2.2 Multiple variables

2.2.1 Many coupled ODEs

The results of the previous section can be generalized to multiple variables indexed by $i = 1, 2, \cdots, n$. The set of coordinates $\{x_i\}$ can be regraded as a point in n dimensional space, and can also be represented as a vector $\vec{x}$ extending from the origin to this point. The generalized equations of motion can now be represented as

$$\dot{x}_i = F_i(\{x_i\}) \text{ for } i = 1, 2, \cdots n\,, \quad \text{or equivalently as} \quad \dot{\vec{x}} = \vec{F}(\vec{x})\,. \tag{2.2.1}$$

The linearized equations take the form

$$\dot{x}_i = \sum_{j=1}^{n} F_{ij} x_j \text{ for } i = 1, 2, \cdots n\,, \quad \text{or equivalently as} \quad \dot{\vec{x}} = \mathbf{F}\vec{x}\,, \tag{2.2.2}$$

in terms of the $n \times n$ matrix formed from n^2 elements $\{F_{ij}\}$.

A particular class of linear equations is obtained from gradient descent in a quadratic potential, which can be written as

$$V(\{x_i\}) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} M_{ij} x_i x_j\,. \tag{2.2.3}$$

It may appear that n^2 elements are needed to specify the potential. This is in fact not the case since after summation over both i and j, only the symmetric part $(M_{ij} + M_{ji})$ contributes as the coefficient of the term $x_i x_j$, while the antisymmetric part $(M_{ij} - M_{ji})$ vanishes. Thus a general quadratic potential can be represented by $n(n+1)/2$ elements forming a symmetric matrix, in which case $F_{ij} = M_{ij}$ in Eq. (2.2.2).

2.2.2 Indexology

The summation convention (introduced to physics by Einstein) is a convenient way to represent sums, such as appearing in Eq. (2.2.3), in compact form. Basically, any index that appears twice has to be summed over all its possible values.[5] With this convention in mind, Eqs. (2.2.2) and (2.2.3) can be written as

$$\dot{x}_i = F_{ij} x_j \quad \text{and} \quad V = \frac{1}{2} M_{ij} x_i x_j\,, \tag{2.2.4}$$

with now implicit sums over i and j. In applying these rules, it is important to keep the following in mind:

[5]The more sophisticated Einstein notation, relevant to general relativity, distinguishes between indices appearing as superscripts (upper) or subscripts (lower) on a variable. We shall not deal with this subtlety here.

- It is very important to ensure that *any index that represents a component along some direction* appears only 0, 1, 2 times on one side of an equation.

- Any such index that appears once on one side of an equation, must also appear once on the other side of the equation.

- Note that there could be other labels, not indexing components of a vector, that are not subject to the summation rule. The labels for eigenvalues that we shall use shortly are an example of this exemption.

Having introduced the index notation, it is useful to be familiar with the following terminology:

- *Scalars* are quantities that do not carry an index, such as the potential V. They can be constructed by contracting (pairing) of entities with indices, such as in $x_i y_i \equiv \vec{x} \cdot \vec{y}$ (the dot product of vectors $\vec{x}$ and $\vec{y}$), or $M_{ii} \equiv \mathrm{tr}\mathbf{M}$ (trace of a matrix).

- *Vectors* carry a single index such as x_i or $\dot{x}_i$.

- *Matrices* such as M_{ij} can for example be constructed from two vectors, as in $x_i x_j$ or $\dot{x}_i x_j$, or also from product of other matrices, as in $A_{ij}B_{jk} = (AB)_{ik} = C_{ik}$ (the component form of the matrix product $\mathbf{A} \cdot \mathbf{B} = \mathbf{C}$).

- *The Kronecker delta-function* δ_{ij} represents the components of the unit matrix, equal to 1 if $i = j$ (along the diagonal) and 0 otherwise (off diagonal). Summing over one index of the delta-function has the effect of replacing it with the other index, as in $\delta_{ij}x_j = x_i$ or $\delta_{ij}M_{jk} = M_{ik}$. Also, note that $\delta_{ii} = n$, where n is the dimensionality of the system.

- We can also construct objects with more indices, such as $x_i x_j x_k$ or $M_{ij}M_{kl}$, sometimes referred to as *tensors* of higher rank (3 and 4 in the two examples).

