# 2. Hamiltonian Systems

The study of Hamiltonian systems starts with the study of linear systems and the associated linear algebra. This will lead to basic results on periodic systems and variational equations of nonlinear systems.

The basic linear algebra introduced in this chapter will be the cornerstone of many of the later results on nonlinear systems. Some of the more advanced results which require a knowledge of multilinear algebra or the theory of analytic functions of a matrix are relegated to Chapter 5 or to the literature.

## 2.1 Linear Equations

A familiarity with the basic theory of linear algebra and linear differential equations is assumed. Let $gl(m, \mathbb{F})$ denote the set of all $m \times m$ matrices with entries in the field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$) and $Gl(m, \mathbb{F})$ the set of all nonsingular $m \times m$ matrices with entries in $\mathbb{F}$. $Gl(m, \mathbb{F})$ is a group under matrix multiplication and is called the general linear group. Let $I = I_m$ and $0 = 0_m$ denote the $m \times m$ identity and zero matrices, respectively. In general, the subscript is clear from the context and so will be omitted.

In this theory a special role is played by the $2n \times 2n$ matrix

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}. \tag{2.1}$$

Note that $J$ is orthogonal and skew-symmetric; i.e.,

$$J^{-1} = J^T = -J. \tag{2.2}$$

Here as elsewhere the superscript $T$ denotes the transpose.

Let $z$ be a coordinate vector in $\mathbb{R}^{2n}$, $\mathbb{I}$ an interval in $\mathbb{R}$, and $S : \mathbb{I} \to gl(2n, \mathbb{R})$ be continuous and symmetric. A linear Hamiltonian system is the system of $2n$ ordinary differential equations

$$\dot{z} = J\frac{\partial H}{\partial z} = JS(t)z = A(t)z, \tag{2.3}$$



K.R. Meyer, D.C. Offin, *Introduction to Hamiltonian Dynamical Systems and the N-Body Problem*, Applied Mathematical Sciences 90,
DOI 10.1007/978-3-319-53691-0_2

where

$$H = H(t,z) = \frac{1}{2} z^T S(t) z, \tag{2.4}$$

$A(t) = JS(t)$. The Hamiltonian $H$ is a quadratic form in $z$ with coefficients that are continuous in $t \in \mathbb{I} \subset \mathbb{R}$. If $S$, and hence $H$, is independent of $t$, then $H$ is an integral for (2.3) by Theorem 1.2.1.

Let $t_0 \in \mathbb{I}$ be fixed. From the theory of differential equations, for each $z_0 \in \mathbb{R}^{2n}$, there exists a unique solution $\phi(t, t_0, z_0)$ of (2.3) for all $t \in \mathbb{I}$ that satisfies the initial condition $\phi(t_0, t_0, z_0) = z_0$. Let $Z(t, t_0)$ be the $2n \times 2n$ fundamental matrix solution of (2.3) that satisfies $Z(t_0, t_0) = \mathrm{I}$. Then $\phi(t, t_0, z_0) = Z(t, t_0) z_0$.

In the case where $S$ and $A$ are constant, we take $t_0 = 0$ and

$$Z(t) = e^{At} = \exp At = \sum_{i=1}^{\infty} \frac{A^n t^n}{n!}. \tag{2.5}$$

A matrix $A \in gl(2n, \mathbb{F})$ is called *Hamiltonian* (or sometimes infinitesimally symplectic), if

$$A^T J + JA = 0. \tag{2.6}$$

The set of all $2n \times 2n$ Hamiltonian matrices is denoted by $sp(2n, \mathbb{R})$.

**Theorem 2.1.1.** *The following are equivalent:*

*(i) $A$ is Hamiltonian,*
*(ii) $A = JR$ where $R$ is symmetric,*
*(iii) $JA$ is symmetric.*

*Moreover, if $A$ and $B$ are Hamiltonian, then so are $A^T$, $\alpha A$ ($\alpha \in \mathbb{F}$), $A \pm B$, and $[A, B] \equiv AB - BA$.*

*Proof.* $A = J(-JA)$ and (2.6) is equivalent to $(-JA)^T = (-JA)$; thus (i) and (ii) are equivalent. Because $J^2 = -I$, (ii) and (iii) are equivalent. Thus the coefficient matrix $A(t)$ of the linear Hamiltonian system (2.1) is a Hamiltonian matrix.

The first three parts of the next statement are easy. Let $A = JR$ and $B = JS$, where $R$ and $S$ are symmetric. Then $[A, B] = J(RJS - SJR)$ and $(RJS - SJR)^T = S^T J^T R^T - R^T J^T S^T = -SJR + RJS$ so $[A, B]$ is Hamiltonian.

In the $2 \times 2$ case,

$$A = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$$

so,

$$A^T J + JA = \begin{bmatrix} 0 & \alpha + \delta \\ -\alpha - \delta & 0 \end{bmatrix}.$$

Thus, a $2 \times 2$ matrix is Hamiltonian if and only if its trace, $\alpha + \delta$, is zero. If you write a second-order equation $\ddot{x} + p(t)\dot{x} + q(t)x = 0$ as a system in the usual way with $\dot{x} = y$, $\dot{y} = -q(t)x - p(t)y$, then it is a linear Hamiltonian system when and only when $p(t) \equiv 0$. The $p(t)\dot{x}$ is usually considered the friction term.

Now let A be a $2n \times 2n$ matrix and write it in block form

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

so,

$$A^T J + JA = \begin{bmatrix} c - c^T & a^T + d \\ -a - d^T & -b + b^T \end{bmatrix}.$$

Therefore, A is Hamiltonian if and only if $a^T + d = 0$ and $b$ and $c$ are symmetric.

The function $[\cdot, \cdot] : gl(m, \mathbb{F}) \times gl(m, \mathbb{F}) \to gl(m, \mathbb{F})$ of Theorem 2.1.1 is called the *Lie product*. The second part of this theorem implies that the set of all $2n \times 2n$ Hamiltonian matrices, $sp(2n, \mathbb{R})$, is a Lie algebra. We develop some interesting facts about Lie algebras of matrices in the Problem section.

A $2n \times 2n$ matrix $T$ is called symplectic with multiplier $\mu$ if

$$T^T J T = \mu J, \tag{2.7}$$

where $\mu$ is a nonzero constant. If $\mu = +1$, then $T$ is simply said to be symplectic. The set of all $2n \times 2n$ symplectic matrices is denoted by $Sp(2n, \mathbb{R})$.

**Theorem 2.1.2.** *If $T$ is symplectic with multiplier $\mu$, then $T$ is nonsingular and*

$$T^{-1} = -\mu^{-1} J T^T J. \tag{2.8}$$

*If $T$ and $R$ are symplectic with multiplier $\mu$ and $\nu$, respectively, then $T^T, T^{-1}$, and $TR$ are symplectic with multipliers $\mu, \mu^{-1}$, and $\mu\nu$, respectively.*

*Proof.* Because the right-hand side, $\mu J$, of (2.7) is nonsingular, $T$ must be also. Formula (2.8) follows at once from (2.7). If $T$ is symplectic, then (2.8) implies $T^T = -\mu J T^{-1} J$; so, $TJT^T = TJ(-\mu J T^{-1} J) = \mu J$. Thus $T^T$ is symplectic with multiplier $\mu$. The remaining facts are proved in a similar manner.

This theorem shows that $Sp(2n, \mathbb{R})$ is a group, a closed subgroup of $Gl(2n, \mathbb{R})$. Weyl says that originally he advocated the name "complex group" for $Sp(2n, \mathbb{R})$, but it became an embarrassment due to the collisions with the word “complex” as in complex numbers. “I therefore proposed to replace it by the corresponding Greek adjective ‘symplectic.’” See page 165 in Weyl (1948).

In the $2 \times 2$ case

$$T = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$$

and so

$$T^T JT = \begin{bmatrix} 0 & \alpha\delta - \beta\gamma \\ -\alpha\delta + \beta\gamma & 0 \end{bmatrix}.$$

So a $2 \times 2$ matrix is symplectic (with multiplier $\mu$) if and only if it has determinant $+1$ (respectively $\mu$). Thus a $2 \times 2$ symplectic matrix defines a linear transformation which is orientation-preserving and area-preserving.

Now let $T$ be a $2n \times 2n$ matrix and write it in block form

$$T = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{2.9}$$

and so

$$T^T JT = \begin{bmatrix} a^T c - c^T a & a^T d - c^T b \\ b^T c - d^T a & b^T d - d^T b \end{bmatrix}.$$

Thus $T$ is symplectic with multiplier $\mu$ if and only if $a^T d - c^T b = \mu I$ and $a^T c$ and $b^T d$ are symmetric. Formula (2.8) gives

$$T^{-1} = \mu^{-1} \begin{bmatrix} d^T & -b^T \\ -c^T & a^T \end{bmatrix}. \tag{2.10}$$

This reminds one of the formula for the inverse of a $2 \times 2$ matrix!

**Theorem 2.1.3.** *The fundamental matrix solution $Z(t, t_0)$ of a linear Hamiltonian system (2.3) is symplectic for all $t, t_0 \in \mathbb{I}$. Conversely, if $Z(t, t_0)$ is a continuously differential function of symplectic matrices, then $Z$ is a matrix solution of a linear Hamiltonian system.*

*Proof.* Let $U(t) = Z(t, t_0)^T JZ(t, t_0)$. Because $Z(t_0, t_0) = I$, it follows that $U(t_0) = J$. Now

$$\dot{U}t) = \dot{Z}^T JZ + Z^T J\dot{Z} = Z^T (A^T J + JA)Z = 0;$$

so, $U(t) \equiv J$.

If $Z^T JZ = J$ for $t \in \mathbb{I}$, then $\dot{Z}^T JZ + Z^T J\dot{Z} = 0$; so, $(\dot{Z}Z^{-1})^T J + J(\dot{Z}Z^{-1}) = 0$. This shows that $A = \dot{Z}Z^{-1}$ is Hamiltonian and $\dot{Z} = AZ$.

**Corollary 2.1.1.** *The constant matrix $A$ is Hamiltonian if and only if $e^{At}$ is symplectic for all $t$.*

Change variables by $z = T(t)u$ in system (2.3). Equation (2.3) becomes

$$\dot{u} = (T^{-1}AT - T^{-1}\dot{T})u. \tag{2.11}$$

In general this equation will not be Hamiltonian, however:

**Theorem 2.1.4.** *If $T$ is symplectic with multiplier $\mu^{-1}$, then* (2.11) *is a Hamiltonian system with Hamiltonian*

$$H(t, u) = \frac{1}{2}u^T(\mu T^T S(t)T + R(t))u,$$

*where*

$$R(t) = JT^{-1}\dot{T}.$$

*Conversely, if* (2.11) *is Hamiltonian for every Hamiltonian system* (2.3)*, then $U$ is symplectic with constant multiplier $\mu$.*

The function $\frac{1}{2}u^T R(t)u$ is called the remainder.

*Proof.* Because $TJT^T = \mu^{-1}J$ for all $t$, $\dot{T}JT^T + TJ\dot{T}^T = 0$ or $(T^{-1}\dot{T})J + J(T^{-1}\dot{T})^T = 0$; so, $T^{-1}\dot{T}$ is Hamiltonian. Also $T^{-1}J = \mu JT^T$; so, $T^{-1}AT = T^{-1}JST = \mu JT^T ST$, and so, $T^{-1}AT = J(\mu T^T ST)$ is Hamiltonian also.

For the converse let (2.11) always be Hamiltonian. By taking $A \equiv 0$ we have that $T^{-1}\dot{T} = B(t)$ is Hamiltonian or $T$ is a matrix solution of the Hamiltonian system

$$\dot{v} = vB(t). \tag{2.12}$$

So, $T(t) = KV(t, t_0)$, where $V(t, t_0)$ is the fundamental matrix solution of (2.12), and $K = T(t_0)$ is a constant matrix. By Theorem 2.1.3, $V$ is symplectic.

Consider the change of variables $z = T(t)u = KV(t, t_0)u$ as a two-stage change of variables: first $w = V(t, t_0)u$ and second $z = Kw$. The first transformation from $u$ to $w$ is symplectic, and so, by the first part of this theorem, preserves the Hamiltonian character of the equations. Because the first transformation is reversible, it would transform the set of all linear Hamiltonian systems onto the set of all linear Hamiltonian systems. Thus the second transformation from $w$ to $z$ must always take a Hamiltonian system to a Hamiltonian system.

