Symplectic Geometry and Wiliamson's Theorem

Mohammad Yusofsani

2018
November 25

1 Motivation

We already know that in order for a linear operator to be diagonalizable by a change of basis, the operator has to be a normal one i.e.:

$$[N, N^{\dagger}] = 0$$

But when investigating the case of simple harmonic oscillator in Hamiltonian mechanics, we see something which at first glance seems contradictory to this assumption. Consider the following Hamiltonian:

$$H(p,q) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$$

The Hamiltonian equations of motion are as follows:

$$\frac{\partial H(p,q)}{\partial x} = -\dot{p}$$
$$\frac{\partial H(p,q)}{\partial p} = \dot{x}$$

Which for a S.H.O leads to the following equations:

$$m\omega^2 x = -\dot{p}$$
$$\frac{p}{m} = \dot{x}$$

This is a system of two coupled first order differential equations. We can write it in the matrix form:

$$\frac{d}{dx}\begin{bmatrix} p \\ x \end{bmatrix} = \begin{bmatrix} 0 & -m\omega^2 \\ \frac{1}{m} & 0 \end{bmatrix}\begin{bmatrix} p \\ x \end{bmatrix}$$

Since there is no t dependence in the matrix of coefficients, the standard steps to decouple this system includes trying to find some coordinate transformation in which this matrix has a diagonal form in a new basis (x', p'). This will obviously decouple the system of equations. After solving the system of independent equations in this basis, we can use the inverse transformation to derive the equations of motion in (x, p) basis. But the problem here is that the matrix of coefficients here is not a normal one:

$$\begin{bmatrix} 0 & -m\omega^2 \\ \frac{1}{m} & 0 \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{m} \\ -m\omega^2 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & \frac{1}{m} \\ -m\omega^2 & 0 \end{bmatrix} \begin{bmatrix} 0 & -m\omega^2 \\ \frac{1}{m} & 0 \end{bmatrix}$$

So is there any way for us to find a transformation that diagonalizes the coefficient matrix? The problem here can be viewed from another point of view by using a different mathematical arsenal. We can then see that the process of diagonalizing matrix of the coefficients is equivalent with diagonalizing the Hamiltonian itself. The geometric tools that I will be using here is called *symplectic geometry* (and some of you migh be already familiar with it). But before that, a few preliminary sections are in order.

2 Preliminaries 1: Some Definitions

Definition 1: Let E be a real vector space. A vector in E will be denoted by z. A symplectic form an E is a mapping $\omega : E \times E \to \mathbb{R}$ which is:

1) Bilinear:

$$\forall z, z_1, z_2, z', z'_1, z'_2 \in E, \forall \alpha_1, \alpha_2 \in \mathbb{R}$$

$$\omega(\alpha_1 z_1 + \alpha_2 z_2, z') = \alpha_1 \omega(z_1, z') + \alpha_2 \omega(z_2, z')$$
$$\omega(z, \alpha_1 z'_1 + \alpha_2 z'_2) = \alpha_1 \omega(z, z'_1) + \alpha_2 \omega(z, z'_2)$$

2) Antisymmetric:

$$\forall z, z' \in E$$

$$\omega(z, z') = -\omega(z', z)$$

(Which implies: $\omega(z, z) = 0$)

3) Non-degenerate:

$$\forall z \in E \ \omega(z, z') = 0$$

if and only if

$$z' = 0$$

Definition 2: A real symplectic space is a pair (E, ω) where E is a real vector space on $\mathbb{R}$ and ω a symplectic form. The dimension of (E, ω) is, by definition, the dimension of E.

A very important example of a finite dimensional symplectic space is the *standard symplectic space* $(\mathbb{R}^{2n}, \sigma)$ where σ the *standard symplectic form* is defined as:

$$\sigma(z, z') = \sum_{j=1}^{n} p_j x'_j - p'_j x_j$$

With $z = (x_1, ...x_n; p_1, ...p_n)$ and $z' = (x'_1, ...x'_n; p'_1, ...p'_n)$

We can already see that the symplectic form can play the role of an inner product in our space. In fact for any Hilbert space X and its dual X^*, we can define a biliear antisymmetric form such that:

$$\forall z = (x, p), z'(x', p') \in X \oplus X^* \zeta : X \oplus X^* \to \mathbb{R}$$

$$\zeta(z, z') = \langle p, x' \rangle - \langle p', x \rangle$$

It is easy to check that ζ is a sympletic form.

Definition 3: Let Φ be the mapping $E \to E^*$ which to every $z \in E$ associates the linear form Φ_z defined by

$$\Phi_z(z') = \omega(z, z')$$

The non-degeneracy of the symplectic form can be restated as follows:

ω is non-degenerate if and only if Φ is a monomorphism $E \to E^*$.

