3. Linear Hamiltonian Systems

3.1 Preliminaries

In this chapter we study Hamiltonian systems that are linear differential equations. Many of the basic facts about Hamiltonian systems and symplectic geometry are easy to understand in this simple context. The basic linear algebra introduced in this chapter is 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 the appendices or to references to the literature. These results are not important for the main development.

We assume a familiarity with the basic theory of linear algebra and linear differential equations. 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 so is called the general linear group. $I = I_m$ and $0 = 0_m$ denote the $m \times m$ identity and zero matrices, respectively. In general, the subscript is clear from the context.

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

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}. \tag{3.1}$$

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

$$J^{-1} = J^T = -J. \tag{3.2}$$

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

$$\dot{z} = J\frac{\partial H}{\partial z} = JS(t)z = A(t)z, \tag{3.3}$$

where

$$H = H(t,z) = \frac{1}{2}z^T S(t)z, \tag{3.4}$$

K.R. Meyer et al., *Introduction to Hamiltonian Dynamical Systems and the N-Body Problem*, Applied Mathematical Sciences 90, DOI 10.1007/978-0-387-09724-4_3,

$A(t) = JS(t)$. H, the Hamiltonian, is a quadratic form in the zs 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 (3.3) by Theorem 1.3.1.

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 $\phi(t, t_o, z_0)$ of (3.3) 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 (3.3) 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

$$Z(t) = e^{At} = \exp At = \sum_{i=1}^{\infty} \frac{A^n t^n}{n!}. \tag{3.5}$$

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

$$A^T J + JA = 0. \tag{3.6}$$

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

Theorem 3.1.1. *The following are equivalent*: *(i) A is Hamiltonian, (ii) $A = JR$ where R is symmetric, and (iii) JA is symmetric.*

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

Proof. $A = J(-JA)$ and (3.6) 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 (3.1) 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 2×2 case,

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

and so,

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

Thus, a 2×2 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

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

and 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. In higher dimensions, being Hamiltonian is more restrictive than just having trace zero.

The function $[\cdot,\cdot] : gl(m,\mathbb{F}) \times gl(m,\mathbb{F}) \to gl(m,\mathbb{F})$ of Theorem 3.1.1 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 μ if

$$T^T JT = \mu J, \tag{3.7}$$

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

Theorem 3.1.2. *If T is symplectic with multiplier μ, then T is nonsingular and*

$$T^{-1} = -\mu^{-1} JT^T J. \tag{3.8}$$

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

Proof. Because the right-hand side, μJ, of (3.7) is nonsingular, T must be also. Formula (3.8) follows at once from (3.7). If T is symplectic, then from (3.8) one gets $T^T = -\mu JT^{-1}J$; so, $TJT^T = TJ(-\mu JT^{-1}J) = \mu J$. Thus T^T is symplectic with multiplier μ. The remaining facts are proved in a similar manner.

This theorem implies that $Sp(2n,\mathbb{R})$ is a group, a 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" in the connotation of complex number. "I therefore proposed to replace it by the corresponding Greek adjective 'symplectic.'" See page 165 in Weyl (1948).

In the 2×2 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 2×2 matrix is symplectic (with multiplier μ) if and only if it has determinant $+1$ (respectively μ). Thus a 2×2 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

$$T = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{3.9}$$

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 μ if and only if $a^T d - c^T b = \mu I$ and $a^T c$ and $b^T d$ are symmetric. Being symplectic is more restrictive in higher dimensions. Formula (3.8) gives

$$T^{-1} = \mu^{-1} \begin{bmatrix} d^T & -b^T \\ -c^T & a^T \end{bmatrix}. \tag{3.10}$$

This reminds one of the formula for the inverse of a 2×2 matrix!

