Chapter 2
Unforced Linear Systems

In this chapter we undertake a systematic study of finite dimensional, unforced, linear, time invariant differential systems. They are defined by a system of ordinary differential equations of the form

$$\dot{x} = Ax \, , \qquad x \in \mathbf{R}^n \, . \tag{2.1}$$

According to the mathematical tradition, (2.1) is called a linear homogeneous system of differential equations (with constant coefficients). In extended notation, (2.1) reads

$$\begin{cases} \dot{x}_1 = a_{11}x_1 + \cdots + a_{1n}x_n \\ \ldots \quad \ldots \quad \ldots \\ \dot{x}_n = a_{n1}x_1 + \cdots + a_{nn}x_n \, . \end{cases}$$

For a general system of ordinary differential equations, the notion of solution is recalled in Appendix A. In force of the special form of (2.1) the solutions enjoy some special properties.

2.1 Prerequisites

In this section we recall some important facts, concerning a system of equations of type (2.1) and its solutions.

Fact 1. *For each initial state x_0 there exists a unique solution $x = \varphi(t)$ of system* (2.1) *such that $\varphi(0) = x_0$; moreover, $\varphi(t)$ is defined for each $t \in \mathbf{R}$.*

Fact 2. *If $v \in \mathbf{R}^n$ ($v \neq 0$) is an eigenvector of A corresponding to the eigenvalue $\lambda \in \mathbf{R}$, then $\varphi(t) = e^{\lambda t}v$ represents the solution of* (2.1) *corresponding to the initial state v.*

A. Bacciotti, *Stability and Control of Linear Systems*, Studies in Systems, Decision and Control 185, https://doi.org/10.1007/978-3-030-02405-5_2

Fact 3. *If $\varphi_1(\cdot)$, $\varphi_2(\cdot)$ are solutions of* (2.1) *and $\alpha_1, \alpha_2 \in \mathbf{R}$, then also $\alpha_1\varphi_1(\cdot) + \alpha_2\varphi_2(\cdot)$ is a solution of* (2.1).

Fact 4. *Let $\varphi_1(\cdot), \ldots, \varphi_k(\cdot)$ be k solutions of* (2.1). *The following statements are equivalent*:

- *there exists $\bar{t} \in \mathbf{R}$ such that the vectors $\varphi_1(\bar{t}), \ldots, \varphi_k(\bar{t})$ are linearly independent in $\mathbf{R}^n$*;
- *the functions $\varphi_1(\cdot), \ldots, \varphi_k(\cdot)$ are linearly independent, as elements of the space $\mathcal{C}(-\infty, +\infty, \mathbf{R}^n)$*;
- *for each $t \in \mathbf{R}$, the vectors $\varphi_1(t), \ldots, \varphi_k(t)$ are linearly independent, as elements of the space $\mathbf{R}^n$.*

When one of the above equivalent conditions holds, we simply say that $\varphi_1(\cdot), \ldots, \varphi_k(\cdot)$ are linearly independent.

Fact 5. *The set of all the solutions of the system* (2.1) *forms a subspace $\mathcal{S}$ of $\mathcal{C}(-\infty, +\infty, \mathbf{R}^n)$. The dimension of $\mathcal{S}$ is finite and, more precisely, it is equal to n. The subspace $\mathcal{S}$ is also called the general integral of system* (2.1).

Notice that system (2.1) makes sense even if we allow that x takes value into the n-dimensional complex space $\mathbf{C}^n$, and that the entries of A are complex numbers: apart from some obvious modifications, all the previous facts remain valid.[1] Actually, to this respect we may list some further properties.

Fact 6. *If the elements of A are real, and if $\varphi(\cdot)$ is a solution of* (2.1) *with nonzero imaginary part, then the conjugate function $\overline{\varphi}(\cdot)$ is a solution of* (2.1), *as well.*

Fact 7. *If the elements of A are real, and if $\varphi(\cdot)$ is a solution of* (2.1) *with nonzero imaginary part, then $\varphi(\cdot)$ and $\overline{\varphi}(\cdot)$ are linearly independent; in addition,*

$$\varphi_1(\cdot) = \frac{\varphi(\cdot) + \overline{\varphi}(\cdot)}{2} \quad \text{and} \quad \varphi_2(\cdot) = \frac{\varphi(\cdot) - \overline{\varphi}(\cdot)}{2\mathrm{i}} \tag{2.2}$$

are two real and linearly independent solutions of (2.1).

If A is a matrix with real elements and with a complex eigenvalue $\lambda = \alpha + \mathrm{i}\,\beta$ ($\beta \neq 0$) associated to an eigenvector $v = u + \mathrm{i}\,w$, we dispose of the complex solution $\varphi(t) = e^{\lambda t}v$. Then, using (2.2), we obtain the representation

$$\varphi_1(t) = e^{\alpha t}[(\cos\beta t)u - (\sin\beta t)w]\ , \quad \varphi_2(t) = e^{\alpha t}[(\cos\beta t)w + (\sin\beta t)u]\ .$$

Remark 2.1 The existence of non-real eigenvalues implies therefore the existence of real oscillatory solutions. In particular, if $\alpha = 0$ and $\beta \neq 0$, the eigenvalues are purely imaginary, and we have periodic solutions with minimal period equal to $2\pi/\beta$. ■

[1]The convenience of extending the search for the solutions to the complex field even if the elements of A are real numbers, is suggested by Fact 2: possible eigenvalues of A represented by conjugate complex numbers (with nonzero imaginary part) generates solutions which, otherwise, would be difficult to identify.

We are now able to conclude that the general integral of system (2.1) can be written as a linear combination

$$\varphi(t) = c_1\varphi_1(t) + \cdots + c_n\varphi_n(t) \tag{2.3}$$

where $c_1, \ldots, c_n$ represent arbitrary constants, and $\varphi_1, \ldots, \varphi_n$ represent n arbitrary solutions, provided that they are linearly independent. If A is real, it is not restrictive to assume that $\varphi_1, \ldots, \varphi_n$ are real valued: hence, Eq. (2.3) describes either the space of all the real solutions when the constants $c_1, \ldots, c_n$ are taken in the real field, or the space of all the complex solutions when the constants $c_1, \ldots, c_n$ are taken in the complex field.

