mathematical methods - week 1

Linear algebra

Georgia Tech PHYS-6124

Homework HW #1 due Tuesday, August 26, 2014

== show all your work for maximum credit,
== put labels, title, legends on any graphs
== acknowledge study group member, if collective effort

Exercise 8.1 *Trace-log of a matrix* 4 points
Exercise 1.2 *Stability, diagonal case* 2 points
Exercise 1.3 *Time-ordered exponentials* 4 points

Bonus points
Exercise 1.4 *Real representation of complex eigenvalues* 4 points

Total of 10 points = 100 % score. Extra points accumulate, can help you later if you miss a few problems.

2014-08-19 Predrag Lecture 1

Grigoriev notes pages 6.2 up to 1) Mechanics inertia tensor, 6.4, 6.5 up to Right and left eigenvectors

Sect. 1.4 *Eigenvalues and eigenvectors*

2014-08-21 Predrag Lecture 2 (with spill over into lecture 3)

Recap from Lecture 1: state Hamilton-Cayley equation, projection operators (1.27), any matrix function is evaluated by spectral decomposition (1.30).
Work through example 1.6

Predrag notes, Grigoriev p. 6.5: Right (columns) and left (rows) eigenvectors

Predrag notes on moment of inertia tensor, (they substitute for Grigoriev p. 6.2 Mechanics, inertia tensor)

Predrag handwritten notes are not on the web, for those streteches you might want to take your own notes in the lecture.

1.1 Literature

> Mopping up operations are the activities that engage most scientists throughout their careers.
>
> — Thomas Kuhn, *The Structure of Scientific Revolutions*

The subject of linear algebra generates innumerable tomes of its own, and is way beyond what we can exhaustively cover. We have added to the course homepage linear operators and matrices reading: Stone and P. Goldbart [1.1], *Mathematics for Physics: A Guided Tour for Graduate Students*, Appendix A. This is an advanced summary where you will find almost everything one needs to know. More pedestrian and perhaps easier to read is Arfken and Weber [1.2] *Mathematical Methods for Physicists: A Comprehensive Guide*, Chapter 3.

Here we recapitulate a few concepts that we found essential. The punch line is Eq. (1.38): Hamilton-Cayley equation $\prod(\mathbf{M} - \lambda_i \mathbf{1}) = 0$ associates with each distinct root λ_i of a matrix $\mathbf{M}$ a projection onto ith vector subspace

$$\mathbf{P}_i = \prod_{j \neq i} \frac{\mathbf{M} - \lambda_j \mathbf{1}}{\lambda_i - \lambda_j} \,.$$

1.2 Matrix-valued functions

We summarize some of the properties of functions of finite-dimensional matrices.

The derivative of a matrix is a matrix with elements

$$\mathbf{A}'(x) = \frac{d\mathbf{A}(x)}{dx} \,, \qquad A'_{ij}(x) = \frac{d}{dx} A_{ij}(x) \,. \tag{1.1}$$

Derivatives of products of matrices are evaluated by the chain rule

$$\frac{d}{dx}(\mathbf{A}\mathbf{B}) = \frac{d\mathbf{A}}{dx}\mathbf{B} + \mathbf{A}\frac{d\mathbf{B}}{dx}\,. \tag{1.2}$$

A matrix and its derivative matrix in general do not commute

$$\frac{d}{dx}\mathbf{A}^2 = \frac{d\mathbf{A}}{dx}\mathbf{A} + \mathbf{A}\frac{d\mathbf{A}}{dx}\,. \tag{1.3}$$

The derivative of the inverse of a matrix, follows from $\frac{d}{dx}(\mathbf{A}\mathbf{A}^{-1}) = 0$:

$$\frac{d}{dx}\mathbf{A}^{-1} = -\frac{1}{\mathbf{A}}\frac{d\mathbf{A}}{dx}\frac{1}{\mathbf{A}}\,. \tag{1.4}$$

A function of a single variable that can be expressed in terms of additions and multiplications generalizes to a matrix-valued function by replacing the variable by the matrix.

In particular, the exponential of a constant matrix can be defined either by its series expansion, or as a limit of an infinite product:

$$e^{\mathbf{A}} = \sum_{k=0}^{\infty}\frac{1}{k!}\mathbf{A}^k\,, \qquad \mathbf{A}^0 = \mathbf{1} \tag{1.5}$$

$$= \lim_{N\to\infty}\left(\mathbf{1} + \frac{1}{N}\mathbf{A}\right)^N \tag{1.6}$$

The first equation follows from the second one by the binomial theorem, so these indeed are equivalent definitions. That the terms of order $O(N^{-2})$ or smaller do not matter follows from the bound

$$\left(1 + \frac{x-\epsilon}{N}\right)^N < \left(1 + \frac{x+\delta x_N}{N}\right)^N < \left(1 + \frac{x+\epsilon}{N}\right)^N\,,$$

where $|\delta x_N| < \epsilon$. If $\lim \delta x_N \to 0$ as $N \to \infty$, the extra terms do not contribute.

Consider now the determinant

$$\det\left(e^{\mathbf{A}}\right) = \lim_{N\to\infty}\left(\det\,(\mathbf{1} + \mathbf{A}/N)\right)^N\,.$$

To the leading order in $1/N$

$$\det\,(\mathbf{1} + \mathbf{A}/N) = 1 + \frac{1}{N}\mathrm{tr}\,\mathbf{A} + O(N^{-2})\,.$$

hence

$$\det e^{\mathbf{A}} = \lim_{N\to\infty}\left(1 + \frac{1}{N}\mathrm{tr}\,\mathbf{A} + O(N^{-2})\right)^N = e^{\mathrm{tr}\,\mathbf{A}} \tag{1.7}$$

Due to non-commutativity of matrices, generalization of a function of several variables to a function is not as straightforward. Expression involving several matrices depend on their commutation relations. For example, the commutator expansion

$$e^{t\mathbf{A}}\mathbf{B}e^{-t\mathbf{A}} = \mathbf{B} + t[\mathbf{A},\mathbf{B}] + \frac{t^2}{2}[\mathbf{A},[\mathbf{A},\mathbf{B}]] + \frac{t^3}{3!}[\mathbf{A},[\mathbf{A},[\mathbf{A},\mathbf{B}]]] + \cdots \tag{1.8}$$

sometimes used to establish the equivalence of the Heisenberg and Schrödinger pictures of quantum mechanics follows by recursive evaluation of t derivatives

$$\frac{d}{dt}\left(e^{t\mathbf{A}}\mathbf{B}e^{-t\mathbf{A}}\right) = e^{t\mathbf{A}}[\mathbf{A},\mathbf{B}]e^{-t\mathbf{A}}\,.$$

