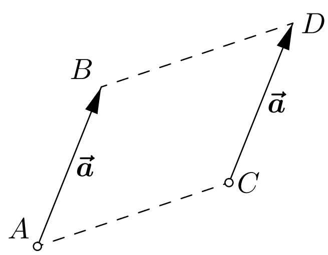
These are compact lecture notes for Math 321 at UW-Madison. Read them carefully, ideally before the lecture, and complete with your own class notes and pictures. Skipping the 'theory' and jumping directly to the exercises is a tried-and-failed strategy that only leads to the typical question 'I have no idea how to get started'. *There are many explicit and implicit exercises within the text that complement the 'theory'. Many of the 'proofs' are actually good 'solved exercises.' The objectives are to review the key concepts, emphasizing geometric understanding and visualization.*

1 Vectors: Geometric Approach

What's a vector? in elementary calculus and linear algebra you probably defined vectors as a list of numbers such as $\vec{\boldsymbol{x}} = (4, 2, 5)$ with special algebraic manipulations rules, but in elementary physics *vectors* were probably defined as *'quantities that have both a magnitude and a direction such as displacements, velocities and forces'* as opposed to *scalars,* such as *mass, temperature and pressure, which only have magnitude.* We begin with the latter point of view because the algebraic version hides geometric issues that are important to physics, namely that physical laws are invariant under a change of coordinates - they do not depend on our particular choice of coordinates - and there is no special system of coordinates, everything is *relative.* Our other motivation is that to truly understand vectors, and math in general, you have to be able to *visualize* the concepts, so rather than developing the geometric interpretation as an after-thought, we start with it.

1.1 Vector addition and multiplication by a scalar

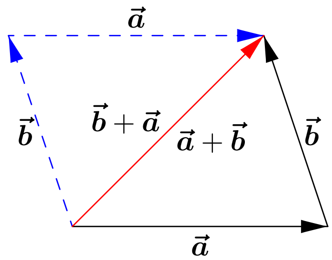
We begin with vectors in 2D and 3D Euclidean spaces, E^2 and E^3 say. E^3 corresponds to our intuitive notion of the space we live in (at human scales). E^2 is any plane in E^3. These are the spaces of classical Euclidean geometry. *There is no special origin or direction in these spaces.* All positions are relative to other reference points and/or directions.

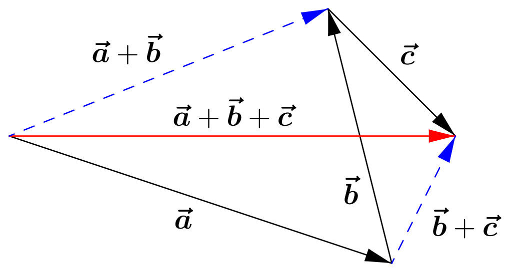


Vectors in those spaces are the set of all possible *displacements* which are *oriented line segments* whose origin does not matter. The vector $\overrightarrow{AB} \equiv \vec{\boldsymbol{a}}$ is the displacement from point A to point B. If we make the same *displacement* $\vec{\boldsymbol{a}}$ starting from another point C we end up at another point D but $\overrightarrow{CD} = \vec{\boldsymbol{a}} = \overrightarrow{AB}$. Two vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are equal if they are opposite legs of a parallelogram and have the same direction.

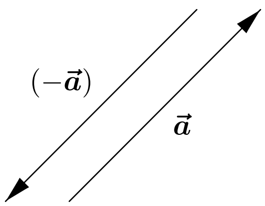
We'll denote by $\mathbf{E}^3$ the set of all possible vectors in E^3, to emphasize that vectors (displacements) and points are distinct concepts. The length of vector $\vec{\boldsymbol{a}}$ is denoted by $|\vec{\boldsymbol{a}}|$. It is a positive real number. When there is no risk of confusion we simply write a for $|\vec{\boldsymbol{a}}|$. When writing by hand, we use an arrow on top $\vec{a}$ or a wiggle underneath instead of the boldface $\vec{\boldsymbol{a}}$. *Velocities, accelerations* and *forces* are also vectors that can be described by oriented line segments, but strictly speaking these are all in different 'spaces': the velocity space, acceleration space, force space, etc. All these spaces are physically connected but we do not add displacements to forces for instance, so they are mathematically distinct.

Vectors add according to the **parallelogram rule**: If we move 1 mile North then 1 mile East, we end up $\sqrt{2}$ miles Northeast of the starting point. The net displacement would be the same if we move 1mi East first, then 1mi North. So vector addition is *commutative:* $\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}} = \vec{\boldsymbol{b}} + \vec{\boldsymbol{a}}$. It is also *associative:* $\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}} + \vec{\boldsymbol{c}} = (\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}}) + \vec{\boldsymbol{c}} = \vec{\boldsymbol{a}} + (\vec{\boldsymbol{b}} + \vec{\boldsymbol{c}})$.



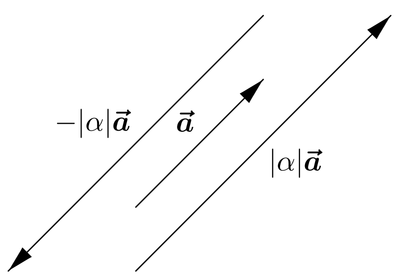


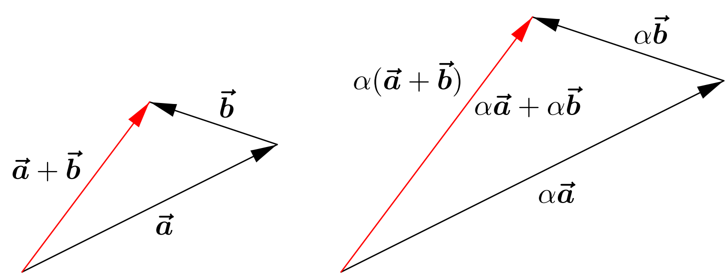
Note that in general $\vec{a}$, $\vec{b}$ and $\vec{c}$ are *not* in the same plane, so the 2D figure is not general, but it is easy enough to visualize associativity in 3D.



To every vector $\vec{a}$ we can associate an opposite vector denoted $(-\vec{a})$ that is the displacement exactly opposite to $\vec{a}$. Vector subtraction $\vec{b}-\vec{a}$ is then defined as the addition of $\vec{b}$ and $(-\vec{a})$. In particular $\vec{a}+(-\vec{a})=0$ corresponds to no net displacement. This is an important difference between points and displacements, there is no special point in our space, but there is one special displacement: the zero vector $\vec{0}$ such that $\vec{a}+(-\vec{a})=\vec{0}=(-\vec{a})+\vec{a}$ and $\vec{a}+\vec{0}=\vec{a}$, for any $\vec{a}$.

The other key operation that characterizes vectors is multiplication by a real number $\alpha\in\mathbb{R}$. Geometrically, $\vec{v}=\alpha\vec{a}$ is a new vector parallel to $\vec{a}$ but of length $|\vec{v}|=|\alpha||\vec{a}|$. The direction of $\vec{v}$ is the same as $\vec{a}$ if $\alpha>0$ and opposite to $\vec{a}$ if $\alpha<0$. Obviously $(-1)\vec{a}=(-\vec{a})$, multiplying $\vec{a}$ by (-1) yields the previously defined opposite of $\vec{a}$. Other geometrically obvious properties are $(\alpha+\beta)\vec{a}=\alpha\vec{a}+\beta\vec{a}$, and $(\alpha\beta)\vec{a}=\alpha(\beta\vec{a})$. A more interesting property is *distributivity* $\alpha(\vec{a}+\vec{b})=\alpha\vec{a}+\alpha\vec{b}$, which geometrically corresponds to *similarity of triangles.*





Generalization of the Vector Space concept

Vector addition and multiplication by a *real* number are the two key operations that *define* a **Vector Space**, provided those operations satisfy the following 8 properties $\forall\vec{a}$, $\vec{b}$ in the vector space and $\forall\alpha$, β in $\mathbb{R}$.

Required vector addition properties:

$$\vec{a}+\vec{b}=\vec{b}+\vec{a}, \tag{1}$$

$$\vec{a}+(\vec{b}+\vec{c})=(\vec{a}+\vec{b})+\vec{c}, \tag{2}$$

$$\vec{a}+\vec{0}=\vec{a}=\vec{0}+\vec{a}, \tag{3}$$

$$\vec{a}+(-\vec{a})=\vec{0}. \tag{4}$$

Required scalar multiplication properties:

$$(\alpha + \beta)\vec{\boldsymbol{a}} = \alpha\vec{\boldsymbol{a}} + \beta\vec{\boldsymbol{a}}, \tag{5}$$

$$(\alpha\beta)\vec{\boldsymbol{a}} = \alpha(\beta\vec{\boldsymbol{a}}), \tag{6}$$

$$\alpha(\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}}) = \alpha\vec{\boldsymbol{a}} + \alpha\vec{\boldsymbol{b}}. \tag{7}$$

$$1\,\vec{\boldsymbol{a}} = \vec{\boldsymbol{a}}, \tag{8}$$

When these properties are used to *define* the vector space they are referred to as *axioms*, *i.e.* the *defining properties.*

The vector space $\mathbb{R}^n$: Consider the set of ordered n-tuplets of real numbers $\boldsymbol{x} \equiv (x_1, x_2, \ldots, x_n)$. These could correspond to lists of student grades on a particular exam, for instance. What kind of operations would we want to do on these lists of student grades? We'll probably want to *add* several grades for *each student* and we'll probably want to *rescale* the grades. So the natural operations on these n-tuplets are *addition* defined by adding the respective components:

$$\boldsymbol{x} + \boldsymbol{y} \equiv (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) = \boldsymbol{y} + \boldsymbol{x}. \tag{9}$$

and multiplication by a real number $\alpha \in \mathbb{R}$ defined as

$$\alpha\boldsymbol{x} \equiv (\alpha x_1, \alpha x_2, \ldots, \alpha x_n). \tag{10}$$

The set of n-tuplets of real numbers equipped with addition and multiplication by a real number as just defined is an important vector space called $\mathbb{R}^n$. The vector spaces $\mathbb{R}^2$ and $\mathbb{R}^3$ will be particularly important to us as they'll soon corresponds to the components of our arrow vectors. But we also use $\mathbb{R}^n$ for very large n when studying systems of equations, for instance.

Exercises:

1. Show that addition and scalar multiplication of n-tuplets satisfy the 8 required properties listed above.

2. Define addition and scalar multiplication of n-tuplets of *complex* numbers and show that all 8 properties are satisfied. That vector space is called $\mathbb{C}^n$.

3. The set of real functions $f(x)$ is also a vector space. Define addition in the obvious way: $f(x) + g(x) \equiv h(x)$ another real function, and scalar multiplication: $\alpha f(x) = F(x)$ yet another real function. Show that all 8 properties are again satisfied.

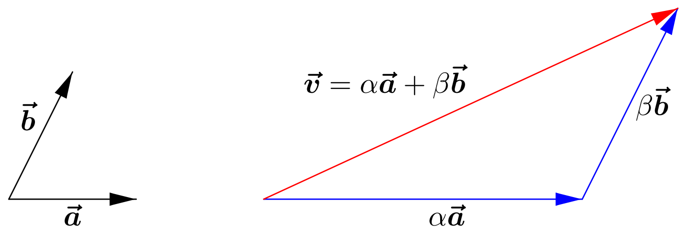
4. Suppose you define addition of n-tuplets $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ as usual but define scalar multiplication according to $\alpha\boldsymbol{x} = (\alpha x_1, x_2, \cdots, x_n)$, that is, only the first component is multiplied by α. Which property is violated? What if you defined $\alpha\boldsymbol{x} = (\alpha x_1, 0, \cdots, 0)$, which property would be violated?

5. From the 8 properties, show that $(0)\vec{\boldsymbol{a}} = \vec{\boldsymbol{0}}$ and $(-1)\vec{\boldsymbol{a}} = (-\vec{\boldsymbol{a}})$, $\forall\vec{\boldsymbol{a}}$, *i.e.* show that multiplication by the scalar 0 yields the neutral element for addition, and multiplication by -1 yields the additive inverse.

1.2 Bases and Components of a Vector

Addition and scalar multiplication of vectors allow us to define the concepts of *linear combination*, *basis*, *components* and *dimension*. These concepts apply to *any* vector space.

A *linear combination* of vectors $\vec{a}$ and $\vec{b}$ is an expression of the form $\alpha\vec{a}+\beta\vec{b}$. This linear combination yields another vector $\vec{v}$. The set of all such vectors, obtained by taking any $\alpha, \beta \in \mathbb{R}$, is itself a vector space (or more correctly a vector 'subspace' if $\vec{a}$ and $\vec{b}$ are two vectors in $\mathbf{E}^3$ for instance). We say that $\vec{a}$ and $\vec{b}$ form a **basis** for that (sub)space. We also say that this is the (sub)space **spanned** by $\vec{a}$ and $\vec{b}$. For a given vector $\vec{v}$, the unique *real* numbers α and β such that $\vec{v} = \alpha\vec{a}+\beta\vec{b}$ are called the **components** of $\vec{v}$ with respect to the basis $\vec{a}$, $\vec{b}$.



Linear Independence: The vectors $\vec{a}_1, \vec{a}_2, \cdots, \vec{a}_k$ for some integer k are *linearly independent* (L.I.) if the *only* way to have

$$\alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \cdots + \alpha_k\vec{a}_k = 0$$

is for all the α's to be zero:

$$\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0.$$

Dimension: The dimension of a vector space is the largest number of linearly independent vectors, n say, in that space. A basis for that space consists of n linearly independent vectors. A vector $\vec{v}$ has n components (some of them possibly zero) with respect to any basis in that space.

Examples:

• Two non-parallel vectors $\vec{a}$ and $\vec{b}$ in $\mathbf{E}^2$ are L.I. and these vectors form a **basis** for $\mathbf{E}^2$. Any given vector $\vec{v}$ in $\mathbf{E}^2$ can be written as $\vec{v} = \alpha\vec{a} + \beta\vec{b}$, for a unique pair (α, β). $\vec{v}$ is the diagonal of the parallelogram $\alpha\vec{a}$, $\beta\vec{b}$. Three or more vectors in $\mathbf{E}^2$ are *linearly dependent.*

• Three non-coplanar vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$ in $\mathbf{E}^3$ are L.I. and those vectors form a basis for $\mathbf{E}^3$. However 4 or more vectors in $\mathbf{E}^3$ are *linearly dependent.* Any given vector $\vec{v}$ can be expanded as $\vec{v} = \alpha\vec{a} + \beta\vec{b} + \gamma\vec{c}$, for a unique triplet of real numbers (α, β, γ). Make sketches to illustrate.

The 8 properties of addition and scalar multiplication imply that if two vectors $\vec{u}$ and $\vec{v}$ are expanded with respect to the same basis $\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$ so

$$\vec{u} = u_1\vec{a}_1 + u_2\vec{a}_2 + u_3\vec{a}_3,$$
$$\vec{v} = v_1\vec{a}_1 + v_2\vec{a}_2 + v_3\vec{a}_3,$$

then

$$\vec{u} + \vec{v} = (u_1 + v_1)\vec{a}_1 + (u_2 + v_2)\vec{a}_2 + (u_3 + v_3)\vec{a}_3,$$
$$\alpha\vec{v} = (\alpha v_1)\vec{a}_1 + (\alpha v_2)\vec{a}_2 + (\alpha v_3)\vec{a}_3,$$

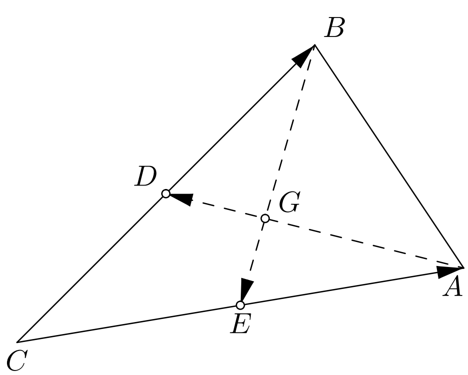
so addition and scalar multiplication are performed component by component and the triplets of real components (v_1, v_2, v_3) are elements of the vector space $\mathbb{R}^3$. A basis $\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$ in $\mathbf{E}^3$ provides a one-to-one correspondence (mapping) between *displacements* $\vec{v}$ in $\mathbf{E}^3$ and *triplets of real numbers* in $\mathbb{R}^3$

$$\vec{v} \in \mathbf{E}^3 \longleftrightarrow (v_1, v_2, v_3) \in \mathbb{R}^3.$$

Exercises

1. Pick two vectors $\vec{\boldsymbol{a}}$, $\vec{\boldsymbol{b}}$ and some arbitrary point A in the plane of your sheet of paper. If the possible displacements from point A to point B are specified by $\overrightarrow{AB} = \alpha\vec{\boldsymbol{a}} + \beta\vec{\boldsymbol{b}}$, sketch the region where B can be if: (i) $0 \le \alpha, \beta \le 1$, (ii) $|\beta| \le |\alpha|$ and $-1 \le \alpha \le 1$.
2. Given $\vec{\boldsymbol{a}}$, $\vec{\boldsymbol{b}}$, show that the set of all $\vec{\boldsymbol{v}} = \alpha\vec{\boldsymbol{a}} + \beta\vec{\boldsymbol{b}}$, $\forall \alpha, \beta \in \mathbb{R}$ is a vector space.
3. Show that the set of all vectors $\vec{\boldsymbol{v}} = \alpha\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}}$, $\forall \alpha \in \mathbb{R}$ and fixed $\vec{\boldsymbol{a}}$, $\vec{\boldsymbol{b}}$ is *not* a vector space.
4. If you defined addition of ordered pairs $\boldsymbol{x} = (x_1, x_2)$ as usual but scalar multiplication by $\alpha\boldsymbol{x} = (\alpha x_1, x_2)$, would it be possible to represent any vector $\boldsymbol{x}$ as a linear combination of two basis vectors $\boldsymbol{a}$ and $\boldsymbol{b}$?
5. Show that the line segment connecting the middle points of two sides of a triangle is parallel to and equal to half of the third side using methods of plane geometry and using vectors.
6. Show that the medians of a triangle intersect at the same point which is 2/3 of the way down from the vertices along each median (a median is a line that connects a vertex to the middle of the opposite side). Do this using both geometrical methods and vectors.
7. Given three point A, B, C, not co-linear, find a point O such that $\overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC} = 0$. Show that the line through A and O cuts BC at its mid-point. Deduce similar results for the other sides of the triangle ABC and therefore that O is the point of intersection of the medians. Sketch. [Hint: $\overrightarrow{OB} = \overrightarrow{OA} + \overrightarrow{AB}$, $\overrightarrow{OC} = \cdots$]
8. Given four points A, B, C, D not co-planar, find a point O such that $\overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC} + \overrightarrow{OD} = 0$. Show that the line through A and O intersects the triangle BCD at its center of area. Deduce similar results for the other faces and therefore that the medians of the tetrahedron $ABCD$, defined as the lines joining each vertex to the center of area of the opposite triangle, all intersect at the same point O which is 3/4 of the way down from the vertices along the medians. Visualize.
9. Find a basis for $\mathbb{R}^n$ (consider the *natural basis:* $\boldsymbol{e}_1 = (1, 0, \cdots, 0)$, $\boldsymbol{e}_2 = (0, 1, 0, \cdots, 0)$, etc.)
10. Find a basis for $\mathbb{C}^n$. What is the dimension of that space?
11. What is the dimension of the vector space of real continuous functions $f(x)$ in $0 < x < 1$?
12. What could be a basis for the vector space of 'nice' functions $f(x)$ in $(0, 1)$? (*i.e.* $0 < x < 1$) (what's a nice function? *smooth* functions are infinitely differentiable, that's nice!)

