expected value $E\{\cdot\}$ is taken over an appropriate probability space. It is well known that this best estimate is given by the conditional expectation

$$\boldsymbol{\gamma}(t) = \widehat{\boldsymbol{\xi}}(t) = E\,\{\boldsymbol{\xi}(t) \mid \Sigma_t\}\,.$$

To describe $\widehat{\boldsymbol{\xi}}(t)$, one introduces the error process $\tilde{\boldsymbol{\xi}}(t) = \boldsymbol{\xi}(t) - \widehat{\boldsymbol{\xi}}(t)$. It turns out that $\tilde{\boldsymbol{\xi}}(t)$ is Gaussian with mean value zero and hence is determined by its $n \times n$ covariance matrix $M(t)$. Kalman and Bucy showed that $M(t)$ satisfies a Riccati equation

$$M' = -M\,B^T(t)\,(S_1 S_1^T)^{-1}(t)\,B(t)\,M + M\,A^T(t) + A(t)\,M + (SS^T)(t)\,.$$

Now, this Riccati equation corresponds to the linear Hamiltonian system

$$\mathbf{z}' = \begin{bmatrix} -A^T(t) & B^T(t)\,(S_1 S_1^T)^{-1}(t)\,B(t) \\ (SS^T)(t) & A(t) \end{bmatrix} \mathbf{z}\,, \tag{10}$$

via the matrix change of variables $M = YX^{-1}$. It turns out that, under standard controllability conditions, the system (10) admits exponential dichotomy. This leads to the conclusion that $M(t)$ tends exponentially fast to a "nonautonomous equilibrium" $M_\infty(t)$, which essentially describes the error process $\tilde{\boldsymbol{\xi}}(t)$, and hence the signal $\boldsymbol{\xi}(t)$ if one takes the estimate $\widehat{\boldsymbol{\xi}}(t)$ to be known.

We will also apply our results concerning the oscillation theory of equation (1) and the spectral theory of the family (4) to the circle of ideas and results centered on the Yakubovich frequency theorem. This theorem was originally formulated and proved by Yakubovich for LQ control processes with periodic coefficients. We will state and prove a more general nonautonomous version of this theorem. We briefly sketch our results in this regard in the next paragraphs.

The point of departure is again the control system (7) combined with a quadratic functional

$$\widetilde{\mathcal{I}}_{\mathbf{x}}(\mathbf{x},\mathbf{u}) = \int_0^\infty \big(\langle \mathbf{x}, G(t)\,\mathbf{x}\rangle + 2\langle \mathbf{x}, g(t)\,\mathbf{u}\rangle + \langle \mathbf{u}, R(t)\,\mathbf{u}\rangle\big)\,dt\,,$$

where the functions A, B, G, g, R are assumed to be bounded and continuous and to have the appropriate dimensions. The functional $\widetilde{\mathcal{I}}_{\mathbf{x}}(\mathbf{x},\mathbf{u})$ differs from the one encountered in the context of the LQ regulator in two respects. First of all, the cross-term $\langle \mathbf{x}, g(t)\,\mathbf{u}\rangle$ is present in the integrand. Second and more importantly, though it is assumed that $G^T(t) = G(t)$ and that $R^T(t) = R(t) > 0$ for all t, it is not assumed that G is positive semidefinite for all t; indeed one is particularly interested in the case when $G(t) < 0$ ($t \in \mathbb{R}$).

We pose the problem of minimizing $\widetilde{\mathcal{I}}_{\mathbf{x}}(\mathbf{x},\mathbf{u})$ subject to (7). Since G is not assumed to be positive semidefinite, this problem need not have a solution. Nevertheless we proceed by applying the Pontryagin maximum principle in a formal way. Introduce the Hamiltonian

$$\widetilde{\mathcal{H}}(t,\mathbf{x},\mathbf{y},\mathbf{u}) = \langle \mathbf{y}, A(t)\,\mathbf{x} + B(t)\,\mathbf{u}\rangle - \frac{1}{2}\,\big(\langle \mathbf{x}, G(t)\,\mathbf{x}\rangle + 2\langle \mathbf{x}, g(t)\,\mathbf{u}\rangle + \langle \mathbf{u}, R(t)\,\mathbf{u}\rangle\big)\,.$$

A minimizing control $\mathbf{u}$ (if it exists) will maximize $\widetilde{\mathcal{H}}$ for each $t \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$, and an appropriate $\mathbf{y} \in \mathbb{R}^n$. This leads to the feedback rule

$$\mathbf{u} = R^{-1}(t)\,B^T(t)\,\mathbf{y} - R^{-1}(t)\,g^T(t)\,\mathbf{x}\,,$$

and via the Hamiltonian equations $\mathbf{x}' = \partial\widetilde{\mathcal{H}}/\partial\mathbf{y}$, $\mathbf{y}' = -\partial\widetilde{\mathcal{H}}/\partial\mathbf{x}$, one is led to the differential system

$$\mathbf{z}' = H(t)\,\mathbf{z}\,, \qquad \text{with } H = \begin{bmatrix} A - B\,R^{-1}g^T & B\,R^{-1}B^T \\ G - g\,R^{-1}g^T & -A^T + g\,R^{-1}B^T \end{bmatrix}. \tag{11}$$

In the case when all the coefficients in (11) are T-periodic, Yakubovich showed that the minimization problem admits a solution if and only if (i) the system (11) has exponential dichotomy (frequency condition) and (ii) certain solutions of (11) have no focal points (nonoscillation condition). We will consider the case when A, B, G, g, R are bounded continuous functions of time and prove a satisfactory generalization of Yakubovich's theorem. It turns out that the frequency condition and the nonoscillation condition (which can be stated as above) imply that the optimal control problem can be solved for all $\mathbf{x} \in \mathbb{R}^n$. The converse statement is not quite true; as a matter of fact, and roughly speaking, the minimizing control must exhibit a uniform continuity condition in order to ensure that the frequency condition and the nonoscillation condition are valid.