Manipulations of such ilk yield

$$e^{(\mathbf{A}+\mathbf{B})/N} = e^{\mathbf{A}/N}e^{\mathbf{B}/N} - \frac{1}{2N^2}[\mathbf{A},\mathbf{B}] + O(N^{-3})\,,$$

and the Trotter product formula: if $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{A} = \mathbf{B}+\mathbf{C}$ are matrices, then

$$e^{\mathbf{A}} = \lim_{N\to\infty}\left(e^{\mathbf{B}/N}e^{\mathbf{C}/N}\right)^N \tag{1.9}$$

1.3 A linear diversion

Linear is good, nonlinear is bad.
—Jean Bellissard

(Notes based of ChaosBook.org/chapters/stability.pdf)

Linear fields are the simplest vector fields, described by linear differential equations which can be solved explicitly, with solutions that are good for all times. The state space for linear differential equations is $\mathcal{M} = \mathbb{R}^d$, and the equations of motion are written in terms of a vector x and a constant stability matrix A as

$$\dot{x} = v(x) = Ax\,. \tag{1.10}$$

Solving this equation means finding the state space trajectory

$$x(t) = (x_1(t), x_2(t), \ldots, x_d(t))$$

passing through a given initial point x_0. If $x(t)$ is a solution with $x(0) = x_0$ and $y(t)$ another solution with $y(0) = y_0$, then the linear combination $ax(t) + by(t)$ with $a, b \in \mathbb{R}$ is also a solution, but now starting at the point $ax_0 + by_0$. At any instant in time, the space of solutions is a d-dimensional vector space, spanned by a basis of d linearly independent solutions.

How do we solve the linear differential equation (1.10)? If instead of a matrix equation we have a scalar one, $\dot{x} = \lambda x$, the solution is $x(t) = e^{t\lambda}x_0$. In order to solve the d-dimensional matrix case, it is helpful to rederive this solution by studying what happens for a short time step δt. If time $t = 0$ coincides with position $x(0)$, then

$$\frac{x(\delta t) - x(0)}{\delta t} = \lambda x(0)\,, \tag{1.11}$$

which we iterate m times to obtain Euler's formula for compounding interest

$$x(t) \approx \left(1 + \frac{t}{m}\lambda\right)^m x(0) \approx e^{t\lambda}x(0)\,. \tag{1.12}$$

The term in parentheses acts on the initial condition $x(0)$ and evolves it to $x(t)$ by taking m small time steps $\delta t = t/m$. As $m \to \infty$, the term in parentheses converges to $e^{t\lambda}$. Consider now the matrix version of equation (1.11):

$$\frac{x(\delta t) - x(0)}{\delta t} = Ax(0)\,. \tag{1.13}$$

A representative point x is now a vector in $\mathbb{R}^d$ acted on by the matrix A, as in (1.10). Denoting by $\mathbf{1}$ the identity matrix, and repeating the steps (1.11) and (1.12) we obtain Euler's formula for the exponential of a matrix:

$$x(t) = J^t x(0)\,, \qquad J^t = e^{tA} = \lim_{m\to\infty}\left(\mathbf{1} + \frac{t}{m}A\right)^m\,. \tag{1.14}$$

We will find this definition for the exponential of a matrix helpful in the general case, where the matrix $A = A(x(t))$ varies along a trajectory.

Now that we have some feeling for the qualitative behavior of eigenvectors and eigenvalues of linear flows, we are ready to return to the nonlinear case. How do we compute the exponential (1.14)?

$$x(t) = f^t(x_0)\,, \quad \delta x(x_0, t) \;=\; J^t(x_0)\,\delta x(x_0, 0)\,. \tag{1.15}$$

The equations are linear, so we should be able to integrate them–but in order to make sense of the answer, we derive this integral step by step. The Jacobian matrix is computed by integrating the equations of variations

$$\dot{x}_i = v_i(x)\,, \quad \dot{\delta x}_i = \sum_j A_{ij}(x)\delta x_j \tag{1.16}$$

Consider the case of a general, non-stationary trajectory $x(t)$. The exponential of a constant matrix can be defined either by its Taylor series expansion or in terms of the Euler limit (1.14):

$$e^{tA} \;=\; \sum_{k=0}^{\infty}\frac{t^k}{k!}A^k \;=\; \lim_{m\to\infty}\left(\mathbf{1} + \frac{t}{m}A\right)^m\,. \tag{1.17}$$

Taylor expanding is fine if A is a constant matrix. However, only the second, tax-accountant's discrete step definition of an exponential is appropriate for the task at hand. For dynamical systems, the local rate of neighborhood distortion $A(x)$ depends on where we are along the trajectory. The linearized neighborhood is deformed along the flow, and the m discrete time-step approximation to J^t is therefore given by a generalization of the Euler product (1.17):

$$\begin{aligned} J^t &= \lim_{m\to\infty}\prod_{n=m}^{1}\left(\mathbf{1} + \delta t A(x_n)\right) = \lim_{m\to\infty}\prod_{n=m}^{1} e^{\delta t\, A(x_n)} \\ &= \lim_{m\to\infty} e^{\delta t\, A(x_m)} e^{\delta t\, A(x_{m-1})}\cdots e^{\delta t\, A(x_2)} e^{\delta t\, A(x_1)}\,, \end{aligned} \tag{1.18}$$

where $\delta t = (t - t_0)/m$, and $x_n = x(t_0 + n\delta t)$. Indexing of the products indicates that the successive infinitesimal deformation are applied by multiplying from the left. The $m \to \infty$ limit of this procedure is the formal integral

$$J^t_{ij}(x_0) = \left[\mathbf{T} e^{\int_0^t d\tau A(x(\tau))}\right]_{ij} , \qquad (1.19)$$

where $\mathbf{T}$ stands for time-ordered integration, *defined* as the continuum limit of the successive multiplications (1.18). This integral formula for J^t is the finite time companion of the differential definition. The definition makes evident important properties of Jacobian matrices, such as their being multiplicative along the flow,

exercise 1.3

$$J^{t+t'}(x) = J^{t'}(x')\, J^t(x), \qquad \text{where } x' = f^t(x_0) , \qquad (1.20)$$

which is an immediate consequence of the time-ordered product structure of (1.18). However, in practice J is evaluated by integrating differential equation along with the ODEs that define a particular flow.