Theorem 3.1.3. *The fundamental matrix solution $Z(t, t_0)$ of a linear Hamiltonian system* (3.3) *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 $U(t_0) = J$. $\dot{U}t) = \dot{Z}^T JZ + Z^T J\dot{Z} = Z^T(A^T J + JA)Z = 0$; 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 3.1.1. *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 (3.3). Equation (3.3) becomes

$$\dot{u} = (T^{-1}AT - T^{-1}\dot{T})u. \tag{3.11}$$

In general this equation will not be Hamiltonian, however:

Theorem 3.1.4. *If T is symplectic with multiplier μ^{-1}, then* (3.11) *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 (3.11) *is Hamiltonian for every Hamiltonian system* (3.3)*, then U is symplectic with constant multiplier μ.*

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, $T^{-1}AT = T^{-1}JST = \mu JT^TST$, and so, $T^{-1}AT = J(\mu T^TST)$ is Hamiltonian also.

Now let (3.11) 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

$$\dot{v} = vB(t). \tag{3.12}$$

So, $T(t) = KV(t, t_0)$, where $V(t, t_0)$ is the fundamental matrix solution of (3.12), and $K = T(t_0)$ is a constant matrix. By Theorem 3.1.3 , 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

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

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 ith, which is the ith row of R and the only nonzero column of R^TC is the ith, which is the ith 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 μ. $R = KJK^TJ = -\mu I$ implies $KJK^T = \mu J$.

This is an example of a change of variables that preserves the Hamiltonian character of the system of equations. The general problem of which changes of variables preserve the Hamiltonian character is discussed in detail in Chapter 6.

The fact that the fundamental matrix of (3.3) is symplectic means that the fundamental matrix must satisfy the identity (3.7). There are many functional relations in (3.7); so, there are functional relations between the solutions. Theorem 3.1.5 given below is just one example of how these relations can be used. See Meyer and Schmidt (1982b) for some other examples.

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

$$\{z_1, z_2\}(t) = z_1^T(t) J z_2(t); \tag{3.13}$$

so, $\{z_1, z_2\} : \mathbb{I} \to \mathbb{R}^{2n}$ 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 and pairwise in involution functions $z_1, \ldots, z_n$ 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 3.1.5. *If a Lagrangian set of solutions of* (3.3) *is known, then a complete set of* $2n$ *linearly independent solutions can be found by quadrature. (See* (3.14)*.)*

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, $C^T JC = 0$; and because they are independent, $C^T C$ is an $n \times n$ nonsingular matrix. Define the $2n \times n$ matrix by $D = JC(C^T C)^{-1}$. Then $D^T JD = 0$ and $C^T JD = -I$, and so $P = (D, C)$ is a symplectic matrix. Therefore,

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

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

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

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,

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

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

$$V(t) = \int_{t_0}^{t} D^T(SD + J\dot{D})dt. \tag{3.14}$$

A symplectic fundamental matrix solution of (3.3) 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 k integrals in involution for a general Hamiltonian system can be used to reduce the number of degrees of freedom by k and, hence, the dimension by $2k$.

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 3.1.6. *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 (3.8) 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}, as is 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 (3.8) these last relations imply that P and O are symplectic. A similar argument gives that P' and O' are symplectic.

Theorem 3.1.7. *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 4.6 and Chapter 5 we give alternate proofs. Let T be symplectic. Formula (3.7) gives $\det(T^TJT) = \det T^T \det J \det T = (\det T)^2 = \det J = 1$ 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 (3.9) for T, formula (3.10) 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

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

3.2 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 $B(u, v) = u^T S v$, where S is any skew-symmetric matrix.

