1 Parabola as a limit of an ellipse

A parabola can be obtained by blowing up the major axis of an ellipse but at the same time shift the ellipse such that its focal point is in the origin.

Proof. Take an ellipse with left focal point in the origin. It is defined via

$$\frac{(x-ae)^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{1}$$

(Draw this.) We want to take the limit of $a \to \infty$ in a smart way, since blunt limit creates two straight lines. The idea is to keep the nearest distance between the path and the left focal point the same. This distance is $a(1-e)$. So the limit is

$$a \to \infty \qquad \text{and} \qquad d \equiv a(1-e) = \text{constant} \tag{2}$$

This can only be true if also $e \to 1$. Using $b^2 = a^2(1-e^2)$, we can rewrite the curve equation as

$$\frac{x^2 - 2aex}{a^2} + \frac{y^2}{a^2(1-e^2)} = 1 - e^2\,, \tag{3}$$

or

$$\frac{x^2 - 2aex}{a^2} + \frac{y^2}{ad(1+e)} = (1-e)(1+e)\,, \tag{4}$$

or

$$\frac{x^2 - 2aex}{a} + \frac{y^2}{d(1+e)} = d(1+e)\,, \tag{5}$$

Now the limit is easy and gives

$$x = \frac{y^2}{4d} - d\,. \tag{6}$$

2 Equation 7.27

Equation 7.27 on page 175 does not seem immediate[1] We start with

$$a\cos\psi = ae + r(\theta)\cos\theta \tag{7}$$

and fill in the equation for $r(\theta)$ to get

$$\frac{a^2}{b^2}(-e + \cos\psi)(1 + e\cos\theta) = \cos\theta\,. \tag{8}$$

[1] At least not to the lecturer, any student that spots a simpler derivation than mine is welcomed to send this to me.

We can rearrange that equation into

$$(1+e\cos\theta)(e\cos\psi-e^2)=e\frac{b^2}{a^2}\cos\theta \tag{9}$$

After using $e^2=1-b^2/a^2$ we rewrite this as

$$e\cos\psi(1+e\cos\theta)-e^2-e\cos\theta=0\,. \tag{10}$$

Equivalently

$$e\cos\psi(1+e\cos\theta)-e\cos\theta-1=e^2-1\,, \tag{11}$$

or

$$(e\cos\psi-1)(1+e\cos\theta)=e^2-1=-\frac{b^2}{a^2} \tag{12}$$

which leads to 7.27.

3 Equation 7.28

Equation 7.28 on page 175 does not seem immediate[2] Here is a derivation that does not require equation 7.27 as an intermediate step. Taking a θ derivative of equation the equation above 7.27 and recalling the definition of $r(\theta)$, I find:

$$\frac{\mathrm{d}\theta}{\mathrm{d}\psi}=\frac{a^2}{b^2}(1+e\cos(\theta))^2\frac{\sin\psi}{\sin\theta}. \tag{13}$$

[2] At least not to the lecturer, any student that spots a simpler derivation than mine is welcomed to send this to me.

Now we rewrite $\sin\psi$:

$$\begin{aligned}
\sin\psi &= \sqrt{1-\cos^2\psi} \\
&= a^{-1}\sqrt{a^2 - a^2\cos^2\psi} \\
&= a^{-1}\sqrt{a^2 - (ae + \frac{b^2\cos\theta}{a(1+e\cos\theta)})^2} \\
&= a^{-1}\sqrt{a^2 - a^2e^2 - \frac{b^4\cos^2\theta}{a^2(1+e\cos\theta)^2} - 2eb^2\frac{\cos\theta}{1+e\cos\theta}} \\
&= \frac{b}{a}\sqrt{1 - \frac{b^2\cos^2\theta}{a^2(1+e\cos\theta)^2} - 2e\frac{\cos\theta}{1+e\cos\theta}} \\
&= \frac{b}{a}\sqrt{1 - 2e\frac{\cos\theta}{1+e\cos\theta} + \frac{e^2\cos\theta^2}{(1+e\cos\theta)^2} - \frac{e^2\cos\theta^2}{(1+e\cos\theta)^2} - \frac{b^2\cos^2\theta}{a^2(1+e\cos\theta)^2}} \\
&= \frac{b}{a}\sqrt{(1 - e\frac{\cos\theta}{1+e\cos\theta})^2 - (e^2 - \frac{b^2}{a^2})\frac{\cos\theta^2}{(1+e\cos\theta)^2}} \\
&= \frac{b}{a}\sqrt{(\frac{1}{1+e\cos\theta})^2 - \frac{\cos\theta^2}{(1+e\cos\theta)^2}} \\
&= \frac{b}{a}\frac{\sin\theta}{(1+e\cos\theta)}.
\end{aligned} \tag{14}$$