Partial solutions to problems 5 and 6:



Let D and E be the midpoints of segments CB and CA. *Geometry:* consider the triangles ABC and EDC. They are similar (why?), so $\overline{ED} = \overline{AB}/2$ and those two segments are parallel. Next consider triangles BAG and EDG. Those are similar too (why?), so $\overline{AG} = 2\overline{GD}$ and $\overline{BG} = 2\overline{GE}$, done! *Vector Algebra:* Let $\overrightarrow{CA} = \vec{\boldsymbol{a}}$ and $\overrightarrow{CB} = \vec{\boldsymbol{b}}$, then (1) $\overrightarrow{ED} = -\vec{\boldsymbol{a}}/2 + \vec{\boldsymbol{b}}/2 = \overrightarrow{AB}/2$, done! (2) $\overrightarrow{AD} = -\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}}/2$ and $\overrightarrow{BE} = -\vec{\boldsymbol{b}} + \vec{\boldsymbol{a}}/2$. Next, $\overrightarrow{CG} = \vec{\boldsymbol{a}} + \alpha\,\overrightarrow{AD} = \vec{\boldsymbol{b}} + \beta\,\overrightarrow{BE}$ for some α, β. Writing this equality in terms of $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}$ yields $\alpha = \beta = 2/3$, done! (Why does this imply that the line through C and G cuts AB at its midpoint?).

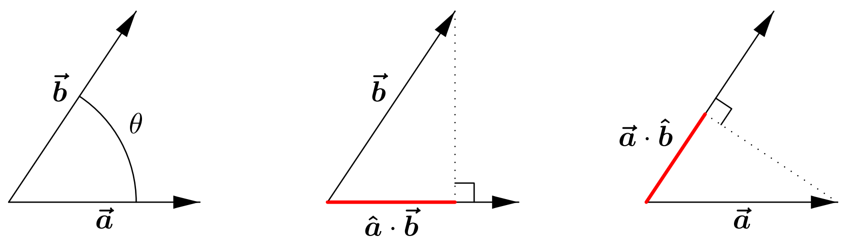
No need to introduce cartesian coordinates, that would be horrible and full of unnecessary algebra. The vector solution refers to a vector basis: $\vec{\boldsymbol{a}} = \overrightarrow{CA}$ and $\vec{\boldsymbol{b}} = \overrightarrow{CB}$ which is perfect for the problem, although it is not orthogonal!

1.3 Dot (*a.k.a.* Scalar or Inner) Product

The geometric definition of the dot product of our arrow vectors is

$$\boxed{\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}} \equiv |\vec{\boldsymbol{a}}|\ |\vec{\boldsymbol{b}}|\ \cos\theta,} \tag{11}$$

where $0 \le \theta \le \pi$ is the angle between the vectors $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}$ when their tails coincide. The dot product is a real number such that $\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}} = 0$ iff $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}$ are *orthogonal* (perpendicular). The $\vec{\boldsymbol{0}}$ vector is considered orthogonal to any vector. The dot product of any vector with itself is the square of its length $\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{a}} = |\vec{\boldsymbol{a}}|^2$. The dot product is directly related to the *perpendicular projections* of $\vec{\boldsymbol{b}}$ onto $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{a}}$ onto $\vec{\boldsymbol{b}}$. The latter are, respectively,

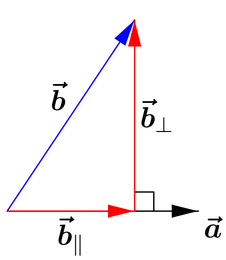


$$|\vec{\boldsymbol{b}}|\cos\theta = \frac{\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}}}{|\vec{\boldsymbol{a}}|} = \vec{\boldsymbol{b}} \cdot \hat{\boldsymbol{a}}, \qquad |\vec{\boldsymbol{a}}|\cos\theta = \frac{\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}}}{|\vec{\boldsymbol{b}}|} = \vec{\boldsymbol{a}} \cdot \hat{\boldsymbol{b}}, \tag{12}$$

where $\hat{\boldsymbol{a}} = \vec{\boldsymbol{a}}/|\vec{\boldsymbol{a}}|$ and $\hat{\boldsymbol{b}} = \vec{\boldsymbol{b}}/|\vec{\boldsymbol{b}}|$ are the *unit* vectors in the $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}$ directions, respectively. A *unit vector* is a vector of length one $|\hat{\boldsymbol{a}}| = 1\ \forall \vec{\boldsymbol{a}} \neq 0$. Curiously, *unit* vectors do not have *physical* units, that is if $\vec{\boldsymbol{a}}$ is a displacement with (physical) units of length, then $\hat{\boldsymbol{a}}$ is a pure *direction* vector. For instance if $\vec{\boldsymbol{a}} =$ "move northeast 3 miles" then $|\vec{\boldsymbol{a}}| = 3$ miles, and $\hat{\boldsymbol{a}} =$ "northeast".
In physics, the work W done by a force $\vec{\boldsymbol{F}}$ on a particle undergoing the displacement $\vec{\boldsymbol{d}}$ is equal to distance $|\vec{\boldsymbol{d}}|$ times the *component of* $\vec{\boldsymbol{F}}$ *in the direction of* $\vec{\boldsymbol{d}}$, but that is equal to the total force $|\vec{\boldsymbol{F}}|$ times the component of the *displacement* $\vec{\boldsymbol{d}}$ in the direction of $\vec{\boldsymbol{F}}$. Both of these statements are contained in the symmetric definition $W = \vec{\boldsymbol{F}} \cdot \vec{\boldsymbol{d}}$, see exercise 1 below.

Parallel and Perpendicular Components

We often want to decompose a vector $\vec{\boldsymbol{b}}$ into *vector* components, $\vec{\boldsymbol{b}}_\parallel$ and $\vec{\boldsymbol{b}}_\perp$, parallel and perpendicular to a vector $\vec{\boldsymbol{a}}$, respectively, such that $\vec{\boldsymbol{b}} = \vec{\boldsymbol{b}}_\parallel + \vec{\boldsymbol{b}}_\perp$ with

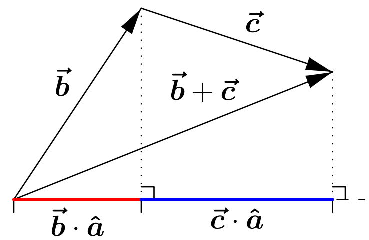


$$\begin{aligned} \vec{\boldsymbol{b}}_\parallel &= (\vec{\boldsymbol{b}} \cdot \hat{\boldsymbol{a}})\,\hat{\boldsymbol{a}} = \frac{\vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{a}}}{\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{a}}}\vec{\boldsymbol{a}} \\ \vec{\boldsymbol{b}}_\perp &= \vec{\boldsymbol{b}} - \vec{\boldsymbol{b}}_\parallel = \vec{\boldsymbol{b}} - (\vec{\boldsymbol{b}} \cdot \hat{\boldsymbol{a}})\hat{\boldsymbol{a}} = \vec{\boldsymbol{b}} - \frac{\vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{a}}}{\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{a}}}\vec{\boldsymbol{a}} \end{aligned} \tag{13}$$

Properties of the dot product

The dot product has the following properties, most of which are immediate:

1. $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$,
2. $(\alpha\vec{a}) \cdot \vec{b} = \vec{a} \cdot (\alpha\vec{b}) = \alpha(\vec{a} \cdot \vec{b})$,
3. $\vec{a} \cdot (\vec{b} + \vec{c}) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}$,
4. $\vec{a} \cdot \vec{a} \geq 0, \quad \vec{a} \cdot \vec{a} = 0 \Leftrightarrow \vec{a} = \vec{0}$,
5. $(\vec{a} \cdot \vec{b})^2 \leq (\vec{a} \cdot \vec{a})\,(\vec{b} \cdot \vec{b})$ *(Cauchy-Schwarz)*



To verify the distributive property $\vec{a} \cdot (\vec{b} + \vec{c}) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}$ geometrically, note that the magnitude of $\vec{a}$ drops out so all we need to check is $\hat{a} \cdot \vec{b} + \hat{a} \cdot \vec{c} = \hat{a} \cdot (\vec{b} + \vec{c})$ or in other words, that the perpendicular projections of $\vec{b}$ and $\vec{c}$ onto a line parallel to $\vec{a}$ add up to the projection of $\vec{b} + \vec{c}$ onto that line. This is obvious from the figure. In general, $\vec{a}$, $\vec{b}$ and $\vec{c}$ are not in the same plane. To visualize the 3D case interpret $\hat{a} \cdot \vec{b}$ as the (signed) distance between two *planes* perpendicular to $\vec{a}$ that pass through the tail and head of $\vec{b}$ and likewise for $\hat{a} \cdot \vec{c}$ and $\hat{a} \cdot (\vec{b} + \vec{c})$. The result follows directly since $\vec{b}$, $\vec{c}$ and $\vec{b} + \vec{c}$ form a triangle. So the picture is the same but the dotted lines represent planes seen from the sides and $\vec{b}$, $\vec{c}$, $\vec{b} + \vec{c}$ are not in the plane of the paper, in general.

Exercises

1. A skier slides down an inclined plane with a total vertical drop of h, show that the work done by gravity is independent of the slope. Use $\vec{F}$ and $\vec{d}$'s and sketch the geometry of this result.

2. Visualize the solutions of $\vec{a} \cdot \vec{x} = \alpha$, where $\vec{a}$ and α are known.

3. Sketch $\vec{c} = \vec{a} + \vec{b}$ then calculate $\vec{c} \cdot \vec{c} = (\vec{a} + \vec{b}) \cdot (\vec{a} + \vec{b})$ and deduce the 'law of cosines'.

4. If $\vec{n}$ is a unit vector, show that $\vec{a}_\perp \equiv \vec{a} - (\vec{a} \cdot \vec{n})\vec{n}$ is orthogonal to $\vec{n}$, $\forall\ \vec{a}$. Sketch.

5. If $\vec{c} = \vec{a} + \vec{b}$ show that $\vec{c}_\perp = \vec{a}_\perp + \vec{b}_\perp$ (defined in previous exercise). Interpret geometrically.

6. $\vec{B}$ is a magnetic field and $\vec{v}$ is the velocity of a particle. We want to decompose $\vec{v} = \vec{v}_\perp + \vec{v}_\parallel$ where $\vec{v}_\perp$ is perpendicular to the magnetic field and $\vec{v}_\parallel$ is parallel to it. Derive vector expressions for $\vec{v}_\perp$ and $\vec{v}_\parallel$.

7. Show that the three normals dropped from the vertices of a triangle perpendicular to their opposite sides intersect at the same point. [Hint: this is similar to problem 6 in section 1.2 but now $\overrightarrow{AD}$ and $\overrightarrow{BE}$ are defined by $\overrightarrow{AD} \cdot \overrightarrow{CB} = 0$ and $\overrightarrow{BE} \cdot \overrightarrow{CA} = 0$ and the goal is to show that $\overrightarrow{CG} \cdot \overrightarrow{AB} = 0$].

8. A and B are two points on a sphere of radius R specified by their longitude and latitude. Find the shortest distance between A and B, traveling on the sphere. [If O is the center of the sphere consider $\overrightarrow{OA} \cdot \overrightarrow{OB}$ to determine their angle].

9. Consider $\vec{v} = \vec{a} + t\vec{b}$ where $t \in \mathbb{R}$. What is the minimum $|\vec{v}|$ and for what t? Solve two ways: geometrically and using calculus.

1.4 Orthonormal basis

Given an arbitrary vector $\vec{v}$ and three non co-planar vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$ in $\mathbf{E}^3$, you can find the three scalars α, β and γ such that

$$\vec{v} = \alpha\vec{a} + \beta\vec{b} + \gamma\vec{c}$$

by a geometric construction (sect. 1.2). The scalars α, β and γ are called the *components* of $\vec{v}$ in the basis $\vec{a}$, $\vec{b}$, $\vec{c}$. Finding those components is much simpler if the basis is *orthogonal*, *i.e.* $\vec{a}\cdot\vec{b} = \vec{b}\cdot\vec{c} = \vec{c}\cdot\vec{a} = 0$. In that case, take the dot product of both sides of the equation $\vec{v} = \alpha\vec{a} + \beta\vec{b} + \gamma\vec{c}$ with each of the 3 basis vectors and show that

$$\alpha = \frac{\vec{a}\cdot\vec{v}}{\vec{a}\cdot\vec{a}}, \quad \beta = \frac{\vec{b}\cdot\vec{v}}{\vec{b}\cdot\vec{b}}, \quad \gamma = \frac{\vec{c}\cdot\vec{v}}{\vec{c}\cdot\vec{c}}.$$

An *orthonormal basis* is even better. That's a basis for which the vectors are mutually orthogonal and of unit norm. Such a basis is often denoted[1] $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$. Its compact definition is

$$\boxed{\vec{e}_i \cdot \vec{e}_j = \delta_{ij}} \tag{14}$$

where $i, j = 1, 2, 3$ and δ_{ij} is the **Kronecker symbol**, $\delta_{ij} = 1$ if $i = j$ and 0 if $i \neq j$.
The components of a vector $\vec{v}$ with respect to the orthonormal basis $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$ in $\mathbf{E}^3$ are the real numbers v_1, v_2, v_3 such that

$$\left\{ \begin{aligned} \vec{v} &= v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3 \equiv \sum_{i=1}^{3} v_i\vec{e}_i \\ v_i &= \vec{e}_i \cdot \vec{v}, \quad \forall i = 1, 2, 3. \end{aligned} \right. \tag{15}$$

If two vectors $\vec{a}$ and $\vec{b}$ are expanded in terms of $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$, *i.e.*

$$\vec{a} = a_1\vec{e}_1 + a_2\vec{e}_2 + a_3\vec{e}_3, \qquad \vec{b} = b_1\vec{e}_1 + b_2\vec{e}_2 + b_3\vec{e}_3,$$

use the properties of the dot product and the orthonormality of the basis to show that

$$\boxed{\vec{a}\cdot\vec{b} = a_1b_1 + a_2b_2 + a_3b_3.} \tag{16}$$

$\triangleright$ Show that this formula is valid *only* for orthonormal bases.
One remarkable property of this formula is that its value is independent of the orthonormal basis. The dot product is a geometric property of the vectors $\vec{a}$ and $\vec{b}$, independent of the basis. This is obvious from the geometric definition (11) but not from its expression in terms of components (16). If $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$ and $\vec{e}_1{}'$, $\vec{e}_2{}'$, $\vec{e}_3{}'$ are two distinct orthogonal bases then

$$\vec{a} = a_1\vec{e}_1 + a_2\vec{e}_2 + a_3\vec{e}_3 = a_1'\vec{e}_1{}' + a_2'\vec{e}_2{}' + a_3'\vec{e}_3{}'$$

but, in general, the components in the two bases are distinct: $a_1 \neq a_1'$, $a_2 \neq a_2'$, $a_3 \neq a_3'$, and likewise for another vector $\vec{b}$, yet

$$\vec{a}\cdot\vec{b} = a_1b_1 + a_2b_2 + a_3b_3 = a_1'b_1' + a_2'b_2' + a_3'b_3'.$$

The simple algebraic form of the dot product is *invariant under a change of orthonormal basis.*

[1] Note about notation: Forget about the notation $\vec{i}$, $\vec{j}$, $\vec{k}$. This is old 19th century notation, it is unfortunately still very common in elementary courses but that old notation will get in the way if you stick to it. We will NEVER use $\vec{i}$, $\vec{j}$, $\vec{k}$, instead we will use $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$ or $\vec{e}_x$, $\vec{e}_y$, $\vec{e}_z$ to denote a set of three orthonormal vectors in 3D euclidean space. We will soon use **indices** i, j and k (next line already!). Those indices are positive integers that can take all the values from 1 to n, the dimension of the space. We spend most of our time in 3D space, so *most* of the time the possible values for these indices i, j and k are 1, 2 and 3. We will use those indices *a lot!*. They should not be confused with those old orthonormal vectors $\vec{i}$, $\vec{j}$, $\vec{k}$ from elementary calculus.

Exercises

1. If $\vec{\boldsymbol{w}} = \sum_{i=1}^{3} w_i \vec{\boldsymbol{e}}_i$, calculate $\vec{\boldsymbol{e}}_j \cdot \vec{\boldsymbol{w}}$ using $\sum$ notation and (14).
2. Why is not true that $\vec{\boldsymbol{e}}_i \cdot \sum_{i=1}^{3} w_i \vec{\boldsymbol{e}}_i = \sum_{i=1}^{3} w_i (\vec{\boldsymbol{e}}_i \cdot \vec{\boldsymbol{e}}_i) = \sum_{i=1}^{3} w_i \delta_{ii} = w_1 + w_2 + w_3$?
3. If $\vec{\boldsymbol{v}} = \sum_{i=1}^{3} v_i \vec{\boldsymbol{e}}_i$ and $\vec{\boldsymbol{w}} = \sum_{i=1}^{3} w_i \vec{\boldsymbol{e}}_i$, calculate $\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{w}}$ using $\sum$ notation and (14).
4. If $\vec{\boldsymbol{v}} = \sum_{i=1}^{3} v_i \vec{\boldsymbol{a}}_i$ and $\vec{\boldsymbol{w}} = \sum_{i=1}^{3} w_i \vec{\boldsymbol{a}}_i$, where the basis $\vec{\boldsymbol{a}}_i$, $i = 1, 2, 3$, is *not* orthonormal, calculate $\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{w}}$.
5. Calculate (i) $\sum_{j=1}^{3} \delta_{ij} a_j$, (ii) $\sum_{i=1}^{3} \sum_{j=1}^{3} \delta_{ij} a_j b_i$, (iii) $\sum_{j=1}^{3} \delta_{jj}$.

Definition of dot product for $\mathbb{R}^n$

The geometric definition of the dot product (11) is great for oriented line segments as it emphasizes the geometric aspects, but the algebraic formula (16) is very useful for calculations. It's also the way to define the dot product for other vector spaces where the concept of 'angle' between vectors may not be obvious *e.g.* what is the angle between the vectors (1,2,3,4) and (4,3,2,1) in $\mathbb{R}^4$?! The dot product (*a.k.a. scalar* product or *inner* product) of the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^n$ is defined as suggested by (16):

$$\boldsymbol{x} \cdot \boldsymbol{y} \equiv x_1 y_1 + x_2 y_2 + \cdots + x_n y_n. \tag{17}$$

Verify that this definition satisfies the first 4 properties of the dot product. To show the *Cauchy-Schwarz* property, you need a bit of Calculus and a classical trick: consider $\boldsymbol{v} = \boldsymbol{x} + \lambda \boldsymbol{y}$, then by prop 4 of the dot product: $(\boldsymbol{x}+\lambda\boldsymbol{y})\cdot(\boldsymbol{x}+\lambda\boldsymbol{y}) \geq 0$, and by props 1,2,3, $F(\lambda) \equiv (\boldsymbol{x}+\lambda\boldsymbol{y})\cdot(\boldsymbol{x}+\lambda\boldsymbol{y}) = \lambda^2 \boldsymbol{y}\cdot\boldsymbol{y} + 2\lambda \boldsymbol{x}\cdot\boldsymbol{y} + \boldsymbol{x}\cdot\boldsymbol{x}$. For given, but arbitrary, $\boldsymbol{x}$ and $\boldsymbol{y}$, this is a quadratic polynomial in λ. That polynomial $F(\lambda)$ has a single minimum at $\lambda = -(\boldsymbol{x}\cdot\boldsymbol{y})/(\boldsymbol{y}\cdot\boldsymbol{y})$. Find that minimum value of $F(\lambda)$ and deduce the Cauchy-Schwarz inequality. Once we know that the definition (17) satisfies Cauchy-Schwarz, $(\boldsymbol{x}\cdot\boldsymbol{y})^2 \leq (\boldsymbol{x}\cdot\boldsymbol{x})\,(\boldsymbol{y}\cdot\boldsymbol{y})$, we can define the length of a vector by $|\boldsymbol{x}| = (\boldsymbol{x}\cdot\boldsymbol{x})^{1/2}$ (this is called the *Euclidean* length since it corresponds to length in Euclidean geometry by Pythagoras's theorem) and the angle θ between two vectors in $\mathbb{R}^n$ by $\cos\theta = (\boldsymbol{x}\cdot\boldsymbol{y})/(|\boldsymbol{x}|\,|\boldsymbol{y}|)$. A vector space for which a dot (or inner) product is defined is called a *Hilbert space* (or an *inner product space*).

