Math Tripos Part IA: Vector Calculus

Michael Li

May 7, 2015

Curves in $\mathbb{R}^3$
Parameterised curves and arc length, tangents and normals to curves in $\mathbb{R}^3$, the radius of curvature. [1]

Integration in $\mathbb{R}^2$ and $\mathbb{R}^3$
Line integrals. Surface and volume integrals: definitions, examples using Cartesian, cylindrical and spherical coordinates; change of variables. [4]

Vector operators
Directional derivatives. The gradient of a real-valued function: definition; interpretation as normal to level surfaces; examples including the use of cylindrical, spherical *and general orthogonal curvilinear* coordinates.

Divergence, curl and ∇^2 in Cartesian coordinates, examples; formulae for these operators (statement only) in cylindrical, spherical *and general orthogonal curvilinear* coordinates. Solenoidal fields, irrotational fields and conservative fields; scalar potentials. Vector derivative identities. [5]

Integration theorems
Divergence theorem, Green's theorem, Stokes's theorem, Green's second theorem: statements; informal proofs; examples; application to fluid dynamics, and to electromagnetism including statement of Maxwell's equations. [5]

Laplace's equation
Laplace's equation in $\mathbb{R}^2$ and $\mathbb{R}^3$: uniqueness theorem and maximum principle. Solution of Poisson's equation by Gauss's method (for spherical and cylindrical symmetry) and as an integral. [4]

Cartesian tensors in $\mathbb{R}^3$
Tensor transformation laws, addition, multiplication, contraction, with emphasis on tensors of second rank. Isotropic second and third rank tensors. Symmetric and antisymmetric tensors. Revision of principal axes and diagonalization. Quotient theorem. Examples including inertia and conductivity. [5]

Contents

1 Derivatives and coordinates

1.1 Derivative of functions

1.1.1 Vector functions $\mathbb{R} \to \mathbb{R}^n$

Definition. A *vector function* is a function $\mathbf{F} : \mathbb{R} \to \mathbb{R}^n$.

Definition (Derivative of vector function)**.** A vector function $\mathbf{F}(x)$ is *differentiable* if

$$\delta\mathbf{F} \stackrel{\text{def}}{=} \mathbf{F}(x + \delta x) - \mathbf{F}(x) = \mathbf{F}'(x)\delta x + o(\delta x)$$

for some $\mathbf{F}'(x)$. $\mathbf{F}'(x)$ is called the *derivative* of $\mathbf{F}(x)$.

Using differential notation, the differentiability condition can be written as

$$\mathrm{d}\mathbf{F} = \mathbf{F}'(x)\ \mathrm{d}x.$$

Given a basis $\mathbf{e}_i$ that is independent of x,we have:

Proposition.

$$\mathbf{F}'(x) = F_i'(x)\mathbf{e}_i. \quad \& \quad \frac{\mathrm{d}}{\mathrm{d}t}(f\mathbf{g}) = \frac{\mathrm{d}f}{\mathrm{d}t}\mathbf{g} + f\frac{\mathrm{d}\mathbf{g}}{\mathrm{d}t}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{g} \cdot \mathbf{h}) = \frac{\mathrm{d}\mathbf{g}}{\mathrm{d}t} \cdot \mathbf{h} + \mathbf{g} \cdot \frac{\mathrm{d}\mathbf{h}}{\mathrm{d}t} \quad \& \quad \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{g} \times \mathbf{h}) = \frac{\mathrm{d}\mathbf{g}}{\mathrm{d}t} \times \mathbf{h} + \mathbf{g} \times \frac{\mathrm{d}\mathbf{h}}{\mathrm{d}t}$$

Note that multiplication order must be retained in the case of the cross product.

1.1.2 Scalar functions $\mathbb{R}^n \to \mathbb{R}$

We also can define derivatives for a different kind of function:

Definition. A *scalar function* is a function $f : \mathbb{R}^n \to \mathbb{R}$.

Definition (Limit of vector)**.** The *limit of vectors* is defined using the norm. So $\mathbf{v} \to \mathbf{c}$ iff $|\mathbf{v} - \mathbf{c}| \to \mathbf{0}$. Similarly, $f(\mathbf{r}) = o(\mathbf{r})$ means $\frac{|f(\mathbf{r})|}{|\mathbf{r}|} \to 0$ as $\mathbf{r} \to \mathbf{0}$.

Definition. A scalar function $f(\mathbf{r})$ is *differentiable* at $\mathbf{r}$ if

$$\delta f \stackrel{\text{def}}{=} f(\mathbf{r} + \delta\mathbf{r}) - f(\mathbf{r}) = (\nabla f) \cdot \delta\mathbf{r} + o(\delta\mathbf{r})$$

for some vector ∇f, the *gradient* of f at $\mathbf{r}$.