Definition 4: A set B of vectors like:

$$B = \{e_1, ...e_n\} \cup \{f_1, ..., f_n\}$$

in E is called "symplectic basis" of (E, ω) if the following conditions hold:

$$\omega(e_i, e_j) = \omega(f_i, f_j) = 0$$

$$\omega(e_i, f_j) = \delta_{ij}$$

Definition 5: The set of all symplectic automorphisms $(E, \omega) \to (E, \omega)$ forms a group $Sp(E, \omega)$.

Proposition: Let (E, ω) and (E', ω') be two symplectic spaces of same dimension $2n$. The symplectic groups $Sp(E, \omega)$ and $Sp(E', \omega')$ are isomorphic.

Proof:(It's not so hard but did not present it in class and so I won't put it in here.)

3 Williamson's Theorem

The idea behind this theorem is fairly simple: We can use the members of symplectic group (i.e. $Sp(n, \mathbb{R})$) to diagonalize any positive-definite symmetric matrix. This should come as no surprise because we have been using orthogonal matrices to diagonalize symmetric matrices so far. Just as orthogonal matrices preserve Euclidean length I, the symplectic matrices preserve J. As we shall see, the action of J is somehow like imaginary i in complex plane. So, we can regard an orthogonal similarity transformation as an action mapping "complex matrices" onto the "real line" by preserving I and symplectic transformation as an action mapping "complex matrices" onto imaginary line by preserving J.

The theorem is as follows:

Williamson's Theorem: Let M be a positive-definite symmetric real $2n \times 2n$ matrix.
(i) There exists $S \in Sp(n)$ such that:

$$S^T M S = \begin{bmatrix} \Lambda & 0 \\ 0 & \Lambda \end{bmatrix}$$

with Λ being a diagonal matrix of rank n.
the diagonal entries λ_j of Λ can be defined using the condition:

$$\pm i\lambda_j \text{ is an eigenvalue of } JM$$

(ii) The sequence $\lambda_1, ..., \lambda_n$ does not depend, up to a reordering of its terms, on the choice of S diagonalizing M.

Proof: We first examine the case of $M = I$. We note that J itself is in the $Sp(n, \mathbb{R})$:

$$J^T = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix} = -J$$

$$J^T J J = -J J^2 = J$$

So, we can easily see that this is the matrix that diagonalizes I:

$$J^T I J = -J I J = -J^2 = I$$

We see that the eigenvalues of $JI = J$ are $\pm i$ which is what we expected them to be. This indicates that it would be a good idea to search for complex eigenvalues and eigenvectors of JM in $\mathbb{C}^{2n}$. Let us define:

$$\langle z, z' \rangle_M = \langle Mz, z' \rangle$$

Since both $\langle ., . \rangle_M$ and symplectic form are nondegenerate, we can find a unique invertible matrix K of order 2n such that:

$$\langle z, Kz' \rangle_M = \sigma(z, z')$$

for all z and z'. That matrix satisfies:

$$K^T M = J = -MK$$

Since skew-product is antisymmetric we must have $K = -K^M$ where $K^M = -M^{-1}K^T M$ is the transpose of K with respect of $\langle ., . \rangle_M$. It follows that eigenvalues of $K = -M^{-1}J$ are of the type $\pm i\lambda_j, \lambda_j > 0$, and hence those of JM^{-1} are $\pm i\lambda_j^{-1}$.The corresponding eigenvectors occurring in conjugate pairs $e'_j + if'_j$ we thus obtain a $\langle , , . \rangle_M$-orthonormal basis $\{e'_j, f'_j\}_{1 \le i,j \le n}$ of $\mathbb{R}_z^{2n}$ such that $Ke'_i = \lambda_i f'_i$ and $Kf'_j = -\lambda_j e'_j$. Note that:

$$K^2 e'_i = -\lambda^2 e'_i$$

$$K^2 f'_j = -\lambda^2 f'_j$$

And the vectors of the basis $\{e'_j, f'_j\}_{1\leq i,j\leq n}$ satisfy the following relations:

$$\sigma(e'_i, e'_j) = \langle e'_i, Ke'_j\rangle_M = \lambda_j \langle e'_i, f'_j\rangle_M = 0$$
$$\sigma(f'_i, f'_j) = \langle f'_i, Kf'_j\rangle_M = -\lambda_j \langle f'_i, e'_j\rangle_M = 0$$
$$\sigma(f'_i, e'_j) = \langle f'_i, Ke'_j\rangle_M = \lambda_j \langle f'_i, e'_j\rangle_M = -\lambda_i \delta_{ij}$$