The frequency theorem has many ramifications and applications, some of which will be considered in this book. Here we mention that the frequency theorem can be used to comment on the Willems concept of dissipativity in the context of control systems. This connection was pointed out and analyzed in the periodic case, by Yakubovich et al. [158]. We will discuss the connection between the frequency theorem and the Willems dissipativity concept when the relevant coefficients are aperiodic functions of time.

The main point here is to interpret the integrand of the functional $\widetilde{\mathcal{I}}_{\mathbf{x}}(\mathbf{x},\mathbf{u})$ as a power function. To explain this, set $\mathbf{x} = \mathbf{0}$ in equation (7). Let $\mathbf{u}\colon [t_1, t_2] \to \mathbb{R}^m$ be an integrable function, and let $\mathbf{x}(t)$ be the corresponding solution of (7) with $\mathbf{x}(t_1) = \mathbf{0}$. Let us write

$$\mathcal{Q}(t,\mathbf{x},\mathbf{u}) = \frac{1}{2}\left(\langle \mathbf{x}, G(t)\,\mathbf{x}\rangle + 2\langle \mathbf{x}, g(t)\,\mathbf{u}\rangle + \langle \mathbf{u}, R(t)\,\mathbf{u}\rangle\right) .$$

Then the net energy entering the system due to the effect of $\mathbf{u}(\cdot)$ is obtained by integrating $\mathcal{Q}(t,\mathbf{x}(t),\mathbf{u}(t))$ in the interval $[t_1, t_2]$. Now one says that the system is dissipative if

$$\int_{t_1}^{t_2} \mathcal{Q}(s,\mathbf{x}(s),\mathbf{u}(s))\,ds \geq 0$$

whenever $t_1 < t_2 \in \mathbb{R}$. That is, "energy must be expended" to move the system from its equilibrium position $\mathbf{x} = \mathbf{0}$.

The basic result which we will prove is that, modulo details, the control system determined by (7) together with $\mathcal{Q}(t,\mathbf{x},\mathbf{u})$ is (strongly) dissipative if and only if the Hamiltonian system (11) satisfies the frequency condition and the nonoscillation condition. So the frequency theorem has deep consequences concerning the structure of LQ control processes.

Outline of the Contents

We end this introduction with a brief outline of the contents of the various chapters which will follow.

The long Chap. 1 contains a discussion of various tools from topological dynamics and from ergodic theory which will be systematically used throughout the book. We discuss the Birkhoff theorem and the Oseledets theorem, the Bebutov construction and some facts concerning flows, the Sacker–Sell–Selgrade theory of exponential dichotomies, and other matters as well.

Chapters 2 and 3 contain the basic theory of the oscillation of the solutions of (1), respectively, as well as a dynamical approach to the spectral theory of the Atkinson problem (4). In Chap. 2, we construct and discuss the rotation number for (1), which is roughly speaking "the average number of focal points" admitted by a so-called conjoined basis of solutions. This quantity can be defined in several ways, using the Gel'fand–Lidskii–Yakubovich argument functions, the Maslov index, and the Barrett–Reid polar angles. In Chap. 3 we complexify the rotation number so as to obtain the Floquet exponent, a quantity which is quite useful in the study of problem (4). We state and prove a basic result, namely, that if (4) satisfies an Atkinson condition, then the rotation number $\alpha = \alpha(\lambda)$ of (4) is constant for λ in an open subinterval $\mathcal{I} \subset \mathbb{R}$ if and only if (4) admits exponential dichotomy for all $\lambda \in \mathcal{I}$.

The Weyl M-matrices, or M-functions, arise in Chap. 3 as a tool used in the study of the spectral theory of (4) and especially in the proof of the theorem relating the constancy of the rotation number to the presence of exponential dichotomy. The M-functions are defined for nonreal values of the parameter λ. However, it is very important to understand their convergence properties in the limit as $\operatorname{Im}\lambda$ tends to zero, and Chap. 4 is dedicated to a study of this issue. In particular, we work out an extension to the Atkinson problem (4) of the classical Kotani theory, which is an important tool in the study of the refined spectral properties of the Schrödinger operator.

The notion of disconjugacy is very important in the context of the Hamiltonian linear differential system (1), because of its significance in the calculus of variations. Chapter 5 is devoted to a discussion of a generalization of the concept of disconjugacy, namely, weak disconjugacy. Under natural and mild auxiliary hypotheses, we prove the existence of a principal solution when (1) is weakly disconjugate. Our approach to the issue of (weak) disconjugacy relies on the systematic use of tools of topological dynamics; these allow a deep understanding of the conditions under which weak disconjugacy holds and also of the properties of the principal solutions.

The book concludes with Chap. 6 (the LQ regulator problem and the Kalman–Bucy filter), Chap. 7 (the nonautonomous version of the Yakubovich frequency theorem), and Chap. 8 (Willems dissipativity for LQ control processes).