For the derivative in a particular direction, $\delta\mathbf{r} = h\mathbf{n}$, with $\mathbf{n}$ a unit vector,

$$f(\mathbf{r} + h\mathbf{n}) - f(\mathbf{r}) = \nabla f \cdot (h\mathbf{n}) + o(h) = h(\nabla f \cdot \mathbf{n}) + o(h),$$

Definition. The *directional derivative* of f along $\mathbf{n}$ is

$$\mathbf{n} \cdot \nabla f = \lim_{h \to 0} \frac{1}{h}[f(\mathbf{r} + h\mathbf{n}) - f(\mathbf{r})],$$

It refers to how fast f changes when we move in the direction of $\mathbf{n}$.

Using this expression, the directional derivative is maximized when $\mathbf{n}$ is parallel to ∇f. So ∇f points in the direction of greatest slope. But we haven't defined the gradient yet. Using summation notation, we have:

Theorem. The gradient is

$$\nabla f = \frac{\partial f}{\partial x_i}\mathbf{e}_i$$

Hence we can write the condition of differentiability as

$$\delta f = \frac{\partial f}{\partial x_i}\delta x_i + o(\delta \mathbf{x}). \quad \text{or} \quad \mathrm{d}f = \nabla f \cdot \mathrm{d}\mathbf{r} = \frac{\partial f}{\partial x_i}\mathrm{d}x_i,$$

which is the chain rule for partial derivatives.

Example. Take $f(x,y,z) = x + e^{xy}\sin z$. Then

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) = (1 + ye^{xy}\sin z, xe^{xy}\sin z, e^{xy}\cos z)$$

At $(x,y,z) = (0,1,0)$, $\nabla f = (1,0,1)$. So f increases/decreases most rapidly for $\mathbf{n} = \pm\frac{1}{\sqrt{2}}(1,0,1)$ with a rate of change of $\pm\sqrt{2}$.

Vector fields $\mathbb{R}^n \to \mathbb{R}^m$

We are now ready to tackle the general case, which are given the fancy name of *vector fields*.

Definition. A *vector field* is a function $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$.

Definition. A vector field $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$ is *differentiable* if

$$\delta \mathbf{F} \overset{def}{=} \mathbf{F}(\mathbf{x} + \delta\mathbf{x}) - \mathbf{F}(\delta\mathbf{x}) = M\delta\mathbf{x} + o(\delta\mathbf{x})$$

for some $n \times m$ matrix M. M is the *derivative* of $\mathbf{F}$.

Given an arbitrary function $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$ that maps $\mathbf{x} \mapsto \mathbf{y}$ and a choice of basis, we can write $\mathbf{F}$ as a set of m functions $y_j = F_j(\mathbf{x})$ such that $\mathbf{y} = (y_1, y_2, \cdots, y_m)$. Then

$$\mathrm{d}y_j = \frac{\partial F_j}{\partial x_i}\mathrm{d}x_i.$$

so the derivative is

Theorem. The derivative of $\mathbf{F}$ is given by

$$M_{ji} = \frac{\partial y_j}{\partial x_i}.$$

Definition. A function is *smooth* if it is infinitely differentiable.

1.2 Chain rule

Theorem (Chain rule)**.** Suppose $g : \mathbb{R}^p \to \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^m$. Suppose the coordinates of vectors in $\mathbb{R}^p, \mathbb{R}^n$ and $\mathbb{R}^m$ are u_a, x_i and y_r respectively. Then:

$$\frac{\partial y_r}{\partial u_a} = \frac{\partial y_r}{\partial x_i}\frac{\partial x_i}{\partial u_a},$$

with summation implied. Equivalently in matrix and operator form:

$$M(f \circ g)_{ra} = M(f)_{ri}M(g)_{ia}. \quad \& \quad \frac{\partial}{\partial u_a} = \frac{\partial x_i}{\partial u_a}\frac{\partial}{\partial x_i}.$$

1.3 Inverse functions

Suppose $g, f : \mathbb{R}^n \to \mathbb{R}^n$ are inverse functions, ie. $g \circ f = f \circ g = \mathrm{id}$. Suppose that $f(\mathbf{x}) = \mathbf{u}$ and $g(\mathbf{u}) = \mathbf{x}$.