If $z = Kw$ transforms all Hamiltonian systems $\dot{z} = JCz$, $C$ constant and symmetric, to a Hamiltonian system $\dot{w} = JDw$, then $JD = K^{-1}JCK$ is Hamiltonian, and $JK^{-1}JCK$ is symmetric for all symmetric $C$. Thus

$$\begin{gathered} JK^{-1}JCK = (JK^{-1}JCK)^T = K^T CJK^{-T}J, \\ C(KJK^T J) = (JKJK^T)C, \\ CR = R^T C, \end{gathered}$$

where $R = KJK^TJ$. Fix $i$, $1 \le i \le 2n$ and take $C$ to be the symmetric matrix that has $+1$ at the $i,i$ position and zero elsewhere. Then the only nonzero row of $CR$ is the $i^{th}$, which is the $i^{th}$ row of $R$ and the only nonzero column of $R^TC$ is the $i^{th}$, which is the $i^{th}$ column of $R^T$. Because these must be equal, the only nonzero entry in $R$ or $R^T$ must be on the diagonal. So $R$ and $R^T$ are diagonal matrices. Thus $R = R^T = \mathrm{diag}(\mathrm{r_1}, \dots, \mathrm{r_{2n}})$, and $RC - CR = 0$ for all symmetric matrices C. But $RC - CR = ((r_j - r_i)c_{ij}) = (0)$. Because $c_{ij}$, $i < j$, is arbitrary, $r_i = r_j$, or $R = -\mu I$ for some constant $\mu$. $R = KJK^TJ = -\mu I$ implies $KJK^T = \mu J$.

This change of variables preserves the Hamiltonian character of a linear system of equations. The general problem of which changes of variables preserve the Hamiltonian character is discussed in detail in Chapter 2.6.

The fact that the fundamental matrix of (2.3) is symplectic means that the fundamental matrix must satisfy the identity (2.7). There are many functional relations in (2.7); so, there are functional relations between the solutions. Theorem 2.1.5 given below is just one example of how these relations can be used. See Meyer and Schmidt (1982) for some other examples.

Let $z_1, z_2 : \mathbb{I} \to \mathbb{R}^{2n}$ be two smooth functions where $\mathbb{I}$ is an interval in $\mathbb{R}^1$; we define the Poisson bracket of $z_1$ and $z_2$ to be

$$\{z_1, z_2\}(t) = z_1^T(t) J z_2(t); \tag{2.13}$$

so, $\{z_1, z_2\} : \mathbb{I} \to \mathbb{R}$ is smooth. The Poisson bracket is bilinear and skew-symmetric. Two functions $z_1$ and $z_2$ are said to be in *involution* if $\{z_1, z_2\} \equiv 0$. A set of $n$ linearly independent functions that are pairwise in involution are said to be a *Lagrangian set.* In general, the complete solution of a $2n$-dimensional system requires $2n$ linearly independent solutions, but for a Hamiltonian system a Lagrangian set of solutions suffices.

**Theorem 2.1.5.** *If a Lagrangian set of solutions of* (2.3) *is known, then a complete set of* $2n$ *linearly independent solutions can be found by quadrature. (See* (2.14)*.)*

*Proof.* Let $C = C(t)$ be the $2n \times n$ matrix whose columns are the $n$ linearly independent solutions. Because the columns are solutions, $\dot{C} = AC$; because they are in involution, $C^TJC = 0$; and because they are independent, $C^TC$ is an $n \times n$ nonsingular matrix. Define the $2n \times n$ matrix by $D = JC(C^TC)^{-1}$. Then $D^TJD = 0$ and $C^TJD = -I$, and so $P = [D, C]$ is a symplectic matrix. Therefore,

$$P^{-1} = \begin{bmatrix} -C^TJ \\ D^TJ \end{bmatrix};$$

change coordinates by $z = P\zeta$ so that

$$\dot{\zeta} = P^{-1}(AP - \dot{P})\zeta = \begin{bmatrix} C^TAD + C^TJ\dot{D} & 0 \\ -D^TAD - D^TJ\dot{D} & 0 \end{bmatrix} \zeta.$$

All the submatrices above are $n \times n$. The one in the upper left-hand corner is also zero, which can be seen by differentiating $C^T JD = -I$ to get $\dot{C}^T JD + C^T J\dot{D} = (AC)^T JD + C^T J\dot{D} = C^T SD + C^T J\dot{D} = 0$. Therefore,

$$\begin{aligned} \dot{u} &= 0, \\ \dot{v} &= -D^T(AD + J\dot{D})u, \end{aligned} \qquad \text{where } \zeta = \begin{bmatrix} u \\ v \end{bmatrix},$$

which has a general solution $u = u_0, v = v_0 - V(t)u_0$, where

$$V(t) = \int_{t_0}^{t} D^T(AD + J\dot{D})dt. \tag{2.14}$$

A symplectic fundamental matrix solution of (2.3) is

$$Z = [D - CV, C].$$

Thus the complete set of solutions is obtained by performing the integration or quadrature in the formula above.

This result is closely related to the general result given in a later chapter which says that when $k$ integrals in involution for a Hamiltonian system of $n$ degrees of freedom are known then the integrals can be used to reduce the number of degrees of freedom by $k$, and hence to a system of dimension $2n - 2k$.

Recall that a nonsingular matrix $T$ has two polar decompositions, $T = PO = O'P'$, where $P$ and $P'$ are positive definite matrices and $O$ and $O'$ are orthogonal matrices. These representations are unique. $P$ is the unique positive definite square root of $TT^T$; $P'$ is the unique positive definite square root of $T^TT$, $O = (TT^T)^{-1/2}T$; and $O' = T(T^TT)^{-1/2}$.

**Theorem 2.1.6.** *If $T$ is symplectic, then the $P, O, P', O'$ of the polar decomposition given above are symplectic also.*

*Proof.* The formula for $T^{-1}$ in (2.8) is an equivalent condition for $T$ to be symplectic. Let $T = PO$. Because $T^{-1} = -JT^TJ$, $O^{-1}P^{-1} = -JO^TP^TJ = (J^TO^TJ)(J^TP^TJ)$. In this last equation, the left-hand side is the product of an orthogonal matrix $O^{-1}$ and a positive definite matrix $P^{-1}$, similarly the right-hand side a product of an orthogonal matrix $J^{-1}OJ$ and a positive definite matrix $J^TPJ$. By the uniqueness of the polar representation, $O^{-1} = J^{-1}O^TJ = -JO^TJ$ and $P^{-1} = J^TPJ = -JP^TJ$. By (2.8) these last relations imply that $P$ and $O$ are symplectic. A similar argument shows that $P'$ and $O'$ are symplectic.

**Theorem 2.1.7.** *The determinant of a symplectic matrix is* $+1$.

*Proof.* Depending on how much linear algebra you know, this theorem is either easy or difficult. In Section 5.3 and Chapter 7 we give alternate proofs.

Let $T$ be symplectic. Formula (2.7) gives $\det(T^T JT) = \det T^T \det J \det T = (\det T)^2 = \det J = 1$ so $\det T = \pm 1$. The problem is to show that $\det T = +1$.

The determinant of a positive definite matrix is positive; so, by the polar decomposition theorem it is enough to show that an orthogonal symplectic matrix has a positive determinant. So let $T$ be orthogonal also.

Using the block representation in (2.9) for $T$, formula (2.10) for $T^{-1}$, and the fact that $T$ is orthogonal, $T^{-1} = T^T$, one has that $T$ is of the form

$$T = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}.$$

Define $P$ by

$$P = \frac{1}{\sqrt{2}} \begin{bmatrix} I & iI \\ I & -iI \end{bmatrix}, \quad P^{-1} = \frac{1}{\sqrt{2}} \begin{bmatrix} I & I \\ -iI & iI \end{bmatrix}.$$

Compute $PTP^{-1} = \mathrm{diag}((\mathrm{a}-\mathrm{bi}), (\mathrm{a}+\mathrm{bi}))$, so

$$\det T = \det PTP^{-1} = \det(a - bi) \det(a + bi) > 0.$$

## 2.2 Symplectic Linear Spaces

What is the matrix $J$? There are many different answers to this question depending on the context in which the question is asked. In this section we answer this question from the point of view of abstract linear algebra. We present other answers later on, but certainly not all.

Let $\mathbb{V}$ be an $m$-dimensional vector space over the field $\mathbb{F}$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. A bilinear form is a mapping $B : \mathbb{V} \times \mathbb{V} \to \mathbb{F}$ that is linear in both variables. A bilinear form is skew-symmetric or alternating if $B(u, v) = -B(v, u)$ for all $u, v \in \mathbb{V}$. A bilinear form $B$ is nondegenerate if $B(u, v) = 0$ for all $v \in \mathbb{V}$ implies $u = 0$. An example of an alternating bilinear form on $\mathbb{F}^m$ is $B(u, v) = u^T Sv$, where $S$ is any skew-symmetric matrix, e.g., $J$.

Let $B$ be a bilinear form and $e_1, \ldots, e_m$ a basis for $\mathbb{V}$. Given any vector $v \in \mathbb{V}$, we write $v = \Sigma \alpha_i e_i$ and define an isomorphism $\Phi : \mathbb{V} \to \mathbb{F}^m : v \to a = (\alpha_1, \ldots, \alpha_m)$.[1] Define $s_{ij} = B(e_i, e_j)$ and $S$ to be the $m \times m$ matrix $S = [s_{ij}]$, the matrix of $B$ in the basis $(e)$. Let $\Phi(u) = b = (\beta_1, \ldots, \beta_m)$; then $B(u, v) = \Sigma\Sigma \alpha_i \beta_j B(e_i, e_j) = b^T Sa$. So in the coordinates defined by the basis $(e_i)$, the bilinear form is just $b^T Sa$ where $S$ is the matrix $[B(e_i, e_j)]$. If $B$ is alternating, then $S$ is skew-symmetric, and if $B$ is nondegenerate, then $S$ is nonsingular and conversely.

If you change the basis by $e_i = \Sigma q_{ij} f_j$ and $Q$ is the matrix $Q = (q_{ij})$, then the bilinear form $B$ has the matrix $R$ in the basis $(f)$, where $S = QRQ^T$.

[1] Remember vectors are column vectors, but sometimes written as row vectors in the text.

One says that $R$ and $S$ are congruent (by $Q$). If $Q$ is any elementary matrix so that premultiplication of $R$ by $Q$ is an elementary row operation, then postmultiplication of $R$ by $Q^T$ is the corresponding column operation. Thus $S$ is obtained from $R$ by performing a sequence of row operations and the same sequence of column operations and conversely.

**Theorem 2.2.1.** *Let $S$ be any skew-symmetric matrix; then there exists a nonsingular matrix $Q$ such that*

$$R = QSQ^T = \mathrm{diag}(\mathrm{K}, \mathrm{K}, \dots, \mathrm{K}, 0, 0, \dots, 0),$$

*where*

$$K = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

*Or given an alternating form $B$ there is a basis for $\mathbb{V}$ such that the matrix of $B$ in this basis is $R$.*

*Proof.* If $S = 0$, we are finished. Otherwise, there is a nonzero entry that can be transferred to the first row by interchanging rows. Perform the corresponding column operations. Now bring the nonzero entry in the first row to the second column (the (1,2) position) by column operations and perform the corresponding row operations.

Scale the first row and the first column so that $+1$ is in the (1,2) and so that $-1$ is in the (2,1) position. Thus the matrix has the $2 \times 2$ matrix $K$ in the upper left-hand corner. Using row operations we can eliminate all the nonzero elements in the first two columns below the first two rows. Performing the corresponding column operation yields a matrix of the form $\mathrm{diag}(\mathrm{K}, \mathrm{S}')$, where $S'$ is an $(m-2) \times (m-2)$ skew-symmetric matrix. Repeat the above argument on $S'$.

Note that the rank of a skew-symmetric matrix is always even; thus, a nondegenerate, alternating bilinear form is defined on an even dimensional space.

A symplectic linear space, or just a symplectic space, is a pair, $(\mathbb{V}, \omega)$ where $\mathbb{V}$ is a $2n$-dimensional vector space over the field $\mathbb{F}$, $\mathbb{R}$ or $\mathbb{C}$, and $\omega$ is a nondegenerate alternating bilinear form on $\mathbb{V}$. The form $\omega$ is called the *symplectic form* or the symplectic inner product. Throughout the rest of this section we shall assume that $\mathbb{V}$ is a symplectic space with symplectic form $\omega$. The standard example is $\mathbb{F}^{2n}$ and $\omega(x, y) = x^T J y$. In this example we shall write $\{x, y\} = x^T J y$ and call the space $(\mathbb{F}^{2n}, J)$ or simply $\mathbb{F}^{2n}$, if no confusion can arise.

A symplectic basis for $\mathbb{V}$ is a basis $v_1, \dots, v_{2n}$ for $\mathbb{V}$ such that $\omega(v_i, v_j) = J_{ij}$, the $i, j^{th}$ entry of $J$. A symplectic basis is a basis so that the matrix of $\omega$ is just $J$. The standard basis $e_1, \dots, e_{2n}$, where $e_i$ is 1 in the $i^{th}$ position and zero elsewhere, is a symplectic basis for $(\mathbb{F}^{2n}, J)$. Given two symplectic spaces $(\mathbb{V}_i, \omega_i), i = 1, 2$, a symplectic isomorphism or an isomorphism is a linear isomorphism $L : \mathbb{V}_1 \to \mathbb{V}_2$ such that $\omega_2(L(x), L(y)) = \omega_1(x, y)$ for all $x, y \in \mathbb{V}_1$; that is, $L$ preserves the symplectic form. In this case we say that the two spaces are symplectically isomorphic or symplectomorphic.