A set formed by n linearly independent solutions of the system (2.1) is called a *fundamental set of solutions*. To each fundamental set of solutions φ_1, ..., φ_n, we associate a *fundamental matrix*

$$\Phi(t) = (\varphi_1(t), \ldots, \varphi_n(t))$$

whose columns are formed by the components of the vectors $\varphi_1(t)$, ..., $\varphi_n(t)$. Notice that if $\Phi(t)$ is a fundamental matrix and Q is a constant, nonsingular matrix, then also $\Phi(t)Q$ is a fundamental matrix. From this remark, it follows easily that, for each $t_0 \in \mathbf{R}$, there exists a unique fundamental matrix such that $\Phi(t_0) = I$. This is also called the *principal fundamental matrix relative to* t_0. The principal fundamental matrix relative to $t_0 = 0$ will be simply called *principal fundamental matrix*.

Let us introduce the constant vector $c = (c_1, \ldots, c_n)^{\mathbf{t}}$. If $\Phi(t)$ is any fundamental matrix, we can rewrite (2.3) as

$$\varphi(t) = \Phi(t)c \tag{2.4}$$

The particular solution satisfying the initial conditions $\varphi(t_0) = x_0$ can be recovered by solving the algebraic system

$$\Phi(t_0)c = x_0$$

with respect to the unknown vector c. If $\Phi(t)$ is the principal fundamental matrix relative to t_0, we simply have $c = x_0$.

2.2 The Exponential Matrix

Let $\mathcal{M}(\mathbf{C})$ be the finite dimensional vector space formed by the square matrices $M = (m_{ij})_{i,j=1,\ldots,n}$ of dimensions $n \times n$ with complex entries, endowed with the Frobenius norm. It is possible to prove that the series

$$\sum_{k=0}^{\infty} \frac{M^k}{k!} \ ,$$

converges for each $M \in \mathcal{M}(\mathbf{C})$ (see for instance [17], p. 83). Its sum is denoted e^M and it is called the *exponential matrix* of M. We list below the main properties of the exponential matrix.

- If the entries of M are real, then the entries of e^M are real.
- $e^0 = I$, where 0 denotes here a matrix whose entries are all equal to zero and I is the identity matrix.
- $e^{N+M} = e^M e^N$, provided that $MN = NM$.
- The eigenvalues of e^M are the complex numbers of the form e^λ, where λ is an eigenvalue of M.
- $e^M M = M e^M$.
- $\det e^M = e^{\operatorname{tr} M}$. As a consequence, $\det e^M \neq 0$ for each M.
- If P is a nonsingular matrix, $e^{P^{-1}MP} = P^{-1} e^M P$.

Let us come back to system (2.1). For each $t \in \mathbf{R}$, all the entries of the matrix e^{tA} are of class C^1. Moreover, the following proposition holds.

Proposition 2.1 *For each $A \in \mathcal{M}(\mathbf{C})$ and each $t \in \mathbf{R}$, we have*

$$\frac{d}{dt} e^{tA} = A e^{tA} \ .$$

Thus, the exponential matrix provides a useful formalism, which allows us to represent the solutions of the system (2.1). Indeed, if $x = \varphi(t)$ is the solution of (2.1) such that $\varphi(t_0) = x_0$, then by using the uniqueness of solutions and the properties of the exponential matrix, we get

$$\varphi(t) = e^{(t-t_0)A} x_0 \ .$$

If $t_0 = 0$, we simply have

$$\varphi(t) = e^{tA} x_0 \tag{2.5}$$

for each $t \in \mathbf{R}$. In other words, computing the exponential matrix is equivalent to compute a fundamental matrix of the system (actually, the principal fundamental matrix).

In the following sections, we will see how to realize an explicit construction of the exponential matrix. The final result will be achieved through several steps. We start by examining some special situations.

2.3 The Diagonal Case

Let

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ 0 & \dots & \dots & 0 \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} = \text{diag}\,(\lambda_1, \dots, \lambda_n)$$

where $\lambda_1, \dots, \lambda_n$ are not necessarily distinct numbers (real or complex).

Remark 2.2 For such a matrix, λ is an eigenvalue if and only if $\lambda = \lambda_i$ for some $i = 1, \dots, n$, and the algebraic multiplicity of λ indicates how many λ_i's are equal to λ. The eigenvectors corresponding to $\lambda_1, \dots, \lambda_n$ can be taken respectively coincident with the vectors of the canonical basis

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, v_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \tag{2.6}$$

■

A fundamental set of solutions of (2.1) can be therefore written in the form

$$\varphi_1(t) = e^{\lambda_1 t} v_1, \ \dots, \varphi_n(t) = e^{\lambda_n t} v_n \, .$$

A system (2.1) defined by a diagonal matrix A is called *decoupled*, since the evolution of each component x_i of x depends on x_i, but not on x_j with $j \neq i$. A system of this type can be trivially solved by integrating separately the single equations. The fundamental set of solutions obtained by this method obviously coincides by the previous one. The same fundamental set of solutions can be obtained also by computing the exponential matrix. Indeed, it is easy to check that for each positive integer k,

$$A^k = \text{diag}\,(\lambda_1^k, \dots, \lambda_n^k) \, ,$$

hence

$$e^{tA} = \text{diag}\,(e^{\lambda_1 t}, \dots, e^{\lambda_n t}) \, .$$

2.4 The Nilpotent Case

If A is nilpotent, there exists a positive integer q such that $A^k = 0$ for each $k \geq q$. Thus, the power series which defines the exponential matrix reduces to a polynomial and can be computed in elementary way. A typical nilpotent matrix (for which $q = n$) is

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \\ . \end{pmatrix} \tag{2.7}$$

The direct computation of the exponential matrix shows that if A has the form (2.7), then

$$e^{tA} = \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \dots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & t & \dots & \frac{t^{n-2}}{(n-2)!} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & t \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} .$$

Alternatively, we can write the corresponding system

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 \\ \dots \\ \dot{x}_n = 0 \end{cases}$$

and solve it by cascaded integration (from down to top). The two approaches obviously lead to the same result. A fundamental set of solutions can be written in the form

$$\varphi_1(t) = v_1 \ , \quad \varphi_2(t) = tv_1 + v_2 \ , \ \dots \ , \quad \varphi_n(t) = \frac{t^{n-1}}{(n-1)!} v_1 + \dots + tv_{n-1} + v_n \tag{2.8}$$

where the vectors $v_1, \dots, v_n$ are as in (2.6) the vector of the canonical basis.