Since the derivative of the identity function is the identity matrix (if you differentiate $\mathbf{x}$ wrt to $\mathbf{x}$, you get 1), we must have

$$M(f \circ g) = I. \;\Rightarrow\; M(g) = M(f)^{-1}.$$

We derive this result more formally by noting

$$\frac{\partial u_b}{\partial u_a} = \delta_{ab}. \Rightarrow \frac{\partial u_b}{\partial x_i}\frac{\partial x_i}{\partial u_a} = \delta_{ab},$$

Example. For $n = 2$, write $u_1 = \rho$, $u_2 = \varphi$ and let $x_1 = \rho\cos\varphi$ and $x_2 = \rho\sin\varphi$. Then the function used to convert between the coordinate systems is $g(u_1, u_2) = (u_2\cos u_1, u_2 \sin u_1)$. Then

$$M(g) = \begin{pmatrix} \partial x_1/\partial\rho & \partial x_1/\partial\varphi \\ \partial x_2/\partial\rho & \partial x_2/\partial\varphi \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\rho\sin\varphi \\ \sin\varphi & \rho\cos\varphi \end{pmatrix}$$

We can then use $M(g)$ to calculate $M(f)$ and associate related basis vectors that point to directions of increasing ρ and φ, obtained by differentiating $\mathbf{r}$ with respect to the variables and then normalizing. Note that $\det M(f)\det M(g) = 1$.

2 Curves and Line

2.1 Parametrised curves, lengths and arc length

Definition. Given a curve C in $\mathbb{R}^n$, a *parametrisation* of it is a continuous and invertible function $\mathbf{r} : D \to \mathbb{R}^n$ for some $D \subseteq \mathbb{R}$ whose image is C. $\mathbf{r}'(u)$ is a vector tangent to the curve at each point. A parametrization is *regular* if $\mathbf{r}'(u) \neq 0$ for all u.

If we change u (and hence $\mathbf{r}$) by a small amount, then the distance $|\delta\mathbf{r}|$ is roughly equal to the change in arclength δs. So $\delta s = |\delta\mathbf{r}| + o(\delta\mathbf{r})$. Then we have

Proposition. Let s denote the arclength of a curve $\mathbf{r}(u)$. Then

$$\frac{\mathrm{d}s}{\mathrm{d}u} = \pm\left|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}u}\right| = \pm|\mathbf{r}'(u)|.$$

The sign depends on whether it is in the direction of increasing or decreasing arclength.

Example. Consider a helix described by $\mathbf{r}(u) = (3\cos u, 3\sin u, 4u)$. Then

$$\frac{\mathrm{d}s}{\mathrm{d}u} = |\mathbf{r}'(u)| = |(-3\sin u, 3\cos u, 4)| = \sqrt{3^2 + 4^2} = 5$$

So $s = 5u$. ie. the arclength from $\mathbf{r}(0)$ and $\mathbf{r}(u)$ is $s = 5u$.

We can change parametrisation of $\mathbf{r}$ by taking an invertible smooth function $u \mapsto \tilde{u}$, and have a new parametrization $\mathbf{r}(\tilde{u}) = \mathbf{r}(\tilde{u}(u))$. Then by the chain rule,

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}u} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\tilde{u}} \times \frac{\mathrm{d}\tilde{u}}{\mathrm{d}u} \quad \& \quad \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\tilde{u}} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}u}\Big/\frac{\mathrm{d}\tilde{u}}{\mathrm{d}u}$$

It is often convenient to use the arclength s as the parameter. Then the tangent vector will always have unit length since the proposition above yields

$$|\mathbf{r}'(s)| = \frac{\mathrm{d}s}{\mathrm{d}s} = 1.$$

Definition (Scalar line element)**.** The *scalar line element* of C is $\mathrm{d}s = \pm|\mathbf{r}'(u)|\mathrm{d}u$.

2.2 Line integrals of vector fields

Definition. The *line integral* of a smooth vector field $\mathbf{F}(\mathbf{r})$ along a path C parametrised by $\mathbf{r}(u)$ with along the direction (orientation) $\mathbf{r}(\alpha) \to \mathbf{r}(\beta)$ is

$$\int_C \mathbf{F}(\mathbf{r}) \cdot \mathrm{d}\mathbf{r} = \int_\alpha^\beta \mathbf{F}(\mathbf{r}(u)) \cdot \mathbf{r}'(u)\mathrm{d}u.$$

We say $\mathrm{d}\mathbf{r} = \mathbf{r}'(u)\mathrm{d}u$ is the *line element* on C. Note that the upper and lower limits of the integral are the end point and start point respectively.

Example. Take $\mathbf{F}(\mathbf{r}) = (xe^y, z^2, xy)$ and we want to find the line integral from $\mathbf{a} = (0,0,0)$ to $\mathbf{b} = (1,1,1)$.