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 $\Phi : \mathbb{V} \to \mathbb{F}^m : v \to a = (\alpha_1, \ldots, \alpha_m)$. 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 $B(u, v) = \Sigma\Sigma \alpha_i \beta_j B(e_i, e_j) = b^T S a$. So in the coordinates defined by the basis (e_i), the bilinear form is just $b^T S a$ 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 $S = QRQ^T$. 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 3.2.1. *Let S be any skew-symmetric matrix; then there exists a nonsingular matrix Q such that*

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

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. Preform 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 preform 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 the 2×2 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 $\mathbb{F}$, $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, and ω is a nondegenerate alternating bilinear form on $\mathbb{V}$. The form ω 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 ω. The standard example is $\mathbb{F}^{2n}$ and $\omega(x, y) = x^T Jy$. In this example we shall write $\{x, y\} = x^T Jy$ 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 $v_1, \ldots, v_{2n}$ for $\mathbb{V}$ such that $\omega(v_i, v_j) = J_{ij}$, the i, jth entry of J. A symplectic basis is a basis so that the matrix of ω is just J. The standard basis $e_1, \ldots, e_{2n}$, where e_i is 1 in the ith 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 3.2.1. *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 3.2.1 there is a basis for $\mathbb{V}$ such that the matrix of ω is $\mathrm{diag}(\mathrm{K}, \ldots, \mathrm{K})$. Interchanging rows $2i$ and $n + 2i - 1$ and the corresponding columns brings the matrix to J. The basis for which the matrix of ω is J is a symplectic basis; so, a symplectic basis exists.

Let $v_1, \ldots, v_{2n}$ be a symplectic basis for $\mathbb{V}$ and $u \in \mathbb{V}$. There exist constants α_i such that $u = \sum \alpha_i v_i$. The linear map $L : \mathbb{V} \to \mathbb{F}^{2n} : u \to (\alpha_1, \ldots, \alpha_{2n})$ 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 αf 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 $v = f^1 u_1 + \cdots + f^m u_m$. 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, for example, Halmos (1958).

Because ω 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 ω 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 3.2.1. *$\mathbb{U}^\perp = \mathbb{U}^{0\#}$. $\dim \mathbb{U} + \dim \mathbb{U}^\perp = \dim \mathbb{V} = 2n$.*

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 ω 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 3.2.1. *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 3.2.1, 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 ω 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 ω is degenerate on all of $\mathbb{V}$. This contradiction yields the second statement.

A Lagrangian space $\mathbb{U}$ is a subspace of $\mathbb{V}$ of dimension n such that ω 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 3.2.2. *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 nonunique. 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 3.2.3. *Let $\mathbb{V} = \mathbb{U} \oplus \mathbb{W}$ be a Lagrange splitting and $x_1, \dots, x_n$ any basis for $\mathbb{U}$. Then there exists a unique basis $y_1, ..., y_n$ of $\mathbb{W}$ such that $x_1, \dots, x_n, y_1, \dots, y_n$ 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 ω is nondegenerate, and this implies $\alpha_i = 0$, because the $x_i s$ are independent. Thus $\phi_1, \dots, \phi_n$ are independent, and so, they form a basis for $\mathbb{W}^0$. Let $y_1, \dots, y_n$ 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

$$\omega(Lx, y) + \omega(x, Ly) = 0 \tag{3.15}$$

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

$$\omega(Lx, Ly) = \omega(x, y) \tag{3.16}$$

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 (3.15) means $x^T(L^TJ+JL)y = 0$ for all x and y. But this implies that L is a Hamiltonian matrix. On the other hand, if L satisfies (3.16) then $x^TL^TJLy = x^TJy$ 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 3.2.4. *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* 3.2.3*; the matrix representation of A in these symplectic coordinates is of the form*

$$\begin{bmatrix} B^T & 0 \\ 0 & -B \end{bmatrix} \quad \left(\textit{respectively, } \begin{bmatrix} B^T & 0 \\ 0 & B^{-1} \end{bmatrix} \right). \tag{3.17}$$

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 (3.17).