Remark 2.3 Notice that zero is the unique eigenvalue of the matrix (2.7); the corresponding proper subspace is one dimensional. Moreover, $Av_1 = 0$ (which means that v_1 is an eigenvector of A), $Av_2 = v_1$, $Av_3 = v_2$ and so on. ■

The general integral of the system defined by the matrix (2.7) can be written as

$$\varphi(t) = c_1\varphi_1(t) + \dots + c_n\varphi_n(t) = d_1 + td_2 + \dots + \frac{t^{n-1}}{(n-1)!} d_n$$

where

$$d_1 = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{n-1} \\ c_n \end{pmatrix}, \quad d_2 = \begin{pmatrix} c_2 \\ c_3 \\ \vdots \\ c_n \\ 0 \end{pmatrix}, \ldots, \quad d_n = \begin{pmatrix} c_n \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} .$$

Notice that $Ad_1 = d_2, Ad_2 = d_3, \ldots, Ad_n = 0$. Notice also that d_1 can be arbitrarily chosen, and that d_n is an eigenvector of A, regardless to the choice of d_1.

Remark 2.4 Combining the methods used for the cases of diagonal and nilpotent matrices, we are able to compute the exponential matrix for each matrix A of the form $\lambda I + T$ where λ is any real number, I is the identity matrix of dimensions $n \times n$, and T is nilpotent. In particular, if T has the form (2.7), then

$$e^{t(\lambda I+T)} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \ldots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & t & \ldots & \frac{t^{n-2}}{(n-2)!} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & t \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix} . \tag{2.9}$$

■

2.5 The Block Diagonal Case

If M is block diagonal, that is

$$M = \begin{pmatrix} M_1 & 0 & \ldots & 0 \\ 0 & M_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & M_k \end{pmatrix} = \operatorname{diag}(M_1, \ldots, M_k) ,$$

then also its exponential matrix is block diagonal

$$e^M = \operatorname{diag}(e^{M_1}, \ldots, e^{M_k}) .$$

The exponential matrix of M is easily obtained, provided that we are able to construct the exponential matrix of every block M_i.

2.6 Linear Equivalence

To address the problem of computing the exponential matrix in the general case, we need to introduce the concept of linear equivalence.

Let us image system (2.1) as the mathematical model of a process evolving in a real vector space V of dimension n, where a basis has been fixed. The state of the system is represented, in this basis, by the n-tuple $x = (x_1, \dots, x_n)^{\mathbf{t}}$.

Assume that a new basis of V is given, and let $y = (y_1, \dots, y_n)^{\mathbf{t}}$ be the components of the state in this new basis. As well known, there exists a nonsingular matrix P such that for each element of V,

$$x = Py.$$

We want to see how (2.1) changes, when the state is represented in the new basis. We have

$$\dot{y} = P^{-1}\dot{x} = P^{-1}APy = By \ . \tag{2.10}$$

We therefore obtain again a linear system, defined by a matrix B which is similar to the given matrix A. Vice versa, two systems of the type (2.1) defined by similar matrices can be always thought of as two representations of the same system in two different systems of coordinates.

Definition 2.1 Two systems

$$\dot{x} = Ax \qquad \text{and} \qquad \dot{y} = By \ , \qquad x \in \mathbf{R}^n, \ y \in \mathbf{R}^n$$

are said to be *linearly equivalent* if A and B are similar, that is if $B = P^{-1}AP$ for some nonsingular matrix P.

The previous definition is actually an equivalence relation. It is clear that each solution $x = \varphi(t)$ of the first system is of the form $\varphi(t) = P\psi(t)$ where $y = \psi(t)$ is a solution of the second one and vice-versa. On the other hand, it is easy to see that

$$e^{tB} = P^{-1}e^{tA}P \qquad (\text{or, equivalently, } e^{tA} = Pe^{tB}P^{-1}) \ . \tag{2.11}$$

Hence, as far as we are interested in solution representation, we can work with any system linearly equivalent to the given one, and finally we can use (2.11) in order to come back to the original coordinates.

The notion of linear equivalence, as well as the notion of similar matrices, can be immediately generalized to the case where $x \in \mathbf{C}^n$. Of course, if A and B are similar matrices, A is real and B contains complex elements, then the matrix P determining the change of basis must contain complex elements, as well.

2.7 The Diagonalizable Case

It is well known that a matrix A of dimensions $n \times n$ is *diagonalizable* (that is, similar to a diagonal matrix) if and only if there exist n linearly independent vectors $v_1, \dots, v_n$ such that each v_i, $i = 1, \dots, n$, is an eigenvector of A. In such a case, we say that the vectors $v_1, \dots, v_n$ constitute a *proper basis* of A. In particular, A is diagonalizable if it admits n distinct eigenvalues.

Denoting by P the matrix whose columns are $v_1, \dots, v_n$ (in this order), we have

$$P^{-1}AP = \text{diag}\,(\lambda_1, \dots, \lambda_n) = D$$

where λ_1 is the eigenvalue of A corresponding to v_1, λ_2 is the eigenvalue of A corresponding to v_2 and so on (it is not required that the numbers $\lambda_1, \dots, \lambda_n$ are distinct).

To compute e^{tA} we can proceed in the following way: first we diagonalize A by means of the change of coordinates determined by P, then we compute e^{tD}, and finally we come back to the original coordinates, making use of (2.11).

Remark 2.5 If A is real but it admits complex eigenvalues, then P and D will have complex elements, as well. However, by construction, the elements of e^{tA} must be real.

Notice that $\Phi(t) = Pe^{tD}$ is a fundamental matrix; its computation do not require to know the inverse of P. However, in general the elements of Pe^{tD} are not real, not even if A is real.

In conclusion, to determine explicitly the elements of the matrix e^{tA} and hence the general integral of (2.1) in the diagonalizable case, it is sufficient to know the eigenvalues of A and the corresponding eigenvectors.