We first integrate along the curve $C_1 : \mathbf{r}(u) = (u, u^2, u^3)$. Then $\mathbf{r}'(u) = (1, 2u, 3u^2)$, and $\mathbf{F}(\mathbf{r}(u)) = (ue^{u^2}, u^6, u^3)$. Then

$$\begin{aligned} \int_{C_1} \mathbf{F} \cdot \mathrm{d}\mathbf{r} &= \int_0^1 \mathbf{F} \cdot \mathbf{r}'(u)\mathrm{d}u = \int_0^1 ue^{u^2} + 2u^7 + 3u^5 \ \mathrm{d}u \\ &= \frac{e}{2} - \frac{1}{2} + \frac{1}{4} + \frac{1}{2} = \frac{e}{2} + \frac{1}{4} \end{aligned}$$

Now we try to integrate along $C_2 : \mathbf{r}(t) = (t,t,t)$. So $\mathbf{r}'(t) = (1,1,1)$.

$$\int_{C_2} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int \mathbf{F} \cdot \mathbf{r}'(t)\mathrm{d}t = \int_0^1 te^t + 2t^2 \ \mathrm{d}t = \frac{5}{3}$$

Note that the line integral depends on the curve C in general, not just $\mathbf{a}, \mathbf{b}$.

We can also use the arclength s as parameter. Since $\mathrm{d}\mathbf{r} = \mathbf{t}\ \mathrm{d}s$, with $\mathbf{t}$ being the unit tangent vector, we have

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_C \mathbf{F} \cdot \mathbf{t}\ \mathrm{d}s.$$

Note that we do not necessarily have to integrate $\mathbf{F} \cdot \mathbf{t}$ with respect to s. We can also integrate a scalar function as a function of s, $\int_C f(s)\ \mathrm{d}s$. By convention, this is calculated in the direction of increasing s. In particular, we have $\int_C 1\ \mathrm{d}s =$ length of C.

Definition (Closed curve)**.** A *closed curve* is a curve with the same start and end point. The line integral along a closed curve is (sometimes) written as $\oint$ and is (sometimes) called the *circulation* of $\mathbf{F}$ around C.

2.3 Sums of curves and integrals

Definition (Piecewise smooth curve)**.** A *piecewise smooth curve* is a curve $C = C_1 + C_2 + \cdots + C_n$ with all C_i smooth with regular parametrisations. The line integral over a piecewise smooth C is

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_{C_1} \mathbf{F} \cdot \mathrm{d}\mathbf{r} + \int_{C_2} \mathbf{F} \cdot \mathrm{d}\mathbf{r} + \cdots + \int_{C_n} \mathbf{F} \cdot \mathrm{d}\mathbf{r}.$$

2.4 Gradients and Differentials

2.4.1 Line integrals and Gradients

Theorem. If $\mathbf{F} = \nabla f(\mathbf{r})$, F is a *conservative vector field*. Then:

$$\int_C \mathbf{F} \cdot \mathbf{r} = f(\mathbf{b}) - f(\mathbf{a}).$$

In particular, the line integral does NOT depend on the curve, but the end points only. This is the vector counterpart of the fundamental theorem of calculus. A special case is when C is a closed curve, then $\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0$.

Proof.

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_C \nabla f \cdot \mathrm{d}\mathbf{r} = \int \nabla f \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}u}\, \mathrm{d}u$$

for any parametrisation of C with $\mathbf{a} = \mathbf{r}(\alpha)$ and $\mathbf{b} = \mathbf{r}(\beta)$. So by the chain rule, this is equal to

$$\int_\alpha^\beta \frac{\mathrm{d}}{\mathrm{d}u}(f(\mathbf{r}(u)))\, \mathrm{d}u = [f(\mathbf{r}(u))]_\alpha^\beta = f(\mathbf{b}) - f(\mathbf{a}).$$

□

Definition (Conservative vector field)**.** If $\mathbf{F} = \nabla f$ for some f, the $\mathbf{F}$ is called a *conservative vector field.*

2.4.2 Differentials

It is convenient to treat differentials $\mathbf{F} \cdot \mathrm{d}\mathbf{r} = F_i \mathrm{d}x_i$ as if they were objects by themselves, which we can integrate along curves if we feel like doing so.

Definition (Exact differential)**.** A differential $\mathbf{F} \cdot \mathrm{d}\mathbf{r}$ is *exact* if there is an f such that $\mathbf{F} = \nabla f$. Then

$$\mathrm{d}f = \nabla f \cdot \mathrm{d}\mathbf{r} = \frac{\partial f}{\partial x_i}\mathrm{d}x_i.$$

Proposition. If $\mathbf{F} = \nabla f$ for some f, then

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}.$$

This is because both are equal to $\partial^2 f/\partial x_i \partial x_j$.