**Corollary 2.2.1.** *Let* $(\mathbb{V}, \omega)$ *be a symplectic space of dimension* $2n$*. Then* $\mathbb{V}$ *has a symplectic basis.* $(\mathbb{V}, \omega)$ *is symplectically isomorphic to* $(\mathbb{F}^{2n}, J)$*, or all symplectic spaces of dimension* $2n$ *are isomorphic.*

*Proof.* By Theorem 2.2.1 there is a basis for $\mathbb{V}$ such that the matrix of $\omega$ is $\mathrm{diag}(\mathrm{K}, \dots, \mathrm{K})$. Interchanging rows $2i$ and $n + 2i - 1$ and the corresponding columns brings the matrix to $J$. The basis for which the matrix of $\omega$ is $J$ is a symplectic basis; so, a symplectic basis exists.

Let $v_1, \dots, v_{2n}$ be a symplectic basis for $\mathbb{V}$ and $u \in \mathbb{V}$. There exist constants $\alpha_i$ such that $u = \sum \alpha_i v_i$. The linear map $L : \mathbb{V} \to \mathbb{F}^{2n} : u \to (\alpha_1, \dots, \alpha_{2n})$ is the desired symplectic isomorphism.

The study of symplectic linear spaces is really the study of one canonical example, e.g., $(\mathbb{F}^{2n}, J)$. Or put another way, $J$ is just the coefficient matrix of the symplectic form in a symplectic basis. This is one answer to the question "What is $J$?".

If $\mathbb{V}$ is a vector space over $\mathbb{F}$, then $f$ is a linear functional if $f : \mathbb{V} \to \mathbb{F}$ is linear, $f(\alpha u + \beta v) = \alpha f(u) + \beta f(v)$ for all $u, v \in \mathbb{V}$, and $\alpha, \beta \in \mathbb{F}$. Linear functionals are sometimes called 1-forms or covectors. If $\mathbb{E}$ is the vector space of displacements of a particle in Euclidean space, then the work done by a force on a particle is a linear functional on $\mathbb{E}$. The usual geometric representation for a vector in $\mathbb{E}$ is a directed line segment. Represent a linear functional by showing its level planes. The value of the linear functional $f$ on a vector $v$ is represented by the number of level planes the vector crosses. The more level planes the vector crosses, the larger is the value of $f$ on v.

The set of all linear functionals on a space $\mathbb{V}$ is itself a vector space when addition and scalar multiplication are just the usual addition and scalar multiplication of functions. That is, if $f$ and $f'$ are linear functionals on $\mathbb{V}$ and $\alpha \in \mathbb{F}$, then define the linear functionals $f + f'$ and $\alpha f$ by the formulas $(f + f')(v) = f(v) + f'(v)$ and $(\alpha f)(v) = \alpha f(v)$. The space of all linear functionals is called the dual space (to $\mathbb{V}$) and is denoted by $\mathbb{V}^*$.

When $\mathbb{V}$ is finite dimensional so is $\mathbb{V}^*$ with the same dimension. Let $u_1, \dots, u_m$ be a basis for $\mathbb{V}$; then for any $v \in \mathbb{V}$, there are scalars $f^1, \dots, f^m$ such that $v = f^1 u_1 + \dots + f^m u_m$. The $f^i$ are functions of $v$ so we write $f^i(v)$, and they are linear. It is not too hard to show that $f^1, \dots, f^m$ forms a basis for $\mathbb{V}^*$; this basis is called the dual basis (dual to $u_1, \dots, u_m$). The defining property of this basis is $f^i(u_j) = \delta^i_j$ (the Kronecker delta function, defined by $\delta^i_j = 1$ if $i = j$ and zero otherwise).

If $\mathbb{W}$ is a subspace of $\mathbb{V}$ of dimension $r$, then define $\mathbb{W}^0 = \{f \in \mathbb{V}^* : f(e) = 0 \text{ for all } e \in \mathbb{W}\}$. $\mathbb{W}^0$ is called the annihilator of $\mathbb{W}$ and is easily shown to be a subspace of $\mathbb{V}^*$ of dimension $m - r$. Likewise, if $\mathbb{W}$ is a subspace of $\mathbb{V}^*$ of dimension $r$ then $\mathbb{W}^0 = \{e \in \mathbb{V} : f(e) = 0 \text{ for all } f \in \mathbb{W}^*\}$ is a subspace of $\mathbb{V}$

of dimension $m - r$. Also $\mathbb{W}^{00} = \mathbb{W}$. See any book on vector space theory for a complete discussion of dual spaces with proofs.

Because $\omega$ is a bilinear form, for each fixed $v \in \mathbb{V}$ the function $\omega(v, \cdot) : \mathbb{V} \to \mathbb{R}$ is a linear functional and so is in the dual space $\mathbb{V}^*$. Because $\omega$ is nondegenerate, the map $\flat : \mathbb{V} \to \mathbb{V}^* : v \to \omega(v, \cdot) = v^\flat$ is an isomorphism. Let $\# : \mathbb{V}^* \to \mathbb{V} : v \to v^\#$ be the inverse of $\flat$. Sharp, $\#$, and flat, $\flat$, are musical symbols for raising and lowering notes and are used here because these isomorphisms are index raising and lowering operations in the classical tensor notation.

Let $\mathbb{U}$ be a subspace of $\mathbb{V}$. Define $\mathbb{U}^\perp = \{v \in \mathbb{V} : \omega(v, \mathbb{U}) = 0\}$. Clearly $\mathbb{U}^\perp$ is a subspace, $\{\mathbb{U}, \mathbb{U}^\perp\} = 0$ and $\mathbb{U} = \mathbb{U}^{\perp\perp}$.

**Lemma 2.2.1.** *$\mathbb{U}^\perp = \mathbb{U}^{0\#}$. $\dim \mathbb{U} + \dim \mathbb{U}^\perp = \dim \mathbb{V} = 2n$.*

*Proof.*

$$\begin{aligned}
\mathbb{U}^\perp &= \{x \in \mathbb{V} : \omega(x, y) = 0 \text{ for all } y \in \mathbb{U}\} \\
&= \{x \in \mathbb{V} : x^\flat(y) = 0 \text{ for all } y \in \mathbb{U}\} \\
&= \{x \in \mathbb{V} : x^\flat \in \mathbb{U}^0\} \\
&= \mathbb{U}^{0\#}.
\end{aligned}$$

The second statement follows from $\dim \mathbb{U} + \dim \mathbb{U}^0 = \dim \mathbb{V}$ and the fact that $\#$ is an isomorphism.

A symplectic subspace $\mathbb{U}$ of $\mathbb{V}$ is a subspace such that $\omega$ restricted to this subspace is nondegenerate. By necessity $\mathbb{U}$ must be of even dimension, and so, $(\mathbb{U}, \omega)$ is a symplectic space.

**Proposition 2.2.1.** *If $\mathbb{U}$ is symplectic, then so is $\mathbb{U}^\perp$, and $\mathbb{V} = \mathbb{U} \oplus \mathbb{U}^\perp$. Conversely, if $\mathbb{V} = \mathbb{U} \oplus \mathbb{W}$ and $\omega(\mathbb{U}, \mathbb{W}) = 0$, then $\mathbb{U}$ and $\mathbb{W}$ are symplectic.*

*Proof.* Let $x \in \mathbb{U} \cap \mathbb{U}^\perp$; so, $\omega(x, y) = 0$ for all $y \in \mathbb{U}$, but $\mathbb{U}$ is symplectic so $x = 0$. Thus $\mathbb{U} \cap \mathbb{U}^\perp = 0$. This, with Lemma 2.2.1, implies $\mathbb{V} = \mathbb{U} \oplus \mathbb{U}^\perp$.

Now let $\mathbb{V} = \mathbb{U} \oplus \mathbb{W}$ and $\omega(\mathbb{U}, \mathbb{W}) = 0$. If $\omega$ is degenerate on $\mathbb{U}$, then there is an $x \in \mathbb{U}$, $x \neq 0$, with $\omega(x, \mathbb{U}) = 0$. Because $\mathbb{V} = \mathbb{U} \oplus \mathbb{W}$ and $\omega(\mathbb{U}, \mathbb{W}) = 0$, this implies $\omega(x, \mathbb{V}) = 0$ or that $\omega$ is degenerate on all of $\mathbb{V}$. This contradiction yields the second statement.

The next lemma is a technical result used much later.

**Lemma 2.2.2.** *Let $\mathbb{U}$ be a subspace of $\mathbb{V}^*$ and $\mathbb{Z} = \mathbb{U}^0/(\mathbb{U}^\# \cap \mathbb{U}^0)$. For $[x], [y] \in \mathbb{Z}$ the bilinear form given by $\{[x], [y]\} = \{x, y\}$ is a well-defined symplectic inner product on $\mathbb{Z}$.*

*Proof.* If $v \in \mathbb{U}^{\#}$ and $u \in \mathbb{U}^0$, then $\{v, u\} = 0$ by definition. Thus if $x, y \in \mathbb{U}^0$ and $\xi, \psi \in \mathbb{U}^{\#} \cap \mathbb{U}^0$ one has

$$\{[x + \xi], [y + \psi]\} = \{x + \xi, y + \psi\} = \{x, y\} = \{[x], [y]\},$$

so the form is well defined.

Now assume that $\{[x], [y]\} = 0$ for all $[y] \in \mathbb{Z}$. Then $\{x, y\} = 0$ for all $y \in \mathbb{U}^0$, or $\{x, \cdot\} \in \mathbb{U}$ and thus $x \in \mathbb{U}^{\#}$ or $[x] = 0$. Thus the form is nondegenerate on $\mathbb{Z}$.

A Lagrangian space $\mathbb{U}$ is a subspace of $\mathbb{V}$ of dimension $n$ such that $\omega$ is zero on $\mathbb{U}$, i.e., $\omega(u, w) = 0$ for all $u, w \in \mathbb{U}$. A direct sum decomposition $\mathbb{V} = \mathbb{U} \oplus \mathbb{W}$ where $\mathbb{U}$, and $\mathbb{W}$ are Lagrangian spaces, is called a Lagrangian splitting, and $\mathbb{W}$ is called the Lagrangian complement of $\mathbb{U}$. In $\mathbb{R}^2$ any line through the origin is Lagrangian, and any other line through the origin is a Lagrangian complement.

**Lemma 2.2.3.** *Let $\mathbb{U}$ be a Lagrangian subspace of $\mathbb{V}$, then there exists a Lagrangian complement of $\mathbb{U}$.*

*Proof.* The example above shows the complement is not unique. Let $\mathbb{V} = \mathbb{F}^{2n}$ and $\mathbb{U} \subset \mathbb{F}^{2n}$. Then $\mathbb{W} = J\mathbb{U}$ is a Lagrangian complement to $\mathbb{U}$. If $x, y \in \mathbb{W}$ then $x = Ju, y = Jv$ where $u, v \in \mathbb{U}$, or $\{u, v\} = 0$. But $\{x, y\} = \{Ju, Jv\} = \{u, v\} = 0$, so $\mathbb{W}$ is Lagrangian. If $x \in \mathbb{U} \cap J\mathbb{U}$ then $x = Jy$ with $y \in \mathbb{U}$. So $x, Jx \in \mathbb{U}$ and so $\{x, Jx\} = -\|x\|^2 = 0$ or $x = 0$. Thus $\mathbb{U} \cap \mathbb{W} = \phi$.

**Lemma 2.2.4.** *Let $\mathbb{V} = \mathbb{U} \oplus \mathbb{W}$ be a Lagrange splitting and $x_1, \ldots, x_n$ any basis for $\mathbb{U}$. Then there exists a unique basis $y_1, \ldots, y_n$ of $\mathbb{W}$ such that $x_1, \ldots, x_n, y_1, \ldots, y_n$ is a symplectic basis for $\mathbb{V}$.*

*Proof.* Define $\phi_i \in \mathbb{W}^0$ by $\phi_i(w) = \omega(x_i, w)$ for $w \in \mathbb{W}$. If $\sum \alpha_i \phi_i = 0$, then $\omega(\sum \alpha_i x_i, w) = 0$ for all $w \in \mathbb{W}$ or $\omega(\sum \alpha_i x_i, \mathbb{W}) = 0$. But because $\mathbb{V} = \mathbb{U} \oplus \mathbb{W}$ and $\omega(\mathbb{U}, \mathbb{U}) = 0$, it follows that $\omega(\sum \alpha_i x_i, \mathbb{V}) = 0$. This implies $\sum \alpha_i x_i = 0$, because $\omega$ is nondegenerate, and this implies $\alpha_i = 0$, because the $x_i s$ are independent. Thus $\phi_1, \ldots, \phi_n$ are independent, and so, they form a basis for $\mathbb{W}^0$. Let $y_1, \ldots, y_n$ be the dual basis in $\mathbb{W}$; so, $\omega(x_i, y_j) = \phi_i(y_j) = \delta_{ij}$.