1.4 Eigenvalues and eigenvectors

10. Try to leave out the part that readers tend to skip.
— Elmore Leonard's Ten Rules of Writing.

Eigenvalues of a $[d \times d]$ matrix $\mathbf{M}$ are the roots of its characteristic polynomial

$$\det(\mathbf{M} - \lambda \mathbf{1}) = \prod (\lambda_i - \lambda) = 0 . \qquad (1.21)$$

Given a nonsingular matrix $\mathbf{M}$, with all $\lambda_i \neq 0$, acting on d-dimensional vectors $\mathbf{x}$, we would like to determine *eigenvectors* $\mathbf{e}^{(i)}$ of $\mathbf{M}$ on which $\mathbf{M}$ acts by scalar multiplication by eigenvalue λ_i

$$\mathbf{M}\, \mathbf{e}^{(i)} = \lambda_i \mathbf{e}^{(i)} . \qquad (1.22)$$

If $\lambda_i \neq \lambda_j$, $\mathbf{e}^{(i)}$ and $\mathbf{e}^{(j)}$ are linearly independent. There are at most d distinct eigenvalues and eigenspaces, which we assume have been computed by some method, and ordered by their real parts, $\mathrm{Re}\, \lambda_i \geq \mathrm{Re}\, \lambda_{i+1}$.

If all eigenvalues are distinct $\mathbf{e}^{(j)}$ are d linearly independent vectors which can be used as a (non-orthogonal) basis for any d-dimensional vector $\mathbf{x} \in \mathbb{R}^d$

$$\mathbf{x} = x_1\, \mathbf{e}^{(1)} + x_2\, \mathbf{e}^{(2)} + \cdots + x_d\, \mathbf{e}^{(d)} . \qquad (1.23)$$

From (1.22) it follows that

$$(\mathbf{M} - \lambda_i \mathbf{1})\, \mathbf{e}^{(j)} = (\lambda_j - \lambda_i)\, \mathbf{e}^{(j)} ,$$

matrix $(\mathbf{M}-\lambda_i\mathbf{1})$ annihilates $\mathbf{e}^{(i)}$, the product of all such factors annihilates any vector, and the matrix $\mathbf{M}$ satisfies its characteristic equation

$$\prod_{i=1}^{d}(\mathbf{M}-\lambda_i\mathbf{1}) = 0\,. \tag{1.24}$$

This humble fact has a name: the Hamilton-Cayley theorem. If we delete one term from this product, we find that the remainder projects $\mathbf{x}$ from (1.23) onto the corresponding eigenspace:

$$\prod_{j\neq i}(\mathbf{M}-\lambda_j\mathbf{1})\mathbf{x} = \prod_{j\neq i}(\lambda_i-\lambda_j)x_i\mathbf{e}^{(i)}\,.$$

Dividing through by the $(\lambda_i-\lambda_j)$ factors yields the *projection operators*

$$\mathbf{P}_i = \prod_{j\neq i}\frac{\mathbf{M}-\lambda_j\mathbf{1}}{\lambda_i-\lambda_j}\,, \tag{1.25}$$

which are *orthogonal* and *complete*:

$$\mathbf{P}_i\mathbf{P}_j = \delta_{ij}\mathbf{P}_j\,, \quad (\text{no sum on } j)\,, \qquad \sum_{i=1}^{r}\mathbf{P}_i = \mathbf{1}\,, \tag{1.26}$$

with the dimension of the ith subspace given by $d_i = \operatorname{tr}\mathbf{P}_i$. For each distinct eigenvalue λ_i of $\mathbf{M}$,

$$(\mathbf{M}-\lambda_j\mathbf{1})\mathbf{P}_j = \mathbf{P}_j(\mathbf{M}-\lambda_j\mathbf{1}) = 0\,, \tag{1.27}$$

the colums/rows of $\mathbf{P}_i$ are the right/left eigenvectors $\mathbf{e}^{(k)}$, $\mathbf{e}_{(k)}$ of $\mathbf{M}$ which (provided $\mathbf{M}$ is not of Jordan type, see example 1.4) span the corresponding linearized subspace, and are a convenient starting seed for tracing out the global unstable/stable manifolds. Once the distinct non-zero eigenvalues $\{\lambda_i\}$ are computed, projection operators are polynomials in $\mathbf{M}$ which need no further diagonalizations or orthogonalizations. Economical description of neighborhoods of equilibria and periodic orbits is afforded by projection operators (1.25), where matrix $\mathbf{M}$ is typically either equilibrium stability matrix A, or periodic orbit Jacobian matrix $\hat{J}$.

It follows from the characteristic equation (1.27) that λ_i is the eigenvalue of $\mathbf{M}$ on $\mathbf{P}_i$ subspace:

$$\mathbf{M}\,\mathbf{P}_i = \lambda_i\mathbf{P}_i \qquad (\text{no sum on } i)\,. \tag{1.28}$$

Using $\mathbf{M} = \mathbf{M}\,\mathbf{1}$ and completeness relation (1.26) we can rewrite $\mathbf{M}$ as

$$\mathbf{M} = \lambda_1\mathbf{P}_1 + \lambda_2\mathbf{P}_2 + \cdots + \lambda_d\mathbf{P}_d\,. \tag{1.29}$$

Any matrix function $f(\mathbf{M})$ takes the scalar value $f(\lambda_i)$ on the $\mathbf{P}_i$ subspace, $f(\mathbf{M})\,\mathbf{P}_i = f(\lambda_i)\,\mathbf{P}_i$, and is thus easily evaluated through its *spectral decomposition*

$$f(\mathbf{M}) = \sum_i f(\lambda_i)\mathbf{P}_i\,. \tag{1.30}$$

This, of course, is the reason why anyone but a fool works with irreducible reps: they reduce matrix (AKA "operator") evaluations to manipulations with numbers.