Putting this back into equation (13) reproduces 7.28.

4 What is a cross section?

I present an alternative definition of a cross section to what can be found in the book. This definition is easier and more insightful.

Definition: Consider an infinitely big beam of incoming particles that are being shot towards a target. All the particle velocities are in exactly the same direction (parallel). The cross section σ is a function defined on the unit 2-sphere surrounding the center of the target[3]. Consider all particles that are being scattered in various directions. We denote these directions collectively by M. So M is a subspace on the 2-sphere: the set of angles of the outgoing velocities. Then the cross section is defined by the following

[3] In other words, σ it is a function of the two Euler angles θ and ϕ that uniquely fix a direction.

equation

$$\int_M \sigma(\theta, \varphi)\sin(\theta)d\theta d\phi = \mathcal{A} \tag{15}$$

where $\mathcal{A}$ is the cross sectional *area* of the part of the imcoming beam that has scattered through M.

Ok, let us explain this in simple words. First note that the conventions of the Euler angles are chosen such that θ is also the scattering angle. For example consider a target that repells incoming particles and is axially symmetric. That means symmetric in the ϕ-direction. Then particles colliding head on are being pushed back where they come from so then $\theta = \pi$. Particles that have an impact parameter that is really large are not going to feel the target and are basically flying straight and then $\theta = 0$. This explains why the cross section for the Rutherford experiment becomes infinite for $\theta = 0$: if we integrate σ from $\theta = 0$ to some finite value of θ, say $\theta = \pi/2$ we must find the area of the imcoming beam that has scattered in those directions. Clearly that is an infitely large part, taking into account all particles at infinity. If you want to picture this: look at the plane perpendicular to the beam. That area is the whole plane aside from a disk in the center. The boundary of that disk (a circle) is formed by particles that scatter at an angle $\pi/2$.

In general all particles that come from some part of the beam, namely the part with area $\mathcal{A}$ will scatter into various directions that form a segment M on the 2-sphere.

Let us apply this definition to a 'hard wall' elastic colision. So we picture a biljart ball of radius R as the target. See the figure below.

Let us compute the cross section. From the picture we can deduce

$$\theta + 2\alpha = \pi\,. \tag{16}$$

Also from the picture we see the following relation between impact parameter p and scattering angle θ:

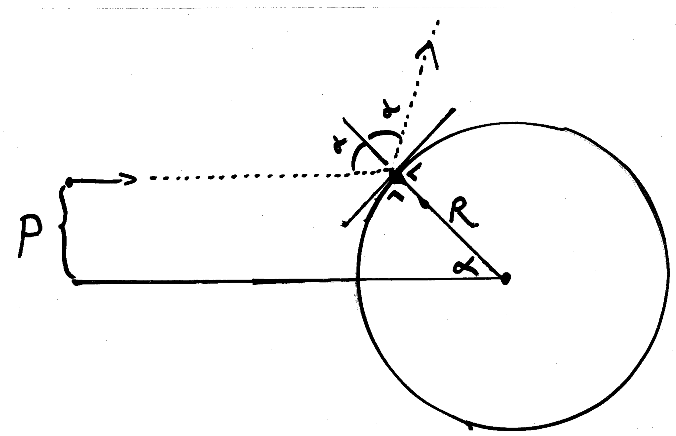
$$R\sin(\alpha) = p \to p(\theta) = R\cos(\theta/2). \tag{17}$$