Note finally that, in this book, methods and results which have been developed in the course of 100 years in the context of linear Hamiltonian systems with constant or periodic coefficients are extended to systems whose coefficients can exhibit a much more general time dependence. Indeed, techniques of topological dynamics and of ergodic theory which have been worked out in recent times permit us to apply new methods and adapt older ones to the study of a rich set of new scenarios which are not possible in the periodic case. In the end we obtain a coherent theory which has been successfully applied to a wide range of problems in the setting of nonautonomous linear Hamiltonian systems.

Firenze, Italy	Russell Johnson
Valladolid, Spain	Rafael Obaya
Valladolid, Spain	Sylvia Novo
Valladolid, Spain	Carmen Núñez
Firenze, Italy	Roberta Fabbri

Acknowledgments

The contents of this book are based on many years of research regarding diverse aspects of the theory of nonautonomous linear Hamiltonian systems. The fact that it exists is due in large part to illuminating discussions with many colleagues and to the support of relatives and friends. To all of them, many thanks.

Russell Johnson would in particular like to thank Mahesh Nerurkar for joint research activity extending over many years. A number of the results in this book can be traced to that collaboration.

Sylvia Novo would like to express her appreciation to Manuel Núñez for his role as a source of mathematical information.

Finally, Carmen Núñez wishes to thank Enrique Marty deeply, for his comprehension and continuous encouragement during the years of elaboration of the book.

Contents

Chapter 1
Nonautonomous Linear Hamiltonian Systems

This chapter is devoted to the general explanation of the framework of the analysis made in this book, and to stating the many foundational facts which will be required. With the aim of being relatively self-contained, precise references where the proofs of the stated properties can be found are included, and at the same time some proofs which the reader may consider elementary or well known, but for which it is not easy to find a completely appropriate reference in the literature, are given.

This long chapter is divided into four sections. The first presents the most fundamental notions and properties of topological dynamics and ergodic theory, including the concept and main characteristics of a skew-product flow, which are fundamental for the book.

The second section summarizes basic results concerning spaces of matrices, the Grassmannian and Lagrangian manifolds, and matrix-valued functions.

Section 1.3 is devoted to the description of the general framework of the book. Under mild conditions on the coefficient matrix, a nonautonomous linear system of ordinary differential equations defines continuous skew-product flows on the trivial and Grassmannian bundles above a compact metric space. Special attention is devoted to the Hamiltonian case, for which two special skew-product flows can be defined. For the first one, which is defined on the Lagrange bundle, the use of generalized polar coordinates simplifies the task of describing the dynamical behavior. The second one, which is closely related to the first, is defined on the bundle given by the set of symmetric matrices. It presents some interesting monotonicity properties.

The last section concerns one of the most fundamental concepts for the development of the analysis made in the book: that of exponential dichotomy, both in the general linear case and in the linear Hamiltonian case. Many of the properties ensured by its presence will be described in detail, and then applied later in the book. The closely related concept of Sacker–Sell spectrum is also discussed, and several aspects of the Sacker–Sell perturbation theory are explained. The section is

R. Johnson et al., *Nonautonomous Linear Hamiltonian Systems: Oscillation, Spectral Theory and Control*, Developments in Mathematics 36,
DOI 10.1007/978-3-319-29025-6_1

completed with the less standard analysis of the behavior of the Grassmannian flows in the presence of exponential dichotomy.

1.1 Some Fundamental Notions

The concepts and properties summarized in this section will be used often throughout the book, many times without reference to these initial pages. Suitable references for all these notions include Nemytskii and Stepanov [110], Ellis [41], Sacker and Sell [133], Cornfeld et al. [35], Walters [148], Mañé [99], and Rudin [128, 129].

1.1.1 Basic Concepts and Properties of Topological Dynamics

Let Ω be a locally compact Hausdorff topological space. Let Σ_Ω and $\Sigma_\mathbb{R}$ represent the Borel sigma-algebras of Ω and $\mathbb{R}$, and let $\Sigma_* = \Sigma_\mathbb{R} \times \Sigma_\Omega$ be the product sigma-algebra; i.e. the intersection of all the sigma-algebras on $\mathbb{R} \times \Omega$ containing the sets $\mathcal{I} \times \mathcal{A}$ for $\mathcal{I} \in \Sigma_\mathbb{R}$ and $\mathcal{A} \in \Sigma_\Omega$. Mild conditions on Ω ensure that Σ_* agrees with the Borel sigma-algebra of $\mathbb{R} \times \Omega$: it is enough to assume that Ω admits a countable basis of open sets (see e.g. Proposition 7.6.2 of Cohn [30]).

It will be convenient to work under the hypothesis that Σ_* is indeed the Borel sigma-algebra of $\mathbb{R}\times\Omega$. So, throughout Sect. 1.1, Ω will represent a locally compact Hausdorff topological space which admits a countable basis of open sets. In fact, throughout the book, any flow will be defined on a set which satisfies, at a minimum, these conditions. Some of the results explained in this section require Ω to be a compact metric space, but this hypothesis will be specified whenever it is assumed.

A map $\sigma\colon \mathbb{R} \times \Omega \to \Omega$ is *Borel measurable* if $\sigma^{-1}(\mathcal{A}) \in \Sigma_*$ for all $\mathcal{A} \in \Sigma_\Omega$. A *global real Borel measurable flow* on Ω is a Borel measurable map $\sigma\colon \mathbb{R} \times \Omega \to \Omega$ such that $\sigma_0 = \mathrm{Id}_\Omega$ and $\sigma_{t+s} = \sigma_t \circ \sigma_s$ for all $s, t \in \mathbb{R}$, where $\sigma_t\colon \Omega \to \Omega,\ \omega \mapsto \sigma(t,\omega)$. The flow is *continuous* if σ satisfies the stronger condition of being a continuous map, in which case each map σ_t is a homeomorphism on Ω with inverse σ_{-t}. The notation (Ω, σ) will be frequently used to represent a real global flow on Ω, and the words *real* and *global* will be omitted when no confusion arises.