We set $e_i = \lambda_i^{-1/2} e'_i$ and $f_j = \lambda_j^{-1/2} f'_j$ so the basis $\{e, f\}_{1\leq i,j\leq n}$ would be symplectic. Now let S be the element of $Sp(n)$ that maps the canonical symplectic basis to $\{e, f\}_{1\leq i,j\leq n}$. The $\langle ., .\rangle_M$-orthogonality of $\{e, f\}_{1\leq i,j\leq n}$ implies diagonalization of $S^T MS$ with $\Lambda = diag[\lambda_1 ... \lambda_n]$. To prove the uniqueness we just have to show that if there exist $S \in Sp(n)$ such that $S^T LS = L'$ with $L = diag[\Lambda, \Lambda]$ and $L' = diag[\Lambda', \Lambda']$ then $\Lambda = \Lambda'$. Since S is symplectic we have $S^T JS = J$ and hence $S^T LS = L'$ is equivalent to $S^{-1} JLS = JL'$ from which follows JL and JL' have the same eigenvalues. These eigenvalues are precisely the complex numbers $\pm i/\lambda_j$

So far we have seen that any symmetric matrix can be decomposed using symplectic matrices. We also know that the diagonal elements are bigger than zero. All the symplectic transformations that diagonalize a symmetric matrix produce the same diagonal elements up to the ordering. So we can make a convention and always use a decreasing order:

Definition: With the ordering convention:

$$\lambda_{\sigma,1} \geq \lambda_{\sigma,2} \geq ... \geq \lambda_{\sigma,n}$$

$(\lambda_{\sigma,1}, \lambda_{\sigma,2}, ..., \lambda_{\sigma,n})$ is called "symplectic spectrum of M" and is denoted by $Spec_\sigma(M)$.

Symplectic spectrum of a matrix M has various properties and we can prove some propositions about it. In what follows, I will state and prove a proposition together with a theorem about it:

Proposition: *Let $Spec_\sigma(M) = (\lambda_{\sigma,1}, ..., \lambda_{\sigma,n})$ be the symplectic spectrum of M.*

(i)$Spec_\sigma(M)$ is a symplectic invariant:

$$Spec_\sigma(S^T MS) = Spec_\sigma(M) \text{ for every } S \in Sp(n)$$

(ii) the sequence $(\lambda_{\sigma,1}^{-1}, ..., \lambda_{\sigma,n}^{-1})$ is the symplectic spectrum of M^{-1}

$$Spec_\sigma(M^{-1}) = (Spec_\sigma(M))^{-1}$$

Proof: We remember from Williamson's theorem that the diagonal form of a matrix is unique up to reordering. So part (i) is the immediate consequence of the definition of the spectrum of M.

We note that the eigenvalues of JM is the same as the eigenvalues of $M^{\frac{1}{2}} JM^{\frac{1}{2}}$ also the eigenvalues of JM^{-1} is the same as $M^{-\frac{1}{2}} JM^{-\frac{1}{2}}$. We have:

$$M^{-\frac{1}{2}} JM^{-\frac{1}{2}} M^{\frac{1}{2}} JM^{\frac{1}{2}} = M^{-\frac{1}{2}} J^2 M^{\frac{1}{2}} = -M^{-\frac{1}{2}} M^{\frac{1}{2}} = -I$$

So:

$$M^{\frac{1}{2}}JM^{\frac{1}{2}} = -(M^{-\frac{1}{2}}JM^{-\frac{1}{2}})^{-1}$$

So we can we can get the eigenvalues of JM^{-1} from JM by the transformation $t \to -\frac{1}{t}$ and this is what we wanted to prove since the symplectic spectra comes from taking the modulii of these eigenvalues.

Theorem: *Let M and M' be two symmetric positive definite matrices of same dimension. We have:*

$$M \leq M' \Rightarrow Spec_\sigma(M) \leq Spec_\sigma(M')$$

Proof: There are a few point to be reviewed about inequality of matrices before getting into actual proof. First, when two matrices A and B have the same eigenvalues we write $A \simeq B$. When the eigenvalues of A are smaller than or equal to those of B (with a common ordering) we write $A \leq B$. When A or B are invertilble we have $AB \simeq BA$. Keeping all of this in mind, the theorem is equivalent to:

$$M \leq M' \Rightarrow (JM')^2 \leq (JM)^2$$

because $Spec_\sigma(M)$ is the modulii of the diagonal form of JM in which diagonal elements are ordered in decreasing order. The inequality $M \leq M'$ is equivalent to $z^T Mz \leq z^T M'z$ for all $z \in \mathbb{R}^{2n}$. So, we can replace z in $z^T Mz \leq z^T M'z$ by $(JM^{\frac{1}{2}})z$ and $(JM'^{\frac{1}{2}})z$. Noting that $J^T = -J$ we have:

$$((JM^{\frac{1}{2}})z)^T M(JM^{\frac{1}{2}})z \leq ((JM^{\frac{1}{2}})z)^T M'(JM^{\frac{1}{2}})z$$

$$\to ((JM^{\frac{1}{2}}))^T M(JM^{\frac{1}{2}}) \leq ((JM^{\frac{1}{2}}))^T M'(JM^{\frac{1}{2}})$$

$$\to M^{\frac{1}{2}}JM'JM^{\frac{1}{2}} \leq M^{\frac{1}{2}}JMJM^{\frac{1}{2}}$$