▷ So what is the angle between $(1, 2, 3, 4)$ and $(4, 3, 2, 1)$?

▷ Can you define a dot product for the vector space of real functions $f(x)$?

The bottom line is that for more complex vector spaces, the dot (or scalar or inner) product is a key mathematical construct that allows us to generalize the concept of 'angle' between vectors and, most importantly, to define 'orthogonal vectors'.

▷ Find a vector orthogonal to $(1, 2, 3, 4)$. Find all the vectors orthogonal to $(1, 2, 3, 4)$.

▷ Decompose (4,2,1,7) into the sum of two vectors one of which is parallel and the other perpendicular to $(1, 2, 3, 4)$.

▷ Show that $\cos nx$ with n integer, is a set of orthogonal functions on $(0, \pi)$. Find formulas for the components of a function $f(x)$ in terms of that orthogonal basis. In particular, find the components of $\sin x$ in terms of the cosine basis in that $(0, \pi)$ interval.

Norm of a vector

The *norm* of a vector, denoted $\|\boldsymbol{a}\|$, is a positive real number that defines its size or 'length' (but not in the sense of the number of its components). For displacement vectors in Euclidean spaces, the norm is the length of the displacement, $\|\vec{\boldsymbol{a}}\| = |\vec{\boldsymbol{a}}|$ *i.e.* the distance between point A and B if $\overrightarrow{AB} = \boldsymbol{a}$. The following properties are geometrically straightforward for length of displacement vectors:

1. $\|\boldsymbol{a}\| \geq 0$ and $\|\boldsymbol{a}\| = 0 \Leftrightarrow \boldsymbol{a} = \vec{\boldsymbol{0}}$,
2. $\|\alpha \boldsymbol{a}\| = |\alpha|\, \|\boldsymbol{a}\|$,
3. $\|\boldsymbol{a} + \boldsymbol{b}\| \leq \|\boldsymbol{a}\| + \|\boldsymbol{b}\|$. (triangle inequality)

Draw the triangle formed by $\vec{\boldsymbol{a}}$, $\vec{\boldsymbol{b}}$ and $\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}}$ to see why the latter is called the *triangle inequality.* For more general vector spaces, these properties become the *defining properties* (*axioms*) that a norm must satisfy. A vector space for which a norm is defined is called a *Banach space.*

Definition of norm for $\mathbb{R}^n$

For other types of vector space, there are many possible definitions for the norm of a vector as long as those definitions satisfy the 3 norm properties. In $\mathbb{R}^n$, the p-norm of vector $\vec{\boldsymbol{x}}$ is defined by the positive number

$$\|\boldsymbol{x}\|_p \equiv \Big(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\Big)^{1/p}, \tag{18}$$

where $p \geq 1$ is a real number. Commonly used norms are the 2-norm $\|\boldsymbol{x}\|_2$ which is the square root of the sum of the squares, the 1-norm $\|\boldsymbol{x}\|_1$ (sum of absolute values) and the infinity norm, $\|\boldsymbol{x}\|_\infty$, defined as the limit as $p \to \infty$ of the above expression.

Note that the 2-norm $\|\boldsymbol{x}\|_2 = (\boldsymbol{x} \cdot \boldsymbol{x})^{1/2}$ and for that reason is also called the *Euclidean norm.* In fact, if a dot product is defined, then a norm can always be defined as the square root of the dot product. In other words, *every Hilbert space is a Banach space,* but the converse is not true.

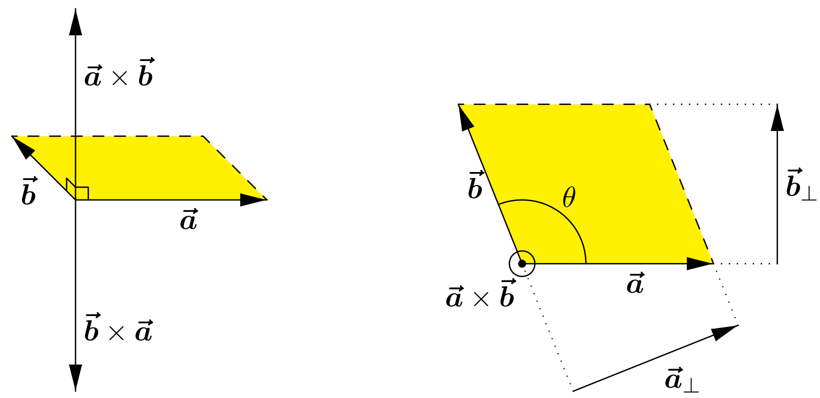
- ▷ Show that the infinity norm $\|\boldsymbol{x}\|_\infty = \max_i |x_i|$.
- ▷ Show that the p-norm satisfies the three norm properties for $p = 1, 2, \infty$.
- ▷ Define a norm for $\mathbb{C}^n$.
- ▷ Define the 2-norm for real functions $f(x)$ in $0 < x < 1$.

1.5 Cross (*a.k.a.* Vector or Area) Product

The cross product is a very useful operation for physical applications (mechanics, electromagnetism), but it is particular to 3D space. The cross product of two vectors $\vec{\boldsymbol{a}}$, $\vec{\boldsymbol{b}}$ is the vector denoted $\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}$ that is (1) orthogonal to both $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}$, (2) has magnitude equal to the area of the parallelogram with sides $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}$, (3) has direction determined by the *right-hand rule* (or the *cork-screw rule*), *i.e.*

$$\begin{gathered}(\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}) \cdot \vec{\boldsymbol{a}} = (\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}) \cdot \vec{\boldsymbol{b}} = 0, \\ |\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}| = |\vec{\boldsymbol{a}}|\, |\vec{\boldsymbol{b}}|\, \sin\theta = \text{area of parallelogram}, \\ \vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}, \vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}} \text{ is right-handed,}\end{gathered} \tag{19}$$

where θ is the angle between $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}$ as defined for the dot product. (Is $\sin\theta \geq 0$?) The following figure illustrates the cross-product with a perspective view (left) and a top view (right), with $\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}$ out of the paper in the top view.

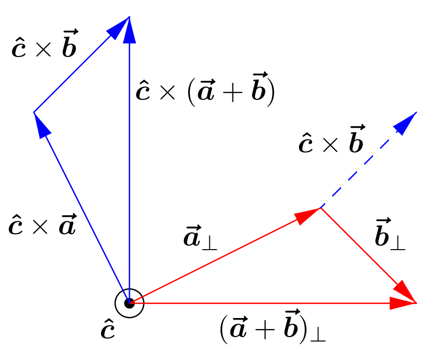


$$\vec{a} \times \vec{b} = \vec{a}_\perp \times \vec{b} = \vec{a} \times \vec{b}_\perp \tag{20}$$

where $\vec{a}_\perp = \vec{a} - (\vec{a} \cdot \hat{b})\hat{b}$ is the vector component of $\vec{a}$ perpendicular to $\vec{b}$ and likewise $\vec{b}_\perp = \vec{b} - (\vec{b} \cdot \hat{a})\hat{a}$ is the vector component of $\vec{b}$ perpendicular to $\vec{a}$ (so the meaning of $\perp$ is relative). The cross-product has the following properties:

1. $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$, (anti-commutativity) $\Rightarrow \vec{a} \times \vec{a} = 0$,
2. $(\alpha \vec{a}) \times \vec{b} = \vec{a} \times (\alpha \vec{b}) = \alpha(\vec{a} \times \vec{b})$,
3. $\vec{c} \times (\vec{a} + \vec{b}) = (\vec{c} \times \vec{a}) + (\vec{c} \times \vec{b})$

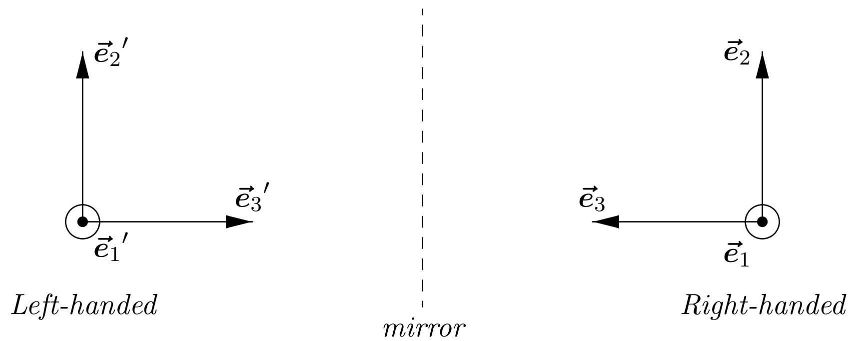
So we manipulate the cross product as we'd expect except for the anti-commutativity, which is a big difference from our other elementary products! The first 2 properties are geometrically obvious from the definition. To show the third property (distributivity) let $\vec{c} = |\vec{c}|\hat{c}$ and get rid of $|\vec{c}|$ by property 2. All three cross products give vectors perpendicular to $\vec{c}$ and furthermore from (20) we have $\hat{c} \times \vec{a} = \hat{c} \times \vec{a}_\perp$, $\hat{c} \times \vec{b} = \hat{c} \times \vec{b}_\perp$ and $\hat{c} \times (\vec{a} + \vec{b}) = \hat{c} \times (\vec{a} + \vec{b})_\perp$, where $\perp$ means vector component perpendicular to $\hat{c}$, that is $\vec{a}_\perp = \vec{a} - (\vec{a} \cdot \hat{c})\hat{c}$, etc. So cross-products with $\vec{c}$ eliminate the components parallel to $\vec{c}$ and *all the action is in the plane perpendicular to* $\vec{c}$.



To visualize the distributivity property it suffices to look at that plane from the top, with $\hat{c}$ pointing out of the paper/screen. Then a cross product with $\hat{c}$ is a *rotation of the perpendicular components by* $\pi/2$ *counterclockwise.* Since $\vec{a}$, $\vec{b}$ and $\vec{a} + \vec{b}$ form a triangle, their perpendicular projections $\vec{a}_\perp$, $\vec{b}_\perp$ and $(\vec{a} + \vec{b})_\perp$ form a triangle and therefore $\hat{c} \times \vec{a}$, $\hat{c} \times \vec{b}$ and $\hat{c} \times (\vec{a} + \vec{b})$ also form a triangle since each of them is simply a counterclockwise rotation by $\pi/2$ of $\vec{a}_\perp$, $\vec{b}_\perp$ and $(\vec{a} + \vec{b})_\perp$, respectively. This demonstrates distributivity.

Orientation of Bases

If we pick an arbitrary unit vector $\vec{e}_1$, then a unit vector $\vec{e}_2$ orthogonal to $\vec{e}_1$, then there are two possible unit vectors $\vec{e}_3$ orthogonal to both $\vec{e}_1$ and $\vec{e}_2$. One choice gives a *right-handed basis* (*i.e.* $\vec{e}_1$ in right thumb direction, $\vec{e}_2$ in right index direction and $\vec{e}_3$ in right major direction). The other choice gives a *left-handed basis.* These two types of bases are *mirror images* of each other as illustrated in the following figure, where $\vec{e}_1{}' = \vec{e}_1$ point straight out of the paper (or screen).



This figure reveals an interesting subtlety of the cross product. For this particular choice of left and right handed bases (other arrangements are possible of course), $\vec{e}_1{}' = \vec{e}_1$ and $\vec{e}_2{}' = \vec{e}_2$ but $\vec{e}_3{}' = -\vec{e}_3$ so $\vec{e}_1 \times \vec{e}_2 = \vec{e}_3$ and $\vec{e}_1{}' \times \vec{e}_2{}' = \vec{e}_3 = -\vec{e}_3{}'$. This indicates that the mirror image of the cross-product is *not* the cross-product of the mirror images. On the opposite, the mirror image of the cross-product $\vec{e}_3{}'$ is *minus* the cross-product of the images $\vec{e}_1{}' \times \vec{e}_2{}'$. We showed this for a special case, but this is general, the cross-product is not invariant under reflection, it changes sign. Physical laws should not depend on the choice of basis, so this implies that they should not be expressed in terms of an *odd* number of cross products. When we write that the velocity of a particle is $\vec{v} = \vec{\omega} \times \vec{r}$, $\vec{v}$ and $\vec{r}$ are 'good' vectors (reflecting as they should under mirror symmetry) but $\vec{\omega}$ is not quite a true vector, it is a *pseudo-vector*. It changes sign under reflection. That is because rotation vectors are themselves defined according to the right-hand rule, so an expression such as $\vec{\omega} \times \vec{r}$ actually contains two applications of the right hand rule. Likewise in the Lorentz force $\vec{F} = q\vec{v} \times \vec{B}$, $\vec{F}$ and $\vec{v}$ are good vectors, but since the definition involves a cross-product, it must be that $\vec{B}$ is a pseudo-vector. Indeed $\vec{B}$ is itself a cross-product so the definition of $\vec{F}$ actually contains two cross-products.

The orientation (right-handed or left-handed) did not matter to us before but, now that we've defined the cross-product with the right-hand rule, we'll typically choose right-handed bases. We don't have to, geometrically speaking, but we need to from an algebraic point of view otherwise we'd need two sets of algebraic formula, one for right-handed bases and one for left-handed bases. In terms of our right-handed cross product definition, we can define a right-handed basis by

$$\vec{e}_1 \times \vec{e}_2 = \vec{e}_3, \tag{21}$$

then deduce geometrically

$$\vec{e}_2 \times \vec{e}_3 = \vec{e}_1, \qquad \vec{e}_3 \times \vec{e}_1 = \vec{e}_2, \tag{22}$$

$$\vec{e}_2 \times \vec{e}_1 = -\vec{e}_3, \quad \vec{e}_1 \times \vec{e}_3 = -\vec{e}_2, \quad \vec{e}_3 \times \vec{e}_2 = -\vec{e}_1. \tag{23}$$

Note that (22) are *cyclic* rotations of the basis vectors in (21), *i.e.* $(\vec{e}_1, \vec{e}_2, \vec{e}_3) \to (\vec{e}_2, \vec{e}_3, \vec{e}_1) \to (\vec{e}_3, \vec{e}_1, \vec{e}_2)$. The orderings of the basis vectors in (23) do not correspond to cyclic rotations. For 3 elements, a *cyclic* rotation corresponds to an *even* number of permutations. For instance to go from $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ to $(\vec{e}_2, \vec{e}_3, \vec{e}_1)$ we can first permute (switch) $\vec{e}_1 \leftrightarrow \vec{e}_2$ to obtain $(\vec{e}_2, \vec{e}_1, \vec{e}_3)$ then permute $\vec{e}_1$ and $\vec{e}_3$. The concept of *even* and *odd* number of permutations is more general. But for three elements it is useful to think in terms of cyclic and acyclic permutations.

If we expand $\vec{a}$ and $\vec{b}$ in terms of the right-handed $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$, then apply the 3 properties of the cross-product *i.e.* in compact summation form

$$\vec{a} = \sum_{i=1}^{3} a_i \vec{e}_i, \quad \vec{b} = \sum_{j=1}^{3} b_j \vec{e}_j, \quad \Rightarrow \vec{a} \times \vec{b} = \sum_{i=1}^{3} \sum_{j=1}^{3} a_i b_j \, (\vec{e}_i \times \vec{e}_j),$$

we obtain

$$\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}} = (a_2 b_3 - a_3 b_2)\vec{\boldsymbol{e}}_1 + (a_3 b_1 - a_1 b_3)\vec{\boldsymbol{e}}_2 + (a_1 b_2 - a_2 b_1)\vec{\boldsymbol{e}}_3. \tag{24}$$

Verify this result explicitly. What would the formula be if the basis was left-handed?
That expansion of the cross product *with respect to a right-handed orthonormal basis* (24) is often remembered using the formal determinant (*i.e.* this is not a true determinant, it's just convenient mnemonics)

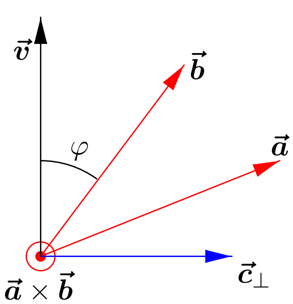
$$\begin{vmatrix} \vec{\boldsymbol{e}}_1 & \vec{\boldsymbol{e}}_2 & \vec{\boldsymbol{e}}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.$$

Double vector product ('Triple vector product')[2]

► *Exercise:* Visualize the vector $(\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}) \times \vec{\boldsymbol{a}}$. Sketch it. What are its geometric properties? What is its magnitude?
Double vector products occur frequently in applications (*e.g.* angular momentum of a rotating body) directly or indirectly (recall above discussion about mirror reflection and cross-products in physics). They have simple expressions

$$\begin{aligned} (\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}) \times \vec{\boldsymbol{c}} &= (\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{c}})\, \vec{\boldsymbol{b}} - (\vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{c}})\, \vec{\boldsymbol{a}}, \\ \vec{\boldsymbol{a}} \times (\vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}}) &= (\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{c}})\, \vec{\boldsymbol{b}} - (\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}})\vec{\boldsymbol{c}}. \end{aligned} \tag{25}$$

One follows from the other after some manipulations and renaming of vectors, but we can remember both at once as: *middle vector times dot product of the other two minus other vector within parentheses times dot product of the other two.*[3]
Let's show this identity for $\vec{\boldsymbol{v}} = (\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}) \times \vec{\boldsymbol{c}}$. *First,* $(\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}})$ is $\perp$ to $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}$, and $\vec{\boldsymbol{v}}$ is $\perp$ to $(\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}})$, therefore $\vec{\boldsymbol{v}}$ is in the $\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}$ plane and can be written $\vec{\boldsymbol{v}} = \alpha\vec{\boldsymbol{a}} + \beta\vec{\boldsymbol{b}}$ for some scalars α and β. *Second,* $\vec{\boldsymbol{v}}$ is also $\perp$ to $\vec{\boldsymbol{c}}$, so $\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{c}} = \alpha(\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{c}}) + \beta(\vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{c}}) = 0$, therefore $\alpha = \mu(\vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{c}})$ and $\beta = -\mu(\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{c}})$ for some scalar μ. Thus $\vec{\boldsymbol{v}} = \mu[(\vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{c}})\vec{\boldsymbol{a}} - (\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{c}})\vec{\boldsymbol{b}}]$ which is almost (25) but we still need to show that $\mu = -1$.