The *orbit* of a point $\omega \in \Omega$ is the set $\{\sigma_t(\omega) \mid t \in \mathbb{R}\}$, and its *positive* (resp. *negative*) *semiorbit* is $\{\sigma_t(\omega) \mid t \in \mathbb{R}_+\}$, where $\mathbb{R}_+ = \{t \in \mathbb{R} \mid t \ge 0\}$ (resp. $\{\sigma_t(\omega) \mid t \in \mathbb{R}_-\}$, where $\mathbb{R}_- = \{t \in \mathbb{R} \mid t \le 0\}$).

Given a Borel measurable flow (Ω, σ), a Borel subset $\mathcal{A} \subseteq \Omega$ (i.e. an element $\mathcal{A}$ of Σ_Ω) is *σ-invariant* (resp. *positively* or *negatively σ-invariant*) if $\sigma_t(\mathcal{A}) = \mathcal{A}$ for all $t \in \mathbb{R}$ (resp. $t \in \mathbb{R}_+$ or $t \in \mathbb{R}_-$). Let $\mathbb{Y}$ be a topological space. If Σ is a sigma-algebra on Ω containing the Borel sets, a map $f\colon \Omega \to \mathbb{Y}$ is *Σ-measurable* if $f^{-1}(\mathcal{B}) \in \Sigma$ for every Borel subset $\mathcal{B} \subseteq \mathbb{Y}$; and f is *Borel measurable* when it is Σ_Ω-measurable. A Borel measurable function $f\colon \Omega \to \mathbb{Y}$ is *σ-invariant* if $f(\sigma_t(\omega)) =$

$f(\omega)$ for all $\omega \in \Omega$ and $t \in \mathbb{R}$. It is obvious that a Borel subset $\mathcal{A}$ is σ-invariant if and only if its characteristic function $\chi_{\mathcal{A}}$ is σ-invariant.

If Σ is a sigma-algebra containing the Borel sets, the concepts of σ-invariant set $\mathcal{A} \in \Sigma$ and σ-invariant Σ-measurable map $f \colon \Omega \to \mathbb{Y}$ are defined analogously. Note that in fact this concept of invariance can be extended to any set or function, since it does not depend on measurability.

All these definitions of *σ-invariance* correspond to *strict σ-invariance*, although the word *strict* will be almost always omitted. A less restrictive definition of invariance, depending on a fixed measure, is given in Sect. 1.1.2.

The flow is *local* if the map σ is defined, Borel measurable, and satisfies the two initially required properties on an open subset $\mathcal{O} \subseteq \mathbb{R} \times \Omega$ containing $\{0\} \times \Omega$. Define $\mathcal{O}_\omega = \{t \in \mathbb{R} \mid (t, \omega) \in \mathcal{O}\}$ for $\omega \in \Omega$. The *orbit* of the point ω for a local flow (Ω, σ) is $\{\sigma_t(\omega) \mid t \in \mathcal{O}_\omega\}$, and it is *globally defined* if $\mathcal{O}_\omega = \mathbb{R}$. The *positive* (resp. *negative*) *semiorbit* of a point ω is the set $\{\sigma_t(\omega) \mid t \in \mathcal{O}_\omega \cap \mathbb{R}_+\}$ (resp. $\{\sigma_t(\omega) \mid t \in \mathcal{O}_\omega \cap \mathbb{R}_-\}$, and it is *globally defined* if $\mathcal{O}_\omega \cap \mathbb{R}_+ = \mathbb{R}_+$ (resp. $\mathcal{O}_\omega \cap \mathbb{R}_- = \mathbb{R}_-$). A (in general Borel) subset $\mathcal{A} \subseteq \Omega$ is *σ-invariant* (resp. *positively* or *negatively σ-invariant*) if it is composed of globally defined orbits (resp. globally defined positive or negative semiorbits).

Finally, replacing $\mathbb{R}$ by $\mathbb{R}_+$ (resp. by $\mathbb{R}_-$) provides the definition of a (global or local) real positive (resp. negative) *semiflow* on Ω. The definitions of positive (resp. negative) semiorbit and (strict) invariance are the obvious ones.

For the remaining definitions and properties discussed in this section, the flow σ is assumed to be continuous.

A compact σ-invariant subset $\mathcal{M} \subseteq \Omega$ is *minimal* if it does not contain properly any other such set; or, equivalently, if each of its positive or negative semiorbits is dense in it. The flow (Ω, σ) is *minimal* or *recurrent* if Ω itself is minimal, which obviously requires Ω to be compact. Note that Zorn's lemma ensures that, if Ω is compact, then it contains at least one minimal subset.

Suppose that the positive semiorbit of a point ω_0 for such a flow is relatively compact. Then the *omega-limit set* of the point (or of its positive semiorbit) is given by those points $\omega \in \Omega$ such that $\omega = \lim_{k \to \infty} \sigma(t_k, \omega_0)$ for some sequence $(t_k) \uparrow \infty$. The omega-limit set is nonempty, compact, connected, and σ-invariant. The concept of *alpha-limit set* is analogous, working now with a negative semiorbit and with sequences $(t_k) \downarrow -\infty$. Clearly, a minimal subset of Ω is the omega-limit set and the alpha-limit set of each of its elements.