Example 2.1 Let us consider the system

$$\begin{cases} \dot{x}_1 = -x_2 \\ \dot{x}_2 = x_1 \end{cases}$$

defined by the matrix

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

The eigenvalues of A are $+\mathrm{i}$, with eigenvector $\begin{pmatrix} \mathrm{i} \\ 1 \end{pmatrix}$, and $-\mathrm{i}$, with eigenvector $\begin{pmatrix} -\mathrm{i} \\ 1 \end{pmatrix}$. It is easy to identify two (complex conjugate) linearly independent solutions

$$\varphi_1(t) = e^{\mathrm{i}t} \begin{pmatrix} \mathrm{i} \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} + \mathrm{i} \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$$

and

$$\varphi_2(t) = e^{-\mathrm{i}\,t}\begin{pmatrix}\mathrm{i}\\1\end{pmatrix} = \begin{pmatrix}-\sin t\\ \cos t\end{pmatrix} - \mathrm{i}\begin{pmatrix}\cos t\\ \sin t\end{pmatrix}\ .$$

Taking their real and imaginary parts we obtain two linearly independent real solutions

$$\psi_1(t) = \begin{pmatrix}-\sin t\\ \cos t\end{pmatrix} \quad \text{and} \quad \psi_2(t) = \begin{pmatrix}\cos t\\ \sin t\end{pmatrix}\ .$$

Alternatively, we can apply the diagonalization procedure. To this end, we need to compute the inverse matrix of

$$P = \begin{pmatrix}-\mathrm{i} & \mathrm{i}\\ 1 & 1\end{pmatrix}$$

given by

$$P^{-1} = -\frac{1}{2\mathrm{i}}\begin{pmatrix}1 & -\mathrm{i}\\ -1 & -\mathrm{i}\end{pmatrix}\ .$$

We easily get

$$D = P^{-1}AP = \begin{pmatrix}\mathrm{i} & 0\\ 0 & -\mathrm{i}\end{pmatrix}\ ,$$

and

$$e^{tD} = \begin{pmatrix}e^{\mathrm{i}\,t} & 0\\ 0 & e^{-\mathrm{i}\,t}\end{pmatrix}\ .$$

Finally,

$$e^{tA} = Pe^{tD}P^{-1} = \begin{pmatrix}\cos t & -\sin t\\ \sin t & \cos t\end{pmatrix}\ .$$

In this case, the exponential matrix could be also obtained directly, by applying the definition; indeed, it is not difficult to see that

$$A^4 = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\ .$$

2.8 Jordan Form

In this section, for any $n \times n$ matrix A, we denote by $\lambda_1, \dots, \lambda_k$ $(1 \le k \le n)$ its distinct eigenvalues. For each eigenvalue λ_i of A, by μ_i and ν_i we denote respectively the algebraic and geometric multiplicity of λ_i $(1 \le \nu_i \le \mu_i)$. Moreover, we will write $\lambda_i = \alpha_i + \mathrm{i}\,\beta_i$.

If A possesses eigenvalues with algebraic multiplicity greater than one and with geometric multiplicity less than the algebraic multiplicity, then A is not diagonalizable. In other words, the number of linearly independent eigenvectors is not sufficient to form a basis of the space. To overcome the difficulty, we resort to generalized eigenvectors. The following theorem holds (see for instance [4]).

Theorem 2.1 *Each matrix A of dimension $n \times n$ is similar to a block-diagonal matrix of the form*

$$J = \begin{pmatrix} C_{1,1} & 0 & \dots & 0 \\ 0 & C_{1,2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & C_{k,\nu_k} \end{pmatrix}$$

where the blocks $C_{i,j}$ are square matrices of the form

$$C_{i,j} = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda_i \end{pmatrix} .$$

Only one eigenvalue appears in each block, but a single eigenvalue can appear in more than one block. More precisely, for each eigenvalue λ_i there are exactly ν_i blocks, and each block is associated to one proper eigenvector. The dimension of a block $C_{i,j}$ equals the length of the chain of generalized eigenvectors originating from the j-th eigenvector associated to λ_i. The eigenvalue λ_i appears exactly μ_i times on the principal diagonal of J.

The matrix J is called a *Jordan form* of A. From our point of view, it is important to remark that each block J has the form $\lambda_i I + T$, where I is the identity matrix (of appropriate dimension), and T is the nilpotent matrix of type (2.7). Taking into account the conclusions of Sect. 2.5, the strategy illustrated for the case of a diagonalizable matrix can be therefore extended to the present situation: we transform the given system (2.1) to the system

$$\dot{y} = Jy \tag{2.12}$$

by means of a suitable change of coordinates, then we find the solutions of (2.12) directly (or, alternatively, we compute e^{tJ}, and we come back to the original coordinates). It remains only the problem of identifying the matrix P which determines the similarity between A and J. To this purpose, as already sketched, we need to determine for each eigenvalue λ_i, a number of linearly independent eigenvectors and generalized eigenvectors equal to μ_i. These vectors must be indexed in accordance to the order of the indices of the eigenvalues and, for each eigenvector, in accordance with the order of generation of the generalized eigenvectors of a same chain.

$$\overbrace{\underbrace{\underbrace{v_{1,1,0}}_{\textit{eigenvector}} \quad v_{1,1,1} \ \cdots}_{\textit{first chain}} \ \underbrace{\underbrace{v_{1,2,0}}_{\textit{eigenvector}} \quad v_{1,2,1} \cdots}_{\textit{second chain}} \ \underbrace{\underbrace{v_{1,3,0}}_{\textit{eigenvector}} \quad v_{1,3,1} \cdots}_{\textit{third chain}} \ \cdots}^{\lambda_1}$$

$$\overbrace{\underbrace{\underbrace{v_{2,1,0}}_{\textit{eigenvector}} \quad v_{2,1,1} \ \cdots}_{\textit{first chain}} \cdots}^{\lambda_2} \quad \cdots$$

The set of all these vectors constitutes a basis of the space, called again a *proper basis*. The columns of the matrix P are formed by the vectors of a proper basis in the aforementioned order, that is

$$\begin{aligned} P &= (v_{1,1,0} \mid v_{1,1,1} \mid \ldots \mid v_{1,2,0} \mid v_{1,2,1} \mid \ldots \mid v_{1,3,0} \mid v_{1,3,1} \mid \ldots \\ &\quad \ldots \mid v_{2,1,0} \mid v_{2,1,1} \mid \ldots) \,. \end{aligned}$$

Another proper basis and another corresponding Jordan form can be obtained by permutations of the order of the eigenvalues or, for each eigenvalue, permutations of the order of the corresponding eigenvectors (but leaving unchanged the order of generation of the generalized eigenvectors). In this sense, the Jordan form is not unique.

After that a proper basis has been constructed and provided that the order of the various indices is correctly settled out, we have all the information we need in order to explicitly write the Jordan form. In fact, we do not need to perform the change of coordinates. However, the computation of P and P^{-1} is inevitable in order to recover e^{tA} in the original coordinates. The computation of P^{-1} can be avoided, if we may limit ourselves to write the fundamental (in general, complex) matrix Pe^{tJ}.

Keeping in mind (2.9), and the procedure illustrated in Sect. 2.1 (Fact 7), we can resume the conclusions achieved so far in the following proposition.