3.3 The Spectra of Hamiltonian and Symplectic Operators

In this section we obtain some canonical forms for Hamiltonian and symplectic matrices in some simple cases. The complete picture is very detailed and would lead us too far afield to develop fully. 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

$$A_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{3.18}$$

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 $T = \mathrm{diag}(1, -1)$ 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 (3.18) 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 λ_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 λ_0 is a zero of a reciprocal polynomial, then so is λ_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 3.3.1. *The characteristic polynomial of a real Hamiltonian matrix is an even polynomial. Thus if λ 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 λ is an eigenvalue of a real symplectic matrix, then so are λ^{-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 3.1.7 $\det T = +1$. $p(\lambda) = \det(T - \lambda I) = \det(T^T - \lambda I) = \det(-JT^{-1}J - \lambda I) = \det(-JT^{-1}J + \lambda JJ) = \det(-T^{-1} + \lambda I) = \det T^{-1}\det(-I + \lambda T) = \lambda^{2n}\det(-\lambda^{-1}I + T) = \lambda^{2n}p(\lambda^{-1})$.

Actually we can prove much more. By (3.6), Hamiltonian matrix A satisfies $A = J^{-1}(-A^T)J$; so, A and $-A^T$ are similar, and the multiplicity of the eigenvalues λ_0 and $-\lambda_0$ are the same. In fact, the whole Jordan block structure will be the same for λ_0 and $-\lambda_0$.

By (3.8), symplectic matrix T satisfies $T^{-1} = J^{-1}T^TJ$; so, T^{-1} and T^T are similar, and the multiplicity of the eigenvalues λ_0 and λ_0^{-1} are the same. The whole Jordan block structure will be the same for λ_0 and λ_0^{-1}.

Consider the linear constant coefficient Hamiltonian system of differential equations

$$\dot{x} = Ax, \tag{3.19}$$

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 λ 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 λ is $\eta(\lambda) = \eta_1(\lambda)$, and the generalized eigenspace is $\eta^\dagger(\lambda)$. If $\{x, y\} = x^TJy = 0$, then x and y are J-orthogonal.

Lemma 3.3.1. *Let λ and μ be eigenvalues of A with $\lambda + \mu \neq 0$, then $\{\eta(\lambda),\ \eta(\mu)\} = 0$. That is, the eigenvectors corresponding to λ and μ are J-orthogonal.*

Proof. Let $Ax = \lambda x$, and $Ay = \mu y$, where $x, y \neq 0$. $\lambda\{x, y\} = \{Ax, y\} = x^TA^TJy = -x^TJAy = -\{x, Ay\} = -\mu\{x, y\}$; and so, $(\lambda + \mu)\{x, y\} = 0$.

Corollary 3.3.1. *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 $S^{-1}AS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n, -\lambda_1, \ldots, -\lambda_n)$.*

Proof. Let $\mathbb{U} = \eta_1(\lambda_1) \cup \cdots \cup \eta_1(\lambda_n)$ and $\mathbb{W} = \eta_1(-\lambda_1) \cup \cdots \cup \eta_1(-\lambda_n)$; by the above, $\mathbb{V} = \mathbb{U} \oplus \mathbb{W}$ is a Lagrange splitting, and A respects this splitting. Choose a symplectic basis for $\mathbb{V}$ by Lemma 3.2.3. Changing to that basis is effected by a symplectic matrix G; i.e., $G^{-1}AG = \mathrm{diag}(B^T, -B)$, where B has eigenvalues $\lambda_1, \ldots, \lambda_n$. Let C be such that $C^{-T}B^TC^T = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and define a symplectic matrix by $Q = \mathrm{diag}(C^T, C^{-1})$. The required symplectic matrix is $S = GQ$.

If complex transformations are allowed, then the two matrices in (3.18) can both be brought to diag(i, –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.