Plugging this into the equation for an axial symmetric cross section

$$\sigma(\theta) = -\frac{p}{\sin(\theta)}\frac{dp}{d\theta} \tag{18}$$

yields:

$$\sigma(\theta) = \frac{R^2}{4}. \tag{19}$$



Figuur 1: *Hard wall collision.*

Now integrating the cross section over the whole S^2 gives:

$$\int_{\theta=0}^{\theta=\pi}\int_{\phi=0}^{\phi=2\pi}\sigma(\theta)\sin(\theta)d\theta d\phi=\pi R^2 \tag{20}$$

exactly as expected! We find the cross section area of the target. Why? Well, any incoming particle that is not directly aimed into the biljart ball will simply miss it and not scatter at all. Recall that the cross section computes the area of the incoming beam that scatters. What does not interact, does not scatter and is not taken into acoount.

Now that we have gotten a feel for what a cross section is, let us contemplate how we actually go about and measure it in an experiment. This then will touch upon the definition given in the text book. One can show that the following definition is equivalent:

Definition 2:

Call the flux[4] of the incoming beam F. Then

$$\sigma(\theta,\phi)\sin(\theta)d\phi d\theta=\frac{Nd\phi d\theta}{\Delta tF}\,, \tag{21}$$

where N is the number of particles scattered into the window $d\phi d\theta$. So the definition says that σ is sort of a probability to measure particles. It is the

[4]Flux is the number of particles crossing an unit area in a unit of time

probability density describing the number of scattered particles flying off in the directions θ, ϕ per unit time (Δt) per unit of flux of the imcoming beam. I will not explain why these definitions are equivalent, but refer top next years course on particle physics or the book of Thomson on Particle Physics (https://www.hep.phy.cam.ac.uk/ thomson/MPP/ModernParticlePhysics.html)

5 Alternative description of Noether's theorem

The point I wish to make is that Noether's theorem is identical to the theorem of a cyclic coordinate, but sometimes seeing the cyclic coordinate requires going to different coordinates. So consider a configuration space with coordinates q^i, but none of these coordinates are cyclic. However there is some (possibly non-linear) combination that is cyclic. Call it λ. So

$$\lambda(q^1, \ldots, q^n)\,. \tag{22}$$

Think of rotational symmetries. In Cartesian coordinates the Lagrangian will depend on x, y, z, but when there is some rotational symmetry around some axis, I can go to cylindrical coordinates and then the angle variable will be cyclic! So when I choose new coordinates $\tilde{q}$ as follows

$$\tilde{q} = (\tilde{q}^1 = \lambda, \tilde{q}^2, \ldots, \tilde{q}^n) \tag{23}$$

such that the first coordinate in the new system is the cyclic one, and the others are some unspecified combination of the original $q's$ that is such that the $\tilde{q}$ cover the configuration space (locally).

Lagrange's equations in the new coordinates are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\tilde{q}}^i}\right) = \frac{\partial L}{\partial \tilde{q}^i}\,. \tag{24}$$

But since $\tilde{q}^1 = \lambda$ is cyclic we have a conserved quantity Q:

$$Q = \frac{\partial L}{\partial \dot{\lambda}}\,. \tag{25}$$

Now we will write this quantity in terms of the old coordinates q by using the chain rule. In general momenta in different coordinates are related as follows

$$\frac{\partial L}{\partial \dot{\tilde{q}}^j} = \sum_i \frac{\partial L}{\partial \dot{q}^i}\frac{\partial \dot{q}^i}{\partial \dot{\tilde{q}}^j} = \sum_i \frac{\partial L}{\partial \dot{q}^i}\frac{\partial q^i}{\partial \tilde{q}^j} \tag{26}$$

where in the last step we relied on

$$\dot{\tilde{q}}^j = \sum_i \frac{\partial \tilde{q}^j}{\partial q^i} \dot{q}^i \,. \tag{27}$$