Finally, assume in addition that Ω is a compact metric space, and let d_Ω represent the distance on Ω. The flow (Ω, σ) is *chain recurrent* if given $\varepsilon > 0$, $t_0 > 0$, and points $\omega, \tilde{\omega} \in \Omega$, there exist points $\omega = \omega_0, \omega_1, \ldots, \omega_m = \tilde{\omega}$ of Ω and real numbers $t_1 > t_0, \ldots, t_m > t_0$ such that $d_\Omega(\sigma_{t_i}(\omega_i), \omega_{i+1}) < \varepsilon$ for $i = 0, \ldots, m-1$. It is easy to check that minimality implies chain recurrence: just take $\omega_0 = \omega$ and $\omega_1 = \tilde{\omega}$ and keep in mind that the positive semiorbit of ω is dense in Ω. It is also easy to check that if (Ω, σ) is chain recurrent, then the set Ω is connected.

1.1.2 Basic Concepts and Properties of Measure Theory

Unless otherwise indicated, any measure appearing in the book is a positive normalized regular Borel measure. Given such a measure m, let Σ_m be the m-completion of the Borel sigma-algebra (see e.g. Theorem 1.36 of [128]), and represent with the same symbol m the extension of the initial measure to Σ_m. As usual, the notation "m-a.e." means *almost everywhere with respect to* m; "for m-a.e. $\omega \in \Omega$" means *for almost every* $\omega \in \Omega$; and $L^1(\Omega, m)$ represents the quotient set of Σ_m-measurable functions $f\colon \Omega \to \mathbb{R}$ with $\int_\Omega |f(\omega)|\, dm < \infty$ (so that two real functions represent the same class if they are m-a.e. equal, in which case they are the same element of $L^1(\Omega, m)$). See Sect. 1.2.4 for the general definitions of L^p spaces of matrix-valued functions on Ω.

Let m be a measure on Ω. Then m is *σ-invariant* if $m(\sigma_t(\mathcal{A})) = m(\mathcal{A})$ for every Borel subset $\mathcal{A} \subseteq \Omega$ and all $t \in \mathbb{R}$, which ensures the same property for every $\mathcal{A} \in \Sigma_m$. A Σ_m-measurable map $f\colon \Omega \to \mathbb{Y}$ (for a topological space $\mathbb{Y}$) is *σ-invariant with respect to m* if, for all $t \in \mathbb{R}$, $f(\sigma_t(\omega)) = f(\omega)$ m-a.e. And a subset $\mathcal{A} \in \Sigma_m$ is *σ-invariant with respect to m* if $\chi_{\mathcal{A}}$ has this property.

The expression "σ-invariant" (for sets, measures, or functions) will often be changed to "invariant" throughout the book, since in most cases no confusion arises.

Proposition 1.2 shows the relation between these concepts of σ-invariance with respect to m and the (strict) ones given in the previous section: it proves that, when moving for instance in the quotient space $L^1(\Omega, m)$, one can always consider that a "σ-invariant function" satisfies the "strict" definition. More information in this regard will be added in Proposition 1.5.

Remark 1.1 Recall that any Σ_m-measurable function $f\colon \Omega \to \mathbb{K}$, for $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, agrees m-a.e. with a Borel measurable one (see [128], Lemma 1 of Theorem 8.12). In addition, if Σ is any sigma-algebra containing the Borel sets, and if a sequence $(f_n\colon \Omega \to \mathbb{K})$ of Σ-measurable functions converges everywhere to a function f, then f is Σ-measurable (see [128], Theorem 1.14). And, as a consequence of this last result, if $(f_n\colon \Omega \to \mathbb{K})$ is a sequence of Σ_m-measurable functions which converges m-a.e. to a function f, then f is Σ_m-measurable.

Proposition 1.2 *Let (Ω, σ) be a Borel measurable flow, and let m be a σ-invariant measure on Ω.*

(i) *Let the Σ_m-measurable function $f\colon \Omega \to \mathbb{K}$ be σ-invariant with respect to m. Then there exists a Σ_m-measurable function $f^*\colon \Omega \to \mathbb{K}$ which is (strictly) σ-invariant such that $f = f^*$ m-a.e.*
(ii) *Let the set $\mathcal{A} \in \Sigma_m$ be σ-invariant with respect to m. Then there exists a (strictly) σ-invariant set $\mathcal{A}^* \in \Sigma_m$ such that $\chi_{\mathcal{A}} = \chi_{\mathcal{A}^*}$ m-a.e.*