A linear operator $L : \mathbb{V} \to \mathbb{V}$ is called Hamiltonian, if

$$\omega(Lx, y) + \omega(x, Ly) = 0 \tag{2.15}$$

for all $x, y \in \mathbb{V}$. A linear operator $L : \mathbb{V} \to \mathbb{V}$ is called symplectic, if

$$\omega(Lx, Ly) = \omega(x, y) \tag{2.16}$$

for all $x, y \in \mathbb{V}$. If $\mathbb{V}$ is the standard symplectic space $(\mathbb{F}^{2n}, J)$ and $L$ is a matrix, then (2.15) means $x^T(L^T J + JL)y = 0$ for all $x$ and y. But this implies that $L$ is a Hamiltonian matrix. On the other hand, if $L$ satisfies (2.16) then

$x^T L^T J L y = x^T J y$ for all $x$ and $y$. But this implies $L$ is a symplectic matrix. The matrix representation of a Hamiltonian (respectively, symplectic) linear operator in a symplectic coordinate system is a Hamiltonian (respectively, symplectic) matrix.

**Lemma 2.2.5.** *Let* $\mathbb{V} = \mathbb{U} \oplus \mathbb{W}$ *be a Lagrangian splitting and* $A : \mathbb{V} \to \mathbb{V}$ *a Hamiltonian (respectively, symplectic) linear operator that respects the splitting; i.e.,* $A : \mathbb{U} \to \mathbb{U}$ *and* $A : \mathbb{W} \to \mathbb{W}$. *Choose any basis of the form given in Lemma* 2.2.4*; the matrix representation of* $A$ *in these symplectic coordinates is of the form*

$$\begin{bmatrix} B^T & 0 \\ 0 & -B \end{bmatrix} \quad \left( \textit{respectively, } \begin{bmatrix} B^T & 0 \\ 0 & B^{-1} \end{bmatrix} \right). \tag{2.17}$$

*Proof.* $A$ respects the splitting and the basis for $\mathbb{V}$ is the union of the bases for $\mathbb{U}$ and $\mathbb{W}$, therefore the matrix representation for A must be in block-diagonal form. A Hamiltonian or symplectic matrix which is in block-diagonal form must be of the form given in (2.17).

## 2.3 Canonical Forms

In this section we obtain some canonical forms for Hamiltonian and symplectic matrices in some simple cases. The complete picture is very detailed and special, so see Sections 5.3 and 5.4 if more extensive material is needed. We start with only real matrices, but sometimes we need to go into the complex domain to finish the arguments. We simply assume that all our real spaces are embedded in a complex space of the same dimension.

If $A$ is Hamiltonian and $T$ is symplectic, then $T^{-1}AT$ is Hamiltonian also. Thus if we start with a linear constant coefficient Hamiltonian system $\dot{z} = Az$ and make the change of variables $z = Tu$, then in the new coordinates the equations become $\dot{u} = (T^{-1}AT)u$, which is again Hamiltonian. If $B = T^{-1}AT$, where $T$ is symplectic, then we say that $A$ and $B$ are symplectically similar. This is an equivalence relation. We seek canonical forms for Hamiltonian and symplectic matrices under symplectic similarity. In as much as it is a form of similarity transformation, the eigenvalue structure plays an important role in the following discussion.

Because symplectic similarity is more restrictive than ordinary similarity, one should expect more canonical forms than the usual Jordan canonical forms. Consider, for example, the two Hamiltonian matrices

$$A_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{2.18}$$

both of which could be the coefficient matrix of a harmonic oscillator. In fact, they are both the real Jordan forms for the harmonic oscillator.

The reflection $T = \operatorname{diag}(1, -1)$ defines a similarity between these two; i.e., $T^{-1}A_1T = A_2$. The determinant of $T$ is not $+1$, therefore $T$ is not symplectic. In fact, $A_1$ and $A_2$ are not symplectically equivalent. If $T^{-1}A_1T = A_2$, then $T^{-1}\exp(A_1t)T = \exp(A_2t)$, and $T$ would take the clockwise rotation $\exp(A_1t)$ to the counterclockwise rotation $\exp(A_2t)$. But, if $T$ were symplectic, its determinant would be $+1$ and thus would be orientation preserving. Therefore, $T$ cannot be symplectic.

Another way to see that the two Hamiltonian matrices in (2.18) are not symplectically equivalent is to note that $A_1 = JI$ and $A_2 = J(-I)$. So the symmetric matrix corresponding to $A_1$ is I, the identity, and to $A_2$ is $-I$. I is positive definite, whereas $-I$ is negative definite. If $A_1$ and $A_2$ where symplectically equivalent, then I and $-I$ would be congruent, which is clearly false.

A polynomial $p(\lambda) = a_m\lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_0$ is even if $p(-\lambda) = p(\lambda)$, which is the same as $a_k = 0$ for all odd k. If $\lambda_0$ is a zero of an even polynomial, then so is $-\lambda_0$; therefore, the zeros of a real even polynomial are symmetric about the real and imaginary axes. The polynomial $p(\lambda)$ is a reciprocal polynomial if $p(\lambda) = \lambda^m p(\lambda^{-1})$, which is the same as $a_k = a_{m-k}$ for all k. If $\lambda_0$ is a zero of a reciprocal polynomial, then so is $\lambda_0^{-1}$; therefore, the zeros of a real reciprocal polynomial are symmetric about the real axis and the unit circle (in the sense of inversion).

**Proposition 2.3.1.** *The characteristic polynomial of a real Hamiltonian matrix is an even polynomial. Thus if $\lambda$ is an eigenvalue of a Hamiltonian matrix, then so are $-\lambda$, $\overline{\lambda}$, $-\overline{\lambda}$.*

*The characteristic polynomial of a real symplectic matrix is a reciprocal polynomial. Thus if $\lambda$ is an eigenvalue of a real symplectic matrix, then so are $\lambda^{-1}$, $\overline{\lambda}$, $\overline{\lambda}^{-1}$*

*Proof.* Recall that $\det J = 1$. Let A be a Hamiltonian matrix; then $p(\lambda) = \det(A - \lambda I) = \det(JA^TJ - \lambda I) = \det(JA^TJ + \lambda JJ) = \det J \det(A + \lambda I)\det J = \det(A + \lambda I) = p(-\lambda)$.

Let $T$ be a symplectic matrix; by Theorem 2.1.7 $\det T = +1$. $p(\lambda) = \det(T - \lambda I) = \det(T^T - \lambda I) = \det(-JT^{-1}J - \lambda I) = \det(-JT^{-1}J + \lambda JJ) = \det(-T^{-1} + \lambda I) = \det T^{-1}\det(-I + \lambda T) = \lambda^{2n}\det(-\lambda^{-1}I + T) = \lambda^{2n}p(\lambda^{-1})$.

Actually we can prove much more. By (2.6), Hamiltonian matrix $A$ satisfies $A = J^{-1}(-A^T)J$; so, $A$ and $-A^T$ are similar, and the multiplicity of the eigenvalues $\lambda_0$ and $-\lambda_0$ are the same. In fact, the whole Jordan block structure will be the same for $\lambda_0$ and $-\lambda_0$.

By (2.8), symplectic matrix $T$ satisfies $T^{-1} = J^{-1}T^TJ$; so, $T^{-1}$ and $T^T$ are similar, and the multiplicity of the eigenvalues $\lambda_0$ and $\lambda_0^{-1}$ are the same. The whole Jordan block structure will be the same for $\lambda_0$ and $\lambda_0^{-1}$.

Consider the linear constant coefficient Hamiltonian system of differential equations

$$\dot{x} = Ax, \tag{2.19}$$

where $A$ is a Hamiltonian matrix and $Z(t) = e^{At}$ is the fundamental matrix solution. By the above it is impossible for all the eigenvalues of A to be in the left half-plane, and, therefore, it is impossible for all the solutions to be exponentially decaying. Thus the origin cannot be asymptotically stable.

Henceforth, let $A$ be a real Hamiltonian matrix and $T$ a real symplectic matrix. First we develop the theory for Hamiltonian matrices and then the theory of symplectic matrices. Because eigenvalues are sometimes complex, it is necessary to consider complex matrices at times, but we are always be concerned with the real answers in the end.

First consider the Hamiltonian case. Let $\lambda$ be an eigenvalue of $A$, and define subspaces of $\mathbb{C}^{2n}$ by $\eta_k(\lambda) = \text{kernel}\,(A - \lambda I)^k$, $\eta^\dagger(\lambda) = \cup_1^{2n} \eta_k(\lambda)$. The eigenspace of $A$ corresponding to the eigenvalue $\lambda$ is $\eta(\lambda) = \eta_1(\lambda)$, and the generalized eigenspace is $\eta^\dagger(\lambda)$. If $\{x, y\} = x^T J y = 0$, then $x$ and $y$ are $J$-orthogonal.

**Lemma 2.3.1.** *Let $\lambda$ and $\mu$ be eigenvalues of $A$ with $\lambda + \mu \neq 0$, then $\{\eta(\lambda),\ \eta(\mu)\} = 0$. That is, the eigenvectors corresponding to $\lambda$ and $\mu$ are $J$-orthogonal.*

*Proof.* Let $Ax = \lambda x$, and $Ay = \mu y$, where $x, y \neq 0$. $\lambda\{x, y\} = \{Ax, y\} = x^T A^T J y = -x^T J A y = -\{x, Ay\} = -\mu\{x, y\}$; and so, $(\lambda + \mu)\{x, y\} = 0$.

**Corollary 2.3.1.** *Let $A$ be a $2n \times 2n$ Hamiltonian matrix with distinct eigenvalues $\lambda_1, \ldots, \lambda_n, -\lambda_1, \ldots, -\lambda_n$; then there exists a symplectic matrix $S$ (possibly complex) such that $S^{-1} A S = \mathrm{diag}(\lambda_1, \ldots, \lambda_\mathrm{n}, -\lambda_1, \ldots, -\lambda_\mathrm{n})$.*

*Proof.* $\mathbb{V} = \mathbb{U} \oplus \mathbb{W}$ is a Lagrange splitting where

$$\mathbb{U} = \eta_1(\lambda_1) \oplus \cdots \oplus \eta_1(\lambda_n), \quad \mathbb{W} = \eta_1(-\lambda_1) \oplus \cdots \oplus \eta_1(-\lambda_n),$$

and $A$ respects this splitting. Choose a symplectic basis for $\mathbb{V}$ by Lemma 2.2.4. Changing to that basis is effected by a symplectic matrix $G$; i.e., $G^{-1} A G = \mathrm{diag}(\mathrm{B}^\mathrm{T}, -\mathrm{B})$, where $B$ has eigenvalues $\lambda_1, \ldots, \lambda_n$. Let $C$ be such that $C^{-T} B^T C^T = \mathrm{diag}(\lambda_1, \ldots, \lambda_\mathrm{n})$ and define a symplectic matrix by $Q = \mathrm{diag}(\mathrm{C}^\mathrm{T}, \mathrm{C}^{-1})$. The required symplectic matrix is $S = GQ$.

If complex transformations are allowed, then the two matrices in (2.18) can both be brought to diag(i, –i) by a symplectic similarity, and thus one is symplectically similar to the other. However, as we have seen they are not similar by a real symplectic similarity. Let us investigate the real case in detail.

A subspace $\mathbb{U}$ of $\mathbb{C}^n$ is called a *complexification* (of a real subspace) if $\mathbb{U}$ has a real basis. If $\mathbb{U}$ is a complexification, then there is a real basis $x_1, \ldots, x_k$ for $\mathbb{U}$, and for any $u \in \mathbb{U}$, there are complex numbers $\alpha_1, \ldots, \alpha_k$ such that $u = \alpha_1 x_1 + \cdots + \alpha_n x_n$. But then $\overline{u} = \overline{\alpha}_1 x_1 + \cdots + \overline{\alpha}_n x_n \in \mathbb{U}$ also.

Conversely, if $\mathbb{U}$ is a subspace such that $u \in \mathbb{U}$ implies $\overline{u} \in \mathbb{U}$, then $\mathbb{U}$ is a complexification. Because if $x_1, \ldots, x_k$ is a complex basis with $x_j = u_j + v_j i$, then $u_j = (x_j + \overline{x}_j)/2$ and $v_j = (x_j - \overline{x}_j)/2i$ are in $\mathbb{U}$, and the totality of

$u_1, \ldots, u_k, v_1, \ldots, v_k$ span $\mathbb{U}$. From this real spanning set, one can extract a real basis. Thus $\mathbb{U}$ is a complexification if and only if $\mathbb{U} = \overline{\mathbb{U}}$ (i.e., $u \in \mathbb{U}$ implies $\overline{u} \in \mathbb{U}$).

Until otherwise said let $A$ be a real Hamiltonian matrix with distinct eigenvalues $\lambda_1, \ldots, \lambda_n, -\lambda_1, \ldots, -\lambda_n$ so 0 is not an eigenvalue. The eigenvalues of A fall into three groups: (1) the real eigenvalues $\pm\alpha_1, \ldots, \pm\alpha_s$, (2) the pure imaginary $\pm\beta_1 i, \ldots, \pm\beta_r i$, and (3) the truly complex $\pm\gamma_1 \pm \delta_1 i, \ldots, \pm\gamma_t \pm \delta_t i$. This defines a direct sum decomposition

$$\mathbb{V} = (\oplus_j \mathbb{U}_j) \oplus (\oplus_j \mathbb{W}_j) \oplus (\oplus_j \mathbb{Z}_j), \tag{2.20}$$

where

$$\mathbb{U}_j = \eta(\alpha_j) \oplus \eta(-\alpha_j)$$

$$\mathbb{W}_j = \eta(\beta_j i) \oplus \eta(-\beta_j i)$$

$$\mathbb{Z}_j = \{\eta(\gamma_j + \delta_j i) \oplus \eta(\gamma_j - \delta_j i)\} \oplus \{\eta(-\gamma_j - \delta_j i) \oplus \eta(-\gamma_j + \delta_j i)\}.$$

Each of the summands in the above is an invariant subspace for $A$. By Lemma 2.3.1, each space is $J$-orthogonal to every other, and so by Proposition 2.2.1 each space must be a symplectic subspace. Because each subspace is invariant under complex conjugation, each is the complexification of a real space. Thus we can choose symplectic coordinates for each of the spaces, and A in these coordinates would be block diagonal. Therefore, the next task is to consider each space separately.