Let $\vec{\boldsymbol{c}} = \vec{\boldsymbol{c}}_\parallel + \vec{\boldsymbol{c}}_\perp$ with $\vec{\boldsymbol{c}}_\parallel$ parallel to $\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}$, hence perpendicular to both $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}$. Then $\vec{\boldsymbol{v}} = (\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}) \times \vec{\boldsymbol{c}} = (\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}) \times \vec{\boldsymbol{c}}_\perp$ and $\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{c}} = \vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{c}}_\perp$ and $\vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{c}} = \vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{c}}_\perp$. *So all the action is in the plane perpendicular to* $\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}} = |\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}|\, \hat{\boldsymbol{n}}$. To determine μ in $\vec{\boldsymbol{v}} = \mu(\vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{c}}_\perp)\vec{\boldsymbol{a}} - \mu(\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{c}}_\perp)\vec{\boldsymbol{b}}$, consider $\vec{\boldsymbol{b}} \times \vec{\boldsymbol{v}} = -\mu\,(\vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{c}}_\perp)\,(\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}) = -\mu|\vec{\boldsymbol{b}}|\,|\vec{\boldsymbol{c}}_\perp|\,\cos(\frac{\pi}{2} - \varphi)\,(\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}) = -\mu|\vec{\boldsymbol{b}}|\,|\vec{\boldsymbol{c}}_\perp|\,\sin\varphi\,(\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}})$, but by direct calculation $\vec{\boldsymbol{b}} \times \vec{\boldsymbol{v}} = |\vec{\boldsymbol{b}}||\vec{\boldsymbol{v}}|\sin\varphi\,\hat{\boldsymbol{n}} = |\vec{\boldsymbol{b}}|\,|\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}|\,|\vec{\boldsymbol{c}}_\perp|\,\sin\varphi\,\hat{\boldsymbol{n}} = |\vec{\boldsymbol{b}}|\,|\vec{\boldsymbol{c}}_\perp|\,\sin\varphi(\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}})$ (with φ *positive counterclockwise about* $\hat{\boldsymbol{n}}$), hence $\mu = -1$.

Exercises

1. Show that $|\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}|^2 + (\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}})^2 = |\vec{\boldsymbol{a}}|^2\,|\vec{\boldsymbol{b}}|^2, \ \forall \vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}$.

2. If three vectors satisfy $\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}} + \vec{\boldsymbol{c}} = 0$, show algebraically and geometrically that $\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}} = \vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}} = \vec{\boldsymbol{c}} \times \vec{\boldsymbol{a}}$. Deduce the 'law of sines' relating the sines of the angles of a triangle and the lengths of its sides.

[2]The double vector product is often called 'triple vector product', there are 3 vectors but only 2 vector products!
[3]This is more useful than the confusing 'BAC-CAB' rule for remembering the 2nd. Try applying the BAC-CAB mnemonic to $(\vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}}) \times \vec{\boldsymbol{a}}$ for confusing fun!

3. Show by vector algebra and geometry that all the vectors $\vec{\boldsymbol{x}}$ such that $\vec{\boldsymbol{a}} \times \vec{\boldsymbol{x}} = \vec{\boldsymbol{b}}$ have the form
$$\vec{\boldsymbol{x}} = \alpha\vec{\boldsymbol{a}} + \frac{\vec{\boldsymbol{b}} \times \vec{\boldsymbol{a}}}{\|\vec{\boldsymbol{a}}\|^2}, \quad \forall\alpha \in \mathbb{R}$$

4. Show the *Jacobi identity:* $\vec{\boldsymbol{a}} \times (\vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}}) + \vec{\boldsymbol{b}} \times (\vec{\boldsymbol{c}} \times \vec{\boldsymbol{a}}) + \vec{\boldsymbol{c}} \times (\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}) = 0$.

5. If $\vec{\boldsymbol{n}}$ is any unit vector, show algebraically *and* geometrically that any vector $\vec{\boldsymbol{a}}$ can be decomposed as
$$\boxed{\vec{\boldsymbol{a}} = (\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{n}})\vec{\boldsymbol{n}} + \vec{\boldsymbol{n}} \times (\vec{\boldsymbol{a}} \times \vec{\boldsymbol{n}}) \equiv \vec{\boldsymbol{a}}_{\parallel} + \vec{\boldsymbol{a}}_{\perp}.} \tag{26}$$
The first component is parallel to $\vec{\boldsymbol{n}}$, the second is perpendicular to $\vec{\boldsymbol{n}}$.

6. A particle of charge q moving at velocity $\vec{\boldsymbol{v}}$ in a magnetic field $\vec{\boldsymbol{B}}$ experiences the Lorentz force $\vec{\boldsymbol{F}} = q\vec{\boldsymbol{v}} \times \vec{\boldsymbol{B}}$. Show that there is no force in the direction of the magnetic field.

7. A left-handed basis $\vec{\boldsymbol{e}}_1{}'$, $\vec{\boldsymbol{e}}_2{}'$, $\vec{\boldsymbol{e}}_3{}'$, is defined by $\vec{\boldsymbol{e}}_i{}' \cdot \vec{\boldsymbol{e}}_j{}' = \delta_{ij}$ and $\vec{\boldsymbol{e}}_1{}' \times \vec{\boldsymbol{e}}_2{}' = -\vec{\boldsymbol{e}}_3{}'$. Show that $(\vec{\boldsymbol{e}}_i{}' \times \vec{\boldsymbol{e}}_j{}') \cdot \vec{\boldsymbol{e}}_k{}'$ has the opposite sign to the corresponding expression for a right-handed basis, $\forall i, j, k$ (the definition of the cross-product remaining its right-hand rule self). Thus deduce that the formula for the components of the cross-product in the left handed basis would all change sign.

8. Prove (25) using the right-handed orthonormal basis $\vec{\boldsymbol{e}}_1 = \vec{\boldsymbol{a}}/|\vec{\boldsymbol{a}}|$, $\vec{\boldsymbol{e}}_3 = (\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}})/|\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}|$ and $\vec{\boldsymbol{e}}_2 = \vec{\boldsymbol{e}}_3 \times \vec{\boldsymbol{e}}_1$. Then $\vec{\boldsymbol{a}} = a_1\vec{\boldsymbol{e}}_1$, $\vec{\boldsymbol{b}} = b_1\vec{\boldsymbol{e}}_1 + b_2\vec{\boldsymbol{e}}_2$, $\vec{\boldsymbol{c}} = c_1\vec{\boldsymbol{e}}_1 + c_2\vec{\boldsymbol{e}}_2 + c_3\vec{\boldsymbol{e}}_3$. Visualize and explain why this is a general result and therefore a proof of the double cross product identity.

9. Prove (25) using the right-handed orthonormal basis $\vec{\boldsymbol{e}}_1 = \vec{\boldsymbol{c}}_{\perp}/|\vec{\boldsymbol{c}}_{\perp}|$, $\vec{\boldsymbol{e}}_3 = (\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}})/|\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}|$ and $\vec{\boldsymbol{e}}_2 = \vec{\boldsymbol{e}}_3 \times \vec{\boldsymbol{e}}_1$. In that basis $\vec{\boldsymbol{a}} = a_1\vec{\boldsymbol{e}}_1 + a_2\vec{\boldsymbol{e}}_2$ and $\vec{\boldsymbol{b}} = b_1\vec{\boldsymbol{e}}_1 + b_2\vec{\boldsymbol{e}}_2$ but what is $\vec{\boldsymbol{c}}$? Show by direct calculation that $\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}} = (a_1b_2 - a_2b_1)\vec{\boldsymbol{e}}_3$ and $\vec{\boldsymbol{v}} = |\vec{\boldsymbol{c}}_{\perp}|(a_1b_2 - a_2b_1)\vec{\boldsymbol{e}}_2 = |\vec{\boldsymbol{c}}_{\perp}|a_1\vec{\boldsymbol{b}} - |\vec{\boldsymbol{c}}_{\perp}|b_1\vec{\boldsymbol{a}}$. Why is this $(\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{c}})\vec{\boldsymbol{b}} - (\vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{c}})\vec{\boldsymbol{a}}$ and thus a proof of the identity?

1.6 Indicial notation

Levi-Civita (*a.k.a.* alternating or permutation) symbol

We have used the Kronecker symbol (14) to express all the dot products $\vec{\boldsymbol{e}}_i \cdot \vec{\boldsymbol{e}}_j = \delta_{ij}$ in a very compact form. There is a similar symbol, ϵ_{ijk}, the *Levi-Civita* or *alternating* symbol, defined as
$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } (i,j,k) \text{ is an even permutation of } (1,2,3), \\ -1 & \text{if } (i,j,k) \text{ is an odd permutation of } (1,2,3), \\ 0 & \text{otherwise,} \end{cases} \tag{27}$$
or, explicitly: $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$ and $\epsilon_{213} = \epsilon_{132} = \epsilon_{321} = -1$, all other $\epsilon_{ijk} = 0$. Recall that for 3 elements an *even* permutation is the same as a *cyclic* permutation, therefore $\epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij}$, $\forall i, j, k$ (why?). The ϵ_{ijk} symbol provides a compact expression for the components of the cross-product of right-handed basis vectors:
$$\boxed{(\vec{\boldsymbol{e}}_i \times \vec{\boldsymbol{e}}_j) \cdot \vec{\boldsymbol{e}}_k = \epsilon_{ijk}.} \tag{28}$$
but since this is the k-component of $(\vec{\boldsymbol{e}}_i \times \vec{\boldsymbol{e}}_j)$ we can also write
$$\boxed{(\vec{\boldsymbol{e}}_i \times \vec{\boldsymbol{e}}_j) = \sum_{k=1}^{3} \epsilon_{ijk}\vec{\boldsymbol{e}}_k.} \tag{29}$$

Note that there is only one non-zero term in the latter sum (but then, why can't we drop the sum?). Verify this result for yourself.

1.6.1 Sigma notation, free and dummy indices

The *expansion* of vectors $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}$ in terms of basis $\vec{\boldsymbol{e}}_1,\vec{\boldsymbol{e}}_2,\vec{\boldsymbol{e}}_3$, $\vec{\boldsymbol{a}} = a_1\vec{\boldsymbol{e}}_1 + a_2\vec{\boldsymbol{e}}_2 + a_3\vec{\boldsymbol{e}}_3$ and $\vec{\boldsymbol{b}} = b_1\vec{\boldsymbol{e}}_1 + b_2\vec{\boldsymbol{e}}_2 + b_3\vec{\boldsymbol{e}}_3$, can be written compactly using the sigma (Σ) notation

$$\vec{\boldsymbol{a}} = \sum_{i=1}^{3} a_i \vec{\boldsymbol{e}}_i, \qquad \vec{\boldsymbol{b}} = \sum_{i=1}^{3} b_i \vec{\boldsymbol{e}}_i. \tag{30}$$

We have introduced the **Kronecker** symbol δ_{ij} and the **Levi-Civita** symbol ϵ_{ijk} in order to write and perform our basic vector operations such as dot and cross products in compact forms, *when the basis is orthonormal and right-handed*, for instance using (14) and (29)

$$\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}} = \sum_{i=1}^{3}\sum_{j=1}^{3} a_i b_j\, \vec{\boldsymbol{e}}_i \cdot \vec{\boldsymbol{e}}_j = \sum_{i=1}^{3}\sum_{j=1}^{3} a_i b_j \delta_{ij} = \sum_{i=1}^{3} a_i b_i \tag{31}$$

$$\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}} = \sum_{i=1}^{3}\sum_{j=1}^{3} a_i b_j\, \vec{\boldsymbol{e}}_i \times \vec{\boldsymbol{e}}_j = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} a_i b_j \epsilon_{ijk}\, \vec{\boldsymbol{e}}_k \tag{32}$$

Note that i and j are *dummy* or *summation* indices in the sums (30) and (31), they do not have a specific value, they have *all* the possible values in their range. It is their *place* in the particular expression and their *range* that matters, not their name

$$\vec{\boldsymbol{a}} = \sum_{i=1}^{3} a_i \vec{\boldsymbol{e}}_i = \sum_{j=1}^{3} a_j \vec{\boldsymbol{e}}_j = \sum_{k=1}^{3} a_k \vec{\boldsymbol{e}}_k = \cdots \neq \sum_{k=1}^{3} a_k \vec{\boldsymbol{e}}_i \tag{33}$$

Indices come in two kinds, the *dummies* and the *free*. Here's an example

$$\vec{\boldsymbol{e}}_i \cdot (\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}})\vec{\boldsymbol{c}} = \left(\sum_{j=1}^{3} a_j b_j\right) c_i, \tag{34}$$

here j is a dummy summation index, but i is *free*, we can pick for it any value $1, 2, 3$. Freedom comes with constraints. If we use i on the left-hand side of the equation, then we have no choice, we must use i for c_i on the right hand side. By convention we try to use i, j, k, l, m, n, to denote indices, which are positive integers. Greek letters are sometimes used for indices.

Mathematical operations impose some naming constraints however. Although, we can use the same index name, i, in the expansions of $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}$, when they appear separately as in (30), we *cannot use the same index name if we multiply them* as in (31) and (32). Bad things will happen if you do, for instance

$$\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}} = \left(\sum_{i=1}^{3} a_i \vec{\boldsymbol{e}}_i\right) \times \left(\sum_{i=1}^{3} b_i \vec{\boldsymbol{e}}_i\right) = \sum_{i=1}^{3} a_i b_i\, \vec{\boldsymbol{e}}_i \times \vec{\boldsymbol{e}}_i = 0 \quad \text{(WRONG!)} \tag{35}$$

1.6.2 Einstein's summation convention

While he was developing the theory of general relativity, Einstein noticed that many of the sums that occur in calculations involve terms where the summation index appears twice. For example, i appears twice in the single sums in (30), i and j appear twice in the double sum in (31) and i, j and k each appear twice in the triple sum in (32). To facilitate such manipulations he dropped the Σ signs and adopted the **summation convention** that **a repeated index implicitly denotes a sum over all values of that index**. In a letter to a friend he wrote *"I have made a great discovery in mathematics; I have suppressed the summation sign every time that the summation must be made over an index which occurs twice"*. Thus with Einstein's summation convention we write

$$\vec{\boldsymbol{a}} = a_i\vec{\boldsymbol{e}}_i, \qquad \vec{\boldsymbol{b}} = b_j\vec{\boldsymbol{e}}_j, \qquad \vec{\boldsymbol{a}}\cdot\vec{\boldsymbol{b}} = a_ib_i, \qquad \vec{\boldsymbol{a}}\times\vec{\boldsymbol{b}} = \epsilon_{ijk}a_ib_j\vec{\boldsymbol{e}}_k, \tag{36}$$

and any repeated index implies a sum over all values of that index. This is a very useful and widely used notation but you have to use it with care and there are cases where it cannot be used. Indices can never be repeated more than twice, if they are, that's probably a mistake as in (35), if not then you are out of luck and need to use Σ's or invent your own notation.

A few common operations in the summation convention: We love to see a δ_{ij} involved in a sum since this collapses that sum. This is called the *substitution rule*, if δ_{ij} appears in a sum, we can forget about it and eliminate the summation index, for example

$$a_i\delta_{ij} = a_j, \qquad \delta_{kl}\delta_{kl} = \delta_{kk} = \delta_{ll} = 3, \qquad \delta_{ij}\epsilon_{ijk} = \epsilon_{iik} = 0 \tag{37}$$

note the second result $\delta_{kk} = 3$ because k is repeated, so there is a sum over all values of k and $\delta_{kk} = \delta_{11} + \delta_{22} + \delta_{33}$. The last result is because ϵ_{ijk} vanishes whenever two indices are the same. That last expression $\delta_{ij}\epsilon_{ijk}$ involves a double sum over i and over j. The δ_{ij} collapses one of those sums. It doesn't matter which index we choose to eliminate since both are dummy indices. Let's compute the l component of $\vec{\boldsymbol{a}}\times\vec{\boldsymbol{b}}$ from (36). We pick l because i, j and k are already taken. The l component is

$$\vec{\boldsymbol{e}}_l\cdot(\vec{\boldsymbol{a}}\times\vec{\boldsymbol{b}}) = \epsilon_{ijk}a_ib_j\vec{\boldsymbol{e}}_k\cdot\vec{\boldsymbol{e}}_l = \epsilon_{ijk}a_ib_j\delta_{kl} = \epsilon_{ijl}a_ib_j = \epsilon_{lmn}a_mb_n \tag{38}$$

what happened on that last step? first, $\epsilon_{ijk} = \epsilon_{kij}$ because (i,j,k) to (k,i,j) is a cyclic permutation which corresponds to an even number of permutation in space of odd dimension (dimension 3, here) and the value of ϵ_{ijk} does not change under even permutations. Then i and j are dummies and we renamed them m and n respectively being careful to keep the order. The final result is worth memorizing: if $\vec{\boldsymbol{c}} = \vec{\boldsymbol{a}}\times\vec{\boldsymbol{b}}$, the l component of $\vec{\boldsymbol{c}}$ is $c_l = \epsilon_{lmn}a_mb_n$, or switching indices to i,j,k

$$\boxed{\vec{\boldsymbol{c}} = \vec{\boldsymbol{a}}\times\vec{\boldsymbol{b}} \iff c_i = \epsilon_{ijk}a_jb_k \iff \vec{\boldsymbol{c}} = \vec{\boldsymbol{e}}_i\,\epsilon_{ijk}a_jb_k.} \tag{39}$$

In the spirit of no pain-no gain, let's write the double cross product identity $(\vec{\boldsymbol{a}}\times\vec{\boldsymbol{b}})\times\vec{\boldsymbol{c}}$ in this indicial notation. Let $\vec{\boldsymbol{v}} = \vec{\boldsymbol{a}}\times\vec{\boldsymbol{b}}$ then the i component of the double cross product $\vec{\boldsymbol{v}}\times\vec{\boldsymbol{c}}$ is $\epsilon_{ijk}v_jc_k$. Now we need the j component of $\vec{\boldsymbol{v}} = \vec{\boldsymbol{a}}\times\vec{\boldsymbol{b}}$. Since i and k are taken we use l, m as new dummy indices, and we have $v_j = \epsilon_{jlm}a_lb_m$. So the i component of the double cross product $(\vec{\boldsymbol{a}}\times\vec{\boldsymbol{b}})\times\vec{\boldsymbol{c}}$ is

$$\epsilon_{ijk}\epsilon_{jlm}a_lb_mc_k. \tag{40}$$

Note that j, k, l and m are repeated, so this expression is a quadruple sum! According to our double cross product identity it should be equal to the i component of $(\vec{\boldsymbol{a}}\cdot\vec{\boldsymbol{c}})\vec{\boldsymbol{b}} - (\vec{\boldsymbol{b}}\cdot\vec{\boldsymbol{c}})\vec{\boldsymbol{a}}$ *for any*

$\vec{a}$, $\vec{b}$, $\vec{c}$. We want the i component of the latter expression since i is a free index in (40), that i component is

$$(a_j c_j) b_i - (b_j c_j) a_i \tag{41}$$

(wait! isn't j repeated 4 times? no, it's not. It's repeated twice in *separate* terms so this is a difference of two sums over j). Since (40) and (41) are equal to each other for any $\vec{a}$, $\vec{b}$, $\vec{c}$, this should be telling us something about ϵ_{ijk}, but to extract that out we need to rewrite (41) in the form $a_l b_m c_k$. How? by making use of our ability to rename dummy variables and adding variables using δ_{ij}. Let's look at the first term in (41), $(a_j c_j) b_i$, here's how to write it in the form $a_l c_k b_m$ as in (40):

$$(a_j c_j) b_i = (a_k c_k) b_i = (\delta_{lk} a_l c_k)(\delta_{im} b_m) = \delta_{lk}\delta_{im} a_l c_k b_m. \tag{42}$$

Do similar manipulations to the second term in (41) to obtain $(b_j c_j) a_i = \delta_{il}\delta_{km} a_l c_k b_m$ and

$$\epsilon_{ijk}\epsilon_{jlm} a_l b_m c_k = (\delta_{lk}\delta_{im} - \delta_{il}\delta_{km}) a_l c_k b_m. \tag{43}$$

Since this equality holds *for any* a_l, c_k, b_m, we must have $\epsilon_{ijk}\epsilon_{jlm} = (\delta_{lk}\delta_{im} - \delta_{il}\delta_{km})$. That's true but it's not written in a nice way so let's clean it up to a form that's easier to reconstruct. First note that $\epsilon_{ijk} = \epsilon_{jki}$ since ϵ_{ijk} is invariant under a cyclic permutation of its indices. So our identity becomes $\epsilon_{jki}\epsilon_{jlm} = (\delta_{lk}\delta_{im} - \delta_{il}\delta_{km})$. We've done that flipping so the summation index j is in first place in both ϵ factors. Now we prefer the lexicographic order (i, j, k) to (j, k, i) so let's rename all the indices being careful to rename the correct indices on both sides. This yields

$$\boxed{\epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}} \tag{44}$$

This takes some digesting. Go through it carefully again. And again, as many times as it takes. The identity (44) is actually pretty easy to remember and verify. First, $\epsilon_{ijk}\epsilon_{ilm}$ is a sum over i but there is never more than one non-zero term (why?). Second, the only possible values for that expression are $+1$, 0 and -1 (why?). The only way to get 1 is to have $(j, k) = (l, m)$ with $j = l \neq k = m$ (why?), but in that case the right hand side of (44) is also 1 (why?). The only way to get -1 is to have $(j, k) = (m, l)$ with $j = m \neq k = l$ (why?), but in that case the right hand side is -1 also (why?). Finally, to get 0 we need $j = k$ or $l = m$ and the right-hand side again vanishes in either case. For instance, if $j = k$ then we can switch j and k in one of the terms and $\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl} = \delta_{jl}\delta_{km} - \delta_{km}\delta_{jl} = 0$.