Proof

(i) The proof of this property is carried out in Lemma 1 of Chapter 1.2 of [35], and included here for the reader's convenience. It follows from Remark 1.1 that there is no loss of generality in assuming that f is Borel measurable. Define the sets $\mathcal{N} = \{(t,\omega) \in \mathbb{R} \times \Omega \mid f(\omega) \neq f(\sigma_t(\omega))\}$, and note that the hypotheses on σ ensure that this set belongs to $\Sigma_* = \Sigma_{\mathbb{R}} \times \Sigma_\Omega$, since the maps $\mathbb{R} \times \Omega \to \mathbb{R}$, $(t,\omega) \mapsto f(\omega)$ and $\mathbb{R} \times \Omega \to \mathbb{R}$, $(t,\omega) \mapsto f(\sigma_t(\omega))$ are Σ_*-measurable. Define now $\mathcal{N}_t = \{\omega \in \Omega \mid (t,\omega) \in \mathcal{N}\}$ for $t \in \mathbb{R}$, and $\mathcal{N}_\omega = \{t \in \mathbb{R} \mid (t,\omega) \in \mathcal{N}\}$ for $\omega \in \Omega$, and note that $\mathcal{N}_t \in \Sigma_\Omega$ for all $t \in \mathbb{R}$ and $\mathcal{N}_\omega \in \Sigma_{\mathbb{R}}$ for all $\omega \in \Omega$ (see Theorem 8.2 of [128]). By definition of σ-invariance with respect to m, $m(\mathcal{N}_t) = 0$ for all $t \in \mathbb{R}$. Define μ as the product measure of m and l on $\Omega \times \mathbb{R}$, where l is the Lebesgue measure on $\mathbb{R}$. Fubini's theorem (see Theorem 8.8 of [128]) ensures that the maps $\omega \mapsto l(\mathcal{N}_\omega)$ and $t \mapsto m(\mathcal{N}_t)$ are Borel, and that $\mu(\mathcal{N}) = \int_\Omega l(\mathcal{N}_\omega)\, dm = \int_{\mathbb{R}} m(\mathcal{N}_t)\, dl = 0$. Therefore the subset $\Omega_f \subseteq \Omega$ of points ω with $l(\mathcal{N}_\omega) = 0$ is Borel, and $m(\Omega_f) = 1$. Suppose that ω and $\sigma_t(\omega)$ belong to Ω_f for a pair $(t,\omega) \in \mathbb{R} \times \Omega$. Then $f(\omega) = f(\sigma_t(\omega))$. In order to prove this assertion, take $s \in \mathbb{R} - \mathcal{N}_{\sigma_t(\omega)}$ such that $s + t \in \mathbb{R} - \mathcal{N}_\omega$, and note that $f(\sigma_t(\omega)) = f(\sigma_s(\sigma_t(\omega))) = f(\sigma_{s+t}(\omega)) = f(\omega)$. Now define

$$f^*(\omega) = \begin{cases} f(\omega) & \text{if there exists } t \in \mathbb{R} \text{ with } \sigma_t(\omega) \in \Omega_f\,, \\ 0 & \text{otherwise}\,, \end{cases}$$

which is Σ_m-measurable, since it agrees with f at least on Ω_f (and hence m-a.e.), and which is σ-invariant in the classical sense.

(ii) Let $g = \chi^*_{\mathcal{A}}$ be the σ-invariant function associated to $\chi_{\mathcal{A}}$ by (i). Then the set $\mathcal{B} = \{\omega \in \Omega \mid g(\omega) \in \{0,1\}\} = 1$ belongs to Σ_m, is σ-invariant, and has full measure for m: $m(\mathcal{B}) = 1$. The set $\mathcal{A}^* = \{\omega \in \Omega \mid g(\omega) = 1\} \subseteq \mathcal{B}$ also belongs to Σ_m and is σ-invariant. In addition, $g(\omega) = \chi_{\mathcal{A}^*}(\omega)$ for all $\omega \in \mathcal{B}$, so that $\chi_{\mathcal{A}} = \chi_{\mathcal{A}^*}$ m-a.e., as asserted.

One of the most fundamental results in measure theory is the Birkhoff ergodic theorem, one of whose simplest versions is now recalled.

Theorem 1.3 *Let (Ω,σ) and m be a Borel measurable flow and a σ-invariant measure on Ω. Given $f \in L^1(\Omega,m)$, there exists a (strictly) σ-invariant set $\Omega_f \in \Sigma_m$ with $m(\Omega_f) = 1$ such that, for all $\omega \in \Omega_f$, the limits*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f(\sigma_s(\omega))\, ds = \lim_{t\to\infty} \frac{1}{2t}\int_{-t}^t f(\sigma_s(\omega))\, ds = \lim_{t\to-\infty} \frac{-1}{t}\int_t^0 f(\sigma_s(\omega))\, ds$$

exist, agree, and take on a real value $\tilde{f}(\omega)$. In addition, $\tilde{f}(\sigma_t(\omega)) = \tilde{f}(\omega)$ for all $\omega \in \Omega_f$ and $t \in \mathbb{R}$, $\tilde{f}$ belongs to $L^1(\Omega,m)$, and $\int_\Omega \tilde{f}(\omega)\, dm = \int_\Omega f(\omega)\, dm$.

Its proof in the case of a discrete flow (given by the iteration of an automorphism on Ω) can be found, for example, in Section II.1 of [99]. The procedure to deduce the result for a real flow from the discrete case is standard: define the automorphism $T(\omega) = \sigma(1,\omega)$ and, given $f \in L^1(\Omega,m)$, define $F(\omega) = \int_0^1 f(\sigma_s(\omega))\, ds$; then, Fubini's theorem ensures that $F \in L^1(\Omega,m)$, and the application of the discrete version of the theorem to this setting provides the sets Ω_f and the function $\tilde{f}$ satisfying the theses of the real version. The details are left to the reader.

Note that the function $\tilde{f}$ provided by the previous theorem can be considered to be σ-invariant in the strict sense: just define it to be 0 outside Ω_f. Note also that the set Ω_f contains a Borel subset with measure 1, which is clearly σ-invariant with respect to m. But in fact this Borel subset of Ω_f can be taken as a (strictly) σ-invariant set, as Proposition 1.5(i) below proves. Therefore, there is no loss of generality in assuming that the set Ω_f itself is Borel.

The following result, whose proof is included for completeness, will be required in Chap. 4. The notation $g\colon \Omega \to [0,\infty]$ is used for *extended-real* functions (which can take the value ∞), and the concept of Σ_m-measurability for such a function is clear.