Differentials can be manipulated using (for constant λ, μ):

Proposition.

$$\mathrm{d}(\lambda f + \mu g) = \lambda \mathrm{d}f + \mu \mathrm{d}g \quad \& \quad \mathrm{d}(fg) = (\mathrm{d}f)g + f(\mathrm{d}g)$$

Example. Consider

$$\int_C 3x^2 y \sin z\ \mathrm{d}x + x^3 \sin z\ \mathrm{d}y + x^3 y \cos z\ \mathrm{d}z.$$

We see that if we integrate the first term with respect to x, we obtain $x^3 y \sin z$. We obtain the same thing if we integrate the second and third term. So this is equal to

$$\int_C \mathrm{d}(x^3 y \sin z) = [x^3 y \sin z]_{\mathbf{a}}^{\mathbf{b}}.$$

2.5 Work and potential energy

Definition. If $\mathbf{F}(\mathbf{r})$ is a force, then $\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r}$ is the *work done* by the force along the curve C.

Consider a point particle moving under $\mathbf{F}(\mathbf{r})$ according to Newton's second law: $\mathbf{F}(\mathbf{r}) = m\ddot{\mathbf{r}}$. We have kinetic energy $K(t)$ and its derivative as:

$$K(t) = \frac{1}{2}m\dot{\mathbf{r}}^2 \quad \& \quad \frac{\mathrm{d}}{\mathrm{d}t}K(t) = m\dot{\mathbf{r}}\cdot\ddot{\mathbf{r}} = \mathbf{F}\cdot\mathbf{r}.$$

Suppose the path of particle is a curve C from $\mathbf{a} = \mathbf{r}(\alpha)$ to $\mathbf{b} = \mathbf{r}(\beta)$, Then

$$K(\beta) - K(\alpha) = \int_\alpha^\beta \frac{\mathrm{d}K}{\mathrm{d}t}\ \mathrm{d}t = \int_\alpha^\beta \mathbf{F}\cdot\dot{\mathbf{r}}\ \mathrm{d}t = \int_C \mathbf{F}\cdot\mathrm{d}\mathbf{r}.$$

So the work done on the particle is the change in kinetic energy.

Definition. Given a conservative force $\mathbf{F} = -\nabla V$, $V(\mathbf{x})$ is the *potential energy*. Then

$$\int_C \mathbf{F}\cdot\mathrm{d}\mathbf{r} = V(\mathbf{a}) - V(\mathbf{b}).$$

So the work done (gain in kinetic energy) is the loss in potential energy. So the total energy $K + V$ is conserved, ie. constant during motion.

In fact, We see that energy is conserved (*only*) for conservative forces.

3 Integration in $\mathbb{R}^2$ and $\mathbb{R}^3$

3.1 Integrals over subsets of $\mathbb{R}^2$

3.1.1 Definition as the limit of as sum

Definition (Surface integral)**.** Let $D \subseteq \mathbb{R}^2$. Let $\mathbf{r}(x, y)$ be in Cartesian coordinates. We can approximate D by N disjoint subsets of simple shapes, eg. triangles, parallelograms. These shapes are labelled by I and have areas δA_i. Then we take the limit as $\ell \to 0$, $N \to \infty$, and the union of the pieces tends to D. For a function $f(\mathbf{r})$, we define the *surface integral* as

$$\int_D f(\mathbf{r})\ \mathrm{d}A = \lim_{\ell\to 0}\sum_I f(\mathbf{r}_i)\delta A_i.$$

where $\mathbf{r}_i$ is some point within each subset A_i. The integral *exists* if the limit is well-defined and exists.

If we put $z = f(x, y)$ and plot out the surface $z = f(x, y)$, then the area integral is the volume under the surface.

3.1.2 Evaluation as multiple integrals

We choose the small sets in the definition to be rectangles, each of size $\delta A_I = \delta x\delta y$ (in cartesian coordinates). We sum over subsets in a narrow horizontal strip of height δy with y and δy held constant. Take the limit as $\delta x \to 0$. We get $\delta y \int_{x_y} f(y, x)\ \mathrm{d}x$ with range $x_y \in \{x : (x, y) \in D\}$.

Sum over all such strips and take $\delta y \to 0$, giving

Proposition.

$$\int_D f(x,y)\ \mathrm{d}A = \int_Y \left(\int_{x_y} f(x,y)\ \mathrm{d}x\right) \mathrm{d}y.$$

with x_y ranging over $\{x : (x,y) \in D\}$.

Similarly:

$$\int_D f(x,y)\ \mathrm{d}A = \int_X \left(\int_{y_x} f(x,y)\ \mathrm{d}y\right) \mathrm{d}x.$$

Theorem (Fubini's theorem). If f is a continuous function and D is a compact (ie. closed and bounded) subset of $\mathbb{R}^2$, then

$$\int\int f\ \mathrm{d}x\ \mathrm{d}y = \int\int f\ \mathrm{d}y\ \mathrm{d}x.$$

Definition. The *area element* is $\mathrm{d}A$.

We have a few easy special cases.