If we now use this for writing out that the momentum associated to $\tilde{q}^1 = \lambda$ is constant, we find

$$Q = \frac{\partial L}{\partial \dot{\lambda}} = \sum_i \frac{\partial L}{\partial \dot{q}^i} \frac{\partial [q^i]_\lambda}{\partial \lambda} \,. \tag{28}$$

which is Noether's theorem. If we want to describe Q entirely in terms of the origional coordinates q^i we need to evaluate the above equation at $\lambda = 0$:

$$Q = \frac{\partial L}{\partial \dot{\lambda}} = \sum_i \frac{\partial L}{\partial \dot{q}^i} \frac{\partial [q^i]_\lambda}{\partial \lambda} |_{\lambda=0} \,. \tag{29}$$

6 Derivation of Liouville's theorem

To be typed.

7 Rotations

Consider the n-dimensional plane $\mathbb{R}^n$ with Cartesian coordinates $x_1, \ldots x_n$. We are interested in linear transformations of points:

$$\begin{pmatrix} x'_1 \\ x'_2 \\ \ldots \\ x'_3 \end{pmatrix} = R \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_3 \end{pmatrix} \tag{30}$$

with R some matrix.

- Argue that all transformations that preserve the distance, d, of the points to the origin $d = \sum_i x_i^2$ are given by matrices that obey $RR^T = 1$ (and hence $R^T R = 1$), where 1 denotes the unity matrix.

- Explain why these are the definitions of what is called a rotation around the origin. In other words, a rotation is given by a matrix that obeys $RR^T = 1$.

- Verify that rotations around the z-axis over an angle θ in $\mathbb{R}$ are given by

$$R_z = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{31}$$

- is this a clockwise or anti-clockwise rotation when viewed from "above"(ie from somehow looking down on the (x, y))-plane.

- Write the rotation matrices for rotations around the x- and y-axis.

Imagine one interested in rotations around the axis $\hat{n}$, which does not coincide with one of the standard axises $\hat{x}, \hat{y}, \hat{z}$. A trick can be this: first one chooses new coordinates such that $\hat{n}$ becomes $\hat{z}$. Then one writes the rotation matrix for a rotation around the z-axis. Then one writes the obtained result in the old coordinates. Let us make this more concrete: Say for instance that the relation between the new $\tilde{x}, \tilde{y}, \tilde{z}$ and old coordinates (x, y, z) is given by a matrix S:

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} = S \begin{pmatrix} x \\ y \\ z \end{pmatrix} . \tag{32}$$

The tilded coordinates are such that now the axis $\hat{n}$ coincided with $\hat{\tilde{z}}$.

- Now show that $R = S R_z S^{-1}$ with R_z given in (31).

- Does this generalise to arbitrary dimensions?

- Now perform this in a concrete example. Say we rotation around the axis defined by the point $(1, 2, 0)$ and the origin $(0, 0, 0)$.

- Verify that your matrix satisfies $RR^T = 1$.

8 Theory of small oscillations

The equations of motion for the approximated system are

$$\mathbb{T}\ddot{\vec{q}} = -\mathbb{V}\vec{q} \tag{33}$$

where the objects $\mathbb{T}$ and $\mathbb{V}$ are the square symmetric matrices appearing in the definition of kinetic and potential energy $\mathbb{T} = t_{ij}$ and $\mathbb{V} = v_{ij}$.

Now our aim is to find a new basis of coordinates such that equations (33) turn into a set of decoupled harmonic oscillators. This would be achieved if

$$\mathbb{T} = 1 , \qquad \mathbb{V} = \mathrm{diag}(\omega_1^2, \omega_2^2, \ldots, \omega_N^2) , \tag{34}$$

where 1 denotes the unity matrix.

Can we find such a basis of new coordinates?