Proposition 1.4 *Let (Ω,σ) be a Borel measurable flow, and let m be a σ-invariant measure on Ω. Let $f\colon \Omega \to [0,\infty)$ be a Σ_m-measurable function. Then, there exists a (strictly) σ-invariant set $\Omega_f \in \Sigma_m$ with $m(\Omega_f) = 1$ such that, for all $\omega \in \Omega_f$, the limits*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f(\sigma_s(\omega))\,ds = \lim_{t\to\infty} \frac{1}{2t}\int_{-t}^t f(\sigma_s(\omega))\,ds = \lim_{t\to-\infty} \frac{-1}{t}\int_t^0 f(\sigma_s(\omega))\,ds$$

exist, agree, and take a value $\tilde{f}(\omega) \in \mathbb{R}\cup\{\infty\}$. In addition, the extended-real function $\tilde{f}\colon \Omega \to [0,\infty]$ is Σ_m-measurable, and it satisfies $\tilde{f}(\sigma_t(\omega)) = \tilde{f}(\omega)$ for all $\omega \in \Omega_f$ and $t \in \mathbb{R}$, and $\int_\Omega \tilde{f}(\omega)\,dm = \int_\Omega f(\omega)\,dm$.

Proof Let $h\colon \Omega \to [0,\infty)$ be a Σ_m-measurable function. For each $k \in \mathbb{N}$, define $h_k = \min(h,k)$, which obviously belongs to $L^1(\Omega,m)$. Hence there exists a function $\tilde{h}_k \in L^1(\Omega,m)$ and a set $\Omega_{h_k} \in \Sigma_m$ with $m(\Omega_{h_k}) = 1$ satisfying the theses of Theorem 1.3. Define $\Omega_h^* = \cap_{k\in\mathbb{N}}\Omega_{h_k}$, which belongs to Σ_m, is σ-invariant, and has full measure for m. Note that the nondecreasing sequence $(h_k(\omega))$ converges to $h(\omega)$ for all $\omega \in \Omega_h^*$, and define $h^*(\omega) \in [0,\infty]$ as the limit of the nondecreasing sequence of σ-invariant functions $(\tilde{h}_k(\omega))$, also for $\omega \in \Omega_h^*$. Then, h^* is Σ_m-measurable (see Remark 1.1) and σ-invariant. In addition, if $h^* \in L^1(\Omega,m)$, then $h \in L^1(\Omega,m)$: apply the Lebesgue monotone convergence theorem and the Birkhoff Theorem 1.3 to get $0 \le \int_\Omega h(\omega)\,dm = \lim_{k\to\infty}\int_\Omega h_k(\omega)\,dm = \lim_{k\to\infty}\int_\Omega \tilde{h}_k(\omega)\,dm = \int_\Omega h^*(\omega)\,dm < \infty$.

Returning to the function f of the statement, note that if $f \in L^1(\Omega,m)$, the assertions follow from Theorem 1.3. Assume hence that $\int_\Omega f(\omega)\,dm = \infty$, and associate to it the sequences (f_k) and $(\tilde{f}_k)$, the set Ω_f^*, and the function f^*, as above. Therefore, $f^* \notin L^1(\Omega,m)$. Clearly, the sets

$$\begin{aligned}\mathcal{A} &= \{\omega \in \Omega_f^* \mid f^*(\omega) = \infty\}\,,\\ \mathcal{A}_j &= \{\omega \in \Omega_f^* \mid j \le f^*(\omega) < j+1\} \quad \text{for } j \ge 0\end{aligned}$$

belong to Σ_m, are σ-invariant and disjoint, and satisfy $\Omega_f^* = \mathcal{A} \cup (\cup_{j=0}^{\infty}\mathcal{A}_j)$. Then, if $\omega \in \mathcal{A}$,

$$\begin{aligned}\liminf_{t\to\infty}\frac{1}{t}\int_0^t f(\sigma_s(\omega))\,ds &\geq \sup_{k\in\mathbb{N}}\lim_{t\to\infty}\frac{1}{t}\int_0^t f_k(\sigma_s(\omega))\,ds\\ &= \sup_{k\in\mathbb{N}}\tilde{f}_k(\omega)\,ds = f^*(\omega) = \infty\,,\end{aligned}$$

so that there exists $\lim_{t\to\infty}(1/t)\int_0^t f(\sigma_s(\omega))\,ds = f^*(\omega) = \infty$. The same property holds for the other two limits of the proposition. Now define

$$g = \sum_{j=0}^{\infty}\frac{1}{j+1}\,\chi_{\mathcal{A}_j} f$$

on Ω_f^*, note that it is Σ_m-measurable, and associate to it the sequences (g_k), $(\tilde{g}_k)$, and the set $\Omega_g^* \subseteq \Omega_f^*$, as at the beginning of the proof. Fix any $k \in \mathbb{N}$ and any $\omega \in \Omega_g^*$ outside $\mathcal{A}$, and take the unique $j \in \mathbb{N}$ such that $\omega \in \mathcal{A}_j \cap \Omega_g^*$. Then $g(\omega) = (1/j+1)f(\omega)$, and hence

$$g_k(\omega) = \frac{1}{j+1}\,\min(f(\omega), k(j+1)) = \frac{1}{j+1}f_{k(j+1)}(\omega)\,.$$