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

$$\mathbb{V} = (\oplus_j \mathbb{U}_j) \oplus (\oplus_j \mathbb{W}_j) \oplus (\oplus_j \mathbb{Z}_j), \tag{3.20}$$

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 3.3.1, each space is J-orthogonal to every other, and so by Proposition 3.2.1 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 3.3.2. *Let A be a 2×2 Hamiltonian matrix with eigenvalues $\pm\alpha$, α real, $\alpha \neq 0$. Then there exists a real 2×2 symplectic matrix S such that*

$$S^{-1}AS = \begin{bmatrix} \alpha & 0 \\ 0 & -\alpha \end{bmatrix}. \tag{3.21}$$

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 3.3.3. *Let A be a real 2×2 Hamiltonian matrix with eigenvalues $\pm\beta i, \beta \neq 0$. Then there exists a real 2×2 symplectic matrix S such that*

$$S^{-1}AS = \begin{bmatrix} 0 & \beta \\ -\beta & 0 \end{bmatrix}, \quad or \quad S^{-1}AS = \begin{bmatrix} 0 & -\beta \\ \beta & 0 \end{bmatrix}. \tag{3.22}$$

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 (3.22), 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 3.3.4. *Let A be a real 2×2 Hamiltonian matrix with eigenvalues $\pm\beta i$, $\beta \neq 0$. Then there exist a 2×2 matrix S and a matrix R such that*

$$S^{-1}AS = \begin{bmatrix} i\beta & 0 \\ 0 & -i\beta \end{bmatrix}, \quad R = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad S^TJS = \pm 2iJ, \quad \overline{S} = SR. \tag{3.23}$$

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 (3.23), 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 $S^TJS = +2iJ$ by allowing an interchange of the signs in (3.23).

Lemma 3.3.5. *Let A be a 4×4 Hamiltonian matrix with eigenvalue $\pm\gamma \pm \delta i$, $\gamma \neq 0$, $\delta \neq 0$. Then there exists a real 4×4 symplectic matrix S such that*

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

where B is a real 2×2 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 3.3.1 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 3.2.4. The result follows from Lemma 3.2.4.

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

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

This completes the case when A has distinct eigenvalues. There are many cases when A has eigenvalues with zero real part; i.e., zero or pure imaginary. These cases are discussed in detail in Section 4.7. In the case where the eigenvalue zero is of multiplicity 2 or 4 the canonical forms are the 2×2 and 4×4 zero matrices and

$$\begin{bmatrix} 0 & \pm 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \pm 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}. \tag{3.24}$$

The corresponding Hamiltonians are

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

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

$$\begin{bmatrix} 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \pm\alpha \\ -\alpha & 0 & 0 & 0 \\ 0 & \mp\alpha & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & \alpha & 0 & 0 \\ -\alpha & 0 & 0 & 0 \\ \pm 1 & 0 & 0 & \alpha \\ 0 & \pm 1 & -\alpha & 0 \end{bmatrix}. \tag{3.25}$$

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 λ 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 λ 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 3.3.6. *If λ and μ 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 λ and μ are J-orthogonal.*

Corollary 3.3.2. *Let T be a $2n \times 2n$ symplectic matrix with distinct eigenvalues $\lambda_1, \ldots, \lambda_n, \lambda_1^{-1}, \ldots, \lambda_n^{-1}$; then there exists a symplectic matrix S (possibly complex) such that*

$$S^{-1} T S = \text{diag}(\lambda_1, \ldots, \lambda_{\text{n}}, \lambda_1^{-1}, \ldots, \lambda_{\text{n}}^{-1}).$$

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

$$\mathbb{V} = (\oplus_j \mathbb{U}_j) \oplus (\oplus_j \mathbb{W}_j) \oplus (\oplus_j \mathbb{Z}_j), \tag{3.26}$$

where

$$\mathbb{U}_j = \eta(\mu_j) \oplus \eta(\mu_j^{-1})$$

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

$$\mathbb{Z}_j = \{\eta(\gamma_j + \delta_j i) \oplus \eta(\gamma_j - \delta_j i)\} \oplus \{\eta(\gamma_j + \delta_j i)^{-1} \oplus \eta(\gamma_j - \delta_j i)^{-1}\}.$$

Each of the summands in (3.26) is invariant for T. By Lemma 3.3.6 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. Therefore, the next task is to consider each space separately.