In what follows we construct this coordinate transformation. We will restrict to linear coordinate transformations of the form:

$$\vec{q}' = S\vec{q}\,, \tag{35}$$

with S some invertible matrix. How do $\mathbb{T}$ and $\mathbb{V}$ transform? We will make use of the fact that energies are unaffected by coordinate transformations (a particle will not suddenly have different kinetic or potential energy if I change spatial coordinates). Let us denote the transformed matrices with a prime. So we have for instance

$$\vec{q'}^T \mathbb{V}' \vec{q'} = \vec{q}^T \mathbb{V} \vec{q}\,, \tag{36}$$

or, equivalently

$$\vec{q}^T S^T \mathbb{V}' S \vec{q} = \vec{q}^T \mathbb{V} \vec{q}\,. \tag{37}$$

Since this is true for all $\vec{q}$ this must mean that

$$S^T \mathbb{V}' S = \mathbb{V} \leftrightarrow \mathbb{V}' = S^{-T} \mathbb{V} S^{-1}\,. \tag{38}$$

Analogously we have

$$\mathbb{T}' = S^{-T} \mathbb{T} S^{-1}\,. \tag{39}$$

Finally, the student needs to know the following theorem in linear algebra. Take M a square matrix. Then there exists an ortogonal transformation P such that

$$PMP^{-1} = D \tag{40}$$

with D a diagonal matrix. The fact that P is orthogonal means $P^T = P^{-1}$. So the above can be rewritten as

$$PMP^T = D \tag{41}$$

Secondly if M is a positive definite matrix then all eigenvalues (elements of D) are positive.

Now let us apply this for the matrix $\mathbb{T}$. Then we know we can diagonalise it. So there exists an S_1 such that

$$\vec{q}' = S_1 \vec{q} \to \mathbb{T}' = D\,, \qquad D = \mathrm{diag}(d_1^2, d_2^2, \ldots) \tag{42}$$

Now also $\mathbb{V}$ changed to $\mathbb{V}'$. But the new $\mathbb{V}'$ will still be symmetric and positive (explain why!).

Now consider yet another transformation:

$$\vec{q}'' = S_2\vec{q}' = S_2 S_1 \vec{q} \tag{43}$$

where

$$S_2 = \sqrt{D}\,. \tag{44}$$

Then

$$\mathbb{T}'' = 1\,, \tag{45}$$

and still $\mathbb{V}''$ is positive and symmetric (explain why!).

Finally we consider a third transformation:

$$\vec{q}''' = S_3\vec{q}'' = S_3 S_2\vec{q}' = S_3 S_2 S_1\vec{q} \tag{46}$$

such that it diagonalises $\mathbb{V}'''$

$$\mathbb{V}''' = \mathrm{diag}(\omega_1^2, \omega_2^2, \ldots) \tag{47}$$

In this new basis the kinetic energy is still the same:

$$\mathbb{T}''' = S_3^{-T}\mathbb{T}'' S_3^{-1} = S_3^{-T} S_3^{-1} = 1\,, \tag{48}$$

since S^3 (and hence S_3^{-1}) is an ortogonal transformation.

We achieved our goal.

The solutions to this set of decoupled harmonic oscillators can be written as

$$\vec{q}'''(t) = \sum_i c_i \vec{a}_i''' \cos(\omega_i t + \gamma_i) \tag{49}$$

where the $\vec{a}_i'''$ are the following vectors

$$\vec{a}_i''' = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{50}$$

with the 1 on position i.