**Lemma 2.3.2.** *Let $A$ be a $2 \times 2$ Hamiltonian matrix with eigenvalues $\pm\alpha$, $\alpha$ real and positive. Then there exists a real $2 \times 2$ symplectic matrix $S$ such that*

$$S^{-1}AS = \begin{bmatrix} \alpha & 0 \\ 0 & -\alpha \end{bmatrix}. \tag{2.21}$$

*The Hamiltonian is $H = \alpha\xi\eta$.*

*Proof.* Let $Ax = \alpha x$, and $Ay = -\alpha y$, where $x$ and $y$ are nonzero. Because $x$ and $y$ are eigenvectors corresponding to different eigenvalues, they are independent. Thus $\{x, y\} \neq 0$. Let $u = \{x, y\}^{-1} y$: so, $x, u$ is a real symplectic basis, $S = [x, u]$ is a real symplectic matrix, and $S$ is the matrix of the lemma.

**Lemma 2.3.3.** *Let $A$ be a real $2 \times 2$ Hamiltonian matrix with eigenvalues $\pm\beta i$, $\beta > 0$. Then there exists a real $2 \times 2$ symplectic matrix $S$ such that*

$$S^{-1}AS = \begin{bmatrix} 0 & \beta \\ -\beta & 0 \end{bmatrix}, \quad \textit{or} \quad S^{-1}AS = \begin{bmatrix} 0 & -\beta \\ \beta & 0 \end{bmatrix}. \tag{2.22}$$

*The Hamiltonian is $H = \frac{\beta}{2}(\xi^2 + \eta^2)$ or $H = -\frac{\beta}{2}(\xi^2 + \eta^2)$.*

*Proof.* Let $Ax = i\beta x$, and $x = u + vi \neq 0$. So $Au = -\beta v$ and $Av = \beta u$. Because $u + iv$ and $u - iv$ are independent, $u$ and $v$ are independent. Thus $\{u, v\} = \delta \neq 0$. If $\delta = \gamma^2 > 0$, then define $S = [\gamma^{-1}u, \gamma^{-1}v]$ to get the first option in (2.22), or if $\delta = -\gamma^2 < 0$, then define $S = [\gamma^{-1}v, \gamma^{-1}u]$ to get the second option.

Sometimes it is more advantageous to have a diagonal matrix than to have a real one; yet you want to keep track of the real origin of the problem. This is usually accomplished by reality conditions as defined in the next lemma.

**Lemma 2.3.4.** *Let $A$ be a real $2 \times 2$ Hamiltonian matrix with eigenvalues $\pm\beta i$, $\beta \neq 0$. Then there exist a $2 \times 2$ matrix $S$ and a matrix $R$ such that*

$$S^{-1}AS = \begin{bmatrix} i\beta & 0 \\ 0 & -i\beta \end{bmatrix}, \quad R = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad S^T JS = \pm 2iJ, \quad \overline{S} = SR. \tag{2.23}$$

*Proof.* Let $Ax = i\beta x$, where $x \neq 0$. Let $x = u + iv$ as in the above lemma. Compute $\{x, \overline{x}\} = 2i\{v, u\} \neq 0$. Let $\gamma = 1/\sqrt{|\ \{v, u\}\ |}$ and $S = [\gamma x, \gamma \overline{x}]$.

If $S$ satisfies (2.23), then $S$ is said to satisfy reality conditions with respect to $R$. The matrix $S$ is no longer a symplectic matrix but is what is called a symplectic matrix with multiplier $\pm 2i$. We discuss these types of matrices later. The matrix $R$ is used to keep track of the fact that the columns of $S$ are complex conjugates. We could require $S^T JS = +2iJ$ by allowing an interchange of the signs in (2.23).

**Lemma 2.3.5.** *Let $A$ be a $2 \times 2$ Hamiltonian matrix with eigenvalue $0$. Then there exists a real $2 \times 2$ symplectic matrix $S$ such that*

$$S^{-1}AS = \begin{bmatrix} 0 & \delta \\ 0 & 0 \end{bmatrix} \tag{2.24}$$

*where $\delta = +1$, $-1$, or $0$.*

*The Hamiltonian is $H = \delta\eta^2$.*

**Lemma 2.3.6.** *Let $A$ be a $4 \times 4$ Hamiltonian matrix with eigenvalue $\pm\gamma \pm \delta i$, $\gamma \neq 0$, $\delta \neq 0$. Then there exists a real $4 \times 4$ symplectic matrix $S$ such that*

$$S^{-1}AS = \begin{bmatrix} B^T & 0 \\ 0 & -B \end{bmatrix},$$

*where $B$ is a real $2 \times 2$ matrix with eigenvalues $+\gamma \pm \delta i$.*

*Proof.* $\mathbb{U} = \eta(\gamma_j + \delta_j i) \oplus \eta(\gamma_j - \delta_j i)$ is the complexification of a real subspace and by Lemma 2.3.1 is Lagrangian. $A$ restricted to this subspace has eigenvalues $+\gamma \pm \delta i$. A complement to $\mathbb{U}$ is $\mathbb{W} = \eta(-\gamma_j + \delta_j i) \oplus \eta(-\gamma_j - \delta_j i)$. Choose any real basis for $\mathbb{U}$ and complete it by Lemma 2.2.5. The result follows from Lemma 2.2.5.

In particular you can choose coordinates so that $B$ is in real Jordan form; so,

$$B = \begin{bmatrix} \gamma & \delta \\ -\delta & \gamma \end{bmatrix}.$$

There are many cases when $n > 1$ and $A$ has eigenvalues with zero real part; i.e., zero or pure imaginary. These cases are discussed in detail in Section 5.4. In the case where the eigenvalue zero is of multiplicity 4 the canonical forms are

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \pm 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}. \tag{2.25}$$

The corresponding Hamiltonians are

$$\xi_2\eta_1, \qquad \xi_2\eta_1 \pm \eta_2^2/2.$$

In the case of a double eigenvalue $\pm\alpha i$, $\alpha \neq 0$, the canonical forms in the $4 \times 4$ case are

$$\begin{bmatrix} 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \pm\alpha \\ -\alpha & 0 & 0 & 0 \\ 0 & \mp\alpha & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & \alpha & 0 & 0 \\ -\alpha & 0 & 0 & 0 \\ \pm 1 & 0 & 0 & \alpha \\ 0 & \pm 1 & -\alpha & 0 \end{bmatrix}. \tag{2.26}$$

The corresponding Hamiltonians are

$$(\alpha/2)(\xi_1^2 + \eta_1^2) \pm (\alpha/2)(\xi_2^2 + \eta_2^2), \qquad \alpha(\xi_2\eta_1 - \xi_1\eta_2) \mp (\xi_1^2 + \xi_2^2)/2.$$

Next consider the symplectic case. Let $\lambda$ be an eigenvalue of $T$, and define subspaces of $\mathbb{C}^{2n}$ by $\eta_k(\lambda) = \text{kernel}\,(T - \lambda I)^k$, $\eta^\dagger(\lambda) = \cup_1^{2n}\eta_k(\lambda)$. The eigenspace of $T$ corresponding to the eigenvalue $\lambda$ is $\eta(\lambda) = \eta_1(\lambda)$, and the generalized eigenspace is $\eta^\dagger(\lambda)$. Because the proof of the next set of lemmas is similar to those given just before, the proofs are left as problems.

**Lemma 2.3.7.** *If $\lambda$ and $\mu$ are eigenvalues of the symplectic matrix $T$ such that $\lambda\mu \neq 1$; then $\{\eta(\lambda), \eta(\mu)\} = 0$. That is, the eigenvectors corresponding to $\lambda$ and $\mu$ are $J$-orthogonal.*

**Corollary 2.3.2.** *Let $T$ be a $2n \times 2n$ symplectic matrix with distinct eigenvalues $\lambda_1, \dots, \lambda_n, \lambda_1^{-1}, \dots, \lambda_n^{-1}$; then there exists a symplectic matrix $S$ (possibly complex) such that*

$$S^{-1}TS = \operatorname{diag}(\lambda_1, \dots, \lambda_{\rm n}, \lambda_1^{-1}, \dots, \lambda_{\rm n}^{-1}).$$

If complex transformations are allowed, then the two matrices

$$\begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}, \quad \alpha^2 + \beta^2 = 1,$$

can both be brought to diag$(\alpha+\beta\mathrm{i}, \alpha-\beta\mathrm{i})$ by a symplectic similarity, and thus, one is symplectically similar to the other. However, they are not similar by a real symplectic similarity. Let us investigate the real case in detail.

Until otherwise said, let $T$ be a real symplectic matrix with distinct eigenvalues $\lambda_1,\dots,\lambda_n,\lambda_1^{-1},\dots,\lambda_n^{-1}$, so 1 is not an eigenvalue. The eigenvalues of $T$ fall into three groups: (1) the real eigenvalues, $\mu_1^{\pm 1},\dots,\mu_s^{\pm 1}$, (2) the eigenvalues of unit modulus, $\alpha\pm\beta_1 i,\dots,\alpha_r\pm\beta_r i$, and (3) the complex eigenvalues of modulus different from one, $(\gamma_1\pm\delta_1 i)^{\pm 1},\dots,(\gamma_t\pm\delta_t i)^{\pm 1}$. This defines a direct sum decomposition

$$\mathbb{V} = (\oplus_j \mathbb{U}_j) \oplus (\oplus_j \mathbb{W}_j) \oplus (\oplus_j \mathbb{Z}_j)\,, \tag{2.27}$$

where

$$\begin{aligned}
\mathbb{U}_j &= \eta(\mu_j)\oplus\eta(\mu_j^{-1})\\
\mathbb{W}_j &= \eta(\alpha_j+\beta_j i)\oplus\eta(\alpha_j-\beta_J i)\\
\mathbb{Z}_j &= \{\eta(\gamma_j+\delta_j i)\oplus\eta(\gamma_j-\delta_j i)\}\oplus\{\eta(\gamma_j+\delta_j i)^{-1}\oplus\eta(\gamma_j-\delta_j i)^{-1}\}.
\end{aligned}$$

Each of the summands in (2.27) is invariant for $T$. By Lemma 2.3.7 each space is $J$-orthogonal to every other, and so each space must be a symplectic subspace. Because each subspace is invariant under complex conjugation, each is the complexification of a real space. Thus we can choose symplectic coordinates for each of the spaces, and $T$ in these coordinates would be block diagonal. The $2\times 2$ case will be considered in detail in the next section and the general case is postponed to Chapter 5.

## 2.4 $Sp(2,\mathbb{R})$

It is time to stop and look back at the simplest case in order to foreshadow complexity that is still to come. A $2\times 2$ matrix $T$ is symplectic when $\det T = +1$ which also means that $T$ is in the special linear group $Sl(2,\mathbb{R})$. So in this lowest dimension $Sp(2,\mathbb{R}) = Sl(2,\mathbb{R})$ and these groups have some interesting properties, so much so that in 1985 the famed algebraist Serge Lang wrote a book whose sole title was $\mathbf{SL}_2(\mathbf{R})$.[2]

One of the topics of this chapter is devoted to the question of when two Hamiltonian or two symplectic matrices $T$ and $S$ are symplectic similar, i.e., when does there exist a symplectic matrix $P$ such that $S = P^{-1}TP$. The Hamiltonian case has been discussed, so let us look at this question in more detail for $2\times 2$ symplectic matrices.

Consider the action $\Phi$ of $Sp(2,\mathbb{R})$ on itself defined by

$$\Phi : Sp(2,\mathbb{R})\times Sp(2,\mathbb{R})\to Sp(2,\mathbb{R}) : (P,T)\mapsto P^{-1}TP,$$

[2] Lang's $\mathbf{SL}_2(\mathbf{R})$ is the same as our $Sl(2,\mathbb{R})$.

and the orbit of $T$ is

$$\mathbb{O}(T) = \{S \in Sp(2, \mathbb{R}) : S = P^{-1}TP \text{ for some } P \in Sp(2, \mathbb{R})\}.$$

Our task is to assign to each orbit a unique angle $\Theta(T)$ defined modulo $2\pi$.

Let us look at the canonical forms for each case as we did for Hamiltonian matrices in the last section. In the $2 \times 2$ case the characteristic polynomial is of the form $\lambda^2 - \operatorname{tr}(T)\lambda + 1$ so the trace determines the eigenvalues. The hyperbolic case is when $|\operatorname{tr}(T)| > 2$ and the eigenvalues are real and not equal to $\pm 1$, the elliptic case is when $|\operatorname{tr}(T)| < 2$ and the eigenvalues are complex conjugates, and the parabolic case is when $|\operatorname{tr}(T)| = 2$.