Formula (44) has a generalization that does not include summation over one index

$$\begin{aligned} \epsilon_{ijk}\epsilon_{lmn} =& \delta_{il}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{kl} + \delta_{in}\delta_{jl}\delta_{km} \\ & -\delta_{im}\delta_{jl}\delta_{kn} - \delta_{in}\delta_{jm}\delta_{kl} - \delta_{il}\delta_{jn}\delta_{km} \end{aligned} \tag{45}$$

note that the first line correspond to (i, j, k) and (l, m, n) matching up to cyclic rotations, while the second line corresponds to (i, j, k) matching with an odd (acyclic) rotation of (l, m, n).

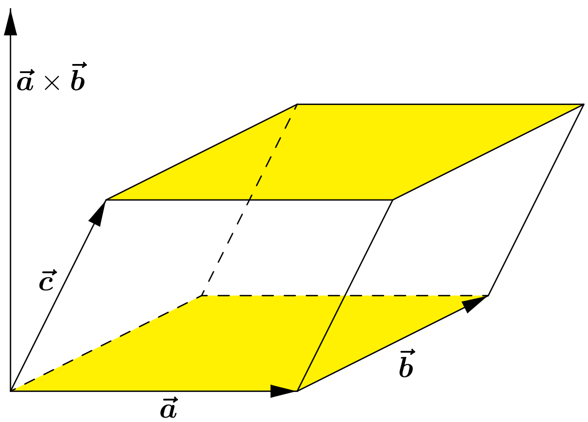
Exercises

1. Explain why $\epsilon_{ijk} = \epsilon_{jki} = -\epsilon_{jik}$ for any integer i, j, k.
2. Using (28) and Einstein's notation show that $(\vec{a} \times \vec{b}) \cdot \vec{e}_k = \epsilon_{ijk} a_i b_j$ and $(\vec{a} \times \vec{b}) = \epsilon_{ijk} a_i b_j \, \vec{e}_k = \epsilon_{ijk} a_j b_k \, \vec{e}_i$.
3. Show that $\epsilon_{ijk}\epsilon_{ljk} = 2\delta_{il}$ by direct deduction and by application of (44).
4. Deduce (44) from (45).

1.7 Mixed (or 'Box') product and Determinant

A mixed product[4] of three vectors has the form $(\vec{a} \times \vec{b}) \cdot \vec{c}$, it involves both a cross and a dot product. The result is a *scalar*. We have already encountered mixed products (*e.g.* eqn. (28)) but their geometric and algebraic properties are so important that they merit their own subsection.

$$\boxed{\begin{array}{c}(\vec{a} \times \vec{b}) \cdot \vec{c} = (\vec{b} \times \vec{c}) \cdot \vec{a} = (\vec{c} \times \vec{a}) \cdot \vec{b} = \\ \vec{a} \cdot (\vec{b} \times \vec{c}) = \vec{b} \cdot (\vec{c} \times \vec{a}) = \vec{c} \cdot (\vec{a} \times \vec{b}) = \\ \pm\text{volume of the parallelepiped spanned by } \vec{a}, \vec{b}, \vec{c}\end{array}} \tag{46}$$



Take $\vec{a}$ and $\vec{b}$ as the base of the parallelepiped then $\vec{a} \times \vec{b}$ is perpendicular to the base and has magnitude equal to the base area. The height is $\hat{z} \cdot \vec{c}$ where $\hat{z}$ is the unit vector perpendicular to the base, *i.e.* parallel to $\vec{a} \times \vec{b}$. So the volume is $(\vec{a} \times \vec{b}) \cdot \vec{c}$. Signwise, $(\vec{a} \times \vec{b}) \cdot \vec{c} > 0$ if $\vec{a}$, $\vec{b}$ and $\vec{c}$, in that order, form a right-handed basis (not orthogonal in general), and $(\vec{a} \times \vec{b}) \cdot \vec{c} < 0$ if the triplet is left-handed. Taking $\vec{b}$ and $\vec{c}$, or $\vec{c}$ and $\vec{a}$, as the base, you get the same volume and sign. The dot product commutes, so $(\vec{b} \times \vec{c}) \cdot \vec{a} = \vec{a} \cdot (\vec{b} \times \vec{c})$, yielding the identity

$$(\vec{a} \times \vec{b}) \cdot \vec{c} = \vec{a} \cdot (\vec{b} \times \vec{c}). \tag{47}$$

We can switch the dot and the cross without changing the result. We have shown (46) geometrically. The properties of the dot and cross products yield many other results such as $(\vec{a}\times\vec{b})\cdot\vec{c} = -(\vec{b}\times\vec{a})\cdot\vec{c}$, etc. We can collect all these results as follows.
A mixed product is one form of a scalar function of three vectors called the *determinant*

$$(\vec{a} \times \vec{b}) \cdot \vec{c} \equiv \det(\vec{a}, \vec{b}, \vec{c}), \tag{48}$$

whose value is the *signed volume* of the parallelepiped with sides $\vec{a}$, $\vec{b}$, $\vec{c}$. The determinant has three fundamental properties

1. it changes sign if *any* two vectors are permuted, *e.g.*

$$\det(\vec{a}, \vec{b}, \vec{c}) = -\det(\vec{b}, \vec{a}, \vec{c}) = \det(\vec{b}, \vec{c}, \vec{a}), \tag{49}$$

2. it is linear in *any* of its vectors *e.g.* $\forall\ \alpha$, $\vec{d}$,

$$\det(\alpha\vec{a} + \vec{d}, \vec{b}, \vec{c}) = \alpha \det(\vec{a}, \vec{b}, \vec{c}) + \det(\vec{d}, \vec{b}, \vec{c}), \tag{50}$$

[4]The mixed product is often called 'triple scalar product'.

3. if the triplet $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$ is right-handed and orthonormal then

$$\det(\vec{e}_1, \vec{e}_2, \vec{e}_3) = 1. \tag{51}$$

Deduce these from the properties of the dot and cross products as well as geometrically. Property (50) is a combination of the distributivity properties of the dot and cross products with respect to vector addition and multiplication by a scalar. For example,

$$\begin{aligned}\det(\alpha\vec{a} + \vec{d}, \vec{b}, \vec{c}) = (\alpha\vec{a} + \vec{d}) \cdot (\vec{b} \times \vec{c}) &= \alpha\big(\vec{a} \cdot (\vec{b} \times \vec{c})\big) + \vec{d} \cdot (\vec{b} \times \vec{c}) \\ &= \alpha \det(\vec{a}, \vec{b}, \vec{c}) + \det(\vec{d}, \vec{b}, \vec{c}).\end{aligned}$$

From these three properties, you deduce easily that the determinant is zero if any two vectors are identical (from prop 1), or if any vector is zero (from prop 2 with $\alpha = 1$ and $\vec{d} = \vec{0}$), and that the determinant does not change if we add a multiple of one vector to another, for example

$$\begin{aligned}\det(\vec{a}, \vec{b}, \vec{a}) &= 0, \\ \det(\vec{a}, \vec{0}, \vec{c}) &= 0, \\ \det(\vec{a} + \beta\vec{b}, \vec{b}, \vec{c}) &= \det(\vec{a}, \vec{b}, \vec{c}).\end{aligned} \tag{52}$$

Geometrically, this last one corresponds to a *shearing* of the parallelepiped, with no change in volume or orientation. One key application of determinants is

$$\det(\vec{a}, \vec{b}, \vec{c}) \neq 0 \;\Leftrightarrow\; \vec{a}, \vec{b}, \vec{c} \text{ form a basis.} \tag{53}$$

If $\det(\vec{a}, \vec{b}, \vec{c}) = 0$ then either one of the vectors is zero or they are co-planar and $\vec{a}$, $\vec{b}$, $\vec{c}$ cannot provide a basis for $\mathbf{E}^3$. This is how the determinant is introduced in elementary linear algebra. But the determinant is so much more! It 'determines' the volume of the parallelepiped and its orientation!

The 3 fundamental properties fully specify the determinant as explored in exercises 5, 6 below. If the vectors are expanded in terms of a right-handed orthonormal basis, *i.e.* $\vec{a} = a_i\vec{e}_i$, $\vec{b} = b_j\vec{e}_j$, $\vec{c} = c_k\vec{e}_k$ (summation convention), then we obtain the following formula for the determinant in terms of the vector components

$$\det(\vec{a}, \vec{b}, \vec{c}) = (\vec{a} \times \vec{b}) \cdot \vec{c} = a_i b_j c_k (\vec{e}_i \times \vec{e}_j) \cdot \vec{e}_k = \epsilon_{ijk}\, a_i b_j c_k. \tag{54}$$

Expanding that expression

$$\epsilon_{ijk}\, a_i b_j c_k = a_1 b_2 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_2 b_1 c_3 - a_3 b_2 c_1 - a_1 b_3 c_2. \tag{55}$$

we recognize the familiar algebraic determinants

$$\det(\vec{a}, \vec{b}, \vec{c}) = \epsilon_{ijk}\, a_i b_j c_k = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}. \tag{56}$$

Note that it does not matter whether we put the vector components along rows or columns. This is a non-trivial and important property of determinants. (see section on matrices).

This familiar determinant has the same three fundamental properties (49), (50), (51) of course

1. it changes sign if *any* two columns (or rows) are permuted, *e.g.*

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = - \begin{vmatrix} b_1 & a_1 & c_1 \\ b_2 & a_2 & c_2 \\ b_3 & a_3 & c_3 \end{vmatrix}, \tag{57}$$

2. it is linear in *any* of its columns (or rows) *e.g.* $\forall\ \alpha$, (d_1, d_2, d_3),

$$\begin{vmatrix} \alpha a_1 + d_1 & b_1 & c_1 \\ \alpha a_2 + d_2 & b_2 & c_2 \\ \alpha a_3 + d_3 & b_3 & c_3 \end{vmatrix} = \alpha \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}, \tag{58}$$

3. finally, the determinant of the *natural basis* is

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1, \tag{59}$$

You deduce easily from these three properties that the det vanishes if any column (or row) is zero or if any two columns (or rows) is a multiple of another, and that the determinant does not change if we add to one column (row) a linear combination of the other columns (rows). These properties allow us to calculate determinants by successive shearings and column-swapping. There is another explicit formula for determinants, in addition to the $\epsilon_{ijk} a_i b_j c_k$ formula, it is the *Laplace (or Cofactor) expansion* in terms of 2-by-2 determinants, *e.g.*

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_1 \end{vmatrix}, \tag{60}$$

where the 2-by-2 determinants are

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1. \tag{61}$$

This formula is nothing but $\vec{\boldsymbol{a}} \cdot (\vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}})$ expressed with respect to a right handed basis. To verify that, compute the components of $(\vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}})$ first, then dot with the components of $\vec{\boldsymbol{a}}$. This cofactor expansion formula can be applied to any column or row, however there are ± 1 factors that appear. We won't go into the straightforward details, but all that follows directly from the column swapping property (57). That's essentially the identities $\vec{\boldsymbol{a}} \cdot (\vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}}) = \vec{\boldsymbol{b}} \cdot (\vec{\boldsymbol{c}} \times \vec{\boldsymbol{a}}) = \cdots$.

Exercises

1. Show that the 2-by-2 determinant $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1$, is the *signed area* of the parallelogram with sides $\vec{\boldsymbol{a}} = a_1 \vec{\boldsymbol{e}}_1 + a_2 \vec{\boldsymbol{e}}_2$, $\vec{\boldsymbol{b}} = b_1 \vec{\boldsymbol{e}}_1 + b_2 \vec{\boldsymbol{e}}_2$. It is positive if $\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}, -\vec{\boldsymbol{a}}, -\vec{\boldsymbol{b}}$ is a counterclockwise cycle, negative if the cycle is clockwise. Sketch (of course).

2. The determinant $\det(\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}, \vec{\boldsymbol{c}})$ of three oriented line segments $\vec{\boldsymbol{a}}$, $\vec{\boldsymbol{b}}$, $\vec{\boldsymbol{c}}$ is a geometric quantity. Show that $\det(\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}, \vec{\boldsymbol{c}}) = |\vec{\boldsymbol{a}}|\,|\vec{\boldsymbol{b}}|\,|\vec{\boldsymbol{c}}| \sin\phi \cos\theta$. Specify ϕ and θ. Sketch.

3. Show that $-|\vec{\boldsymbol{a}}|\,|\vec{\boldsymbol{b}}|\,|\vec{\boldsymbol{c}}| \le \det(\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}, \vec{\boldsymbol{c}}) \le |\vec{\boldsymbol{a}}|\,|\vec{\boldsymbol{b}}|\,|\vec{\boldsymbol{c}}|$. When do the equalities apply? Sketch.

4. Use properties (49) and (50) to show that

$$\det(\alpha\vec{\boldsymbol{a}} + \lambda\vec{\boldsymbol{d}}, \beta\vec{\boldsymbol{b}} + \mu\vec{\boldsymbol{e}}, \vec{\boldsymbol{c}}) =$$

$$\alpha\beta\det(\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}, \vec{\boldsymbol{c}}) + \alpha\mu\det(\vec{\boldsymbol{a}}, \vec{\boldsymbol{e}}, \vec{\boldsymbol{c}}) + \beta\lambda\det(\vec{\boldsymbol{d}}, \vec{\boldsymbol{b}}, \vec{\boldsymbol{c}}) + \lambda\mu\det(\vec{\boldsymbol{d}}, \vec{\boldsymbol{e}}, \vec{\boldsymbol{c}}).$$

5. Use properties (49) and (51) to show that $\det(\vec{\boldsymbol{e}}_i, \vec{\boldsymbol{e}}_j, \vec{\boldsymbol{e}}_k) = \epsilon_{ijk}$.

6. Use property (50) and exercise 5 above to show that if $\vec{\boldsymbol{a}} = a_i\vec{\boldsymbol{e}}_i$, $\vec{\boldsymbol{b}} = b_i\vec{\boldsymbol{e}}_i$, $\vec{\boldsymbol{c}} = c_i\vec{\boldsymbol{e}}_i$ *(summation convention)* then $\det(\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}, \vec{\boldsymbol{c}}) = \epsilon_{ijk}a_ib_jc_k$.

7. Prove the identity $(\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}) \cdot (\vec{\boldsymbol{c}} \times \vec{\boldsymbol{d}}) = (\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{c}})(\vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{d}}) - (\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{d}})(\vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{c}})$ using both vector identities and indicial notation.

8. Express $(\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}) \cdot (\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}})$ in terms of dot products of $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}$.

9. Show that $(\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{a}})(\vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{b}}) - (\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}})^2$ is the square of the area of the parallelogram spanned by $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}$.

10. If A is the area the parallelogram with sides $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}$, show that

$$A^2 = \begin{vmatrix} \vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{a}} & \vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}} \\ \vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}} & \vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{b}} \end{vmatrix}.$$

11. If $\det(\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}, \vec{\boldsymbol{c}}) \neq 0$, then any vector $\vec{\boldsymbol{v}}$ can be expanded as $\vec{\boldsymbol{v}} = \alpha\vec{\boldsymbol{a}} + \beta\vec{\boldsymbol{b}} + \gamma\vec{\boldsymbol{c}}$. Find explicit expressions for the components α, β, γ in terms of $\vec{\boldsymbol{v}}$ and the basis vectors $\vec{\boldsymbol{a}}$, $\vec{\boldsymbol{b}}$, $\vec{\boldsymbol{c}}$ in the general case when the latter are *not* orthogonal. [Hint: project on cross products of the basis vectors, then collect the mixed products into determinants and deduce *Cramer's rule*].

12. Given three vectors $\vec{\boldsymbol{a}}_1$, $\vec{\boldsymbol{a}}_2$, $\vec{\boldsymbol{a}}_3$ such that $D = \vec{\boldsymbol{a}}_1 \cdot (\vec{\boldsymbol{a}}_2 \times \vec{\boldsymbol{a}}_3) \neq 0$, define

$$\vec{\boldsymbol{a}}'_1 = (\vec{\boldsymbol{a}}_2 \times \vec{\boldsymbol{a}}_3)/D, \quad \vec{\boldsymbol{a}}'_2 = (\vec{\boldsymbol{a}}_3 \times \vec{\boldsymbol{a}}_1)/D, \quad \vec{\boldsymbol{a}}'_3 = (\vec{\boldsymbol{a}}_1 \times \vec{\boldsymbol{a}}_2)/D. \tag{62}$$

This is the **reciprocal basis** of the basis $\vec{\boldsymbol{a}}_1$, $\vec{\boldsymbol{a}}_2$, $\vec{\boldsymbol{a}}_3$.

(i) Show that $\vec{\boldsymbol{a}}_i \cdot \vec{\boldsymbol{a}}'_j = \delta_{ij}$, $\forall\ i, j = 1, 2, 3$.

(ii) Show that if $\vec{\boldsymbol{v}} = \sum_{i=1}^3 v_i\,\vec{\boldsymbol{a}}_i$ and $\vec{\boldsymbol{v}} = \sum_{i=1}^3 v'_i\,\vec{\boldsymbol{a}}'_i$, then $v_i = \vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{a}}'_i$ and $v'_i = \vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{a}}_i$. So the components in one basis are obtained by projecting onto the other basis.

13. If $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}$ are linearly independent and $\vec{\boldsymbol{c}}$ is any arbitrary vector, find α, β and γ such that $\vec{\boldsymbol{c}} = \alpha\vec{\boldsymbol{a}} + \beta\vec{\boldsymbol{b}} + \gamma(\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}})$. Express α, β and γ in terms of dot products only. [Hint: find α and β first, then use $\vec{\boldsymbol{c}}_{\parallel} = \vec{\boldsymbol{c}} - \vec{\boldsymbol{c}}_{\perp}$.]

14. Express $(\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}) \cdot \vec{\boldsymbol{c}}$ in terms of dot products of $\vec{\boldsymbol{a}}$, $\vec{\boldsymbol{b}}$ and $\vec{\boldsymbol{c}}$ only. [Hint: solve problem 13 first.]