Proposition 2.2 *The generic element* $\varphi_{r,s}(t)$ *of the matrix* e^{tA} $(r, s = 1, \ldots, n)$ *reads as*

$$\varphi_{r,s}(t) = \sum_{i=1}^{k} (Z_{r,s})_i(t) e^{\lambda_i t}$$

where each term $(Z_{r,s})_i(t)$ *is a polynomial (in general, with complex coefficients) whose degree is (strictly) less than the algebraic multiplicity of* λ_i*, and* λ_i *is an eigenvalue of* A $(i = 1, \ldots, k)$.

If A *is real, the generic element* $\varphi_{r,s}(t)$ *of the matrix* e^{tA} *can be put in the form*

$$\varphi_{r,s}(t) = \sum_{i=1}^{k} e^{\alpha_i t} \left[(p_{r,s})_i(t) \cos \beta_i t + (q_{r,s})_i(t) \sin \beta_i t \right] \tag{2.13}$$

where $(p_{r,s})_i$ *and* $(q_{r,s})_i$ *are polynomials with real coefficients whose degree is (strictly) less than the algebraic multiplicity of* λ_i *(of course, the previous formula includes also the contributions of the real eigenvalues, for which* $\beta_i = 0$*).*

2.9 Asymptotic Estimation of the Solutions

To our purposes, one of the main applications of the conclusions of the previous section is the estimation of the asymptotic behavior of the solutions of (2.1) for $t \to +\infty$.

Lemma 2.1 *For each* $\varepsilon > 0$ *and each integer* $m \in \mathbf{N}$ *there exists a constant* $k > 0$ *such that* $t^m < ke^{\varepsilon t}$*, for each* $t \geq 0$.

Proof The proof can be carried on by mathematical induction. If $m = 1$ we can take $k = \frac{1}{\varepsilon}$. Indeed, setting

$$f(t) = \frac{e^{\varepsilon t}}{\varepsilon} - t$$

we have $f(0) = \frac{1}{\varepsilon}$ and $f'(t) = e^{\varepsilon t} - 1 > 0$ for $t > 0$. Let us assume that the result holds for $m - 1$, with $k = \bar{k}$. The function

$$f(t) = ke^{\varepsilon t} - t^m$$

is such that

$$f(0) = k \quad \text{and} \quad f'(t) = k\varepsilon e^{\varepsilon t} - mt^{m-1} = m\left(\frac{k\varepsilon}{m}e^{\varepsilon t} - t^{m-1}\right) > 0$$

for $t > 0$, provided that we choose $k = \frac{m\bar{k}}{\varepsilon}$. ■

Let α_0 be the maximum of the real parts α_i of the eigenvalues λ_i of the matrix A $(i = 1, \ldots, k)$ and let α be any real number greater than α_0:

$$\alpha > \alpha_0 \geq \alpha_i \quad \text{for each} \quad (i = 1, \ldots, k)\ .$$

Since $|\sin \beta_i t| \leq 1$ and $|\cos \beta_i t| \leq 1$ for each $i = 1, \ldots, k$, starting from (2.13) and using repeatedly the triangular inequality we get, for $t \geq 0$,

$$|\varphi_{r,s}(t)| \leq \sum_{i=1}^{k} e^{\alpha_i t}\left(|(p_{r,s})_i(t)| + |(q_{r,s})_i(t)|\right) \leq \sum_{i=1}^{k}(Q_{r,s})_i(t)e^{\alpha_i t}$$

where $(Q_{r,s})_i$ is a polynomial whose coefficients are nonnegative real numbers, which majorize the absolute values of the corresponding coefficients of the polynomials

$(p_{r,s})_i(t)$ and $(q_{r,s})_i(t)$. Even if not essential for the subsequent developments, we note that the degree of $(Q_{r,s})_i$ is less than the algebraic multiplicity of λ_i.

Let $0 < \varepsilon < \alpha - \alpha_0$. By Lemma 2.1, there are constants $k_{r,s}$ such that $|\varphi_{r,s}(t)| \leq k_{r,s}e^{(\alpha_0+\varepsilon)t} \leq k_{r,s}e^{\alpha t}$ for each $t \geq 0$. Hence $\|e^{tA}\| = \sqrt{\sum_{r,s}\varphi^2_{r,s}(t)} \leq \sqrt{\sum_{r,s}k^2_{r,s}}e^{\alpha t}$ and, finally,

$$\|e^{tA}\| \leq k_0 e^{\alpha t} \qquad \forall t \geq 0$$

where k_0 is a new constant.

Note that if all the eigenvalues λ_i whose real parts are exactly equal to α_0 (i.e., $\alpha_i = \operatorname{Re}\lambda_i = \alpha_0$) have algebraic multiplicity coincident with the geometric multiplicity, then the previous inequality holds even when α is replaced by α_0. Indeed, in this case the corresponding polynomials $(p_{r,s})_i(t)$ and $(q_{r,s})_i(t)$ reduce to constants. Hence, the term $(Q_{r,s})_i(t)e^{\alpha_i t}$ can be directly majorized by $e^{\alpha_0 t}$, apart from a multiplicative constant, without need of using Lemma 2.1. Concerning the eigenvalues λ_i for which $\alpha_i = \operatorname{Re}\lambda_i < \alpha_0$, we may apply Lemma 2.1 with $\varepsilon = \alpha_0 - \alpha_i$. The corresponding terms $(Q_{r,s})_i(t)e^{\alpha_i t}$ can therefore be majorized, apart form some multiplicative constants, by $e^{\alpha_i t}e^{\varepsilon t} = e^{\alpha_0 t}$. Summing up, we can state the following proposition.

Proposition 2.3 *Let A be a real matrix. For each $\alpha > \alpha_0$, there exists $k_0 > 0$ such that*

$$\|e^{tA}\| \leq k_0 e^{\alpha t} \qquad \forall t \geq 0\,. \tag{2.14}$$

If all the eigenvalues of A with real part equal to α_0 have the algebraic multiplicity coincident with the geometric multiplicity, then in (2.14) we can take $\alpha = \alpha_0$.

From (2.14) it follows

$$\|e^{tA}c\| \leq k_0\|c\|e^{\alpha t}\,, \quad t \geq 0 \tag{2.15}$$

for each real constant vector c.