By (1.27) every column of $\mathbf{P}_i$ is proportional to a right eigenvector $\mathbf{e}^{(i)}$, and its every row to a left eigenvector $\mathbf{e}_{(i)}$. In general, neither set is orthogonal, but by the idempotence condition (1.26), they are mutually orthogonal,

$$\mathbf{e}_{(i)} \cdot \mathbf{e}^{(j)} = c\,\delta_i^j\,. \tag{1.31}$$

The non-zero constant c is convention dependent and not worth fixing, unless you feel nostalgic about Clebsch-Gordan coefficients. We shall set $c = 1$. Then it is convenient to collect all left and right eigenvectors into a single matrix as follows.

Example 1.1. ***Fundamental matrix.*** *If A is constant in time, the system (1.16) is autonomous, and the solution is*

$$x(t) = e^{A\,t}x(0)\,,$$

where $\exp(A\,t)$ is defined by the Taylor series for $\exp(x)$. As the system is linear, the sum of any two solutions is also a solution. Therefore, given d independent initial conditions, $x_1(0), x_2(0), \ldots x_d(0)$ we can write the solution for an arbitrary initial condition based on its projection on to this set,

$$x(t) = \mathbf{F}(t)\,\mathbf{F}(0)^{-1}x(0) = e^{At}x(0)\,,$$

where $\mathbf{F}(t) = (x_1(t), x_2(t), \ldots, x_d(t))$ is a fundamental matrix of the system.

Fundamental matrix (take 1). *As the system is a linear, a superposition of any two solutions to $x(t) = J^t x(0)$ is also a solution. One can take any d independent initial states, $x^{(1)}(0)$, $x^{(2)}(0)$, $\ldots, x^{(d)}(0)$, assemble them as columns of a matrix $\Phi(0)$, and formally write the solution for an arbitrary initial condition projected onto this basis,*

$$x(t) = \Phi(t)\Phi(0)^{-1}x(0) \tag{1.32}$$

where $\Phi(t) = [x^{(1)}(t), x^{(2)}(t), \cdots, x^{(d)}(t)]$. $\Phi(t)$ is called the fundamental matrix *of the system, and the Jacobian matrix $J^t = \Phi(t)\Phi(0)^{-1}$ can thus be fashioned out of d trajectories $\{x^{(j)}(t)\}$. Numerically this works for sufficiently short times.*

Fundamental matrix (take 2). *The set of solutions $x(t) = J^t(x_0)x_0$ for a system of homogeneous linear differential equations $\dot{x}(t) = A(t)x(t)$ of order 1 and dimension d forms a d-dimensional vector space. A basis $\{\mathbf{e}^{(1)}(t), \ldots, \mathbf{e}^{(d)}(t)\}$ for this vector space is called a* fundamental system. *Every solution $x(t)$ can be written as*

$$x(t) = \sum_{i=1}^{d} c_i\,\mathbf{e}^{(i)}(t)\,.$$

The $[d\times d]$ matrix $\mathbf{F}_{ij}^{-1} = \mathbf{e}_i^{(j)}$ whose columns are the right eigenvectors of J^t

$$\mathbf{F}(t)^{-1} = (\mathbf{e}^{(1)}(t), \ldots, \mathbf{e}^{(d)}(t))\,, \qquad \mathbf{F}(t)^T = (\mathbf{e}_{(1)}(t), \ldots, \mathbf{e}_{(d)}(t)) \tag{1.33}$$

is the inverse of a fundamental matrix.

Jacobian matrix. *The Jacobian matrix $J^t(x_0)$ is the linear approximation to a differentiable function $f^t(x_0)$, describing the orientation of a tangent plane to the function at a given point and the amount of local rotation and shearing caused by the transformation. The inverse of the Jacobian matrix of a function is the Jacobian matrix of the inverse function. If f is a map from d-dimensional space to itself, the Jacobian matrix is a square matrix, whose determinant we refer to as the 'Jacobian.'*

The Jacobian matrix can be written as transformation from basis at time t_0 to the basis at time t_1,

$$J^{t_1-t_0}(x_0) = \mathbf{F}(t_1)\mathbf{F}(t_0)^{-1} . \tag{1.34}$$

Then the matrix form of (1.31) is $\mathbf{F}(t)\mathbf{F}(t)^{-1} = \mathbf{1}$, i.e., for zero time the Jacobian matrix is the identity. *(J. Halcrow)*

exercise 1.4

Example 1.2. ***Linear stability of 2-dimensional flows:*** *For a 2-dimensional flow the eigenvalues λ_1, λ_2 of A are either real, leading to a linear motion along their eigenvectors, $x_j(t) = x_j(0)\exp(t\lambda_j)$, or form a complex conjugate pair $\lambda_1 = \mu + i\omega\,, \lambda_2 = \mu - i\omega\,$, leading to a circular or spiral motion in the $[x_1, x_2]$ plane, see example 1.3.*

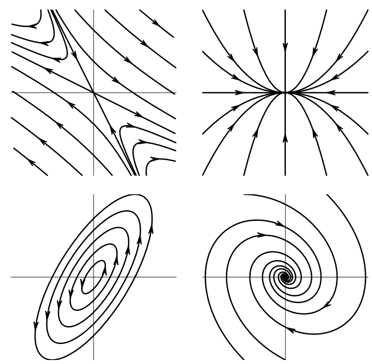

Figure 1.1: Streamlines for several typical 2-dimensional flows: saddle (hyperbolic), in node (attracting), center (elliptic), in spiral.

These two possibilities are refined further into sub-cases depending on the signs of the real part. In the case of real $\lambda_1 > 0$, $\lambda_2 < 0$, x_1 grows exponentially with time, and x_2 contracts exponentially. This behavior, called a saddle, *is sketched in figure 1.1, as are the remaining possibilities: in/out nodes, inward/outward spirals, and the center. The magnitude of out-spiral $|x(t)|$ diverges exponentially when $\mu > 0$, and in-spiral contracts into $(0,0)$ when $\mu < 0$; whereas, the phase velocity ω controls its oscillations.*

If eigenvalues $\lambda_1 = \lambda_2 = \lambda$ are degenerate, the matrix might have two linearly independent eigenvectors, or only one eigenvector, see example 1.4. We distinguish two cases: (a) A can be brought to diagonal form and (b) A can be brought to Jordan form, *which (in dimension 2 or higher) has zeros everywhere except for the repeating eigenvalues on the diagonal and some 1's directly above it. For every such Jordan $[d_\alpha \times d_\alpha]$ block there is only one eigenvector per block.*

We sketch the full set of possibilities in figures 1.1 and 1.2.