15. Provide an algorithm to compute the volume of the parallelepiped ($\vec{\boldsymbol{a}}$, $\vec{\boldsymbol{b}}$, $\vec{\boldsymbol{c}}$) by taking only dot products. [Hint: 'rectify' the parallelepiped $(\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}, \vec{\boldsymbol{c}}) \to (\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}_{\perp}, \vec{\boldsymbol{c}}_{\perp}) \to (\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}_{\perp}, \vec{\boldsymbol{c}}_{\perp\perp})$ where $\vec{\boldsymbol{b}}_{\perp}$ and $\vec{\boldsymbol{c}}_{\perp}$ are perpendicular to $\vec{\boldsymbol{a}}$, and $\vec{\boldsymbol{c}}_{\perp\perp}$ is perpendicular to both $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}_{\perp}$. Explain geometrically why these transformations do not change the volume. Explain why these transformations do not change the determinant by using the properties of determinants.]

16. (*) If V is the volume of the parallelepiped with sides $\vec{a}$, $\vec{b}$, $\vec{c}$ show that

$$V^2 = \begin{vmatrix} \vec{a}\cdot\vec{a} & \vec{a}\cdot\vec{b} & \vec{a}\cdot\vec{c} \\ \vec{b}\cdot\vec{a} & \vec{b}\cdot\vec{b} & \vec{b}\cdot\vec{c} \\ \vec{c}\cdot\vec{a} & \vec{c}\cdot\vec{b} & \vec{c}\cdot\vec{c} \end{vmatrix}.$$

Do this in several ways: (i) from problem 13, (ii) using indicial notation and the formula (45).

1.8 Points, Lines, Planes, etc.

Points and vectors are different. We do not add points, but vector addition is defined. However, once a reference point has been picked, called the *origin* O, any point P is uniquely determined by specifying the vector $\vec{r} \equiv \overrightarrow{OP}$. This special vector is called the *position vector*. It is often denoted $\vec{x}$ also. Position vectors have a special meaning because they are tied to a specific origin. Representing points by position vectors is extremely useful as all vector operations are then available to us.
Examples:

- The center of mass $\vec{r}_c$ of a system of N particles of mass m_i located at position $\vec{r}_i$, $i = 1, \ldots, N$ is defined by $M\,\vec{r}_c = \sum_{i=1}^N m_i\vec{r}_i$ where $M = \sum_{i=1}^N m_i$ is the total mass. This is a coordinate-free expression for the center of mass. In particular, if all the masses are equal then for $N = 2$, $\vec{r}_c = (\vec{r}_1 + \vec{r}_2)/2$, for $N = 3$, $\vec{r}_c = (\vec{r}_1 + \vec{r}_2 + \vec{r}_3)/3$.

▷ Show that the center of gravity of three points of equal mass is at the point of intersection of the medians of the triangle formed by the three points.

- The vector equation of a line parallel to $\vec{a}$ passing through a point $\vec{r}_0$ is

$$\boxed{(\vec{r} - \vec{r}_0) \times \vec{a} = 0} \Leftrightarrow \boxed{\vec{r} = \vec{r}_0 + \alpha\vec{a}, \quad \forall\alpha \in \mathbb{R}.} \tag{63}$$

- The equation of a plane through $\vec{r}_0$ parallel to $\vec{a}$ and $\vec{b}$ (with $\vec{a} \times \vec{b} \neq 0$), or (equivalently) perpendicular to $\vec{n} \equiv \vec{a} \times \vec{b}$ is

$$\boxed{(\vec{r} - \vec{r}_0) \cdot \vec{n} = 0} \Leftrightarrow \boxed{\vec{r} = \vec{r}_0 + \alpha\vec{a} + \beta\vec{b}, \quad \forall\alpha, \beta \in \mathbb{R}.} \tag{64}$$

- The equation of a sphere of center $\vec{r}_c$ and radius R is

$$|\vec{r} - \vec{r}_c| = R \Leftrightarrow \vec{r} = \vec{r}_c + R\,\vec{n}, \quad \forall\vec{n}\ s.t.\ |\vec{n}| = 1. \tag{65}$$

▷ Find vector equations for the line passing through the two points $\vec{r}_1$, $\vec{r}_2$ and the plane through the three points $\vec{r}_1$, $\vec{r}_2$, $\vec{r}_3$.

▷ What is the distance between the point $\vec{r}_1$ and the plane through $\vec{r}_0$ perpendicular to $\vec{a}$?

▷ What is the distance between the point $\vec{r}_1$ and the plane through $\vec{r}_0$ parallel to $\vec{a}$ and $\vec{b}$?

▷ What is the distance between the line parallel to $\vec{a}$ that passes through point A and the line parallel to $\vec{b}$ that passes through point B?

▷ A particle was at point P_1 at time t_1 and is moving at the constant velocity $\vec{v}_1$. Another particle was at P_2 at t_2 and is moving at the constant velocity $\vec{v}_2$. How close did the particles get to each other and at what time? What conditions are needed for a collision?

1.9 Vector function of a scalar variable

The position vector of a moving particle is a vector function $\vec{\boldsymbol{r}}(t)$ of the scalar time t. The derivative of a vector function is defined as usual

$$\frac{d\vec{\boldsymbol{r}}}{dt}(t) = \lim_{h\to 0} \frac{\vec{\boldsymbol{r}}(t+h) - \vec{\boldsymbol{r}}(t)}{h}. \tag{66}$$

The derivative of the position vector is of course the instantaneous velocity vector $\vec{\boldsymbol{v}}(t) = d\vec{\boldsymbol{r}}/dt$. The position vector $\vec{\boldsymbol{r}}(t)$ describes a curve in three-dimensional space, the particle trajectory, and $\vec{\boldsymbol{v}}(t)$ is tangent to that curve. The derivative of the velocity vector is the acceleration vector $\vec{\boldsymbol{a}}(t) = d\vec{\boldsymbol{v}}/dt$. We'll often use Newton's notation for time derivatives: $d\vec{\boldsymbol{r}}/dt \equiv \dot{\boldsymbol{r}}$, $d^2\vec{\boldsymbol{r}}/dt^2 = \ddot{\boldsymbol{r}}$, etc.
We need to know how to manipulate derivatives of vector functions. It is easy to show that the derivative of a sum of vectors is the sum of the derivatives,

$$\frac{d}{dt}(\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}}) = \frac{d\vec{\boldsymbol{a}}}{dt} + \frac{d\vec{\boldsymbol{b}}}{dt}.$$

For the various products, we can show by the 'standard' product derivative trick recalled in class that

$$\frac{d}{dt}(\alpha\vec{\boldsymbol{a}}) = \frac{d\alpha}{dt}\vec{\boldsymbol{a}} + \alpha\frac{d\vec{\boldsymbol{a}}}{dt}, \qquad \frac{d}{dt}(\vec{\boldsymbol{a}}\cdot\vec{\boldsymbol{b}}) = \frac{d\vec{\boldsymbol{a}}}{dt}\cdot\vec{\boldsymbol{b}} + \vec{\boldsymbol{a}}\cdot\frac{d\vec{\boldsymbol{b}}}{dt},$$

$$\frac{d}{dt}(\vec{\boldsymbol{a}}\times\vec{\boldsymbol{b}}) = \frac{d\vec{\boldsymbol{a}}}{dt}\times\vec{\boldsymbol{b}} + \vec{\boldsymbol{a}}\times\frac{d\vec{\boldsymbol{b}}}{dt},$$

then

$$\frac{d}{dt}[(\vec{\boldsymbol{a}}\times\vec{\boldsymbol{b}})\cdot\vec{\boldsymbol{c}}] = (\frac{d\vec{\boldsymbol{a}}}{dt}\times\vec{\boldsymbol{b}})\cdot\vec{\boldsymbol{c}} + (\vec{\boldsymbol{a}}\times\frac{d\vec{\boldsymbol{b}}}{dt})\cdot\vec{\boldsymbol{c}} + (\vec{\boldsymbol{a}}\times\vec{\boldsymbol{b}})\cdot\frac{d\vec{\boldsymbol{c}}}{dt},$$

therefore

$$\frac{d}{dt}\det(\vec{\boldsymbol{a}},\vec{\boldsymbol{b}},\vec{\boldsymbol{c}}) = \det(\frac{d\vec{\boldsymbol{a}}}{dt},\vec{\boldsymbol{b}},\vec{\boldsymbol{c}}) + \det(\vec{\boldsymbol{a}},\frac{d\vec{\boldsymbol{b}}}{dt},\vec{\boldsymbol{c}}) + \det(\vec{\boldsymbol{a}},\vec{\boldsymbol{b}},\frac{d\vec{\boldsymbol{c}}}{dt}).$$

All of these are as expected but the formula for the derivative of a determinant is worth noting because it generalizes to any dimension. [5]

- ▷ Show that if $\vec{\boldsymbol{e}}(t)$ is any vector with constant norm, *i.e.* $\vec{\boldsymbol{e}}(t)\cdot\vec{\boldsymbol{e}}(t) = Constant\ \forall\ t$, then

$$\vec{\boldsymbol{e}}(t)\cdot\frac{d\vec{\boldsymbol{e}}}{dt}(t) = 0, \quad \forall t. \tag{67}$$

  The derivative of a constant norm vector is orthogonal to the vector.

- ▷ If $r = \|\vec{\boldsymbol{r}}\|$, where $\vec{\boldsymbol{r}} = \vec{\boldsymbol{r}}(t)$, show that $dr/dt = \hat{\boldsymbol{r}}\cdot\dot{\boldsymbol{r}} \equiv \hat{\boldsymbol{r}}\cdot\vec{\boldsymbol{v}}$.

- ▷ If $\vec{\boldsymbol{v}}(t) = d\vec{\boldsymbol{r}}/dt$, show that $d(\vec{\boldsymbol{r}}\times\vec{\boldsymbol{v}})/dt = \vec{\boldsymbol{r}}\times d\vec{\boldsymbol{v}}/dt$. In mechanics, $\vec{\boldsymbol{r}}\times m\vec{\boldsymbol{v}}$ is the *angular momentum* of the particle of mass m and velocity $\vec{\boldsymbol{v}}$ with respect to the origin.

We now illustrate all these concepts and results by considering the basic problems of classical mechanics: motion of a particle and motion of a rigid body.

[5] For determinants in $\mathbb{R}^3$ it reads

$$\frac{d}{dt}\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} \dot{a}_1 & b_1 & c_1 \\ \dot{a}_2 & b_2 & c_2 \\ \dot{a}_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} a_1 & \dot{b}_1 & c_1 \\ a_2 & \dot{b}_2 & c_2 \\ a_3 & \dot{b}_3 & c_3 \end{vmatrix} + \begin{vmatrix} a_1 & b_1 & \dot{c}_1 \\ a_2 & b_2 & \dot{c}_2 \\ a_3 & b_3 & \dot{c}_3 \end{vmatrix}$$

and of course we could also take the derivatives along rows instead of columns.

1.10 Motion of a particle

In classical mechanics, the motion of a particle of mass m is governed by Newton's law

$$\vec{\boldsymbol{F}} = m\vec{\boldsymbol{a}}, \tag{68}$$

where $\vec{\boldsymbol{F}}$ is the resultant of the forces acting on the particle and $\vec{\boldsymbol{a}}(t) = d\vec{\boldsymbol{v}}/dt = d^2\vec{\boldsymbol{r}}/dt^2$ is its acceleration, with $\vec{\boldsymbol{r}}(t)$ its position vector. Newton's law is a vector equation.

Free motion

If $\vec{\boldsymbol{F}} = 0$ then $d\vec{\boldsymbol{v}}/dt = 0$ so the velocity of the particle is constant, $\vec{\boldsymbol{v}}(t) = \vec{\boldsymbol{v}}_0$ say, and its position is given by the vector differential equation $d\vec{\boldsymbol{r}}/dt = \vec{\boldsymbol{v}}_0$ whose solution is $\vec{\boldsymbol{r}}(t) = \vec{\boldsymbol{r}}_0 + t\vec{\boldsymbol{v}}_0$ where $\vec{\boldsymbol{r}}_0$ is a constant of integration which corresponds to the position of the particle at time $t = 0$. The particle moves in a straight line through $\vec{\boldsymbol{r}}_0$ parallel to $\vec{\boldsymbol{v}}_0$.

Constant acceleration

$$\frac{d^2\vec{\boldsymbol{r}}}{dt^2} = \frac{d\vec{\boldsymbol{v}}}{dt} = \vec{\boldsymbol{a}}(t) = \vec{\boldsymbol{a}}_0 \tag{69}$$

where $\vec{\boldsymbol{a}}_0$ is a time-independent vector. Integrating we find

$$\vec{\boldsymbol{v}}(t) = \vec{\boldsymbol{a}}_0 t + \vec{\boldsymbol{v}}_0, \qquad \vec{\boldsymbol{r}}(t) = \vec{\boldsymbol{a}}_0\frac{t^2}{2} + \vec{\boldsymbol{v}}_0 t + \vec{\boldsymbol{r}}_0 \tag{70}$$

where $\vec{\boldsymbol{v}}_0$ and $\vec{\boldsymbol{r}}_0$ are *vector* constants of integration. They are easily interpreted as the velocity and position at $t = 0$.

Uniform rotation

If a particle rotates with angular velocity ω about an *axis* parallel to $\vec{\boldsymbol{n}}$ that passes through the point $\vec{\boldsymbol{r}}_a$ (we take $\|\vec{\boldsymbol{n}}\| = 1$ and define $\omega > 0$ for right-handed rotation about $\vec{\boldsymbol{n}}$, $\omega < 0$ for left-handed rotation) then its velocity is

$$\vec{\boldsymbol{v}}(t) = \vec{\boldsymbol{\omega}} \times (\vec{\boldsymbol{r}}(t) - \vec{\boldsymbol{r}}_a)$$

where $\vec{\boldsymbol{\omega}} \equiv \omega\vec{\boldsymbol{n}}$ is the *rotation vector.*
▷ Show that $\|\vec{\boldsymbol{r}}(t) - \vec{\boldsymbol{r}}_a\|$ remains constant. Calculate the particle acceleration if $\vec{\boldsymbol{\omega}}$ and $\vec{\boldsymbol{r}}_a$ are constants and interpret geometrically. Find the force required to sustain such a motion.

Motion due to a central force

A force $\vec{\boldsymbol{F}} = -F(r)\,\hat{\boldsymbol{r}}$ where $r = \|\vec{\boldsymbol{r}}\|$ that always points toward the origin (if $F(r) > 0$, away if $F(r) < 0$) and depends only on the distance to the origin is called a *central force*. The gravitational force for planetary motion and the Coulomb force in electromagnetism are of that kind. Newton's law for a particle submitted to such a force is

$$m\frac{d\vec{\boldsymbol{v}}}{dt} = -F(r)\,\hat{\boldsymbol{r}}$$

where $\vec{\boldsymbol{r}}(t) = r\hat{\boldsymbol{r}}$ is the position vector of the particle, hence both r and $\hat{\boldsymbol{r}}$ are functions of time, and $\vec{\boldsymbol{v}} = d\vec{\boldsymbol{r}}/dt$. Motion due to such a force has two conserved quantities.

- *Conservation of angular momentum* $\forall\ F(r)$

$$\vec{\boldsymbol{r}} \times \frac{d\vec{\boldsymbol{v}}}{dt} = 0 \Leftrightarrow \frac{d}{dt}\left(\vec{\boldsymbol{r}} \times \vec{\boldsymbol{v}}\right) = 0 \Leftrightarrow \vec{\boldsymbol{r}} \times \vec{\boldsymbol{v}} \equiv L_0 \hat{\boldsymbol{z}} \tag{71}$$

where $L_0 > 0$ and $\hat{\boldsymbol{z}}$ are constants. So the motion remains in the plane orthogonal to $\hat{\boldsymbol{z}}$. Now $L_0\hat{\boldsymbol{z}}dt = \vec{\boldsymbol{r}} \times \vec{\boldsymbol{v}}dt = \vec{\boldsymbol{r}} \times d\vec{\boldsymbol{r}} = 2da\hat{\boldsymbol{z}}$ where da is the triangular area swept by $\vec{\boldsymbol{r}}$ in time dt. This yields

Kepler's law: *The radius vector sweeps equal areas in equal times.*

- *Conservation of energy: kinetic + potential*

$$m\frac{d\vec{\boldsymbol{v}}}{dt} \cdot \vec{\boldsymbol{v}} + F(r)\,\hat{\boldsymbol{r}} \cdot \vec{\boldsymbol{v}} = 0 \Leftrightarrow \frac{d}{dt}\left(m\frac{\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{v}}}{2} + V(r)\right) = 0,$$

where $dV(r)/dr \equiv F(r)$ as by the chain rule $dV(r)/dt = (dV/dr)(dr/dt) = (dV/dr)\,\hat{\boldsymbol{r}} \cdot \vec{\boldsymbol{v}}$. This implies that

$$\left(m\frac{\|\vec{\boldsymbol{v}}\|^2}{2} + V(r)\right) = E_0 \tag{72}$$

where E_0 is a constant. The first term $m\|\vec{\boldsymbol{v}}\|^2/2$ is the kinetic energy and the second $V(r)$ is the potential energy which is defined up to an arbitrary constant. The constant E_0 is the total conserved energy. Note that $V(r)$ and E_0 can be negative but $m\|\vec{\boldsymbol{v}}\|^2/2 \geq 0$, so the physically admissible r domain is that were $V(r)$ is less than E_0.

1.11 Motion of a system of particles

Consider N particles of mass m_i at positions $\vec{\boldsymbol{r}}_i$, $i = 1, \ldots, N$. The net force acting on particle number i is $\vec{\boldsymbol{F}}_i$ and Newton's law for each particle reads $m_i\ddot{\vec{\boldsymbol{r}}}_i = \vec{\boldsymbol{F}}_i$. Summing over all i's yields

$$\sum_{i=1}^{N} m_i\ddot{\vec{\boldsymbol{r}}}_i = \sum_{i=1}^{N} \vec{\boldsymbol{F}}_i.$$

Great cancellations occur on both sides. On the left side, let $\vec{\boldsymbol{r}}_i = \vec{\boldsymbol{r}}_c + \vec{\boldsymbol{s}}_i$, where $\vec{\boldsymbol{r}}_c$ is the center of mass and $\vec{\boldsymbol{s}}_i$ is the position vector of particle i with respect to the center of mass, then

$$\sum_i m_i\vec{\boldsymbol{r}}_i = \sum_i m_i\left(\vec{\boldsymbol{r}}_c + \vec{\boldsymbol{s}}_i\right) = M\vec{\boldsymbol{r}}_c + \sum_i m_i\vec{\boldsymbol{s}}_i \;\Rightarrow \sum_i m_i\vec{\boldsymbol{s}}_i = 0,$$

as, by definition of the center of mass $\sum_i m_i\vec{\boldsymbol{r}}_i = M\vec{\boldsymbol{r}}_c$, where $M = \sum_i m_i$ is the total mass. If the masses m_i are constants then $\sum_i m_i\vec{\boldsymbol{s}}_i = 0 \Rightarrow \sum_i m_i\dot{\vec{\boldsymbol{s}}}_i = 0 \Rightarrow \sum_i m_i\ddot{\vec{\boldsymbol{s}}}_i = 0$. In that case, $\sum_i m_i\ddot{\vec{\boldsymbol{r}}}_i = \sum_i m_i\left(\ddot{\vec{\boldsymbol{r}}}_c + \ddot{\vec{\boldsymbol{s}}}_i\right) = \sum_i m_i\ddot{\vec{\boldsymbol{r}}}_c = M\ddot{\vec{\boldsymbol{r}}}_c$. On the right-hand side, by action-reaction, all internal forces cancel out and the resultant is therefore the sum of all external forces only $\sum_i \vec{\boldsymbol{F}}_i = \sum_i \vec{\boldsymbol{F}}_i^{(e)} = \vec{\boldsymbol{F}}^{(e)}$. Therefore,

$$M\ddot{\vec{\boldsymbol{r}}}_c = \vec{\boldsymbol{F}}^{(e)} \tag{73}$$

where M is the total mass and $\vec{\boldsymbol{F}}^{(e)}$ is the resultant of all external forces acting on all the particles. The motion of the center of mass of a system of particles is that of a single particle of mass M with position vector $\vec{\boldsymbol{r}}_c$ under the action of the sum of all external forces. This is a fundamental theorem of mechanics.