**Lemma 2.4.1 (Hyperbolic).** *Let $T$ be a $2\times 2$ symplectic matrix with eigenvalues $\mu^{\pm 1}$, $\mu$ real and $\mu \neq 1$. Then there exists a real $2\times 2$ symplectic matrix $S$ such that*

$$S^{-1}TS = \begin{bmatrix} \mu & 0 \\ 0 & \mu^{-1} \end{bmatrix}.$$

*Proof.* Let $Tx = \mu x$ and $Ty = \mu^{-1}y$. Since $x$ and $y$ are eigenvectors corresponding to different eigenvalues they are independent and thus $\{x, y\} \neq 0$. Let $u = \{x, y\}^{-1}y$ so $S = [x, u]$ is the symplectic matrix of the lemma.

In this case define $\Theta(T) = 0$ if $\mu > 0$ and $\Theta(T) = \pi$ if $\mu < 0$.

**Lemma 2.4.2 (Elliptic).** *Let $T$ be a real $2\times 2$ symplectic matrix with eigenvalues $\alpha \pm \beta i$, with $\alpha^2 + \beta^2 = 1$, and $\beta > 0$. Then there exists a real $2 \times 2$ symplectic matrix $S$ such that*

$$S^{-1}TS = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \quad \text{or} \quad S^{-1}TS = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}. \tag{2.28}$$

*Proof.* Let $x = u + iv$ be an eigenvector corresponding to $\lambda = \alpha + i\beta$, so $Tu = \alpha u - \beta v$ and $Tv = \alpha v + \beta u$. Since $u + iv$ and $u - iv$ are independent so are $u$ and $v$ so $\{u, v\} = \delta \neq 0$. If $\delta = \gamma^2 > 0$ then define $S = [\gamma^{-1}u, \gamma^{-1}v]$ to get the first option. If $\delta = -\gamma^2 < 0$ then define $S = [\gamma^{-1}v, \gamma^{-1}v)]$ to get the second option.

Note that these two matrices in (2.28) are orthogonal/rotation matrices. In this case the eigenvalues are of the form $e^{\pm i\theta} = \alpha \pm \beta i$ and for a general matrix one cannot choose the plus or the minus sign, but here one can. The image of $(1, 0)^T$ by the second matrix in (2.28) is $(\alpha, \beta)^T$ which is in the upper half-plane so one should take $0 < \theta < \pi$ modulo $2\pi$ for this second matrix. Likewise for the first matrix one should take $-\pi < \theta < 0$ modulo $2\pi$. Thus we have define $\Theta(T) = \theta$ in this case.

Since a symplectic change of variables is orientation preserving the two matrices is (2.28) are not symplectically similar.

**Lemma 2.4.3 (Parabolic).** *Let $T$ be a real $2 \times 2$ symplectic matrix with eigenvalue $+1$ (respectively $-1$). Then there exists a real $2 \times 2$ symplectic matrix $S$ such that*

$$S^{-1}TS = \begin{bmatrix} 1 & \delta \\ 0 & 1 \end{bmatrix}, \quad \left( \textit{respectively,} \quad S^{-1}TS = \begin{bmatrix} -1 & -\delta \\ 0 & -1 \end{bmatrix} \right) \tag{2.29}$$

*where $\delta = +1,\ -1,$ or $0$.*

*Proof.* First let the eigenvalue be $+1$. If $T$ is the identity matrix then clearly $\delta = 0$. Assume otherwise so $T$ has the repeated eigenvalue 1 and its standard Jordan canonical form is $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. In other words there are independent vectors $x, y$ such that $Ax = x$, $Ay = x + y$. Since they are independent $\{x, y\} = \delta\gamma^2$ where $\delta = 1$, or $-1$ and $\gamma \neq 0$.

Define $u = \delta\gamma^{-1}x$, $v = \gamma^{-1}y$, so $\{u, v\} = 1$ and $T$ in these coordinates is of the form in the Lemma.

The eigenvalue $-1$ is similar.

In this case define $\Theta(T) = 0$ if $T$ is the identity matrix, $\Theta(T) = \pi$ if $\delta = +1$ and $\Theta(T) = -\pi$ if $\delta = -1$

Each $T \in Sp(2, \mathbb{R})$ lies in one of the orbits defined in the above lemmas, so $\Theta$ is defined and continuous on $Sp(2, \mathbb{R})$.

Except in the parabolic case when the eigenvalues are $\pm 1$ the orbits are closed. When $\delta \neq 0$ the orbit of $\begin{bmatrix} \pm 1 & \delta \\ 0 & \pm 1 \end{bmatrix}$ is not closed but has $\pm I$ ($I$ the identity matrix) as a limit.

Let $\Gamma : [0, 1] \to Sp(2, \mathbb{R})$ be a smooth path of symplectic matrices with $\Gamma(0) = I$, the identity matrix. Let

$$\gamma : [0, 1] \to S^1 : t \mapsto \exp(i\Theta(\Gamma(t)))$$

where $S^1$ the unit circle in $\mathbb{C}$. Now lift this map to the universal cover $\mathbb{R}$ of $S^1 = \mathbb{R}/\mathbb{Z}$ to get the map $\bar{\gamma} : [0, 1] \to \mathbb{R}$ with the property that $\bar{\gamma}(0) = 0$. Now define the rotation of the path $\Gamma$ by

$$\mathrm{rot}(\Gamma) = \bar{\gamma}(1).$$

*Note that if $\Gamma^*$ is another such path with the same end points which is homotopic to $\Gamma$ then* $\mathrm{rot}(\Gamma)$=$\mathrm{rot}(\Gamma^*)$.

Let $\lfloor r \rfloor = \max\{n \in \mathbb{Z} : n \leq r\}$ be the floor function. The path $\Gamma$ is nondegenerate if $\Gamma(1)$ does not have the eigenvalue 1. For a nondegenerate path $\Gamma$ we define its Conley-Zehnder index $\chi = \chi(\Gamma)$ as follows

(i) If $|\mathrm{tr}(\Gamma(1))| \leq 2$, define $\chi = 2\lfloor \mathrm{rot}(\Gamma(1)) \rfloor + 1$,

(ii) If $|\mathrm{tr}(\Gamma(1))| > 2$, define $\chi = 2\lfloor \mathrm{rot}(\Gamma(1)) \rfloor$.

## 2.5 Floquet-Lyapunov Theory

In this section we introduce some of the vast theory of periodic Hamiltonian systems. A classical detailed discussion of periodic systems can be found in the two-volume set by Yakubovich and Starzhinskii (1975).

Consider a periodic, linear Hamiltonian system

$$\dot{z} = J\frac{\partial H}{\partial z} = JS(t)z = A(t)z, \tag{2.30}$$

where

$$H = H(t,z) = \frac{1}{2}z^T S(t)z, \tag{2.31}$$

and $A(t) = JS(t)$. Assume that $A$ and $S$ are continuous and $T$-periodic; i.e.

$$A(t+T) = A(t), \quad S(t+T) = S(t) \quad \text{for all } t \in \mathbb{R}$$

for some fixed $T > 0$. The Hamiltonian, $H$, is a quadratic form in $z$ with coefficients which are continuous and $T$-periodic in $t \in \mathbb{R}$. Let $Z(t)$ be the fundamental matrix solution of (2.30) that satisfies $Z(0) = I$.

**Lemma 2.5.1.** *$Z(t+T) = Z(t)Z(T)$ for all $t \in \mathbb{R}$.*

*Proof.* Let $X(t) = Z(t+T)$ and $Y(t) = Z(t)Z(T)$. $\dot{X}(t) = \dot{Z}(t+T) = A(t+T)Z(t+T) = A(t)X(t)$; so, $X(t)$ satisfies (2.30) and $X(0) = Z(T)$. $Y(t)$ also satisfies (2.30) and $Y(0) = Z(T)$. By the uniqueness theorem for differential equations, $X(t) \equiv Y(t)$.

The above lemma only requires (2.30) to be periodic, not necessarily Hamiltonian. Even though the equations are periodic the fundamental matrix need not be so, and the matrix $Z(T)$ is the measure of the nonperiodicity of the solutions. $Z(T)$ is called the *monodromy matrix* of (2.30), and the eigenvalues of $Z(T)$ are called the (characteristic) multipliers of (2.30). The multipliers measure how much solutions are expanded, contracted, or rotated after a period. The monodromy matrix is symplectic by Theorem 2.1.3, and so the multipliers are symmetric with respect to the real axis and the unit circle by Proposition 2.3.1. Thus the origin cannot be asymptotically stable.

In order to understand periodic systems we need some information on logarithms of matrices. The complete proof is long, therefore the proof has been relegated to Section 5.2. Here we shall prove the result in the case when the matrices are diagonalizable.

A matrix $R$ has a logarithm if there is a matrix $Q$ such that $R = \exp Q$, and we write $Q = \log R$. The logarithm is not unique in general, even in the real case, because $I = \exp O = \exp 2\pi J$. Thus the identity matrix $I$ has logarithm $O$ the zero matrix and also $2\pi J$.

If $R$ has a logarithm, $R = \exp Q$, then $R$ is nonsingular and has a square root $R^{1/2} = \exp(Q/2)$. The matrix

$$R = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}$$

has no real square root and hence no real logarithm.

**Theorem 2.5.1.** *Let $R$ be a nonsingular matrix; then there exists a matrix $Q$ such that $R = \exp Q$. If $R$ is real and has a square root, then $Q$ may be taken as real. If $R$ is symplectic, then $Q$ may be taken as Hamiltonian.*

*Proof.* We only prove this result in the case when $R$ is symplectic and has distinct eigenvalues because in this case we only need to consider the canonical forms of Section 2.3. See Section 5.2 for a complete discussion of logarithms of symplectic matrices.

Consider the cases. First

$$\log \begin{bmatrix} \mu & 0 \\ 0 & \mu^{-1} \end{bmatrix} = \begin{bmatrix} \log \mu & 0 \\ 0 & -\log \mu \end{bmatrix}$$

is a real logarithm when $\mu > 0$ and complex when $\mu < 0$. A direct computation shows that $\mathrm{diag}(\mu, \mu^{-1})$ has no real square root when $\mu < 0$.

If $\alpha$ and $\beta$ satisfy $\alpha^2 + \beta^2 = 1$, then let $\theta$ be the solution of $\alpha = \cos\theta$ and $\beta = \sin\theta$ so that

$$\log \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} = \begin{bmatrix} 0 & \theta \\ -\theta & 0 \end{bmatrix}.$$

Lastly, $\log \mathrm{diag}(\mathrm{B}^{\mathrm{T}}, \mathrm{B}^{-1}) = \mathrm{diag}(\log \mathrm{B}^{\mathrm{T}}, -\log \mathrm{B})$ where

$$B = \begin{bmatrix} \gamma & \delta \\ -\delta & \gamma \end{bmatrix},$$

and

$$\log B = \log \rho \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & \theta \\ -\theta & 0 \end{bmatrix},$$

is real where $\rho = \sqrt{(\gamma^2 + \delta^2)}$, and $\gamma = \rho\cos\theta$ and $\delta = \rho\sin\theta$.

The monodromy matrix $Z(T)$ is nonsingular and symplectic so there exists a Hamiltonian matrix $K$ such that $Z(T) = \exp(KT)$. Define $X(t)$ by $X(t) = Z(t)\exp(-tK)$ and compute

$$\begin{aligned} X(t+T) &= Z(t+T)\exp K(-t-T) \\ &= Z(t)Z(T)\exp(-KT)\exp(-Kt) \\ &= Z(t)\exp(-Kt) \\ &= X(t). \end{aligned}$$

Therefore, $X(t)$ is $T$-periodic. Because $X(t)$ is the product of two symplectic matrices, it is symplectic. In general, $X$ and $K$ are complex even if $A$ and $Z$ are real. To ensure a real decomposition, note that by Lemma 2.5.1, $Z(2T) = Z(T)Z(T)$; so, $Z(2T)$ has a real square root. Define $K$ as the real solution of $Z(2T) = \exp(2KT)$ and $X(t) = Z(t)\exp(-Kt)$. Then $X$ is $2T$ periodic.