2.10 The Scalar Equation of Order *n*

The scalar differential equation (with constant coefficients, $n > 1$)

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}y' + a_n y = 0 \tag{2.16}$$

can be thought of as a particular case of (2.1). Indeed, setting

$$y = x_1,\ y' = x_2,\ \ldots,\ y^{(n-1)} = x_n$$

and using (2.16) we have

$$\begin{aligned} x_1' &= y' = x_2 \\ x_2' &= y'' = x_3 \\ &\dots\dots\dots\dots \\ x_n' &= y^{(n)} = -a_1 x_n - \ldots - a_n x_1 \end{aligned}$$

that is, with vector notation,

$$\dot{x} = Cx \tag{2.17}$$

where we set $x = (x_1, \ldots, x_n)^{\mathbf{t}}$, and

$$C = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ -a_n & -a_{n-1} & \ldots & \ldots & -a_1 \end{pmatrix} . \tag{2.18}$$

By *solution* of (2.16) we obviously mean a n-times continuously differentiable function $y(t) : \mathbf{R} \to \mathbf{R}$ for which (2.16) is identically satisfied for each $t \in \mathbf{R}$. The problem of determining the general integral (that is, the set of all the solutions) of (2.16) is clearly equivalent to the problem of finding n linearly independent solutions of (2.17).

A matrix exhibiting the structure (2.18) is called a *companion matrix*. More precisely, we say that (2.18) is the companion matrix associated to the Eq. (2.16).

The special structure of C displayed by (2.18) allows us to identify immediately an important algebraic object, invariant under similarity. Indeed, using mathematical induction, it is easy to check that the characteristic polynomial of C is $p_C(\lambda) = (-1)^n \left[\lambda^n + a_1\lambda^{n-1} + \cdots + a_n\right]$. We are especially interested in the eigenvalues of C, which are the roots of $p_C(\lambda)$; hence, the coefficient $(-1)^n$ can be neglected. In fact, it is customary (even is slightly confusing) to call $(-1)^n p_C(\lambda)$ the *characteristic polynomial of the differential equation* (2.16). From now on, we adopt the notation

$$p_{ch}(\lambda) = (-1)^n p_C(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n . \tag{2.19}$$

Note that $p_{ch}(\lambda)$ is monic for each n, and that it can be immediately written, without need of transforming (2.16) in the equivalent system (2.17), by replacing formally y by λ and reinterpreting the orders of the derivatives as powers. It is also customary to say that

$$p_{ch}(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n = 0 \tag{2.20}$$

is the *characteristic equation* of the differential equation (2.16), and that its solutions are the *characteristic roots* of (2.16).

Now we change the point of view. Let A be an arbitrary $n \times n$ real matrix, and let

$$p_A(\lambda) = (-1)^n \left[\lambda^n + a_1 \lambda^{n-1} + \cdots + a_n\right] . \tag{2.21}$$

its characteristic polynomial. Write a matrix C_A of the form (2.18), reporting in the last row the coefficients $a_1, \ldots, a_n$ taken from (2.21). In this way, A and C_A will have the same characteristic polynomial and hence the same eigenvalues (with the same algebraic multiplicity). The matrix C_A is called the *the companion matrix associated to A*. Unfortunately, in general, A and C_A need not to be similar. For instance, the characteristic polynomial of the identity matrix is $p_I(\lambda) = \sum_{i=0}^{n}(-1)^i \binom{n}{i}\lambda^i$. We may write the associated companion matrix C_I; however the identity matrix I is not similar to C_I, the class of equivalence of I under the similarity relation being just the singleton $\{I\}$. It follows that not all the systems of linear differential equations in $\mathbf{R}^n$ can be reduced by linear transformations to a scalar equation of order n.

The following theorem provides conditions ensuring that a given matrix A is similar to its associated matrix C_A in companion form. This theorem has its own interest from an algebraic point of view, but it is also very important for our future developments.

Theorem 2.2 *Let A be a square matrix of dimensions $n \times n$. The following properties are equivalent.*

(i) *A is similar to its associated matrix in companion form.*
(ii) rank $(A - \lambda I) = n - 1$ *for each eigenvalue λ of A.*
(iii) *The geometric multiplicity of each eigenvalue of A is equal to* 1.
(iv) *The characteristic polynomial of A coincides with its minimal polynomial.*
(v) *There exists a vector $v \neq 0$ such that the n vectors*

$$v, Av, A^2 v, \ldots, A^{n-1} v$$

are linearly independent.

The complete proof of Theorem 2.2 can be found for instance in [22]. To our future purposes, the equivalence between *(i)* and *(v)* is especially important.[2] A vector v enjoying the property stated in *(v)* is said to be *cyclic* for A.

Thus, if the $n \times n$ matrix A satisfies one of the assumptions of Theorem 2.2, solving the linear system defined by A is actually equivalent to solve a differential equation of order n of the form (2.16). Property *(iii)* of Theorem 2.2 implies in particular that for each eigenvalue λ of A there is a unique eigenvector v and hence a unique chain of possible generalized eigenvectors engendered by v.

[2]For reader's convenience, a proof of this equivalence will be given in the next section.

Of course, statements *(i),...,(v)* are fulfilled by any matrix C assigned in companion form. It follows that the general integral of (2.16) can be obtained as a linear combination of the n functions

$$\begin{array}{l} e^{\lambda_1 t},\ te^{\lambda_1 t}, \dots, t^{\mu_i - 1} e^{\lambda_1 t} \\ \dots\dots\dots\dots\dots\dots\dots \\ e^{\lambda_k t},\ te^{\lambda_k t}, \dots, t^{\mu_k - 1} e^{\lambda_k t} \end{array}$$

where $\lambda_1, \dots, \lambda_k$ are the distinct roots of the characteristic Eq. (2.20) and $\mu_1, \dots, \mu_k$ the respective algebraic multiplicities.

Example 2.2 Let us consider in detail the case of a linear equation of order 2

$$y'' + ay' + by = 0 \,. \tag{2.22}$$

To write the general integral $y(t)$, first we need to discuss the characteristic equation

$$\lambda^2 + a\lambda + b = 0 \,. \tag{2.23}$$

If (2.23) has two distinct real solutions λ_1, λ_2, then we have

$$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} \,. \tag{2.24}$$

If (2.23) has a unique real solution $\lambda_1 = \lambda_2 = \lambda$ of multiplicity 2, then we have

$$y(t) = (c_1 + tc_2) e^{\lambda t} \,. \tag{2.25}$$

Finally, if (2.23) has complex (not real) conjugate solutions[3] $\alpha \pm \mathrm{i}\,\beta$, then we have

$$y(t) = (c_1 \cos \beta t + c_2 \sin \beta t) e^{\alpha t} \,. \tag{2.26}$$

The behavior of the solutions for $t \geq 0$ depends on the signs of λ_1 and λ_2 in the case (2.24) and, respectively, an the signs of λ and α in the cases (2.25) (2.26).