Lemma 3.3.7. *Let T be a 2×2 symplectic matrix with eigenvalues $\mu^{\pm 1}, \mu$ real, and $\mu \neq 1$. Then there exists a real 2×2 symplectic matrix S such that*

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

Lemma 3.3.8. *Let T be a real 2×2 symplectic matrix with eigenvalues $\alpha \pm \beta i$, $\alpha^2 + \beta^2 = 1$, and $\beta \neq 0$. Then there exists a real 2×2 symplectic matrix S such that*

$$S^{-1} T S = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \quad \text{or} \quad S^{-1} T S = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}. \tag{3.27}$$

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 3.3.9. *Let T be a real 2×2 symplectic matrix with eigenvalues $\alpha \pm \beta i$, $\alpha^2 + \beta^2 = 1$, and $\beta \neq 0$. Then there exists a 2×2 matrix S and a matrix R such that*

$$S^{-1}TS = \begin{bmatrix} \alpha+\beta i & 0 \\ 0 & \alpha-\beta i \end{bmatrix}, \qquad R = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

$S^TJS = \pm 2iJ$, and $\overline{S} = SR$.

Lemma 3.3.10. *Let T be a 4×4 symplectic matrix with eigenvalues $(\gamma \pm \delta i)^{\pm 1}$, $\gamma^2+\delta^2 \neq 1$, and $\delta \neq 0$. Then there exists a real 4×4 symplectic matrix S such that*

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

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

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

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

This completes the case when T has distinct eigenvalues.

3.4 Periodic Systems and Floquet–Lyapunov Theory

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

Consider a periodic, linear Hamiltonian system

$$\dot z = J\frac{\partial H}{\partial z} = JS(t)z = A(t)z, \tag{3.28}$$

where

$$H = H(t,z) = \frac{1}{2}z^TS(t)z, \tag{3.29}$$

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 the zs with coefficients which are continuous and T-periodic in $t \in \mathbb{R}$. Let $Z(t)$ be the fundamental matrix solution of (3.28) that satisfies $Z(0) = I$.

Lemma 3.4.1. *$Z(t+T) = Z(t)Z(T)$ 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 (3.28) and $X(0) = Z(T)$. $Y(t)$ also satisfies (3.28) and $Y(t) = Z(T)$. By the uniqueness theorem for differential equations, $X(t) \equiv Y(t)$.

The above lemma only requires (3.28) 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 nonperodicity of the solutions. $Z(T)$ is called the monodromy matrix of (3.28), and the eigenvalues of $Z(T)$ are called the (characteristic) multipliers of (3.28). The multipliers measure how much solutions are expanded, contracted, or rotated after a period. The monodromy matrix is symplectic by Theorem 3.1.3, and so the multipliers are symmetric with respect to the real axis and the unit circle by Proposition 3.3.1. 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 4.3. 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 $I = \exp O = \exp 2\pi J$. 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 3.4.1. *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 s 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 consider the canonical forms of Section 3.3. See Section 4.3 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 α and β satisfy $\alpha^2 + \beta^2 = 1$, then let θ 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 3.4.1, $Z(2T) = Z(T)Z(T)$; 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 3.4.2. *(The Floquet–Lyapunov theorem) The fundamental matrix solution $Z(t)$ of the Hamiltonian* (3.28) *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 (3.28) 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

$$\dot{w} = Kw. \tag{3.30}$$

Corollary 3.4.1. *The symplectic periodic change of variables $z = X(t)w$ transforms the periodic Hamiltonian system* (3.28) *to the constant Hamiltonian system* (3.30)*. 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 (3.28) 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 $2\pi i/T$.