There are also nice cancellations occurring for the motion about the center of mass. This involves considering angular momentum and torques about the center of mass. Taking the cross-product of Newton's law, $m_i \ddot{\vec{\boldsymbol{r}}}_i = \vec{\boldsymbol{F}}_i$, with $\vec{\boldsymbol{s}}_i$ for each particle and summing over all particles gives

$$\sum_i \vec{\boldsymbol{s}}_i \times m_i \ddot{\vec{\boldsymbol{r}}}_i = \sum_i \vec{\boldsymbol{s}}_i \times \vec{\boldsymbol{F}}_i.$$

On the left hand side, $\vec{\boldsymbol{r}}_i \equiv \vec{\boldsymbol{r}}_c + \vec{\boldsymbol{s}}_i$ and the definition of center of mass implies $\sum_i m_i \vec{\boldsymbol{s}}_i = 0$. Therefore

$$\sum_i \vec{\boldsymbol{s}}_i \times m_i \ddot{\vec{\boldsymbol{r}}}_i = \sum_i \vec{\boldsymbol{s}}_i \times m_i (\ddot{\vec{\boldsymbol{r}}}_c + \ddot{\vec{\boldsymbol{s}}}_i) = \sum_i \vec{\boldsymbol{s}}_i \times m_i \ddot{\vec{\boldsymbol{s}}}_i = \frac{d}{dt}\left(\sum_i \vec{\boldsymbol{s}}_i \times m_i \dot{\vec{\boldsymbol{s}}}_i\right).$$

This last expression is the rate of change of the total angular momentum about the center of mass

$$\vec{\boldsymbol{L}}_c \equiv \sum_{i=1}^{N} \left(\vec{\boldsymbol{s}}_i \times m_i \dot{\vec{\boldsymbol{s}}}_i\right).$$

On the right hand side, one can argue that the (internal) force exerted by particle j on particle i is in the direction of the relative position of j with respect to i, $\vec{\boldsymbol{f}}_{ij} \equiv \alpha_{ij}(\vec{\boldsymbol{r}}_i - \vec{\boldsymbol{r}}_j)$. By action-reaction the force from i onto j is $\vec{\boldsymbol{f}}_{ji} = -\vec{\boldsymbol{f}}_{ij} = -\alpha_{ij}(\vec{\boldsymbol{r}}_i - \vec{\boldsymbol{r}}_j)$, and the net contribution to the torque from the internal forces will cancel out: $\vec{\boldsymbol{r}}_i \times \vec{\boldsymbol{f}}_{ij} + \vec{\boldsymbol{r}}_j \times \vec{\boldsymbol{f}}_{ji} = 0$. This is true with respect to any point and in particular, with respect to the center of mass $\vec{\boldsymbol{s}}_i \times \vec{\boldsymbol{f}}_{ij} + \vec{\boldsymbol{s}}_j \times \vec{\boldsymbol{f}}_{ji} = 0$. Hence, for the motion about the center of mass we have

$$\frac{d\vec{\boldsymbol{L}}_c}{dt} = \vec{\boldsymbol{T}}_c^{(e)} \tag{74}$$

where $\vec{\boldsymbol{T}}^{(e)} = \sum_i \vec{\boldsymbol{s}}_i \times \vec{\boldsymbol{F}}_i$ is the net torque about the center of mass due to external forces only. This is another fundamental theorem, that the rate of change of the total angular momentum about the center of mass is equal to the total torque due to the external forces only.

$\triangleright$ If $\vec{\boldsymbol{f}}_{ij} = \alpha\,(\vec{\boldsymbol{r}}_i - \vec{\boldsymbol{r}}_j)$ and $\vec{\boldsymbol{f}}_{ji} = \alpha\,(\vec{\boldsymbol{r}}_j - \vec{\boldsymbol{r}}_i)$, show algebraically and geometrically that $\vec{\boldsymbol{s}}_i \times \vec{\boldsymbol{f}}_{ij} + \vec{\boldsymbol{s}}_j \times \vec{\boldsymbol{f}}_{ji} = 0$, where $\vec{\boldsymbol{s}}$ is the position vector from the center of mass.

1.12 Motion of a rigid body

The two vector differential equations for motion of the center of mass and evolution of the angular momentum about the center of mass are sufficient to fully determine the motion of a rigid body.
A rigid body is such that all lengths and angles are preserved within the rigid body. If A, B and C are any three points of the rigid body, then $\overrightarrow{AB} \cdot \overrightarrow{AC} = constant$.

Kinematics of a rigid body

Consider a right-handed orthonormal basis, $\vec{\boldsymbol{e}}_1(t)$, $\vec{\boldsymbol{e}}_2(t)$, $\vec{\boldsymbol{e}}_3(t)$ tied to the body. These vectors are functions of time t because they are frozen into the body so they rotate with it. However the basis remains orthonormal as all lengths and angles are preserved. Hence $\vec{\boldsymbol{e}}_i(t) \cdot \vec{\boldsymbol{e}}_j(t) = \delta_{ij}\ \forall\ i,j = 1,2,3$, and $\forall t$ and differentiating with respect to time

$$\frac{d\vec{\boldsymbol{e}}_i}{dt} \cdot \vec{\boldsymbol{e}}_j + \vec{\boldsymbol{e}}_i \cdot \frac{d\vec{\boldsymbol{e}}_j}{dt} = 0. \tag{75}$$

In particular, as seen in an earlier exercise, the derivative of a unit vector is orthogonal to the vector: $\vec{e}_l \cdot \dot{e}_l = 0$, $\forall l = 1, 2, 3$. So we can write

$$\frac{d\vec{e}_l}{dt} \equiv \vec{\omega}_l \times \vec{e}_l, \quad \forall\, l = 1, 2, 3 \tag{76}$$

as this guarantees that $\vec{e}_l \cdot \dot{e}_l = 0$ for any $\vec{\omega}_l$. Substituting this expression into (75) yields

$$(\vec{\omega}_i \times \vec{e}_i) \cdot \vec{e}_j + \vec{e}_i \cdot (\vec{\omega}_j \times \vec{e}_j) = 0,$$

and rewriting the mixed products

$$(\vec{e}_i \times \vec{e}_j) \cdot \vec{\omega}_i = (\vec{e}_i \times \vec{e}_j) \cdot \vec{\omega}_j. \tag{77}$$

Now let

$$\vec{\omega}_l \equiv \sum_k \omega_{kl} \vec{e}_k = \omega_{1l}\vec{e}_1 + \omega_{2l}\vec{e}_2 + \omega_{3l}\vec{e}_3,$$

so ω_{kl} is the k component of vector $\vec{\omega}_l$. Substituting in (77) gives

$$\sum_k \epsilon_{ijk}\omega_{ki} = \sum_k \epsilon_{ijk}\omega_{kj} \tag{78}$$

where as before $\epsilon_{ijk} \equiv (\vec{e}_i \times \vec{e}_j) \cdot \vec{e}_k$. The sums over k have at most one non-zero term. This yields the three equations

$$\begin{aligned} (i,j,k) = (1,2,3) &\longrightarrow \omega_{31} = \omega_{32} \\ (i,j,k) = (2,3,1) &\longrightarrow \omega_{12} = \omega_{13} \\ (i,j,k) = (3,1,2) &\longrightarrow \omega_{23} = \omega_{21}. \end{aligned} \tag{79}$$

The second equation, for instance, says that the first component of $\vec{\omega}_2$ is equal to the first component of $\vec{\omega}_3$. Now ω_{ll} is arbitrary according to (76) (why?), so we can choose to define ω_{11}, the first component of $\vec{\omega}_1$, for instance, equal to the first components of the other two vectors that are equal to each other, *i.e.* $\omega_{11} = \omega_{12} = \omega_{13}$. Likewise, pick $\omega_{22} = \omega_{23} = \omega_{21}$ and $\omega_{33} = \omega_{31} = \omega_{32}$. This choice implies that

$$\vec{\omega}_1 = \vec{\omega}_2 = \vec{\omega}_3 \equiv \vec{\omega} \tag{80}$$

The vector $\vec{\omega}(t)$ is the *Poisson vector* of the rigid body.
The Poisson vector $\vec{\omega}(t)$ gives the rate of change of *any* vector tied to the body. Indeed, if A and B are any two points of the body then the vector $\vec{c} \equiv \overrightarrow{AB}$ can be expanded with respect to the body basis $\vec{e}_1(t)$, $\vec{e}_2(t)$, $\vec{e}_3(t)$

$$\vec{c}(t) = c_1\vec{e}_1(t) + c_2\vec{e}_2(t) + c_3\vec{e}_3(t),$$

but the components $c_i \equiv \vec{c}(t) \cdot \vec{e}_i(t)$ are constants because all lengths and angles, and therefore all dot products, are time-invariant. Thus

$$\frac{d\vec{c}}{dt} = \sum_{i=1}^{3} c_i \frac{d\vec{e}_i}{dt} = \sum_{i=1}^{3} c_i\,(\vec{\omega} \times \vec{e}_i) = \vec{\omega} \times \vec{c}.$$

This is true for any vector tied to the body (*material vectors*), implying that the Poisson vector is unique for the body.

Dynamics of rigid body

The center of mass of a rigid body moves according to the sum of the external forces as for a system of particles. A continuous rigid body can be considered as a continuous distribution of 'infinitesimal' masses dm

$$\sum_{i=1}^{N} m_i \vec{s}_i \longrightarrow \int_V \vec{s}\, dm$$

where the three-dimensional integral is over all points $\vec{s}$ in the domain V of the body (dm is the 'measure' of the infinitesimal volume element dV, or in other words $dm = \rho dV$, where $\rho(\vec{s})$ is the mass density at point $\vec{s}$).
For the motion about the center of mass, the position vectors $\vec{s}_i$ are frozen into the body hence $\dot{\boldsymbol{s}}_i = \vec{\omega} \times \vec{s}_i$ for any point of the body. The total angular momentum for a rigid system of particles then reads

$$\vec{L} = \sum_i m_i \vec{s}_i \times \dot{\boldsymbol{s}}_i = \sum_i m_i \vec{s}_i \times (\vec{\omega} \times \vec{s}_i) = \sum_i m_i \left(\|\vec{s}_i\|^2 \vec{\omega} - \vec{s}_i \left(\vec{s}_i \cdot \vec{\omega} \right) \right). \tag{81}$$

and for a continuous rigid body

$$\vec{L} = \int_V \left(\|\vec{s}\|^2 \vec{\omega} - \vec{s} \left(\vec{s} \cdot \vec{\omega} \right) \right) dm. \tag{82}$$

The Poisson vector is unique for the body, so it does not depend on $\vec{s}$ and we should be able to take it out of the sum, or integral. That's easy for the $\|\vec{s}\|^2 \vec{\omega}$ term, but how can we get $\vec{\omega}$ out of the $\int \vec{s} \left(\vec{s} \cdot \vec{\omega} \right) dm$ term?! We need to introduce the concepts of *tensor product* and *tensors* to do this. But we better talk about matrices first.

1.13 Cartesian Coordinates

So far, we have avoided using coordinates in order to emphasize the geometric, *i.e.* coordinate independent, aspects of vectors and vector operations, but coordinates are crucial for calculations. A cartesian system of coordinates consist of three oriented orthogonal lines, the x, y and z coordinate axes, passing through a point O, the origin. The orientation of the axes is usually chosen to correspond to the right-hand rule. A point P is then specified by its coordinates x, y, z with respect to the origin along each axis. To describe displacements, *i.e.* vectors, we need a set of basis vectors. It makes sense to take a basis that is aligned with the coordinate axes. Therefore we pick unit vectors $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, $\hat{\boldsymbol{z}}$ aligned with each of the axes, respectively, and pointing in the positive direction.[6] A point with coordinates (x, y, z) then has position vector $\vec{r} = x\hat{\boldsymbol{x}} + y\hat{\boldsymbol{y}} + z\hat{\boldsymbol{z}}$ or $\vec{r} = x\vec{e}_x + y\vec{e}_y + z\vec{e}_z$. Note that points have *coordinates* and vectors have *components*. It is often more convenient to use subscripts writing (x_1, x_2, x_3) in lieu of (x, y, z). In that notation, the position vectors reads $\vec{r} = x_1\hat{\boldsymbol{x}}_1 + x_2\hat{\boldsymbol{x}}_2 + x_3\hat{\boldsymbol{x}}_3$ or $\vec{r} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3$

Exercises

Express the lines, planes and spheres of section 1.8 in terms of Cartesian coordinates.

[6]Although we call $\hat{\boldsymbol{x}}$, for instance, a *unit vector*, in the sense that $\|\hat{\boldsymbol{x}}\| = 1$, unit vectors have no *physical* units. The position vector $\vec{r}$ and its components x, y, z have units of length. The unit vectors in physics are pure numbers indicating directions. "Direction vector" is a better term.

2 Matrices

2.1 Orthogonal transformations

Consider two orthonormal bases $\vec{e}_i$ and $\vec{e}_i{}'$ in 3D euclidean space so $i = 1, 2, 3$. A vector $\vec{v}$ can be expanded in terms of each bases as $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3$ and $\vec{v} = v_1'\vec{e}_1{}' + v_2'\vec{e}_2{}' + v_3'\vec{e}_3{}'$. What are the connections between the two sets of components (v_1, v_2, v_3) and (v_1', v_2', v_3')?
We can find the relations between these coordinates using geometry but it is much easier and systematic to use vectors and vector operations. Although the components are different $(v_1, v_2, v_3) \neq (v_1', v_2', v_3')$, the geometric vector $\vec{v}$ is independent of the choice of basis, thus

$$\vec{v} = \sum_{i=1}^{3} v_i'\vec{e}_i{}' = \sum_{i=1}^{3} v_i\vec{e}_i,$$

and the relationship between the two sets of coordinates are then

$$v_i' = \vec{e}_i{}' \cdot \vec{v} = \sum_{j=1}^{3}(\vec{e}_i{}' \cdot \vec{e}_j)\, v_j \equiv \sum_{j=1}^{3} Q_{ij}\, v_j, \tag{83}$$

and, likewise,

$$v_i = \vec{e}_i \cdot \vec{v} = \sum_{j=1}^{3}(\vec{e}_i \cdot \vec{e}_j{}')\, v_j' \equiv \sum_{j=1}^{3} Q_{ji}\, v_j', \tag{84}$$

where we defined

$$\boxed{Q_{ij} \equiv \vec{e}_i{}' \cdot \vec{e}_j} \tag{85}$$

These Q_{ij} coefficients are the *direction cosines*, they equal the cosine of the angle between the direction vectors $\vec{e}_i'$ and $\vec{e}_j$. *A priori*, there are 9 such coefficients. However, orthonormality of both bases imply many constraints. These constraints follow from eqns. (83), (84) which must hold for any (v_1, v_2, v_3) and (v_1', v_2', v_3'). Substituting (84) into (83) (watching out for dummy indices!) yields

$$v_i' = \sum_{k=1}^{3}\sum_{j=1}^{3} Q_{ik}Q_{jk}\, v_j', \quad \forall v' \;\Rightarrow\; \sum_{k=1}^{3} Q_{ik}Q_{jk} = \delta_{ij}. \tag{86}$$

Likewise, substituting (83) into (84) gives

$$v_i = \sum_{k=1}^{3}\sum_{j=1}^{3} Q_{ki}Q_{kj}\, v_j, \quad \forall v \;\Rightarrow\; \sum_{k=1}^{3} Q_{ki}Q_{kj} = \delta_{ij}. \tag{87}$$

These two relationships have simple geometric interpretations. Indeed $Q_{ij} = \vec{e}_i{}' \cdot \vec{e}_j$ can be interpreted as the j component of $\vec{e}_i{}'$ in the $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ basis, as well as the i component of $\vec{e}_j$ in the $\{\vec{e}_1{}', \vec{e}_2{}', \vec{e}_3{}'\}$ basis. Therefore we can write

$$\vec{e}_i{}' = \sum_{k=1}^{3}(\vec{e}_i{}' \cdot \vec{e}_k)\, \vec{e}_k = \sum_{k=1}^{3} Q_{ik}\, \vec{e}_k = Q_{i1}\vec{e}_1 + Q_{i2}\vec{e}_2 + Q_{i3}\vec{e}_3,$$

and

$$\vec{e}_j = \sum_{k=1}^{3}(\vec{e}_j \cdot \vec{e}_k{}')\, \vec{e}_k{}' = \sum_{k=1}^{3} Q_{kj}\, \vec{e}_k{}' = Q_{1j}\vec{e}_1{}' + Q_{2j}\vec{e}_2{}' + Q_{3j}\vec{e}_3{}'.$$

Then

$$\boxed{\vec{e}_i{}' \cdot \vec{e}_j{}' \equiv \sum_{k=1}^{3} Q_{ik} Q_{jk} = \delta_{ij},} \tag{88}$$

and

$$\boxed{\vec{e}_i \cdot \vec{e}_j \equiv \sum_{k=1}^{3} Q_{ki} Q_{kj} = \delta_{ij}.} \tag{89}$$

These are the **orthogonality conditions** (orthonormality, really) satisfied by the Q_{ij}.

2.2 Definitions and basic matrix operations

The 9 coefficients Q_{ij} in (83) are the elements of a 3-by-3 *matrix* Q, *i.e.* a 3-by-3 table with the first index i corresponding to the row index and the second index j to the column index. That Q was a very special *i.e.* orthogonal matrix. More generally a 3-by-3 real matrix A is a table of 9 real numbers

$$A \equiv [A_{ij}] \equiv \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}. \tag{90}$$

Matrices are denoted by a capital letter, *e.g.* A and Q and by square brackets [ ]. By convention, vectors in $\mathbb{R}^3$ are defined as 3-by-1 matrices *e.g.*

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_1 \\ x_3 \end{bmatrix},$$

although for typographical reasons we'll often write $\boldsymbol{x} = (x_1, x_2, x_3)$ but not $[x_1, x_2, x_3]$ which would denote a 1-by-3 matrix, or *row* vector. The term *matrix* is similar to *vectors* in that it implies precise rules for manipulations of these objects (for vectors these are the two fundamental addition and scalar multiplication operations with specific properties, see Sect. 1.1).

Matrix-vector multiply

Equation (83) shows how matrix-vector multiplication should be defined. The matrix vector product $A\boldsymbol{x}$ (A 3-by-3, $\boldsymbol{x} \in \mathbb{R}^3$) is a vector $\boldsymbol{b}$ in $\mathbb{R}^3$ whose i-th component is the dot-product of row i of matrix A with the column $\boldsymbol{x}$,

$$A\boldsymbol{x} = \boldsymbol{b} \quad \Leftrightarrow \quad b_i = \sum_j A_{ij} x_j.$$

The product is performed *row-by-column*. This product is defined only if the number of columns of A is equal to the number of rows of $\boldsymbol{x}$. A 2-by-1 vector cannot be multiplied by a 3-by-3 matrix.