Example 1.3. ***Complex eigenvalues: in-out spirals.*** *As $\mathbf{M}$ has only real entries, it will in general have either real eigenvalues, or complex conjugate pairs of eigenvalues. Also the corresponding eigenvectors can be either real or complex. All coordinates used in defining a dynamical flow are real numbers, so what is the meaning of a* complex *eigenvector?*

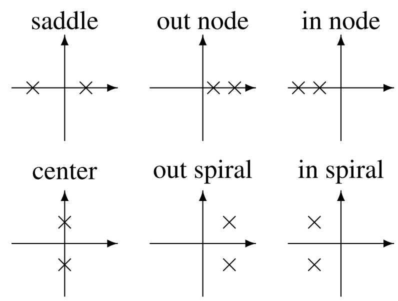

Figure 1.2: Qualitatively distinct types of exponents $\{\lambda_1, \lambda_2\}$ of a $[2\times 2]$ Jacobian matrix.

If λ_k, λ_{k+1} eigenvalues that lie within a diagonal $[2\times 2]$ sub-block $\mathbf{M}' \subset \mathbf{M}$ form a complex conjugate pair, $\{\lambda_k, \lambda_{k+1}\} = \{\mu + i\omega, \mu - i\omega\}$, the corresponding complex eigenvectors can be replaced by their real and imaginary parts, $\{\mathbf{e}^{(k)}, \mathbf{e}^{(k+1)}\} \to \{\mathrm{Re}\,\mathbf{e}^{(k)}, \mathrm{Im}\,\mathbf{e}^{(k)}\}$. In this 2-dimensional real representation, $\mathbf{M}' \to A$, the block A is a sum of the rescaling$\times$identity and the generator of $SO(2)$ rotations in the $\{\mathrm{Re}\,\mathbf{e}^{(1)}, \mathrm{Im}\,\mathbf{e}^{(1)}\}$ plane.

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

Trajectories of $\dot{\mathbf{x}} = A\,\mathbf{x}$, given by $\mathbf{x}(t) = J^t\,\mathbf{x}(0)$, where (omitting $\mathbf{e}^{(3)}, \mathbf{e}^{(4)}, \cdots$ eigen-directions)

$$J^t = e^{tA} = e^{t\mu} \begin{pmatrix} \cos\omega t & -\sin\omega t \\ \sin\omega t & \cos\omega t \end{pmatrix} , \tag{1.35}$$

spiral in/out around $(x, y) = (0, 0)$, see figure 1.1, with the rotation period T and the radial expansion /contraction multiplier along the $\mathbf{e}^{(j)}$ eigen-direction per a turn of the spiral:

exercise 1.4

$$T = 2\pi/\omega\,, \qquad \Lambda_{radial} = e^{T\mu}\,. \tag{1.36}$$

We learn that the typical turnover time scale in the neighborhood of the equilibrium $(x, y) = (0, 0)$ is of order $\approx T$ (and not, let us say, $1000\,T$, or $10^{-2}T$).

Example 1.4. ***Degenerate eigenvalues.*** *While for a matrix with generic real elements all eigenvalues are distinct with probability 1, that is not true in presence of symmetries, or spacial parameter values (bifurcation points). What can one say about situation where d_α eigenvalues are degenerate, $\lambda_\alpha = \lambda_i = \lambda_{i+1} = \cdots = \lambda_{i+d_\alpha-1}$? Hamilton-Cayley (1.24) now takes form*

$$\prod_{\alpha=1}^{r} (\mathbf{M} - \lambda_\alpha \mathbf{1})^{d_\alpha} = 0\,, \qquad \sum_\alpha d_\alpha = d\,. \tag{1.37}$$

We distinguish two cases:

M can be brought to diagonal form. *The characteristic equation (1.37) can be replaced by the minimal polynomial,*

$$\prod_{\alpha=1}^{r} (\mathbf{M} - \lambda_\alpha \mathbf{1}) = 0\,, \tag{1.38}$$

where the product includes each distinct eigenvalue only once. Matrix $\mathbf{M}$ *acts multiplicatively*

$$\mathbf{M}\,\mathbf{e}^{(\alpha,k)} = \lambda_i \mathbf{e}^{(\alpha,k)}\,, \tag{1.39}$$

on a d_α*-dimensional subspace spanned by a linearly independent set of basis eigenvectors* $\{\mathbf{e}^{(\alpha,1)}, \mathbf{e}^{(\alpha,2)}, \cdots, \mathbf{e}^{(\alpha,d_\alpha)}\}$*. This is the easy case. Luckily, if the degeneracy is due to a finite or compact symmetry group, relevant* $\mathbf{M}$ *matrices can always be brought to such Hermitian, diagonalizable form.*

$\mathbf{M}$ can only be brought to upper-triangular, Jordan form. *This is the messy case, so we only illustrate the key idea in example 1.5.*

Example 1.5. ***Decomposition of 2-dimensional vector spaces:*** *Enumeration of every possible kind of linear algebra eigenvalue / eigenvector combination is beyond what we can reasonably undertake here. However, enumerating solutions for the simplest case, a general* $[2\times2]$ *non-singular matrix*

$$\mathbf{M} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}.$$

takes us a long way toward developing intuition about arbitrary finite-dimensional matrices. The eigenvalues

$$\lambda_{1,2} = \frac{1}{2}\mathrm{tr}\,\mathbf{M} \pm \frac{1}{2}\sqrt{(\mathrm{tr}\,\mathbf{M})^2 - 4\det\mathbf{M}} \tag{1.40}$$

are the roots of the characteristic (secular) equation (1.21):

$$\begin{aligned}\det(\mathbf{M}-\lambda\,\mathbf{1}) &= (\lambda_1-\lambda)(\lambda_2-\lambda) \\ &= \lambda^2 - \mathrm{tr}\,\mathbf{M}\,\lambda + \det\mathbf{M} \;=\; 0\,.\end{aligned}$$