Since $\sigma_s(\omega) \in \mathcal{A}_j \cap \Omega_g^*$ for all $s \in \mathbb{R}$,

$$\begin{aligned}\tilde{g}_k(\omega) &= \lim_{t\to\infty}\frac{1}{t}\int_0^t g_k(\sigma_s(\omega))\,ds = \frac{1}{j+1}\lim_{t\to\infty}\frac{1}{t}\int_0^t f_{k(j+1)}(\sigma_s(\omega))\,ds\\ &= \frac{1}{j+1}\tilde{f}_{k(j+1)}(\omega) \leq \frac{1}{j+1}f^*(\omega) \leq 1\end{aligned}$$

for all $k \in \mathbb{N}$. Note that g_k vanishes outside $\cup_{j=1}^{\infty}\mathcal{A}_j$. Hence $\int_\Omega g_k(\omega)\,dm = \int_\Omega \tilde{g}_k(\omega)\,dm \leq 1$, so that the Lebesgue dominated convergence theorem ensures that $g \in L^1(\Omega, m_0)$. Let $\tilde{g}$ and $\Omega_g \subseteq \Omega_f^*$ be the σ-invariant function and subset associated to g by Theorem 1.3, with $m(\Omega_g) = 1$. Then for all ω in the σ-invariant set $\mathcal{A}_j \cap \Omega_g, f(\omega) = (j+1)\,g(\omega)$ and hence

$$\begin{aligned}\lim_{t\to\infty}\frac{1}{t}\int_0^t f(\sigma_s(\omega))\,ds &= \lim_{t\to\infty}\frac{1}{2t}\int_{-t}^t f(\sigma_s(\omega))\,ds\\ &= \lim_{t\to-\infty}\frac{-1}{t}\int_t^0 f(\sigma_s(\omega))\,ds = (j+1)\,\tilde{g}(\omega) = (j+1)\,\chi_{\mathcal{A}_j}\tilde{g}(\omega)\,.\end{aligned}$$

Define $\Omega_f = \mathcal{A} \cup \Big((\cup_{j=0}^{\infty} \mathcal{A}_j) \cap \Omega_g \Big)$, and note that it belongs to Σ_m and satisfies $m(\Omega_f) = 1$. This σ-invariant set and the Σ_m-measurable and σ-invariant function

$$\tilde{f} = \begin{cases} f^*(\omega) & \text{if } \omega \in \mathcal{A} \\ \displaystyle\sum_{j=0}^{\infty} (j+1)\chi_{\mathcal{A}_j} \tilde{g} & \text{if } \omega \in (\cup_{j=0}^{\infty} \mathcal{A}_j) \cap \Omega_g \end{cases} \tag{1.1}$$

satisfy the statements regarding the limits. In addition, for all $\omega \in \Omega_f$,

$$\tilde{f}(\omega) = \lim_{t\to\infty} \frac{1}{t} \int_0^t f(\sigma_s(\omega))\, ds \geq \lim_{t\to\infty} \frac{1}{t} \int_0^t f_n(\sigma_s(\omega))\, ds = \tilde{f}_n(\omega)\,,$$

so that $\tilde{f}(\omega) \geq f^*(\omega)$ on Ω_f. Hence, $\int_\Omega \tilde{f}(\omega)\, dm \geq \int_\Omega f^*(\omega)\, dm = \infty = \int_\Omega f(\omega)\, dm$, which completes the proof.

As in the case of Theorem 1.3, the function $\tilde{f}$ provided by Proposition 1.4 can be considered to be σ-invariant in the strict sense, and Proposition 1.5(i), which is proved immediately below, ensures that the set Ω_f contains a Borel subset with measure 1 which is σ-invariant with respect to m.

Proposition 1.5 *Let (Ω, σ) be a Borel measurable flow, and let m be a σ-invariant measure on Ω.*

(i) *Let $\mathcal{A} \in \Sigma_m$ be a (strictly) σ-invariant set with $m(\mathcal{A}) = 1$. Then $\mathcal{A}$ contains a (strictly) σ-invariant Borel set $\mathcal{B}$ with $m(\mathcal{B}) = 1$.*
(ii) *Let $f\colon \Omega \to \mathbb{R}$ be Σ_m-measurable and σ-invariant with respect to m_0. Then there exists $g\colon \Omega \to \mathbb{R}$ which is Borel and (strictly) σ-invariant such that $g = f$ m-a.e.*

Proof

(i) It suffices to prove that for all $n \in \mathbb{N}$ there exists a σ-invariant Borel set $\mathcal{B}_n \subseteq \mathcal{A}$ with $m(\mathcal{B}_n) \geq m(\mathcal{A}) - 1/n$, and then take $\mathcal{B} = \cup_{n\geq 1} \mathcal{B}_n$.

Fix $n \in \mathbb{N}$, and note that the regularity of the measure m implies the existence of a compact set $\mathcal{K}_n \subseteq \mathcal{A}$ with $m(\mathcal{A} - \mathcal{K}_n) \leq 1/n$. The Borel measurability of the flow ensures that the map $\mathbb{R} \times \Omega \to \mathbb{R}\,,\ (t,\omega) \mapsto \chi_{\mathcal{K}_n}(\sigma(t,\omega))$ is Borel measurable, and hence Fubini's theorem guarantees that the maps $h_n^j\colon \Omega \to \mathbb{R}$ given by

$$h_n^j(\omega) = \sum_{i=-j}^{j} \frac{1}{|i|^2 + 1} \int_j^{j+1} \chi_{\mathcal{K}_n}(\sigma_t(\omega))\, dt$$

are Borel measurable (see e.g. Theorem 8.8 of [128]). Clearly, $h_n^j \leq h_n^{j+1}$, so that the limit $h_n(\omega) = \lim_{j\to\infty} h_n^j(\omega)$ exists for all $\omega \in \Omega$, and the (bounded)