For instance, if $\lambda_1, \lambda_2 < 0$ [respectively, $\lambda < 0, \alpha < 0$] the energy initially stored in the system (measured by the initial conditions) is dissipated:

---

[3]The expression (2.26) results from the application of formulæ (2.2). Alternatively, (2.26) can be obtained starting from the complex version of (2.24)

$$k_1 e^{\lambda t} + k_2 e^{\overline{\lambda} t}$$

using the fact that $e^{\alpha \pm \mathrm{i}\,\beta} = e^{\alpha t}(\cos \beta t \pm \mathrm{i}\, \sin \beta t)$ and setting

$$c_1 = k_1 + k_2 \qquad c_2 = -\mathrm{i}\,(k_1 - k_2) \,.$$

In particular, c_1 and c_2 turn out to be real if we take $k_2 = \overline{k}_1$.

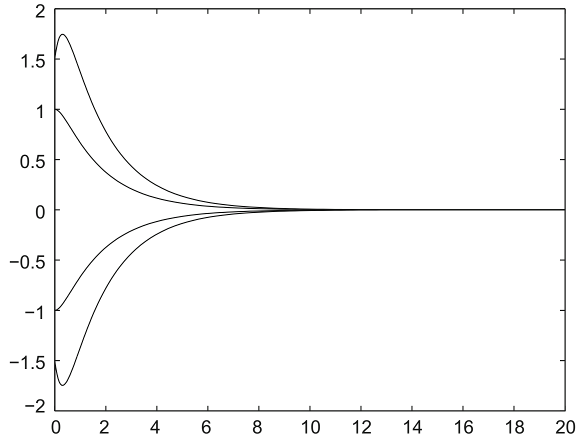

Fig. 2.1 Dissipation in the case of real characteristic roots

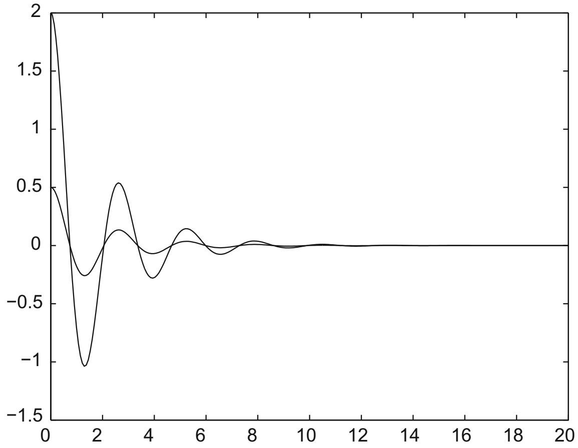

Fig. 2.2 Dissipation in the case of complex characteristic roots

- monotonically in the cases (2.24) and (2.25), after possible initial picks, whose occurrence depends on the choice of c_1 and c_2 (Fig. 2.1);
- with oscillatory decay in the case (2.26) (Fig. 2.2).

The case $a = 0$ and $b > 0$ is a particular instance of (2.26) with $\alpha = 0$ and $\beta = \sqrt{b}$: the solutions are periodic with minimal period $2\pi/\beta$. The general integral takes the form

$$y(t) = c_1 \cos \beta t + c_2 \sin \beta t = \rho \cos(\beta t + \theta) \tag{2.27}$$

where ρ and $\theta \in [0, 2\pi)$ are identified by the relations $c_1 = \rho \cos \theta, c_2 = \rho \sin \theta$. The numbers $\rho = \sqrt{c_1^2 + c_2^2}$ and θ are called *amplitude* and *phase* of the periodic function (2.27). The inverse of the minimal period is called the *frequency*. Note that in (2.27), the frequency depends on b while the amplitude depends on the initial conditions and remains constant. In other words, in this case we have conservation of the energy. The reader can easily check that these conclusions agree with those of Example 2.1.

Remark 2.6 Let us denote by D the derivative operator. Formally, (2.16) can be rewritten as

$$L(D)y = (D^n + a_1 D^{n-1} + \cdots + a_n)y = 0$$

where $(-1)^n L(D) = p_{ch}(D)$. Notice that $L(D)$ acts as a linear operator.

2.11 The Companion Matrix

In this section we show that a matrix A is similar to the associated matrix in companion form if and only if there exists a cyclic vector for A, that is a vector $v \neq 0$ such that $v, Av, \ldots, A^{n-1}v$ form a basis of $\mathbf{R}^n$ (this proves the equivalence of statements *(i)* and *(v)* of Theorem 2.2).

Lemma 2.2 *Let C be a matrix in companion form, and let $\lambda_1, \ldots, \lambda_n$ its eigenvalues (not necessarily distinct). Then, C is similar to a matrix of the form*

$$M = \begin{pmatrix} \lambda_1 & 1 & 0 \ldots & 0 & 0 \\ 0 & \lambda_2 & 1 \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \ldots & \lambda_{n-1} & 1 \\ 0 & 0 & 0 \ldots & \ldots & \lambda_n \end{pmatrix}. \tag{2.28}$$

Proof Let us start with Eq. (2.16). Let us show that by means of suitable linear substitutions, (2.16) can be transformed in a system of first order linear equations defined by the matrix (2.28). Let us set

$$\xi_1 = y, \quad \xi_2 = p_{2,1}y + y', \ldots, \quad \xi_n = p_{n,1}y + \cdots + p_{n,n-1}y^{(n-2)} + y^{(n-1)}$$

where the coefficients $p_{i,j}$ are recovered by the relations

$$p_{2,1}y + y' = (-\lambda_1 + D)y$$
$$p_{3,1}y + p_{3,2}y' + y'' = (-\lambda_1 + D)(-\lambda_2 + D)y$$
$$\ldots\ldots\ldots\ldots$$
$$p_{n,1}y + \cdots + p_{n,n-1}y^{(n-2)} + y^{(n-1)} = (-\lambda_1 + D) \cdot \cdots \cdot (-\lambda_{n-1} + D)y$$