Definition (Separable function). A function $f(x,y)$ is separable if it can be written as $f(x,y) = h(y)g(x)$.

Proposition. Take separable $f(x,y) = h(y)g(x)$ and D be a rectangle $\{(x,y) : a \leq x \leq b, c \leq y \leq d\}$. Then

$$\int_D f(x,y)\ \mathrm{d}x\ \mathrm{d}y = \left(\int_a^b g(x)\ \mathrm{d}x\right)\left(\int_c^d h(y)\ \mathrm{d}y\right)$$

3.2 Change of variables for an integral in $\mathbb{R}$

Proposition. Suppose we have a change of variables $(x,y) \leftrightarrow (u,v)$ that is smooth and invertible, with regions D, D' in one-to-one correspondence. Then

$$\int_D f(x,y)\ \mathrm{d}x\ \mathrm{d}y = \int_D f(x(u,v), y(u,v))|J|\ \mathrm{d}u\ \mathrm{d}v,$$

where J is the *Jacobian*:

$$J = \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix}$$

is the Jacobian. In other words,

$$\mathrm{d}x\ \mathrm{d}y = |J|\ \mathrm{d}u\ \mathrm{d}v.$$

Proof. Since we are writing $(x(u,v), y(u,v))$, we are actually transforming from (u,v) to (x,y) and not the other way round.

Suppose we start with an area $\delta A' = \delta u \delta v$ in the (u,v) plane. Then by Taylors' theorem, we have

$$\delta x = x(u + \delta u, v + \delta v) - x(u,v) \approx \frac{\partial x}{\partial u}\delta u + \frac{\partial x}{\partial v}\delta v.$$

We have a similar expression for δy and we obtain

$$\begin{pmatrix} \delta x \\ \delta y \end{pmatrix} \approx \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} \begin{pmatrix} \delta u \\ \delta v \end{pmatrix}$$

Recall from Vectors and Matrices that the determinant of the matrix is how much it scales up an area. So the area formed by δx and δy is $|J|$ times the area formed by δu and δv. Hence

$$\mathrm{d}x\ \mathrm{d}y = |J|\ \mathrm{d}u\ \mathrm{d}v.$$

□

3.3 Generalization to $\mathbb{R}^3$

3.3.1 Definitions

We will do exactly the same thing as we just did, but with one more dimension:

Definition (Volume integral)**.** Consider a volume $V \subseteq \mathbb{R}^3$ with position vector $\mathbf{r} = (x, y, z)$. With the same exact terms, assume that as $\ell \to 0$ and $N \to \infty$, the union of the small subsets tend to V. Then

$$\int_V f(\mathbf{r})\ \mathrm{d}V = \lim_{\ell \to 0} \sum_I f(\mathbf{r}_I^*) \delta V_I,$$

where $\mathbf{r}_I^*$ is any chosen point in each small subset.

We evaluate it the same way, taking two of the variables δx, δy to 0 and summing over the third one before evaluating the whole of (x, y).

Often, $f(\mathbf{r})$ is the density of some quantity, and is usually denoted as ρ.

Definition (Volume element)**.** The *volume element* is $\mathrm{d}V = \mathrm{d}x\ \mathrm{d}y\ \mathrm{d}z$.

3.3.2 Change of variables in $\mathbb{R}^3$

Proposition.

$$\int_V f\ \mathrm{d}x\ \mathrm{d}y\ \mathrm{d}z = \int_V f|J|\ \mathrm{d}u\ \mathrm{d}v\ \mathrm{d}w \quad \& \quad J = \frac{\partial(x, y, z)}{\partial(u, v, w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix}$$

Example. Suppose $f(\mathbf{r})$ is spherically symmetric and V is a sphere of radius a centered on the origin. Then

$$\begin{aligned} \int_V f\ \mathrm{d}V &= \int_{r=0}^{a} \int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} f(r) r^2 \sin\theta\ \mathrm{d}r\ \mathrm{d}\theta\ \mathrm{d}\varphi \\ &= \int_0^a \mathrm{d}r \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\varphi\ r^2 f(r) \sin\theta \\ &= \int_0^a \mathrm{d}r\ r^2 f(r) \Big[-\cos\theta\Big]_0^\pi \Big[\varphi\Big]_0^{2\pi} \\ &= 4\pi \int_0^a f(r) r^2\ \mathrm{d}r. \end{aligned}$$

This is a useful general result. We understand it as the sum of spherical shells of thickness δr and volume $4\pi r^2 \delta r$.