Distinct eigenvalues case *has already been described in full generality. The left/right eigenvectors are the rows/columns of projection operators*

$$P_1 = \frac{\mathbf{M}-\lambda_2\mathbf{1}}{\lambda_1-\lambda_2}\,, \qquad P_2 = \frac{\mathbf{M}-\lambda_1\mathbf{1}}{\lambda_2-\lambda_1}\,, \qquad \lambda_1 \neq \lambda_2\,. \tag{1.41}$$

Degenerate eigenvalues. *If* $\lambda_1 = \lambda_2 = \lambda$*, we distinguish two cases: (a)* $\mathbf{M}$ *can be brought to diagonal form. This is the easy case. (b)* $\mathbf{M}$ *can be brought to* Jordan form*, with zeros everywhere except for the diagonal, and some 1's directly above it; for a* $[2\times2]$ *matrix the Jordan form is*

$$\mathbf{M} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \qquad \mathbf{e}^{(1)} = \begin{pmatrix} 1 & 0 \end{pmatrix}, \;\; \mathbf{v}^{(2)} = \begin{pmatrix} 0 & 1 \end{pmatrix}.$$

$\mathbf{v}^{(2)}$ *helps span the 2-dimensional space,* $(\mathbf{M}-\lambda)^2\mathbf{v}^{(2)} = 0$*, but is not an eigenvector, as* $\mathbf{M}\mathbf{v}^{(2)} = \lambda\mathbf{v}^{(2)} + \mathbf{e}^{(1)}$*. For every such Jordan* $[d_\alpha \times d_\alpha]$ *block there is only one eigenvector per block. Noting that*

$$\mathbf{M}^m = \begin{pmatrix} \lambda^m & m\lambda^{m-1} \\ 0 & \lambda^m \end{pmatrix},$$

we see that instead of acting multiplicatively on $\mathbb{R}^2$, Jacobian matrix $J^t = \exp(t\mathbf{M})$

$$e^{t\mathbf{M}} \begin{pmatrix} u \\ v \end{pmatrix} = e^{t\lambda} \begin{pmatrix} u + tv \\ v \end{pmatrix} \tag{1.42}$$

picks up a power-low correction. That spells trouble (logarithmic term $\ln t$ if we bring the extra term into the exponent).

Example 1.6. ***Projection operator decomposition in 2 dimensions:*** *Let's illustrate how the distinct eigenvalues case works with the $[2\times 2]$ matrix*

$$\mathbf{M} = \left(\begin{array}{cc} 4 & 1 \\ 3 & 2 \end{array} \right).$$

Its eigenvalues $\{\lambda_1, \lambda_2\} = \{5, 1\}$ are the roots of (1.40):

$$\det(\mathbf{M} - \lambda \mathbf{1}) = \lambda^2 - 6\,\lambda + 5 = (5 - \lambda)(1 - \lambda) = 0\,.$$

That $\mathbf{M}$ satisfies its secular equation (Hamilton-Cayley theorem) can be verified by explicit calculation:

$$\left(\begin{array}{cc} 4 & 1 \\ 3 & 2 \end{array} \right)^2 - 6 \left(\begin{array}{cc} 4 & 1 \\ 3 & 2 \end{array} \right) + 5 \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) = \left(\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right).$$

Associated with each root λ_i is the projection operator *(1.41)*

$$P_1 \;=\; \frac{1}{4}(\mathbf{M} - \mathbf{1}) \;=\; \frac{1}{4} \left(\begin{array}{cc} 3 & 1 \\ 3 & 1 \end{array} \right) \tag{1.43}$$

$$P_2 \;=\; \frac{1}{4}(\mathbf{M} - 5 \cdot \mathbf{1}) \;=\; \frac{1}{4} \left(\begin{array}{cc} 1 & -1 \\ -3 & 3 \end{array} \right). \tag{1.44}$$

Matrices $\mathbf{P}_i$ are orthonormal and complete, The dimension of the ith subspace is given by $d_i = \operatorname{tr} \mathbf{P}_i$; in case at hand both subspaces are 1-dimensional. From the characteristic equation it follows that $\mathbf{P}_i$ satisfies the eigenvalue equation $\mathbf{M}\,\mathbf{P}_i = \lambda_i \mathbf{P}_i$. Two consequences are immediate. First, we can easily evaluate any function of $\mathbf{M}$ by spectral decomposition, for example

$$\mathbf{M}^7 - 3 \cdot \mathbf{1} = (5^7 - 3)\mathbf{P}_1 + (1 - 3)\mathbf{P}_2 = \left(\begin{array}{cc} 58591 & 19531 \\ 58593 & 19529 \end{array} \right).$$

Second, as $\mathbf{P}_i$ satisfies the eigenvalue equation, its every column is a right eigenvector, and every row a left eigenvector. Picking first row/column we get the eigenvectors:

$$\begin{aligned} \{\mathbf{e}^{(1)}, \mathbf{e}^{(2)}\} &= \{\left(\begin{array}{cc} 1 & 1 \end{array} \right), \left(\begin{array}{cc} 1 & -3 \end{array} \right)\} \\ \{\mathbf{e}_{(1)}, \mathbf{e}_{(2)}\} &= \{\left(\begin{array}{cc} 3 & 1 \end{array} \right), \left(\begin{array}{cc} 1 & -1 \end{array} \right)\}, \end{aligned}$$

with overall scale arbitrary. The matrix is not symmetric, so $\{\mathbf{e}^{(j)}\}$ do not form an orthogonal basis. The left-right eigenvector dot products $\mathbf{e}_{(j)} \cdot \mathbf{e}^{(k)}$, however, are orthogonal as in (1.31), by inspection. (Continued in example 1.8.)