Problems

1. Supply proofs to the lemmas and corollaries 3.3.6 to 3.3.10.
2. Prove that the two symplectic matrices in formula (3.27) in Lemma 3.3.8 are not symplectically similar.
3. Consider a quadratic form $H = (1/2)x^T Sx$, 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.)
4. Consider the linear fractional (or Möbius transformation)
$$\Phi : z \to w = \frac{1+z}{1-z}, \Phi^{-1} : w \to z = \frac{w-1}{w+1}.$$
   a) Show that Φ 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 Φ 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 Φ to each of the canonical forms for Hamiltonian matrices to obtain canonical forms for symplectic matrices.
5. 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 $T = \mathrm{diag}(\mathrm{PR}, \mathrm{P}^{-\mathrm{T}}\mathrm{R})$).
6. Let M and V be as in Problem 4.
   a) Show that if V has one negative eigenvalue, then some solutions of (*) in Problem 4 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 (**).

7. 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.,

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

8. Let G be a set. A product on G is a function from $G \times G$ into G. A product is usually written using infix notation; so, if the product is denoted by $\circ$ then one writes $a \circ b$ instead of $\circ(a,b)$. Addition and multiplication of real numbers define products on the reals, but the inner product of two vectors does not define a product because the inner product of two vectors is a scalar not a vector.
A group is a set G with a product $\circ$ on G that satisfies (i) there is a unique element $e \in G$ such that $a \circ e = e \circ a = a$ for all $a \in G$, (ii) for every $a \in G$ there is a unique element $a^{-1} \in G$ such that $a \circ a^{-1} = a^{-1} \circ a = e$, (iii) $(a \circ b) \circ c = a \circ (b \circ c)$ for all $a, b, c \in$ G. e is called the identity, a^{-1} the inverse of a, and the last property is the associative law.
Show that the following are groups.
   a) $G = \mathbb{R}$, the reals, and $\circ = +$, addition of real numbers. (What is e? Ans. 0.)
   b) $G = \mathbb{C}$, the complex numbers, and $\circ = +$, addition of complex numbers. (What is a^{-1}? Ans -a.)
   c) $G = \mathbb{R}\backslash\{0\}$, the nonzero reals, and $\circ = \cdot$, multiplication of reals.
   d) $G = Gl(n,\mathbb{R})$, the set of all $n \times n$ real, nonsingular matrices, and $\circ = \cdot$ matrix multiplication.
9. Using the notation of the previous problem show that the following are not groups.
   a) $G = \mathbb{E}^3$, 3-dimensional geometric vectors, and $\circ = \times$, the vector cross product.
   b) $G = \mathbb{R}^+$, the positive reals, and $\circ = +$, addition.
   c) $G = \mathbb{R}$, and $\circ = \cdot$, real multiplication.
10. A subgroup of a group G is a subset $H \subset G$, which is a group with the same product. A matrix Lie group is a closed subgroup of $Gl(m,\mathbb{F})$. Show that the following are matrix Lie groups.
   a) $Gl(m,\mathbb{F})$ = general linear group = all $n \times n$ nonsingular matrices
   b) $Sl(m,\mathbb{F})$ = special linear group = set of all $A \in Gl(m,\mathbb{F})$ with $\det A = 1$.
   c) $O(m,\mathbb{F})$ = orthogonal group = set of all $m \times m$ orthogonal matrices.
   d) $\mathrm{So}(m,\mathbb{F})$ = special orthogonal group = $O(m,\mathbb{F}) \cap Sl(m,\mathbb{F})$.
   e) $Sp(2n,\mathbb{F})$ = symplectic group = set of all $2n \times 2n$ symplectic matrices.
11. 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.

12. 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.
13. 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.
14. 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 $\mathcal{A} = sp(2n, \mathbb{R})$
15. Consider the map $\Phi : \mathcal{A} \to \mathcal{G} : A \mapsto e^A = \sum_0^\infty A^n/n!$. Show that Φ 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 Φ is $A \mapsto I + A$. Think implicit function theorem.)
16. 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.)