and D is the derivation operator. We have

$$\begin{aligned}
\xi_1' &= y' = y' + p_{2,1}y - p_{2,1}y = (-\lambda_1 + D)y + \lambda_1 y = \lambda_1\xi_1 + \xi_2 \\
\xi_2' &= (p_{2,1}y + y')' = \\
&= D(-\lambda_1 + D)y + \lambda_2(-\lambda_1 + D)y - \lambda_2(-\lambda_1 + D)y = \\
&= (-\lambda_2 + D)(-\lambda_1 + D)y + \lambda_2\xi_2 = \lambda_2\xi_2 + \xi_3 \\
&\qquad \ldots\ldots\ldots\ldots\ldots \\
\xi_n' &= D(p_{n,1}y + \cdots + p_{n,n-1}y^{(n-2)} + y^{(n-1)}) = \\
&= D(-\lambda_{n-1} + D) \cdot \cdots \cdot (-\lambda_1 + D)y = \\
&= (-\lambda_n + D) \cdot \cdots \cdot (-\lambda_1 + D)y + \lambda_n(-\lambda_{n-1} + D) \cdot \cdots \cdot (-\lambda_1 + D)y\ .
\end{aligned}$$

The term $(-\lambda_n + D) \cdot \cdots \cdot (-\lambda_1 + D)y$ vanishes, since it coincides with (2.16). Hence we get

$$\xi_n' = \lambda_n\xi_n\ .$$

The statement easily follows.

We emphasize that (2.28) is not a Jordan form of C (it coincides with the Jordan form of C, only in the case where $\lambda_1 = \lambda_2 = \cdots = \lambda_n$). Let P be the matrix such that $C = P^{-1}MP$. Then P has the form

$$\begin{pmatrix}
1 & 0 & 0 & \ldots & 0 & 0 \\
p_{2,1} & 1 & 0 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots \\
p_{n,1} & p_{n,2} & p_{n,3} & \ldots & p_{n,n-1} & 1
\end{pmatrix}$$

where the numbers p_{ij} are the same as in the proof of Lemma 2.2. Let us remark that the companion form is not the unique way to rewrite (2.16) as a system of first order equations. We can take for instance

$$\begin{cases}
z_1 = a_{n-1}y + a_{n-2}y' + \cdots + a_1 y^{(n-2)} + y^{(n-1)} \\
z_2 = a_{n-2}y + a_{n-3}y' + \cdots + a_1 y^{(n-3)} + y^{(n-2)} \\
\ldots\ldots\ldots\ldots \\
z_{n-1} = a_1 y + y' \\
z_n = y\ .
\end{cases}$$

Then we have

$$\begin{cases} z_1' & = a_{n-1}y' + a_{n-2}y'' + \cdots + a_1 y^{(n-1)} + y^{(n)} = \\ & = -a_n y = -a_n z_n \\ z_2' & = a_{n-2}y' + \cdots + a_1 y^{(n-2)} + y^{(n-1)} = \\ & = a_{n-2}y' + \cdots + a_1 y^{(n-2)} + z_1 - \\ & -(a_{n-1}y + a_{n-2}y' + \cdots + a_1 y^{(n-2)}) = \\ & = z_1 - a_{n-1}y = z_1 - a_{n-1}z_n \\ & \ldots\ldots\ldots\ldots \\ z_n' & = y' = z_{n-1} - a_1 y = z_{n-1} - a_1 z_n \,. \end{cases}$$

This system corresponds to the matrix

$$C^{\mathbf{t}} = \begin{pmatrix} 0 & 0 & \ldots & 0 & -a_n \\ 1 & 0 & \ldots & 0 & -a_{n-1} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -a_1 \end{pmatrix} .$$

Since all these substitutions are linear and invertible, we have actually proved that C and $C^{\mathbf{t}}$ are similar (as a matter of fact, this is true for every square matrix).

We are now able to conclude the proof. Let us assume that A is similar to its companion form C_A. We know by Lemma 2.2 that A is similar to the matrix M given by (2.28), as well.

Let $w = (0, \ldots, 0, 1)^{\mathbf{t}}$. The result of the multiplication Mw is a vector coinciding with the last column of M. Let us perform the iterated multiplications $M^2w = M(Mw)$, $M^3w = M(M^2w), \ldots$ and let us form a new matrix whose columns are given by the vectors $w, Mw, \ldots, M^{n-1}w$, in this order:

$$\begin{pmatrix} 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 1 & \ldots & * \\ 0 & 1 & \lambda_n + \lambda_{n-1} & \ldots & * \\ 1 & \lambda_n & \lambda_n^2 & \ldots & * \end{pmatrix}$$

where we denoted by $*$ some unessential functions of $\lambda_1, \ldots, \lambda_n$. This matrix is not singular, which means that $w, Mw, \ldots, M^{n-1}w$ are linearly independent. Now let P be the matrix transforming A in M, and let $v = Pw$. We have

$$(w|Mw|\ldots|M^{n-1}w) = (P^{-1}Pw|P^{-1}APw|\ldots|P^{-1}A^{n-1}Pw) = \\ = P^{-1}(v|Av|\ldots|A^{n-1}v)$$

which yields the desired conclusion. Vice versa, we finally prove that A is similar to $C_A^{\mathbf{t}}$ (the transpose of the companion form of A) provided that the condition *(v)* of Theorem 2.2 holds. Setting $R = (v, Av, \ldots, A^{n-1}v)$, we have to prove that $R^{-1}AR = C_A^{\mathbf{t}}$ or, equivalently,

$$AR = (Av|A^2v|\ldots|A^n v) = RC_A^{\mathbf{t}} =$$
$$= (v|Av|\ldots|A^{n-1}v) \cdot \begin{pmatrix} 0 & 0 & \ldots & 0 & -a_n \\ 1 & 0 & \ldots & 0 & -a_{n-1} \\ 0 & 1 & \ldots & 0 & -a_{n-2} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -a_1 \end{pmatrix} .$$

The computations are not difficult (for the last column we need to apply the Cayley-Hamilton Theorem). We already know that a matrix in companion form and its transpose are similar, but we can also proceed in a direct way. Indeed, it is sufficient to remark that a matrix in companion form satisfies *(v)* with $v = (0, \ldots, 0, 1)^{\mathbf{t}}$. By repeating the same computations as before, we recover the required similarity.

Chapter Summary

This chapter is devoted to the mathematical problem of representing the solutions of a homogeneous system of linear differential equations by means of suitable explicit formulæ. This corresponds to the study of the qualitative behavior of a system when the evolution depends only on the internal forces and the external inputs are switched off. It is actually the first step in the investigation of the properties of a system.