Example 1.7. ***Computing matrix exponentials.*** *If A is diagonal (the system is uncoupled), then e^{tA} is given by*

$$\exp\begin{pmatrix} \lambda_1 t & & & \\ & \lambda_2 t & & \\ & & \ddots & \\ & & & \lambda_d t \end{pmatrix} = \begin{pmatrix} e^{\lambda_1 t} & & & \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ & & & e^{\lambda_d t} \end{pmatrix}.$$

If A is diagonalizable, $A = FDF^{-1}$, where D is the diagonal matrix of the eigenvalues of A and F is the matrix of corresponding eigenvectors, the result is simple: $A^n = (FDF^{-1})(FDF^{-1})\dots(FDF^{-1}) = FD^nF^{-1}$. Inserting this into the Taylor series for e^x gives $e^{At} = Fe^{Dt}F^{-1}$.

But A may not have d linearly independant eigenvectors, making F singular and forcing us to take a different route. To illustrate this, consider $[2\times2]$ matrices. For any linear system in $\mathbb{R}^2$, there is a similarity transformation

$$B = U^{-1}AU\,,$$

where the columns of U consist of the generalized eigenvectors of A such that B has one of the following forms:

$$B = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}, \qquad B = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \qquad B = \begin{pmatrix} \mu & -\omega \\ \omega & \mu \end{pmatrix}.$$

These three cases, called normal forms, correspond to A having (1) distinct real eigenvalues, (2) degenerate real eigenvalues, or (3) a complex pair of eigenvalues. It follows that

$$e^{Bt} = \begin{pmatrix} e^{\lambda t} & 0 \\ 0 & e^{\mu t} \end{pmatrix}, \qquad e^{Bt} = e^{\lambda t}\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \qquad e^{Bt} = e^{at}\begin{pmatrix} \cos bt & -\sin bt \\ \sin bt & \cos bt \end{pmatrix},$$

and $e^{At} = Ue^{Bt}U^{-1}$. What we have done is classify all $[2\times2]$ matrices as belonging to one of three classes of geometrical transformations. The first case is scaling, the second is a shear, and the third is a combination of rotation and scaling. The generalization of these normal forms to $\mathbb{R}^d$ is called the Jordan normal form. *(J. Halcrow)*

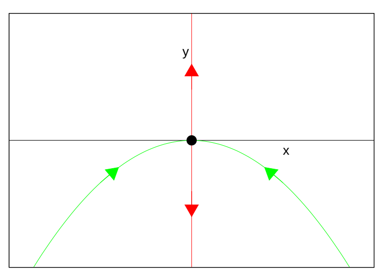

Figure 1.3: The stable/unstable manifolds of the equilibrium $(x_q, x_q) = (0, 0)$ of 2-dimensional flow (1.45).

Example 1.8. ***A simple stable/unstable manifolds pair:*** *Consider the 2-dimensional ODE system*

$$\frac{dx}{dt} = -x, \quad \frac{dy}{dt} = y + x^2\,, \tag{1.45}$$

The flow through a point $x(0) = x_0, y(0) = y_0$ *can be integrated*

$$x(t) = x_0\, e^{-t}, \quad y(t) = (y_0 + x_0^2/3)\, e^{t} - x_0^2\, e^{-2t}/3\,. \tag{1.46}$$

Linear stability of the flow is described by the stability matrix

$$\mathbf{A} = \begin{pmatrix} -1 & 0 \\ 2x & 1 \end{pmatrix}. \tag{1.47}$$

The flow is hyperbolic, with a real expanding/contracting eigenvalue pair $\lambda_1 = 1$*,* $\lambda_2 = -1$*, and area preserving. The right eigenvectors at the point* (x, y)

$$\mathbf{e}^{(1)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \mathbf{e}^{(2)} = \begin{pmatrix} 1 \\ -x \end{pmatrix}. \tag{1.48}$$

can be obtained by acting with the projection operators (see example 1.5 Decomposition of 2-dimensional vector spaces*)*

$$\mathbf{P}_i = \frac{\mathbf{A} - \lambda_j \mathbf{1}}{\lambda_i - \lambda_j} \; : \qquad \mathbf{P}_1 = \begin{pmatrix} 0 & 0 \\ x & 1 \end{pmatrix}, \quad \mathbf{P}_2 = \begin{pmatrix} 1 & 0 \\ -x & 0 \end{pmatrix} \tag{1.49}$$

on an arbitrary vector. Matrices $\mathbf{P}_i$ *are orthonormal and complete.*

The flow has a degenerate pair of equilibria at $(x_q, y_q) = (0, 0)$*, with eigenvalues (stability exponents),* $\lambda_1 = 1$*,* $\lambda_2 = -1$*, eigenvectors* $\mathbf{e}^{(1)} = (0, 1)$*,* $\mathbf{e}^{(2)} = (1, 0)$*. The unstable manifold is the* y *axis, and the stable manifold is given by (see figure 1.3)*

$$y_0 + \frac{1}{3}x_0^2 = 0 \Rightarrow y(t) + \frac{1}{3}x(t)^2 = 0\,. \tag{1.50}$$

(N. Lebovitz)

1.4.1 Yes, but how do you really do it?

As $\mathbf{M}$ has only real entries, it will in general have either real eigenvalues (over-damped oscillator, for example), or complex conjugate pairs of eigenvalues (under-damped oscillator, for example). That is not surprising, but also the corresponding eigenvectors can be either real or complex. All coordinates used in defining the flow are real numbers, so what is the meaning of a *complex* eigenvector?

If two eigenvalues form a complex conjugate pair, $\{\lambda_k, \lambda_{k+1}\} = \{\mu + i\omega, \mu - i\omega\}$, they are in a sense degenerate: while a real λ_k characterizes a motion along a line, a complex λ_k characterizes a spiralling motion in a plane. We determine this plane by replacing the corresponding complex eigenvectors by their real and imaginary parts, $\{\mathbf{e}^{(k)}, \mathbf{e}^{(k+1)}\} \rightarrow \{\mathrm{Re}\,\mathbf{e}^{(k)}, \mathrm{Im}\,\mathbf{e}^{(k)}\}$, or, in terms of projection operators:

$$\mathbf{P}_k = \frac{1}{2}(\mathbf{R} + i\mathbf{Q})\,, \qquad \mathbf{P}_{k+1} = \mathbf{P}_k^*\,,$$

where $\mathbf{R} = \mathbf{P}_k + \mathbf{P}_{k+1}$ is the subspace decomposed by the kth complex eigenvalue pair, and $\mathbf{Q} = (\mathbf{P}_k - \mathbf{P}_{k+1})/i$, both matrices with real elements. Substitution