Identity Matrix

There is a unique matrix such that $I\boldsymbol{x} = \boldsymbol{x}$, $\forall \boldsymbol{x}$. For $\boldsymbol{x} \in \mathbb{R}^3$, show that

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{91}$$

Matrix-Matrix multiply

Two successive transformation of orthogonal coordinates, *i.e.*

$$x'_i = \sum_{j=1}^{3} A_{ij}\,x_j, \quad \text{then} \quad x''_i = \sum_{j=1}^{3} B_{ij}\,x'_j$$

can be combined into one transformation from x_j to x''_i

$$x''_i = \sum_{j=1}^{3}\sum_{k=1}^{3} B_{ik}A_{kj}\,x_j \equiv \sum_{j=1}^{3} C_{ij}\,x_j$$

where

$$(BA)_{ij} \equiv C_{ij} = \sum_{k=1}^{3} B_{ik}A_{kj}. \tag{92}$$

This defines matrix multiplication. The product of two matrices BA is a matrix, C say, whose (i,j) element is the dot product of row i of B with column j of A. As for matrix-vector multiplication, the product of two matrices is done *row-by-column*. This requires that the length of the rows of the first matrix equals the length of the columns of the second, *i.e.* the number of *columns* of the first must match the number of *rows* of the second. The product of a 3-by-3 matrix and a 2-by-2 matrix is not defined. In general, $BA \neq AB$, matrix multiplication does not commute. You can visualize this by considering two successive rotation of axes, one by angle α about $\vec{\boldsymbol{e}}_3$, followed by one by β about $\vec{\boldsymbol{e}}'_2$. This is not the same as rotating by β about $\vec{\boldsymbol{e}}_2$, then by α about $\vec{\boldsymbol{e}}'_3$. You can also see it algebraically

$$(BA)_{ij} = \sum_k B_{ik}A_{kj} \neq \sum_k A_{ik}B_{kj} \equiv (AB)_{ij}.$$

Matrix transpose

The transformation (84) involves the sum $\sum_j A_{ji}x'_j$ that is similar to the matrix vector multiply except that the multiplication is column-by-column! To write this as a matrix-vector multiply, we define the *transpose matrix* A^T whose row i correspond to column i of A. If the (i,j) element of A is A_{ij} then the (i,j) element of A^T is A_{ji}

$$(A^T)_{ij} = (A)_{ji}.$$

Then

$$x_i = \sum_{j=1}^{3} A_{ji}x'_j \quad \Leftrightarrow \quad \boldsymbol{x} = A^T\boldsymbol{x}'. \tag{93}$$

A *symmetric matrix* A is such that $A = A^T$, but an *anti-symmetric matrix* A is such that $A = -A^T$. Verify (as done in class) that *the transpose of a product is equal to the product of the transposes in reverse order* $(AB)^T = B^T A^T$.

Orthogonal Matrices

Arbitrary matrices are typically denoted A, while orthogonal matrices are typically denoted Q in the literature. In matrix notation, the orthogonality conditions (88), (89) read

$$Q^TQ = QQ^T = I. \tag{94}$$

A matrix that satisfy these relationships is called an *orthogonal matrix* (it should have been called *orthonormal*). A *proper* orthogonal matrix has determinant equal to 1 and corresponds to a pure rotation. An improper orthogonal matrix has determinant -1. It corresponds to a combination of rotations and an *odd* number of reflections. The *product* of orthogonal matrices is an orthogonal matrix.
This is useful as any 3-by-3 proper orthogonal matrix can be decomposed into the product of three elementary rotations. There are several ways to define these elementary rotations but a common one that corresponds to spherical coordinates is to (1) rotate by α about $\vec{\boldsymbol{e}}_3$, (2) rotate by β about $\vec{\boldsymbol{e}}_2{}'$, (3) rotate by γ about $\vec{\boldsymbol{e}}_3{}''$. The 3 angles α, β, γ are called **Euler angles**. Hence a general 3-by-3 orthogonal matrix A can always be written as

$$Q = \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{95}$$

To define an arbitrary orthogonal matrix, we can then simply pick any three arbitrary (Euler) angles α, β, γ and construct an orthonormal matrix using (95). Another important procedure to do this is the **Gram-Schmidt** procedure: pick any three $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ and orthonormalize them, *i.e.*

(1) First, define $\boldsymbol{q}_1 = \boldsymbol{a}_1/\|\boldsymbol{a}_1\|$ and $\boldsymbol{a}_2' = \boldsymbol{a}_2 - (\boldsymbol{a}_2 \cdot \boldsymbol{q}_1)\boldsymbol{q}_1$, $\boldsymbol{a}_3' = \boldsymbol{a}_3 - (\boldsymbol{a}_3 \cdot \boldsymbol{q}_1)\boldsymbol{q}_1$,

(2) next, define $\boldsymbol{q}_2 = \boldsymbol{a}_2'/\|\boldsymbol{a}_2'\|$ and $\boldsymbol{a}_3'' = \boldsymbol{a}_3' - (\boldsymbol{a}_3' \cdot \boldsymbol{q}_2)\boldsymbol{q}_2$,

(3) finally, define $\boldsymbol{q}_3 = \boldsymbol{a}_3''/\|\boldsymbol{a}_3''\|$.

The vectors $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, $\boldsymbol{q}_3$ form an orthonormal basis. This procedure generalizes not only to any dimension but also to other vector spaces, *e.g.* to construct orthogonal polynomials.

Exercises

1. Give explicit examples of 2-by-2 and 3-by-3 symmetric and antisymmetric matrices.
2. If $\boldsymbol{x}^T = [x_1, x_2, x_3]$, calculate $\boldsymbol{x}^T\boldsymbol{x}$ and $\boldsymbol{x}\boldsymbol{x}^T$.
3. Show that $\boldsymbol{x}^T\boldsymbol{x}$ and $\boldsymbol{x}\boldsymbol{x}^T$ are symmetric (explicitly and by matrix manipulations).
4. If A is a square matrix of appropriate size, what is $\boldsymbol{x}^TA\boldsymbol{x}$?
5. Show that the product of two orthogonal matrices is an orthogonal matrix. Interpret geometrically.
6. What is the general form of a 3-by-3 orthogonal *and* symmetric matrix?
7. What is the orthogonal matrix corresponding to a reflection about the $x - z$ plane? What is its determinant?
8. What is the most general form of a 2-by-2 orthogonal matrix?

9. Suppose that you would like to rotate an object (*i.e.* a set of points) about a given axis by an angle γ. Can you explain how to do this? [Hint: (1) Translation: express coordinates of any point $\vec{r}$ with respect to any point $\vec{r}_0$ on the rotation axis: $\vec{r}-\vec{r}_0$. (2) Perform two elementary rotations to align the vertical axis with the rotation axis, *i.e.* find the Euler angles α and β. Express the coordinates of $\vec{r}-\vec{r}_0$ in that new set of coordinates. (3) Rotate the *vector* by γ, this is equivalent to multiplying by the *transpose of the matrix corresponding to rotation of axes by* γ. Then you need to re-express the coordinates in terms of the original axes! that's a few multiplication by transpose of matrices you already have].

10. What is the rotation matrix corresponding to rotation by π about $\vec{e}_2$?

11. What is the matrix corresponding to (right-handed) rotation by angle α about the direction $\vec{e}_1+\vec{e}_2+\vec{e}_3$?

12. Find the components of a vector $\vec{a}$ rotated by angle γ about the direction $\vec{e}_1+2\vec{e}_2+2\vec{e}_3$.

13. Pick three non-trivial but arbitrary vectors in $\mathbb{R}^3$ (*e.g.* using Matlab's `randn(3,3)`), then construct an orthonormal basis $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, $\boldsymbol{q}_3$ using the Gram-Schmidt procedure. Verify that the matrix $Q=[\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3]$ is orthogonal. Note in particular that the rows are orthogonal eventhough you orthogonalized the columns only.

14. Pick *two* arbitrary vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ in $\mathbb{R}^3$ and orthogonalize them to construct $\boldsymbol{q}_1$, $\boldsymbol{q}_2$. Consider the 3-by-2 matrix $Q=[\boldsymbol{q}_1,\boldsymbol{q}_2]$ and compute QQ^T and Q^TQ. Can you explain the results?

2.3 Determinant of a matrix

See earlier discussion of determinants (section on mixed product). The determinant of a matrix has the explicit formula $\det(A)=\epsilon_{ijk}A_{i1}A_{j2}A_{k3}$, the only non-zero terms are for (i,j,k) equal to a permutation of $(1,2,3)$. We can deduce several fundamental properties of determinants from that formula. We can reorder $A_{i1}A_{j2}A_{k3}$ into $A_{1l}A_{2m}A_{3n}$ using an even number of permutations if (i,j,k) is an even perm of (1,2,3) and an odd number for odd permutations. So

$$\det(A)=\epsilon_{ijk}A_{i1}A_{j2}A_{k3}=\epsilon_{lmn}A_{1l}A_{2m}A_{3n}=\det(A^T). \tag{96}$$

Another useful result is that

$$\epsilon_{ijk}A_{il}A_{jm}A_{kn}=\epsilon_{ijk}\epsilon_{lmn}A_{i1}A_{j2}A_{k3} \tag{97}$$

Then it is easy to prove that $\det(AB)=\det(A)\det(B)$:

$$\det(AB)=\epsilon_{ijk}A_{il}B_{l1}A_{jm}B_{m2}A_{kn}B_{n3}=\epsilon_{ijk}\epsilon_{lmn}A_{i1}A_{j2}A_{k3}B_{l1}B_{m2}B_{n3}=\det(A)\det(B) \tag{98}$$

One nice thing is that these results and manipulations generalize straightforwardly to any dimension.

2.4 Three views of $A\boldsymbol{x}=\boldsymbol{b}$

2.4.1 Column View

► View $\boldsymbol{b}$ as a linear combination of the columns of A.
Write A as a row of columns, $A=[\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3]$, where $\boldsymbol{a}_1^T=[a_{11},a_{21},a_{31}]$ etc., then

$$\boldsymbol{b}=A\boldsymbol{x}=x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+x_3\boldsymbol{a}_3$$

and $\boldsymbol{b}$ is a linear combination of the columns $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$. If $\boldsymbol{x}$ is unknown, the linear system of equations $A\boldsymbol{x} = \boldsymbol{b}$ will have a solution for any $\boldsymbol{b}$ if and only if the columns form a basis, *i.e.* iff $\det(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3) \equiv \det(A) \neq 0$. If the determinant is zero, then the 3 columns are in the same plane and the system will have a solution only if $\boldsymbol{b}$ is also in that plane.
As seen in earlier exercises, we can find the components (x_1, x_2, x_3) by thinking geometrically and projecting on the *reciprocal basis e.g.*

$$x_1 = \frac{\boldsymbol{b} \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3)}{\boldsymbol{a}_1 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3)} \equiv \frac{\det(\boldsymbol{b}, \boldsymbol{a}_2, \boldsymbol{a}_3)}{\det(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)}. \tag{99}$$

Likewise

$$x_2 = \frac{\det(\boldsymbol{a}_1, \boldsymbol{b}, \boldsymbol{a}_3)}{\det(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)}, \quad x_3 = \frac{\det(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{b})}{\det(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)}.$$

This is a nifty formula. Component x_i equals the determinant where vector i is replaced by $\boldsymbol{b}$ divided by the determinant of the basis vectors. You can deduce this directly from the algebraic properties of determinants, for example,

$$\det(\boldsymbol{b}, \boldsymbol{a}_2, \boldsymbol{a}_3) = \det(x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3, \boldsymbol{a}_2, \boldsymbol{a}_3) = x_1 \det(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3).$$

This is **Cramer's rule** and it generalizes to any dimension, however computing determinants in higher dimensions can be very costly and the next approach is computationally much more efficient.

2.4.2 Row View:

▶ View $\boldsymbol{x}$ as the intersection of planes perpendicular to the rows of A.

View A as a column of rows, $A = [\vec{\boldsymbol{n}}_1, \vec{\boldsymbol{n}}_2, \vec{\boldsymbol{n}}_3]^T$, where $\vec{\boldsymbol{n}}_1^T = [a_{11}, a_{12}, a_{13}]$ is the first *row* of A, etc., then

$$\boldsymbol{b} = A\boldsymbol{x} = \begin{bmatrix} \vec{\boldsymbol{n}}_1^T \\ \vec{\boldsymbol{n}}_2^T \\ \vec{\boldsymbol{n}}_3^T \end{bmatrix} \boldsymbol{x} \quad \Leftrightarrow \begin{cases} \vec{\boldsymbol{n}}_1 \cdot \boldsymbol{x} = b_1 \\ \vec{\boldsymbol{n}}_2 \cdot \boldsymbol{x} = b_2 \\ \vec{\boldsymbol{n}}_3 \cdot \boldsymbol{x} = b_3 \end{cases}$$

and $\boldsymbol{x}$ is seen as the position vector of the intersection of three planes. Recall that $\vec{\boldsymbol{n}} \cdot \boldsymbol{x} = C$ is the equation of a plane perpendicular to $\vec{\boldsymbol{n}}$ and passing through a point $\boldsymbol{x}_0$ such that $\vec{\boldsymbol{n}} \cdot \boldsymbol{x}_0 = C$, for instance the point $\boldsymbol{x}_0 = C\vec{\boldsymbol{n}}/\|\vec{\boldsymbol{n}}\|$.
To find $\boldsymbol{x}$ such that $A\boldsymbol{x} = \boldsymbol{b}$, for given $\boldsymbol{b}$ and A, we can combine the equations in order to eliminate unknowns, *i.e.*

$$\begin{cases} \vec{\boldsymbol{n}}_1 \cdot \boldsymbol{x} = b_1 \\ \vec{\boldsymbol{n}}_2 \cdot \boldsymbol{x} = b_2 \\ \vec{\boldsymbol{n}}_3 \cdot \boldsymbol{x} = b_3 \end{cases} \Leftrightarrow \begin{cases} \vec{\boldsymbol{n}}_1 \cdot \boldsymbol{x} &= b_1 \\ (\vec{\boldsymbol{n}}_2 - \alpha_2\vec{\boldsymbol{n}}_1) \cdot \boldsymbol{x} &= b_2 - \alpha_2 b_1 \\ (\vec{\boldsymbol{n}}_3 - \alpha_3\vec{\boldsymbol{n}}_1) \cdot \boldsymbol{x} &= b_3 - \alpha_3 b_1 \end{cases}$$

where we pick α_2 and α_3 such that the new normal vectors $\vec{\boldsymbol{n}}_2' = \vec{\boldsymbol{n}}_2 - \alpha_2\vec{\boldsymbol{n}}_1$ and $\vec{\boldsymbol{n}}_3' = \vec{\boldsymbol{n}}_3 - \alpha_3\vec{\boldsymbol{n}}_1$ have a zero 1st component *i.e.* $\vec{\boldsymbol{n}}_2' = (0, a_{22}', a_{23}')$, $\vec{\boldsymbol{n}}_3' = (0, a_{32}', a_{33}')$. At the next step, one defines a $\vec{\boldsymbol{n}}_3'' = \vec{\boldsymbol{n}}_3' - \beta_3\vec{\boldsymbol{n}}_2'$ picking β_3 so that the 1st and 2nd components of $\vec{\boldsymbol{n}}_3''$ are zero, *i.e.* $\vec{\boldsymbol{n}}_3'' = (0, 0, a_{33}'')$. And the resulting system of equations is then easy to solve by *backward substitution.* This is **Gaussian Elimination** which in general requires swapping of equations to avoid dividing by small numbers. We could also pick the α's and β's to orthogonalize the $\vec{\boldsymbol{n}}$'s, just as in the Gram-Schmidt procedure. That is better in terms of roundoff error and does not require equation swapping but is computationally twice as expensive as Gaussian elimination.

2.4.3 Linear Transformation of vectors into vectors

► View $\boldsymbol{b}$ as a linear transformation of $\boldsymbol{x}$.

Here A is a 'black box' that transforms the vector input $\boldsymbol{x}$ into the vector output $\boldsymbol{b}$. This is the most general view of $A\boldsymbol{x} = \boldsymbol{b}$. The transformation is linear, this means that

$$A(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) = \alpha(A\boldsymbol{x}) + \beta(A\boldsymbol{y}), \quad \forall\, \alpha, \beta \in \mathbb{R}, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n \tag{100}$$

This can be checked directly from the explicit definition of matrix-vector multiply:

$$\sum_k A_{ik}(\alpha x_k + \beta y_k) = \sum_k \alpha A_{ik} x_k + \sum_k \beta A_{ik} y_k.$$

This linearity property is a key property because if A is really a black box (*e.g.* the "matrix" is not actually known, it's just a machine that takes a vector and spits out another vector) we can figure out the effect of A onto any vector $\boldsymbol{x}$ once we know $A\vec{\boldsymbol{e}}_1$, $A\vec{\boldsymbol{e}}_2$, ..., $A\vec{\boldsymbol{e}}_n$.
This transformation view of matrices leads to the following extra rules of matrix manipulations.
Matrix-Matrix addition

$$A\boldsymbol{x} + B\boldsymbol{x} = (A+B)\boldsymbol{x} \;\Leftrightarrow\; \sum_k A_{ik}x_k + \sum_k B_{ik}x_k = \sum_k (A_{ik} + B_{ik})x_k,\ \forall x_k \tag{101}$$

so matrices are added *components by components* and $A+B = B+A$, $(A+B)+C = A+(B+C)$. The zero matrix is the matrix whose entries are all zero.
Matrix-scalar multiply

$$A(\alpha\boldsymbol{x}) = (\alpha A)\boldsymbol{x} \Leftrightarrow \sum_k A_{ik}(\alpha x_k) = \sum_k (\alpha A_{ik})x_k,\ \forall \alpha, x_k \tag{102}$$

so multiplication by a scalar is also done component by component and $\alpha(\beta A) = (\alpha\beta)A = \beta(\alpha A)$. In other words, matrices can be seen as elements of a vector space! This point of view is also useful in some instances (in fact, computer languages like C and Fortran typically store matrices as long vectors. Fortran stores it column by column, and C row by row). The set of *orthogonal* matrices does NOT form a vector space because the sum of two orthogonal matrices is not, in general, an orthogonal matrix. The set of orthogonal matrices is a *group*, the *orthogonal group* $O(3)$ (for 3-by-3 matrices). The *special orthogonal* group $SO(3)$ is the set of all 3-by-3 proper orthogonal matrices, *i.e.* orthogonal matrices with determinant =+1 that correspond to pure rotation, not reflections. The motion of a rigid body about its center of inertia is a motion in $SO(3)$, not $\mathbb{R}^3$. $SO(3)$ is the *configuration space* of a rigid body.

Exercises

▷ Pick a random 3-by-3 matrix A and a vector $\boldsymbol{b}$, ideally in matlab using its `A=randn(3,3)`, `b=randn(3,1)`. Solve $Ax = b$ using Cramer's rule and Gaussian Elimination. Ideally again in matlab, unless punching numbers into your calculator really turns you on. Matlab knows all about matrices and vectors. To compute $\det(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3) = \det(A)$ and $\det(\boldsymbol{b}, \boldsymbol{a}_2, \boldsymbol{a}_3)$ in matlab, simply use `det(A), det(b,A(:,2),A(:,3))`. Type `help matfun`, or `help elmat`, and or `demos` for a peek at all the goodies in matlab.

2.5 Eigenvalues and Eigenvectors

Problem: Given a matrix A, find $\boldsymbol{x} \neq 0$ and λ such that

$$A\boldsymbol{x} = \lambda \boldsymbol{x}. \tag{103}$$

These special vectors are *eigenvectors* for A. They are simply shrunk or elongated by the transformation A. The scalar λ is the eigenvalue. The eigenvalue problem can be rewritten

$$(A - \lambda I)\boldsymbol{x} = 0$$

where I is the identity matrix of the same size as A. This will have a non-zero solution iff

$$\det(A - \lambda I) = 0. \tag{104}$$

This is the *characteristic* equation for λ. If A is n-by-n, it is a polynomial of degree n in λ called the *characteristic polynomial*.

3 Vectors and Matrices in n dimensions

3.1 The vector space $\mathbb{R}^n$

3.2 Matrices in $\mathbb{R}^m \times \mathbb{R}^n$

3.3 Determinants