Similarly:

$$M'^{\frac{1}{2}}JM'JM'^{\frac{1}{2}} \leq M'^{\frac{1}{2}}JMJM'^{\frac{1}{2}}$$

We note that:

$$M^{\frac{1}{2}}JM'JM^{\frac{1}{2}} \simeq MJM'J$$

$$M'^{\frac{1}{2}}JMJM'^{\frac{1}{2}} \simeq M'JMJ \simeq MJM'J$$

Now we can rewrite the inequalities:

$$MJM'J \leq M^{\frac{1}{2}}JMJM^{\frac{1}{2}}$$

$$M'^{\frac{1}{2}}JM'JM'^{\frac{1}{2}} \leq MJM'J$$

So we have:

$$M'^{\frac{1}{2}}JM'JM'^{\frac{1}{2}} \leq M^{\frac{1}{2}}JMJM^{\frac{1}{2}}$$

Now we have:

$$M'^{\frac{1}{2}}JM'JM'^{\frac{1}{2}} \simeq (M'J)^2$$

$$M^{\frac{1}{2}}JMJM^{\frac{1}{2}} \simeq (MJ)^2$$

Which yield to:

$$(M'J)^2 \leq (MJ)^2$$

Which is what we wanted to prove.

4 Back to the Motivation

Now with the Williamson's theorem being said, we can look back to the problem of simple harmonic oscillator of the beginning of the lecture notes. It is important to note that the Hamiltonian itself is in the form of a diagonal matrix:

$$H(p,q) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2 = \frac{\omega}{2}\begin{bmatrix} p & x \end{bmatrix}\begin{bmatrix} \frac{1}{\gamma} & 0 \\ 0 & \gamma \end{bmatrix}\begin{bmatrix} p \\ x \end{bmatrix} = \langle z, Mz \rangle$$

with $\gamma = \frac{\omega}{2}$, $M = \begin{bmatrix} \frac{1}{\gamma} & 0 \\ 0 & \gamma \end{bmatrix}$ and $z = \begin{bmatrix} p \\ x \end{bmatrix}$

Williamson's theorem tells us that there exist a matrix $S \in Sp(2)$ such that

$$S^T MS = I$$

with $\pm i\lambda$ being the eigenvalues of JM. But we note that:

$$JM = \begin{bmatrix} 0 & -\gamma \\ \frac{1}{\gamma} & 0 \end{bmatrix}$$

Which (up to a proportionality constant) is the same as the matrix that we were trying to diagonalize in the first section. Calculating the eigenvalues of JM we see that they are $\pm i$. It is also easy to show that the following symplectic matrix diagonalizes M with diagonal elements being equal to 1 as we expected from Williamson's theorem:

$$S = \begin{bmatrix} \gamma^{1/2} & 0 \\ 0 & \gamma^{-1/2} \end{bmatrix}$$

If we go on and diagonalize JM (as we wanted to in the first section) we see that since JM is not a normal matrix, we can't diagonalize it using a unitary similarity transformation. But there exist a similarity transformation such that:

$$PJMP^{-1} = (JM)_d$$

One choice for such a P is:

$$P = \begin{bmatrix} \frac{i}{\gamma} & 1 \\ \frac{-i}{\gamma} & 1 \end{bmatrix}$$

So:

$$PJMP^{-1} = \frac{1}{detP} \begin{bmatrix} 2\gamma & 0 \\ 0 & -2\gamma \end{bmatrix}$$

But one interesting fact here is when we apply P on the position and momentum we see:

$$P \begin{bmatrix} p \\ x \end{bmatrix} = \begin{bmatrix} x + \frac{ip}{\gamma} \\ x - \frac{ip}{\gamma} \end{bmatrix}$$

This now coordinate system in which Hamiltonian equation of motion of S.H.O decouples, is (up to a normalization constant) the classical counterparts of ladder operators a and $a^\dagger$. If we try to solve quantum mechanical S.H.O in Heisenberg picture, we can see that equations of motion for these (now time dependent operators) are decoupled.