$$\begin{pmatrix} \mathbf{P}_k & \mathbf{P}_{k+1} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}\begin{pmatrix} \mathbf{R} & \mathbf{Q} \end{pmatrix},$$

brings the $\lambda_k \mathbf{P}_k + \lambda_{k+1}\mathbf{P}_{k+1}$ complex eigenvalue pair in the spectral decomposition into the real form,

$$\left(\mathbf{P}_k \quad \mathbf{P}_{k+1} \right) \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^* \end{pmatrix} \left(\mathbf{P}_k \quad \mathbf{P}_{k+1} \right) = \left(\mathbf{R} \quad \mathbf{Q} \right) \begin{pmatrix} \mu & -\omega \\ \omega & \mu \end{pmatrix} \left(\mathbf{R} \quad \mathbf{Q} \right) , \tag{1.51}$$

where we have dropped the superscript $^{(k)}$ for notational brevity.

To summarize, spectrally decomposed matrix $\mathbf{M}$ acts along lines on subspaces corresponding to real eigenvalues, and as a $[2\times 2]$ rotation in a plane on subspaces corresponding to complex eigenvalue pairs.

Commentary

Remark 1.1. Projection operators. The construction of projection operators given in appendix 1.4.1 is taken from refs. [1.3, 1.4]. Who wrote this down first we do not know, lineage certainly goes all the way back to Lagrange polynomials [1.5], but projection operators tend to get drowned in sea of algebraic details. Arfken and Weber [1.2] ascribe spectral decomposition (1.30) to Sylvester. Halmos [1.6] is a good early reference - but we like Harter's exposition [1.7, 1.8, 1.9] best, for its multitude of specific examples and physical illustrations.

References

[1.1] M. Stone and P. Goldbart, *Mathematics for Physics: A Guided Tour for Graduate Students* (Cambridge Univ. Press, Cambridge, 2009).

[1.2] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists: A Comprehensive Guide*, 6 ed. (Academic Press, New York, 2005).

[1.3] P. Cvitanović, Group theory for Feynman diagrams in non-Abelian gauge theories, Phys. Rev. D **14**, 1536 (1976).

[1.4] P. Cvitanović, Classical and exceptional Lie algebras as invariance algebras, 1977, www.nbi.dk/GroupTheory, (Oxford University preprint 40/77, unpublished).

[1.5] K. Hoffman and R. Kunze, *Linear Algebra*, Second ed. (Prentice-Hall, Englewood Cliffs, NJ, 1971).

[1.6] P. R. Halmos, *Finite-Dimensional Vector Spaces*, Second ed. (D. Van Nostrand, Princeton, 1958).

[1.7] W. G. Harter, Algebraic theory of ray representations of finite groups, J. Math. Phys. **10**, 739 (1969).

[1.8] W. G. Harter and N. Dos Santos, Double-group theory on the half-shell and the two-level system. I. Rotation and half-integral spin states, Am. J. Phys. **46**, 251 (1978).

[1.9] W. G. Harter, *Principles of Symmetry, Dynamics, and Spectroscopy* (Wiley, New York, 1993).

Exercises

1.1. **Trace-log of a matrix.** Prove that

$$\det M = e^{\mathrm{tr}\,\ln M}.$$

for an arbitrary nonsingular finite dimensional matrix M, $\det M \neq 0$.

1.2. **Stability, diagonal case.** Verify that for a diagonalizable matrix A the exponential is also diagonalizable

$$J^t = e^{tA} = \mathbf{U}^{-1} e^{tA_D} \mathbf{U}\,, \quad A_D = \mathbf{U}\mathbf{A}\mathbf{U}^{-1}\,. \tag{1.52}$$

1.3. **Time-ordered exponentials.** Given a time dependent matrix $A(t)$, show that the time-ordered exponential

$$J(t) = \mathbf{T}e^{\int_0^t d\tau A(\tau)}$$

may be written as

$$J(t) \;=\; \mathbf{1} + \sum_{m=1}^{\infty} \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{m-1}} dt_m\, A(t_1)\, A(t_2) \cdots A(t_m)\,. \tag{1.53}$$

(Hint: for a warmup, consider summing elements of a finte-dimensional symmetric matrix $S = S^{\top}$. Use the symmetry to sum over each matrix element once; (1.53) is a continuous limit generalization, for an object symmetric in m variables.) Verify, by using this representation, that $J(t)$ satisfies the equation

$$\dot{J}(t) = A(t)J(t),$$

with the initial condition $J(0) = 1$.

1.4. **Real representation of complex eigenvalues.** (Verification of example 1.3.) λ_k, λ_{k+1} eigenvalues form a complex conjugate pair, $\{\lambda_k, \lambda_{k+1}\} = \{\mu + i\omega, \mu - i\omega\}$. Show that

(a) corresponding projection operators are complex conjugates of each other,

$$\mathbf{P} = \mathbf{P}_k\,, \qquad \mathbf{P}^* = \mathbf{P}_{k+1}\,,$$

where we denote $\mathbf{P}_k$ by $\mathbf{P}$ for notational brevity.

(b) $\mathbf{P}$ can be written as

$$\mathbf{P} = \frac{1}{2}(\mathbf{R} + i\mathbf{Q})\,,$$

where $\mathbf{R} = \mathbf{P}_k + \mathbf{P}_{k+1}$ and $\mathbf{Q}$ are matrices with real elements.

(c)

$$\begin{pmatrix} \mathbf{P}_k \\ \mathbf{P}_{k+1} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \begin{pmatrix} \mathbf{R} \\ \mathbf{Q} \end{pmatrix}.$$

(d) The $\cdots + \lambda_k \mathbf{P}_k + \lambda_k^* \mathbf{P}_{k+1} + \cdots$ complex eigenvalue pair in the spectral decomposition (1.29) is now replaced by a real $[2\times2]$ matrix

$$\cdots + \begin{pmatrix} \mu & -\omega \\ \omega & \mu \end{pmatrix} \begin{pmatrix} \mathbf{R} \\ \mathbf{Q} \end{pmatrix} + \cdots$$

or whatever you find the clearest way to write this real representation.