3.4 Further generalizations

3.4.1 Integration in R^n

$\int_D f(x_1, x_2, \cdots x_n)\ \mathrm{d}x_1\ \mathrm{d}x_2 \cdots\ \mathrm{d}x_n$ is simply the integration over an n-dimensional volume. The change of variable formula is

Proposition.

$$\int_D f(x_1, x_2, \cdots x_n)\ \mathrm{d}x_1\ \mathrm{d}x_2 \cdots\ \mathrm{d}x_n = \int_{D'} f(\{x_i(\mathbf{u})\})|J|\ \mathrm{d}u_1\ \mathrm{d}u_2 \cdots\ \mathrm{d}u_n.$$

3.4.2 Change of variables for $n = 1$

In the $n = 1$ case, the Jacobian is $\frac{\mathrm{d}x}{\mathrm{d}u}$. However, we use the following formula for change of variables:

$$\int_D f(x)\ \mathrm{d}x = \int_{D'} f(x(u)) \left|\frac{\mathrm{d}x}{\mathrm{d}u}\right|\ \mathrm{d}u.$$

We introduce the modulus because of our natural convention about the lower bound being ”smaller”.

This is not easily generalized to higher dimensions, so we don’t employ the same trick in other cases.

3.4.3 Vector valued integrals

We can define $\int_V \mathbf{F}(\mathbf{r})\ \mathrm{d}V$ in a similar way to $\int_V f(\mathbf{r})\ \mathrm{d}V$as the limit of a sum over small contributions of volume. In practice, we integrate them componentwise. If

$$\mathbf{F}(\mathbf{r}) = F_i(\mathbf{r})\mathbf{e}_i,$$

then

$$\int_V \mathbf{F}(\mathbf{r})\ \mathrm{d}V = \int_V (F_i(\mathbf{r})\ \mathrm{d}V)\mathbf{e}_i.$$

4 Surfaces and surface integrals

4.1 Surfaces and Normal

Let f be a smooth function on $\mathbb{R}^3$ and a constant c. Then $f(\mathbf{r}) = c$ defines a smooth surface. Consider any curve $\mathbf{r}(u)$ on S. Then by the chain rule, if we differentiate $f(\mathbf{r}) = c$ with respect to u:

$$\frac{\mathrm{d}}{\mathrm{d}u}[f(\mathbf{r}(u))] = \nabla f \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}u} = 0.$$

This means that ∇f is always perpendicular to the tangent, $\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}u}$. Since this is true for *any* curve $\mathbf{r}(u)$:

Proposition. ∇f is the normal to the surface $f(\mathbf{r}) = c$.

Example. Take the sphere $f(\mathbf{r}) = x^2 + y^2 + z^2 = c$ for $c > 0$. Then $\nabla f = 2(x, y, z) = 2\mathbf{r}$, which is clearly normal to the sphere.

Definition (Boundary)**.** A surface S can be defined to have a *boundary* ∂S consisting of a piecewise smooth curve.

A surface is *bounded* if it can be contained in a solid sphere, *unbounded* otherwise. A bounded surface with no boundary is called *closed* (eg. sphere).

Definition (Orientable surface)**.** At each point, there is a unit normal $\mathbf{n}$ that's unique up to a sign.

If we can find a consistent choice of $\mathbf{n}$ that varies smoothly across S, then we say ∂S is *orientable*, and the choice of sign is the *orientation* of the surface.

Most surfaces we encounter are orientable.

4.2 Parametrized surfaces and area

A surface S can also be parametrised by $\mathbf{r}(u, v)$ with parameters u, v. S is swept out as u and v vary in some region D. By the chain rule,

$$\delta r = \frac{\partial \mathbf{r}}{\partial u}\delta u + \frac{\partial \mathbf{r}}{\partial v}\delta v + o(\delta u, \delta v).$$

Definition. A parametrization is *regular* if for all u, v, there are always two independent tangent directions:

$$\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \neq 0,$$

The parametrizations we use will all be regular.

The small changes $\delta u, \delta v$ make a small parallelogram on S. Up to o-terms, it has a vector area of

$$\delta \mathbf{S} = \frac{\partial \mathbf{r}}{\delta u} \times \frac{\partial \mathbf{r}}{\partial v} = \mathbf{n}\ \delta S.$$

The order of u, v gives the choice of unit normal. By summing and taking limits,we have the *scalar area element*:

$$\mathrm{d}S = \left|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right|\ \mathrm{d}u\ \mathrm{d}v$$

Example. Let S be part of a sphere of radius a with $0 \leq \theta \leq \alpha$. then

$$\mathbf{r}(\theta, \varphi) = (a\cos\varphi\sin\theta, a\sin\theta\sin\varphi, a\cos\theta) = a\mathbf{e}_r.$$

Then

$$\frac{\partial \mathbf{r}}{\partial \theta} = a\mathbf{e}_\theta. \quad \& \quad \frac{\partial \mathbf{r}}{\partial \varphi} = a\sin\theta\mathbf{e}_\varphi. \Rightarrow \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi} = a^2\sin\theta\,\mathbf{e}_r.$$

So

$$\mathrm{d}S = a^2\sin\theta\ \mathrm{d}\theta\ \mathrm{d}\varphi.$$

Then the area is

$$\int_0^{2\pi}\int_0^{\alpha} a^2\sin\theta\ \mathrm{d}\theta\ \mathrm{d}\varphi = 2\pi a^2(1 - \cos\alpha).$$

4.3 Surface integral of vector fields

Let S be a smooth surface parametrized by $\mathbf{r}(u,v)$, where (u,v) takes values in D.