Now let us see what this implies in the original coordinates. The old coordinates and the new coordinates are related as follows

$$\vec{q} = (S_3 S_2 S_1)^{-1}\vec{q}''' \tag{51}$$

so we can write

$$\vec{q}(t) = \sum_i c_i \vec{a}_i \cos(\omega_i t + \gamma_i) \tag{52}$$

where $\vec{a} = (S_3 S_2 S_1)^{-1} \vec{a}'''$. Let us see what kind of relations the $\vec{a}_i$ obey. Either by putting the above equation back into the original equation of motion, or by realising algebraicely what we did, one can see that the $\vec{a}_i$ are solutions to

$$\left(-\omega_i^2 \mathbb{T} + \mathbb{V}\right) \vec{a}_i = 0\,. \tag{53}$$

The normalisation of the $\vec{a}_i$ do not matter too much since they can be absorbed in the coefficients c_i. But in case one cares, it goes as follows. The $\vec{a}_i'''$ clearly obeyed:

$$\vec{a}_i'''^T \vec{a}_j''' = \delta_{ij}\,. \tag{54}$$

This gives

$$\vec{a}_i^T (S_3 S_2 S_1)^T S_3 S_2 S_1 \vec{a}_j = \delta_{ij} \to \vec{a}_i^T S_1^T S_2^T S_3^T S_3 S_2 S_1 \vec{a}_j = \delta_{ij} \tag{55}$$

using that S^3 is orthogonal and that $S_2 = \sqrt{D}$ we find:

$$\vec{a}_i^T S_1^T D S_1 \vec{a}_j = \delta_{ij} \tag{56}$$

Finally using the very definition of S_1 we find:

$$\vec{a}_i^T \mathbb{T} \vec{a}_j = \delta_{ij}\,. \tag{57}$$

So what does one do in practice? One does not seek to do coordinate transformations. Instead one follows 3 steps:

1. Compute the frequencies ω_i by solving the determinant ($\det\left(-\omega_i^2 \mathbb{T} + \mathbb{V}\right) = 0$).
2. Then we find $\vec{a}_i$, $\left(-\omega_i^2 \mathbb{T} + \mathbb{V}\right) \vec{a}_i = 0$.
3. Then the general solution is given by (52).

A Bilinear forms

Consider kinetic energy

$$T = \tfrac{1}{2} \sum_{ij} a_{ij} \dot{q}^i \dot{q}^j\,. \tag{58}$$

Let's work in the context of chapter 15, where in the linearised approximation a_{ij} is not dependent on q, although it typically is. This is just to simplify things (otherwise I need to discuss tensor fields instead of just tensors). The object a_{ij} can be written as a square matrix. In algebra you have seen square matrices as a method to write down a linear map. A linear map M maps vectors to other vectors in a linear way. Of typical interest is what happens to a matrix when I change basis of my vector space. Say I change basis using a transformation S. This means that my new vectors, denoted with a prime are written as

$$v' = Sv \,, \tag{59}$$

for all v. This just means that in the new basis the components of the same vector are now given by a different expression (obtained by multiplying with a matrix S). So how does the matrix M change? Well that is easy. We just use the definition of M. It maps vectors to other vectors. So we have that for two vectors w, v, where

$$w = Mv \,, \tag{60}$$

we must have

$$w' = M'v' \,. \tag{61}$$

This allows us to find M', because we can rewrite the above equation as

$$Sw = M'Sv \quad \rightarrow w = S^{-1}M'Sv = Mv \,. \tag{62}$$

Hence we find

$$S^{-1}M'S = M \rightarrow \boxed{M' = SMS^{-1}} \,. \tag{63}$$

This you must have seen in a course on linear algebra.

Instead of linear maps there also exist things called bilinear forms and they can also be represented by square matrices. The object a_{ij} in the kinetic energy turns out to be such a thing. What is a bilinear form? It is defined as an object a that acts linearly on two vectors and maps it to a number. That is:

$$a(v, w) \in \mathbb{R} \tag{64}$$

where a acts linear in both its arguments. For example:

$$\begin{aligned} a(2v_1 + 3v_2, w_1 + 4w_2) &= 2a(v_1, w_1 + 4w_2) + 3a(v_2, w_1 + 4w_2) \\ &= 2a(v_1, w_1) + 8a(v_1, w_2) + 3a(v_2, w_1) + 12a(v_2, w_2) \,. \end{aligned} \tag{65}$$