**Theorem 2.5.2.** *(The Floquet–Lyapunov theorem) The fundamental matrix solution $Z(t)$ of the Hamiltonian* (2.30) *that satisfies $Z(0) = I$ is of the form $Z(t) = X(t)\exp(Kt)$, where $X(t)$ is symplectic and $T$-periodic and $K$ is Hamiltonian. Real $X(t)$ and $K$ can be found by taking $X(t)$ to be $2T$-periodic if necessary.*

Let $Z, X$, and $K$ be as above. In Equation (2.30) make the symplectic periodic change of variables $z = X(t)w$; so,

$$\begin{aligned}\dot{z} &= \dot{X}w + X\dot{w} = (\dot{Z}e^{-Kt} - Ze^{-Kt}K)w + Ze^{-Kt}\dot{w}\\ &= AZe^{-Kt}w - Ze^{-Kt}Kw + Ze^{-Kt}\dot{w}\\ &= Az = AXw = AZe^{-Kt}w\end{aligned}$$

and hence

$$-Ze^{-Kt}Kw + Ze^{-Kt}\dot{w} = 0$$

or

$$\dot{w} = Kw. \tag{2.32}$$

**Corollary 2.5.1.** *The symplectic periodic change of variables $z = X(t)w$ transforms the periodic Hamiltonian system* (2.30) *to the constant Hamiltonian system* (2.32)*. Real $X$ and $K$ can be found by taking $X(t)$ to be $2T$-periodic if necessary.*

The eigenvalues of $K$ are called the (characteristic) exponents of (2.30) where $K$ is taken as $\log(Z(T)/T)$ even in the real case. The exponents are the logarithms of the multipliers and so are defined modulo $2\pi i/T$.

## 2.6 Symplectic Transformations

It is now time to turn to the nonlinear systems and generalize the notion of symplectic transformations. Some examples and applications are given in this chapter and many more specialized examples are given in Chapter 8.

Let $\Xi : O \to \mathbb{R}^{2n} : (t, z) \to \zeta = \Xi(t, z)$ be a smooth function where $O$ is some open set in $\mathbb{R}^{2n+1}$; $\Xi$ is called a symplectic function (or transformation or map etc.) if the Jacobian of $\Xi$ with respect to $z$, $D_2\Xi(t, z) = \partial\Xi/\partial z$, is a symplectic matrix at every point of $(t, z) \in O$. Sometimes we use the notation $D_2\Xi$ for the Jacobian of $\Xi$, and sometimes the notation $\partial\Xi/\partial z$ is used. In the first case we think of the Jacobian $D_2\Xi$ as a map from $O$ into the space $\mathcal{L}(\mathbb{R}^{2n}, \mathbb{R}^{2n})$ of linear operators from $\mathbb{R}^{2n}$ to $\mathbb{R}^{2n}$, and in the second case, we think of $\partial\Xi/\partial z$ as the matrix

$$\frac{\partial \Xi}{\partial z} = \begin{bmatrix} \dfrac{\partial \Xi_1}{\partial z_1} & \cdots & \dfrac{\partial \Xi_1}{\partial z_{2n}} \\ \vdots & & \vdots \\ \dfrac{\partial \Xi_{2n}}{\partial z_1} & \cdots & \dfrac{\partial \Xi_{2n}}{\partial z_{2n}} \end{bmatrix}.$$

Thus $\Xi$ is symplectic if and only if

$$\frac{\partial \Xi}{\partial z} J \frac{\partial \Xi}{\partial z}^T = J. \tag{2.33}$$

Recall that if a matrix is symplectic then so is its transpose, therefore we could just as easily transpose the first factor in (2.33). Because the product of two symplectic matrices is symplectic, the composition of two symplectic maps is symplectic by the chain rule of differentiation. Because a symplectic matrix is invertible, and its inverse is symplectic, the inverse function theorem implies that a symplectic map is locally invertible and its inverse, $Z(t, \zeta)$, is symplectic where defined. Because the determinant of a symplectic matrix is $+1$, the transformation is orientation and volume-preserving.

If the transformation $z \to \zeta = \Xi(t, z)$ is considered a change of variables, then one calls $\zeta$ symplectic or canonical coordinates. Consider a nonlinear Hamiltonian system

$$\dot{z} = J\nabla_z H(t, z), \tag{2.34}$$

where $H$ is defined and smooth in some open set $O \subset \mathbb{R}^{2n+1}$. Make a symplectic change of variables from $z$ to $\zeta$ by

$$\zeta = \Xi(t, z) \quad \text{with inverse } z = Z(t, \zeta) \tag{2.35}$$

(so $\zeta \equiv \Xi(t, Z(t, \zeta))$, $z \equiv Z(t, \Xi(t, z))$). Let $\mathbb{O} \in \mathbb{R}^{2n+1}$ be the image of $O$ under this transformation. Then the Hamiltonian $H(t, z)$ transforms to the function $\hat{H}(t, \zeta) = H(t, Z(t, \zeta))$. Later we abuse notation and write $H(t, \zeta)$ instead of introducing a new symbol, but now we are careful to distinguish $H$ and $\hat{H}$. The equation (2.34) transforms to

$$\begin{aligned}
\dot{\zeta} &= \frac{\partial \Xi}{\partial t}(t,z) + \frac{\partial \Xi}{\partial z}(t,z)\dot{z} \\
&= \frac{\partial \Xi}{\partial t}(t,z) + \frac{\partial \Xi}{\partial z}(t,z) J \left(\frac{\partial H}{\partial z}(t,z)\right)^T \\
&= \frac{\partial \Xi}{\partial t}(t,z) + \frac{\partial \Xi}{\partial z}(t,z) J \left(\frac{\partial \hat{H}}{\partial \zeta}(t,\zeta)\frac{\partial \Xi}{\partial z}(t,z)\right)^T \\
&= \frac{\partial \Xi}{\partial t}(t,z) + J\left(\frac{\partial \hat{H}}{\partial \zeta}\right)^T \\
&= \left.\frac{\partial \Xi}{\partial t}(t,z)\right|_{z=Z(t,\zeta)} + J\nabla_\zeta \hat{H}(t,\zeta).
\end{aligned} \tag{2.36}$$

The notation in the second to last term in (2.36) means that you are to take the partial derivative with respect to $t$ first and then substitute in $z = Z(t,\zeta)$. If the change of coordinates, $\Xi$, is independent of $t$, then the term $\partial\Xi/\partial t$ is missing in (2.36); so, the equation in the new coordinates is simply $\dot{\zeta} = J\nabla_\zeta \hat{H}$, a Hamiltonian system with Hamiltonian $\hat{H}$. In this case one simply substitutes the change of variables into the Hamiltonian $H$ to get the new Hamiltonian $\hat{H}$. The Hamiltonian character of the equations is preserved. Actually the system (2.36) is still Hamiltonian even if $\Xi$ depends on $t$, provided $\mathbb{O}$ is a nice set, as we show.

For each fixed $t$, let the set $\mathbb{O}_t = \{\zeta : (t,\zeta) \in \mathbb{O}\}$ be a ball in $\mathbb{R}^{2n}$. We show that there is a smooth function $R : \mathbb{O} \to \mathbb{R}^1$ such that

$$\left.\frac{\partial \Xi}{\partial t}(t,z)\right|_{z=Z(t,\zeta)} = J\nabla_\zeta R(t,\zeta). \tag{2.37}$$

$R$ is called the remainder function. Therefore, in the new coordinates, the equation (2.36) is Hamiltonian with Hamiltonian $R(t,\zeta) + H(t,\zeta)$. (In the case where $\mathbb{O}_t$ is not a ball, the above holds locally; i.e., at each point of $p \in \mathbb{O}$ there is a function $R$ defined in a neighborhood of $p$ such that (2.37) holds in the neighborhood, but $R$ may not be globally defined as a single-valued function on all of $\mathbb{O}$.) By Corollary 7.3.1, we must show that $J$ times the Jacobian of the left-hand side of (2.37) is symmetric. That is, we must show

$$\Gamma = \Gamma^T,$$

where

$$\Gamma(t,\zeta) = J\left.\frac{\partial^2 \Xi}{\partial t \partial z}(t,z)\right|_{z=Z(t,\zeta)} \frac{\partial Z}{\partial \zeta}(t,\zeta).$$

Differentiating (2.33) with respect to $t$ gives

$$\begin{aligned}\frac{\partial^2 \Xi^T}{\partial t \partial z}(t,z) J \frac{\partial \Xi}{\partial z}(t,z) + \frac{\partial \Xi^T}{\partial z}(t,z) J \frac{\partial^2 \Xi}{\partial t \partial z}(t,z) = 0 \\ \frac{\partial \Xi^{-T}}{\partial z}(t,z) \frac{\partial^2 \Xi^T}{\partial t \partial z}(t,z) J + J \frac{\partial^2 \Xi}{\partial t \partial z}(t,z) \frac{\partial \Xi^{-1}}{\partial z}(t,z) = 0.\end{aligned} \tag{2.38}$$

Substituting $z = Z(t,\zeta)$ into (2.38) and noting that $(\partial \Xi^{-1}/\partial z)(t, Z(t,\zeta)) = \partial Z(t,\zeta)$ yields $-\Gamma^T + \Gamma = 0$. Thus we have shown the following.

**Theorem 2.6.1.** *A symplectic change of variables on $\mathbb{O}$ takes a Hamiltonian system of equations into a Hamiltonian system.*

A partial converse is also true. If a change of variables preserves the Hamiltonian form of all Hamiltonian equations, then it is symplectic. We do not need this result and leave it as an exercise.

### 2.6.1 The Variational Equations

Let $\phi(t,\tau,\zeta)$ be the general solution of (2.34); so, $\phi(\tau,\tau,\zeta) = \zeta$, and let $X(t,\tau,\zeta)$ be the Jacobian of $\phi$ with respect to $\zeta$; i.e.,

$$X(t,\tau,\zeta) = \frac{\partial \phi}{\partial \zeta}(t,\tau,\zeta).$$

$X(t,\tau,\zeta)$ is called the monodromy matrix. Substituting $\phi$ into (2.34) and differentiating with respect to $\zeta$ gives

$$\dot{X} = JS(t,\tau,\zeta)X, \qquad S(t,\tau,\zeta) = \frac{\partial^2 H}{\partial x^2}(t, \phi(t,\tau,\zeta)). \tag{2.39}$$

Equation (2.39) is called the variational equation and is a linear Hamiltonian system. Differentiating the identity $\phi(\tau,\tau,\zeta) = \zeta$ with respect to $\zeta$ gives $X(\tau,\tau,\zeta) = I$, the $2n \times 2n$ identity matrix; so, $X$ is a fundamental matrix solution of the variational equation. By Theorem 2.1.3, $X$ is symplectic.

**Theorem 2.6.2.** *Let $\phi(t,\tau,\zeta)$ be the general solution of the Hamiltonian system (2.34). Then for fixed $t$ and $\tau$, the map $\zeta \to \phi(t,\tau,\zeta)$ is symplectic. Conversely, if $\phi(t,\tau,\zeta)$ is the general solution of a differential equation $\dot{z} = f(t,z)$, where $f$ is defined and smooth on $I \times O$, $I$ an interval in $\mathbb{R}$ and $O$ a ball in $\mathbb{R}^{2n}$, and the map $\zeta \to \phi(t,\tau,\zeta)$ is always symplectic, then the differential equation $\dot{z} = f(t,z)$ is Hamiltonian.*

*Proof.* The direct statement was proved above; now consider the converse. Let $\phi(t,\tau,\zeta)$ be the general solution of $\dot{z} = f(t,z)$, and let $X(t,\tau,\zeta)$ be the Jacobian of $\phi$. $X$ satisfies

$$X^T(t,\tau,\zeta)JX(t,\tau,\zeta) = J$$

Differentiate this with respect to $t$ (first argument) and set $t = \tau$ so $\dot{X}^T(\tau,\tau,\zeta)J + J\dot{X}(\tau,\tau,\zeta) = 0$ so $\dot{X}(\tau,\tau,\zeta)$ is Hamiltonian.

$X$ also satisfies $\dot{X}(t,\tau,\zeta) = f_z(t,\phi(t,\tau,\zeta))X(t,\tau,\zeta)$ Set $t = \tau$ to get $\dot{X}(\tau,\tau,\zeta) = f_z(\tau,\zeta)$, so $f_z(\tau,\zeta)$ is Hamiltonian, and thus $-J\dot{X}$ is symmetric. But $X(t,\tau,\zeta) = \partial f/\partial z(t,\phi(t,\tau,\zeta))$; so,$-J\partial f/\partial z$ is symmetric. Because $O$ is a ball, $-Jf$ is a gradient of a function $H$ by Corollary 7.3.1. Thus $f(t,z) = J\nabla H(t,z)$.

This theorem says that the flow defined by an autonomous Hamiltonian system is volume-preserving. So, in particular, there cannot be an asymptotically stable equilibrium point, periodic solution, etc. This makes the stability theory of Hamiltonian systems difficult and interesting. In general, it is difficult to construct a symplectic transformation with nice properties using definition (2.33). The theorem above gives one method of assuring that a transformation is symplectic, and this is the basis of the method of Lie transforms explored in Chapter 9.