- We already encountered the *gradient operator* in the contexts of descent in a scalar potential $V(\{x_i\})$. The operation of taking gradient can be represented by the components of the derivative vector $\nabla_i \equiv \frac{\partial}{\partial x_i} \equiv \partial_i$.

- *Vector fields*, such as the previously encountered force $F_i(\{x_i\})$ are vectors whose magnitude and direction vary in coordinate space. The *divergence* of a vector field is the scalar quantity
$$\mathrm{div}\vec{F} = \nabla \cdot \vec{F} = \partial_i F_i \,. \tag{2.2.5}$$
Note that for $F_i = -\partial_i V$, we find
$$\mathrm{div}\vec{F} = \partial_i F_i = -\partial_i \partial_i V \equiv -\nabla^2 V \,, \tag{2.2.6}$$
involving the *Laplacian operator* $\nabla^2 = \sum_i \frac{\partial^2}{\partial x_i^2} = \partial_i \partial_i$.

- The chain rule can also be compactly expressed in this notation as

$$\frac{dV(\{x_i\})}{dt} = \partial_i V \frac{dx_i}{dt} = \dot{x}_i \partial_i V \,, \quad \text{and} \quad \frac{dF_j(\{x_i\})}{dt} = \partial_i F_j \frac{dx_i}{dt} = \dot{x}_i \partial_i F_j \,, \tag{2.2.7}$$

  acting on a scalar and vector respectively.

2.2.3 Eigenvectors and eigenvalues

As noted before, directions along which the solution proceeds as a single exponential, as in Eq. (2.1.7) correspond to eigenvectors of the matrix $\mathbf{F}$ in Eq. (2.2.2). For an $n \times n$ matrix, there are n such eigenvectors that we shall label as $\vec{e}^{\,\alpha}$ for $\alpha = 1, 2, \cdots, n$, such that

$$F_{ij} e_j^\alpha = \lambda_\alpha e_i^\alpha \,, \quad \text{for} \quad \alpha = 1, 2, \cdots, n \,, \tag{2.2.8}$$

with e_i^α indicating the components of $\vec{e}^{\,\alpha}$, and λ_α as the corresponding eigenvalue. (Note that the index α on the right hand side of the above equation appears twice, but is not summed over, as $\{\lambda_\alpha\}$ do not represent components of a vector, but instead label the solutions of Eq. (2.2.8)).

We noted earlier that eigenvalues of a symmetric matrix with real entries $M_{ij} = M_{ji}$ are real numbers. Let us prove this as an exercise in the summation convention. Multiply both sides of Eq. (2.2.8) with $(e_i \beta)^*$ and sum over i to get

$$(e_i^\beta)^* M_{ij} e_j^\alpha = \lambda_\alpha (e_i^\beta)^* e_i^\alpha \,. \tag{2.2.9}$$

Taking complex conjugates of the above equation, and taking advantage of $M_{ij}^* = M_{ji}$ allows us to rearrange the equation as

$$(e_j^\alpha)^* M_{ji} e_i^\beta = \lambda_\alpha^* e_i^\beta (e_i^\alpha)^* \,. \tag{2.2.10}$$

Noting $M_{ji} e_i^\beta = \lambda_\beta e_j^\beta$, the above equation can be recast as

$$(\lambda_\beta - \lambda_\alpha^*)\, e_i^\beta (e_i^\alpha)^* = 0 \,. \tag{2.2.11}$$

For $\beta = \alpha$, the second term $\sum_i |e_i^\alpha|^2$, the squared magnitude of a (possibly complex) eigenvector $\vec{e}^{\,\alpha}$ is explicitly positive. We must therefore have $\lambda_\alpha = \lambda_\alpha^*$, requiring real eigenvalues. In fact both the real and imaginary parts of the vector $\vec{e}^{\,\alpha}$ are eigenvectors, and without loss of generality we can limit discussion to real eigenvectors and drop the complex conjugate sign.

For $\alpha \neq \beta$ (and assuming non-degenerate eigenvalues $\lambda_\alpha \neq \lambda_\beta$), we are then lead to another important result, that $\vec{e}^{\,\beta} \cdot \vec{e}^{\,\alpha} = 0$. The eigenvectors of a real symmetric matrix thus form an orthogonal set in the n-dimensional space. The magnitude of the eigenvectors is arbitrary, but it is useful to make them all equal to unity, such that they form an orthonormal set with $\vec{e}^{\,\beta} \cdot \vec{e}^{\,\alpha} = \delta_{\alpha\beta}$.