Similar to the linear map M, we can ask how a matrix associated to a bilinear form changes when the basis of the vector space is changed. Again we just use the definition of a bilinear form:

$$a'(v', w') = a(v, w) \,. \tag{66}$$

The reason we can write the above equation is that the image of a bilinear form is a number, something that is unaffected by a change of basis. Rewriting the above equation in components means:

$$\sum_{ij} a'_{ij} v'_i, w'_j = \sum_{kl} a_{kl} v_k, w_l \,. \tag{67}$$

Or in matrix language:

$$v'^T a' w' = v^T a w \,. \tag{68}$$

Now we use the transformation properties of the vectors ($v' = Sv$, $w' = Sw$):

$$v^T S^T a' S w = v^T a w \,. \tag{69}$$

Since w and v are arbitrary we have

$$S^T a' S = a \quad \rightarrow \boxed{a' = S^{-T} a S^{-1}} \,. \tag{70}$$

This is clearly different from a transformation of a linear map, although both linear maps and bilinear forms are described by square matrices. A matrix by itself does not tell you whether you are looking at a linear map or a bilinear form. It is crucial to know this.

Now consider special kind of transformations, called ortogonal transformations. Then $SS^T = 1$ or, equivalently, $S^{-1} = S^T$. Then linear maps and bilinear forms transform identically! This can be inferred from (63) and (70).

Now consider special bilinear forms. Namely those that are symmetrical:

$$a = a^T \tag{71}$$

Note that this property is not changed by going to a different basis (check this). You have learned in a course on linear algebra that a symmetric matrix associated to a linear map (!) can be diagonalised using an orthogonal transformation. So there exists a matrix O, that obeys OO^T, such that

$$M' = D = OMO^{-1} \,, \tag{72}$$

with D diagonal. Since we just explained that, under orthogonal transformations bilinear forms and linear maps have identical transformation properties we can also show that *any symmetric bilinear form can be diagonalised using an orthogonal transformation!*

$$a' = D \,, \tag{73}$$

with D diagonal. Now with bilinear maps we can even simplify the matrix expression more. So we assume we brought a' into the above diagonal form. Call the non-zero diagonal elements d_i. Then define a further, non-orthogonal, transformation S as follows

$$S = \mathrm{diag}(s_1, \ldots, s_n)\,, \tag{74}$$

where the s_i are chosen as follows:

- When $d_i > 0$, we take $s_i = \sqrt{d_i}$.
- When $d_i < 0$, we take $s_i = -\sqrt{d_i}$.
- When $d_i = 0$, we take $s_i = 1$.

Call a'' the transformation of a'. Then one can easily show that (check this!) that

$$a'' = \mathrm{diag}(\epsilon_i)\,, \tag{75}$$

where ϵ_i is either $0, 1$ or -1. This is called Sylvesters theorem. If a is positive definite then a'' has only $\epsilon_i = 1$ such that $a'' = 1$ (the unity matrix).

The matrix a used in the formula for the kinetic energy is an example of a bilinear form. The reason is that the value of the kinetic energy is independent of the basis used to describe the q^i. In other words:

$$T' = \tfrac{1}{2}\sum_{ij} a'_{ij}\dot{q}'^i\dot{q}'^j = \tfrac{1}{2}\sum_{ij} a_{ij}\dot{q}^i\dot{q}^j = T\,. \tag{76}$$

It is also positive definite since T has to be positive, for all choices of velocities.

So, in the context of chapter 15, we reason like this: once a basis of q's is found such that $a' = 1$ (unity matrix), we can still use orthogonal transformations to simplify the matrix of the bilinear form $\mathbb{V}$. The reason is that under an orthognal transformation the unity matrix is mapped to itself, since, if $a = 1$ then $a' = OaO^T = OO^T = 1$. Since $\mathbb{V}$ is a positive definite bilinear form (explain!) we can then use orthogonal transformations to bring $\mathbb{V}$ to a diagonal matrix with positive entries.