### 2.6.2 Poisson Brackets

Let $F(t,z)$ and $G(t,z)$ be smooth, and recall the definition of the Poisson bracket $\{F,G\}_z(t,z) = \nabla_z F(t,z)^T J\nabla_z G(t,z)$. Here we subscript the bracket to remind us it is a coordinate dependent definition. Let $\hat{F}(t,\zeta) = F(t,Z(t,\zeta))$ and $\hat{G}(t,\zeta) = G(t,Z(t,\zeta))$ where $Z(t,\zeta)$ is symplectic for fixed $t$; so,

$$\begin{aligned}
\{\hat{F},\hat{G}\}_\zeta(t,\zeta) &= \nabla_\zeta \hat{F}(t,\zeta)^T J\nabla_\zeta \hat{G}(t,\zeta)\\
&= \left(\frac{\partial Z^T}{\partial\zeta}(t,\zeta)\nabla_z F(t,Z(t,\zeta))\right)^T J\frac{\partial Z^T}{\partial\zeta}\nabla_z G(t,Z(t,\zeta))\\
&= \nabla_z F(t,Z(t,\zeta))^T\frac{\partial Z}{\partial\zeta}J\frac{\partial Z^T}{\partial\zeta}(t,\zeta)\nabla_z G(t,Z(t,\zeta))\\
&= \nabla_z F(t,Z(t,\zeta))^T J\nabla_z G(t,Z(t,\zeta))\\
&= \{F,G\}_z(t,Z(t,\zeta)).
\end{aligned}$$

This shows that the Poisson bracket operation is invariant under symplectic changes of variables. That is, you can commute the operations of computing Poisson brackets and making a symplectic change of variables.

**Theorem 2.6.3.** *Poisson brackets are preserved by a symplectic change of coordinates.*

Let $\zeta_i = \Xi_i(t,z)$ be the $i$th component of the transformation. In components, Equation (2.33) says

$$\{\Xi_i, \Xi_j\} = J_{ij}, \tag{2.40}$$

where $J = (J_{ij})$.

If the transformation (2.35) is given in the classical notation

$$Q_i = Q_i(q,p), \quad P_i = P_i(q,p), \tag{2.41}$$

then (2.40) becomes

$$\{Q_i, Q_j\} = 0, \qquad \{P_i, P_j\} = 0, \qquad \{Q_i, P_j\} = \delta_{ij}, \tag{2.42}$$

where $\delta_{ij}$ is the Kronecker delta.

**Theorem 2.6.4.** *The transformation* (2.35) *is symplectic if and only if* (2.40) *holds, or the transformation* (2.41) *is symplectic if and only if* (2.42) *holds.*

## 2.7 Symplectic Scaling

If instead of satisfying (2.33) a transformation $\zeta = \Xi(t,z)$ satisfies

$$J = \mu \frac{\partial \Xi}{\partial z} J \frac{\partial \Xi}{\partial z}^T,$$

where $\mu$ is some nonzero constant, then $\zeta = \Xi(t,z)$ is called a symplectic transformation (map, change of variables, etc.) with multiplier $\mu$. Equation (2.34) becomes

$$\dot{\zeta} = \mu J \nabla_\zeta H(t,\zeta) + J \nabla_\zeta R(t,\zeta),$$

where all the symbols have the same meaning as in Section 2.6. In the time-independent case, you simply multiply the Hamiltonian by $\mu$.

As an example consider scaling the universal gravitational constant $G$. When the $N$-body problem was introduced in Section 3.1, the equations contained the universal gravitational constant $G$. Later we set $G = 1$. This can be accomplished by a symplectic change of variables with multiplier. The change of variables $q = \alpha q', p = \alpha p'$ is symplectic with multiplier $\alpha^{-2}$, and so the Hamiltonian of the $N$-body problem, (3.5), becomes

$$H = \sum_{i=1}^{N} \frac{\|p_i'\|^2}{2m_i} - \sum_{1 \le i < j \le N} \frac{G}{\alpha^3} \frac{m_i m_j}{\|q_i' - q_j'\|}.$$

If we take $\alpha^3 = G$, then in the prime coordinates the gravitational constant will be 1. $q$ has the dimensions of distance, and $p$ has the dimensions of distance-mass/time; and so the change of variables can be done by changing the units of distance only. A better way to make the universal gravitational constant unity is to change the unit of mass. The scaling given here is simply an example.

### 2.7.1 Equations Near an Equilibrium Point

Consider a Hamiltonian that has a critical point at the origin; so,

$$H(z) = \frac{1}{2} z^T S z + K(z),$$

where $S$ is the Hessian of $H$ at $z = 0$, and $K$ vanishes along with its first and second partial derivatives at the origin. The change of variables $z = \epsilon w$ is a symplectic change of variables with multiplier $\epsilon^{-2}$; so, the Hamiltonian becomes

$$H(w) = \frac{1}{2} w^T S w + \epsilon^{-2} K(\epsilon w) = \frac{1}{2} w^T S w + O(\epsilon).$$

In the above, the classical notation, $O(\epsilon)$, of perturbation theory is used. Because $K$ is at least third order at the origin, there is a constant $C$ such that $\left|\epsilon^{-2} K(\epsilon w)\right| \leq C\epsilon$ for $w$ in a neighborhood of the origin and $\epsilon$ small, which is written $\epsilon^{-2} K(\epsilon w) = O(\epsilon)$. The equations of motion become

$$\dot{w} = Aw + O(\epsilon), \qquad A = JS. \tag{2.43}$$

If $\|w\|$ is about 1 and $\epsilon$ is small, then $z$ is small. Thus the above transformation is useful in studying the equations near the critical point. To the lowest order in $\epsilon$ the equations are linear; so, close to the critical point the linear terms are the most important terms. This is an example of what is called scaling variables, and $\epsilon$ is called the scale parameter. To avoid the growth of symbols, one often says: scale by $z \to \epsilon z$ which means replace $z$ by $\epsilon z$ everywhere. This would have the effect of changing $w$ back to $z$ in (2.43). It must be remembered that scaling is really changing variables.

## 2.8 Problems

1. Consider a quadratic form $H = (1/2)x^T S x$, where $S = S^T$ is a real symmetric matrix. The index of the quadratic form $H$ is the dimension of the largest linear space where $H$ is negative. Show that the index of $H$ is the same as the number of negative eigenvalues of $S$. Show that if $S$ is nonsingular and $H$ has odd index, then the linear Hamiltonian system $\dot{x} = JSx$ is unstable. (Hint: Show that the determinant of $JS$ is negative.)

2. Consider the linear fractional (or Möbius) transformation

$$\Phi : z \to w = \frac{1+z}{1-z}, \quad \Phi^{-1} : w \to z = \frac{w-1}{w+1}.$$

   a) Show that $\Phi$ maps the left half-plane into the interior of the unit circle. What are $\Phi(0), \Phi(1), \Phi(i), \Phi(\infty)$?
   b) Show that $\Phi$ maps the set of $m \times m$ matrices with no eigenvalue $+1$ bijectively onto the set of $m \times m$ matrices with no eigenvalue $-1$.
   c) Let $B = \Phi(A)$ where A and $B$ are $2n \times 2\mathrm{n}$. Show that $B$ is symplectic if and only if A is Hamiltonian.
   d) Apply $\Phi$ to each of the canonical forms for Hamiltonian matrices to obtain canonical forms for symplectic matrices.
3. Consider the system (*) $M\ddot{q} + Vq = 0$, where $M$ and $V$ are $n \times n$ symmetric matrices and $M$ is positive definite. From matrix theory there is a nonsingular matrix $P$ such that $P^T MP = I$ and an orthogonal matrix $R$ such that $R^T(P^T VP)R = \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_\mathrm{n})$. Show that the above equation can be reduced to $\ddot{p} + \Lambda p = 0$. Discuss the stability and asymptotic behavior of these systems. Write (*) as a Hamiltonian system with Hamiltonian matrix $A = J\mathrm{diag}(\mathrm{V}, \mathrm{M}^{-1})$. Use the above results to obtain a symplectic matrix $T$ such that

$$T^{-1}AT = \begin{bmatrix} 0 & I \\ -\Lambda & 0 \end{bmatrix}.$$

   (Hint: Try $T = \mathrm{diag}(\mathrm{PR}, \mathrm{P}^{-\mathrm{T}}\mathrm{R})$).
4. Let $M$ and $V$ be as in Problem 5.
   a) Show that if $V$ has one negative eigenvalue, then some solutions of (*) in Problem 5 tend to infinity as $t \to \pm\infty$.
   b) Consider the system (**) $M\ddot{q} + \nabla U(q) = 0$, where $M$ is positive definite and $U : \mathbb{R}^n \to \mathbb{R}$ is smooth. Let $q_0$ be a critical point of $U$ such that the Hessian of $U$ at $q_0$ has one negative eigenvalue (so $q_0$ is not a local minimum of $U$). Show that $q_0$ is an unstable critical point for the system (**).
5. Let $H(t, z) = \frac{1}{2} z^T S(t) z$ and $\zeta(t)$ be a solution of the linear system with Hamiltonian H. Show that

$$\frac{d}{dt} H = \frac{\partial}{\partial t} H;$$

   i.e.,

$$\frac{d}{dt} H(t, \zeta(t)) = \frac{\partial}{\partial t} H(t, \zeta(t)).$$

6. The general linear group, $Gl(m, \mathbb{F})$, is the set of all $m \times m$ nonsingular matrices. A matrix Lie group is a closed subgroup of $Gl(m, \mathbb{F})$. Show that the following are matrix Lie groups.

   a) $Sl(m, \mathbb{F})$ = special linear group = set of all $A \in Gl(m, \mathbb{F})$ with $\det A = 1$.
   b) $O(m, \mathbb{F})$ = orthogonal group = set of all $m \times m$ orthogonal matrices.
   c) $\mathrm{So}(m, \mathbb{F})$ = special orthogonal group = $O(m, \mathbb{F}) \cap Sl(m, \mathbb{F})$.
   d) $Sp(2n, \mathbb{F})$ = symplectic group = set of all $2n \times 2n$ symplectic matrices.
7. Show that the following are Lie subalgebras of $gl(m, \mathbb{F})$, see Problem 2 in Chapter 1.
   a) $sl(m, \mathbb{F})$ = set of $m \times m$ matrices with trace = 0. ($sl$ = special linear.)
   b) $o(m, \mathbb{F})$ = set of $m \times m$ skew-symmetric matrices. ($o$ = orthogonal.)
   c) $sp(2n, \mathbb{F})$ = set of all $2n \times 2n$ Hamiltonian matrices.
8. Let $Q(n, \mathbb{F})$ be the set of all quadratic forms in $2n$ variables with coefficients in $\mathbb{F}$, so $q \in \mathcal{Q}(n, \mathbb{R})$, if $q(x) = \frac{1}{2}x^T Sx$, where $S$ is a $2n \times 2n$ symmetric matrix and $x \in \mathbb{F}^{2n}$.
   a) Prove that $\mathcal{Q}(n, \mathbb{F})$ is a Lie algebra, where the product is the Poisson bracket.
   b) Prove that $\Psi : \mathcal{Q}(n, \mathbb{F}) \to sp(2n, \mathbb{F}) : q(x) = \frac{1}{2}x^T Sx \to JS$ is a Lie algebra isomorphism.
9. Show that the matrices

$$\begin{bmatrix} -1 & 1 \\ \\ 0 & -1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -2 & 0 \\ \\ 0 & -1/2 \end{bmatrix}$$

   have no real logarithm.
10. Prove the theorem: $e^{At} \in \mathcal{G}$ for all $t$ if and only if $A \in \mathcal{A}$ in the following cases:
   a) When $\mathcal{G} = Gl(m, \mathbb{R})$ and $\mathcal{A} = gl(m, \mathbb{R})$
   b) When $\mathcal{G} = Sl(m, \mathbb{R})$ and $\mathcal{A} = sl(m, \mathbb{R})$
   c) When $\mathcal{G} = O(m, \mathbb{R})$ and $\mathcal{A} = so(m, \mathbb{R})$
   d) When $\mathcal{G} = Sp(2n, \mathbb{R})$ and $\mathcal{A} = sp(2n, \mathbb{R})$
11. Consider the map $\Phi : \mathcal{A} \to \mathcal{G} : A \mapsto e^A = \sum_0^\infty A^n/n!$. Show that $\Phi$ is a diffeomorphism of a neighborhood of $0 \in \mathcal{A}$ onto a neighborhood of $I \in \mathcal{G}$ in the following cases:
   a) When $\mathcal{G} = Gl(m, \mathbb{R})$ and $\mathcal{A} = gl(m, \mathbb{R})$
   b) When $\mathcal{G} = Sl(m, \mathbb{R})$ and $\mathcal{A} = sl(m, \mathbb{R})$
   c) When $\mathcal{G} = O(m, \mathbb{R})$ and $\mathcal{A} = so(m, \mathbb{R})$
   d) When $\mathcal{G} = Sp(2n, \mathbb{R})$ and $\mathcal{A} = sp(2n, \mathbb{R})$

   (Hint: The linearization of $\Phi$ is $A \mapsto I + A$. Think implicit function theorem.)
12. Show that $Gl(m, \mathbb{R})$ (respectively $Sl(m, \mathbb{R})$, $O(m, \mathbb{R})$, $Sp(2n, \mathbb{R})$) is a differential manifold of dimension $m^2$ (respectively, $m^2$, $m(m-1)/2$, $(2n^2 + n)$). (Hint: Use the problem above and group multiplication to move neighborhoods around.)
13. Show that if you scale time by $t \to \mu t$, then you should scale the Hamiltonian by $H \to \mu^{-1}H$.

Springer
http://www.springer.com/978-3-319-53690-3

Introduction to Hamiltonian Dynamical Systems and the
N-Body Problem
Meyer, K.; Offin, D.C.
2017, XIII, 384 p. 40 illus., 9 illus. in color., Hardcover
ISBN: 978-3-319-53690-3