To solve the set of linear ODEs $\dot{x}_i(t) = M_{ij} x_j(t)$, with the initial condition $x_i(t = 0) = x_i(0) = x_i^0$:

- Find the eigenvectors $\vec{e}^{\alpha}$ and the corresponding eigenvalues λ_α.
- Compute the coordinates of the starting point in the basis formed by the eigenvectors, i.e. $a_\alpha(0) = x_i(0)e_i^\alpha$.
- Each component in the eigenvector basis will evolve as a simple exponential with the corresponding eigenvalue, i.e. $a_\alpha(t) = a_\alpha(0)e^{\lambda_\alpha t}$.
- In terms of these components the location at time t is given by

$$x_i(t) = \sum_\alpha a_\alpha(t)e_i^\alpha = \sum_\alpha a_\alpha(0)e^{\lambda_\alpha t}e_i^\alpha = x_j(0)\sum_\alpha e_j^\alpha e^{\lambda_\alpha t}e_i^\alpha \equiv U_{ij}(t)x_j(0)\,, \qquad (2.2.12)$$

  where we have introduced the linear operator $U_{ij}(t) = \sum_\alpha e_i^\alpha e^{\lambda_\alpha t}e_j^\alpha$ whose action (multiplication) on the initial vector leads to the position at time t.
- For displacements around a stable equilibrium point, the solution in Eq. (2.2.12) must not diverge for any choice of initial condition. For this to hold, all eigenvalues of the matrix must be negative.[6] If the matrix is obtained from gradient descent in the potential $V = K_{ij}x_ix_j/2$, stability requires all eigenvalues of the matrix to be positive. (The change of sign is due to the negative sign from gradient descent, $F_i = -\partial_i V$.) Such a matrix is called *positive definite* and $K_{ij}x_ix_j > 0$ for any displacement $\vec{x}$.

2.2.4 Functions of a matrix

In the same way that a function of a variable $f(x)$ can be constructed through its Taylor series, functions $f(\mathbf{M})$ of a matrix $\mathbf{M}$ can be defined through the corresponding Taylor series, e.g.

$$\exp(M) = \mathbf{1} + \mathbf{M} + \frac{\mathbf{M}^2}{2} + \cdots = \sum_{n=0}^{\infty}\frac{\mathbf{M}^n}{n!}\,. \qquad (2.2.13)$$

Individual components of the matrix are obtained using standard rules of multiplication of matrices, e.g.

$$\exp(\mathbf{M})_{ij} = \delta_{ij} + M_{ij} + \frac{M_{ik}M_{kj}}{2} + \cdots\,. \qquad (2.2.14)$$

Upon acting on an eigenvector,

$$\mathbf{M}^2\vec{e}^{\alpha} = \mathbf{M}\mathbf{M}\vec{e}^{\alpha} = \lambda_\alpha\mathbf{M}\vec{e}^{\alpha} = \lambda_\alpha^2\vec{e}^{\alpha}\,, \text{ and similarly } \mathbf{M}^n\vec{e}^{\alpha} = \lambda_\alpha^n\vec{e}^{\alpha}\,. \qquad (2.2.15)$$

Thus $\vec{e}^{\alpha}$ are eigenvectors of any function $f(\mathbf{M})$ of the matrix $\mathbf{M}$ with corresponding eigenvalues being $f(\lambda_\alpha)$. The action of the matrix function $f(\mathbf{M})$ on any vector $\vec{v}$ can then be calculated by the same procedure as used in calculating $x_i(t)$ in the previous section:

- Compute the coordinates of the vector $\vec{v}$ in the basis formed by the eigenvectors, as $a_\alpha = v_ie_i^\alpha$.

[6] For non-symmetric matrices with complex eigenvalues, the real parts of all eigenvalues must be negative.

- Under the action of $f(\mathbf{M})$, each component in the eigenvector basis is multiplied by $f(\lambda_\alpha)$, i.e. $f(\mathbf{M})a_\alpha\vec{e}^{\,\alpha} = f(\lambda_\alpha)a_\alpha\vec{e}^{\,\alpha}$.

- From the components in the eigenvector basis we can reconstruct the coordinates in the original basis as

$$[f(\mathbf{M})v]_i = a_\alpha f(\lambda_\alpha)e_i^\alpha = v_j e_j^\alpha f(\lambda_\alpha)e_i^\alpha \equiv f(\mathbf{M})_{ij}v_j\,. \tag{2.2.16}$$

- Thus quite generally the elements of a matrix (in any basis) can be computed in terms of a sum over its eigenvectors and eigenvalues as

$$f(\mathbf{M})_{ij} = \sum_\alpha e_i^\alpha f(\lambda_\alpha)e_j^\alpha\,. \tag{2.2.17}$$

- Note that the trace of $f(\mathbf{M})$ is obtained as

$$f(\mathbf{M})_{ii} = \sum_\alpha f(\lambda_\alpha)e_i^\alpha e_i^\alpha = \sum_\alpha f(\lambda_\alpha)\,. \tag{2.2.18}$$

  since $e_i^\alpha e_i^\alpha = \vec{e}^{\,\alpha}\cdot\vec{e}^{\,\alpha} = 1$.

We can now see that the time evolution operation in Eq. (2.2.12) is carried out by the matrix $\mathbf{U}(t) = \exp(t\mathbf{M})$. Indeed this amount to solving the linear set of ODEs as

$$\frac{d\vec{x}}{dt} = \mathbf{M}\vec{x} \quad \Longrightarrow \quad \vec{x}(t) = \exp\left(t\mathbf{M}\right)\vec{x}(0)\,, \tag{2.2.19}$$

treating the vector of ODEs similar to one for a scalar x. However, treating matrices in functions and in equations as in the case of scalars has to be done very carefully, and fails in dealing with *non-commuting* matrices. The commuting property of two scalar quantities $XY = YX$ does not expend to matrices, and generically $\mathbf{X}\cdot\mathbf{Y} \neq \mathbf{Y}\cdot\mathbf{X}$. The Taylor series of a function of two variables must then be ordered appropriately as, for example $2\mathbf{X}\cdot\mathbf{Y} \neq \mathbf{X}\cdot\mathbf{Y} + \mathbf{Y}\cdot\mathbf{X}$.

Suppose we want to solve the ODE in Eq. (2.2.19) for a scalar $x(t)$, but with M that changes from M_1 after a time t_1 to M_2. After a subsequent time interval of t_2, we find

$$x(t_1+t_2) = \exp(t_2M_2)x(t_1) = \exp(t_2M_2)\exp(t_1M_1)x(0) = \exp(t_1M_1+t_2M_2)x(0)\,. \tag{2.2.20}$$

For the matrix version, the last step cannot be performed for non-commuting matrices, as

$$\vec{x}(t_1{+}t_2) = \exp(t_2\mathbf{M}_2)x(t_1) = \exp(t_2\mathbf{M}_2)\exp(t_1\mathbf{M}_1)x(0) \neq \exp(t_1\mathbf{M}_1{+}t_2\mathbf{M}_2)x(0)\,. \tag{2.2.21}$$

For a time-varying $\mathbf{M}_n$, the time evolution operator must strictly follow the ordering of matrices acting on the initial vector, i.e.

$$\mathbf{U}(t_1+t_2+\cdots t_N) = \exp(t_N\mathbf{M}_N)\exp(t_{N-1}\mathbf{M}_{N-1})\cdots\exp(t_2\mathbf{M}_2)\exp(t_1\mathbf{M}_1)\,. \tag{2.2.22}$$

In field theory, this is referred to as *path ordering* or *time ordering* of operators.

Recap

- A coupled linear set of ODEs can be represented by a matrix $\mathbf{M}$.
- Eigenvectors of $\mathbf{M}$ determine directions along which time evolution is in the form of a single exponential with the eigenvalue as the rate.
- For symmetric matrices the eigenvalues are real and the eigenvectors can be arranged as an orthonormal set, such that with $\vec{e}^{\beta} \cdot \vec{e}^{\alpha} = \delta_{\alpha\beta}$.
- Functions of a matrix can be represented using a decomposition in terms of eigenvectors and eigenvalues.