Definition (Surface integral)**.** The *surface integral* or *flux* of a vector field $\mathbf{F}(\mathbf{r})$ over S is defined by

$$\int_S \mathbf{F}(\mathbf{r})\cdot \mathrm{d}\mathbf{S} = \int_S \mathbf{F}(\mathbf{r})\cdot \mathbf{n}\,\mathrm{d}S = \int_D \mathbf{F}(\mathbf{r}(u,v))\cdot\left(\frac{\partial \mathbf{r}}{\partial u}\times\frac{\partial \mathbf{r}}{\partial v}\right)\,\mathrm{d}u\,\mathrm{d}v.$$

Intuitively, this is the total amount of $\mathbf{F}$ passing through S. For example, if $\mathbf{F}$ is the electric field, the flux is the amount of electric field passing through a surface. If we change orientation, we changed the sign of the integral.

What happens when we change parametrization? Let $\mathbf{r}(u,v)$ and $\mathbf{r}(\tilde{u},\tilde{v})$ be two regular parametrizations for the surface. By the chain rule,

$$\frac{\partial \mathbf{r}}{\partial u} = \frac{\partial \mathbf{r}}{\partial \tilde{u}}\frac{\partial \tilde{u}}{\partial u} + \frac{\partial \mathbf{r}}{\partial \tilde{v}}\frac{\partial \tilde{v}}{\partial u} \qquad \& \qquad \frac{\partial \mathbf{r}}{\partial v} = \frac{\partial \mathbf{r}}{\partial \tilde{u}}\frac{\partial \tilde{u}}{\partial v} + \frac{\partial \mathbf{r}}{\partial \tilde{v}}\frac{\partial \tilde{v}}{\partial v}$$

So

$$\frac{\partial \mathbf{r}}{\partial u}\times\frac{\partial \mathbf{r}}{\partial v} = \frac{\partial(\tilde{u},\tilde{v})}{\partial(u,v)}\frac{\partial \mathbf{r}}{\partial \tilde{u}}\times\frac{\partial \mathbf{r}}{\partial \tilde{v}}, \quad \& \quad \mathrm{d}\tilde{u}\,\mathrm{d}\tilde{v} = \frac{\partial(\tilde{u},\tilde{v})}{\partial(u,v)}\,\mathrm{d}u\,\mathrm{d}v,$$

where $\frac{\partial(\tilde{u},\tilde{v})}{\partial(u,v)}$ is the Jacobian.

We recover the formula

$$\mathrm{d}S = \left|\frac{\partial \mathbf{r}}{\partial u}\times\frac{\partial \mathbf{r}}{\partial v}\right|\,\mathrm{d}u\,\mathrm{d}v = \left|\frac{\partial \mathbf{r}}{\partial \tilde{u}}\times\frac{\partial \mathbf{r}}{\partial \tilde{v}}\right|\,\mathrm{d}\tilde{u}\,\mathrm{d}\tilde{v}.$$

provided (u,v) and $(\tilde{u},\tilde{v})$ have the same orientation.

4.4 Change of variables in $\mathbb{R}^2$ and $\mathbb{R}^3$ revisited

In this section, we derive the formulae in a slightly different way.

Change of variable formula in $\mathbb{R}^2$

We derive the 2D change of variable formula from the 3D surface integral formula.

Consider a subset S of the plane $\mathbb{R}^2$ parametrized by $\mathbf{r}(x(u,v),y(u,v))$. We can embed it to $\mathbb{R}^3$ as $\mathbf{r}(x(u,v),y(u,v),0)$. Then

$$\frac{\partial \mathbf{r}}{\partial u}\times\frac{\partial \mathbf{r}}{\partial v} = (0,0,J),$$

with J being the Jacobian. Therefore

$$\int_S f(\mathbf{r})\,\mathrm{d}S = \int_D f(\mathbf{r}(u,v))|\frac{\partial \mathbf{r}}{\partial u}\times\frac{\partial \mathbf{r}}{\partial v}|\,\mathrm{d}u\,\mathrm{d}v = \int_D f(\mathbf{r}(u,v))|J|\,\mathrm{d}u\,\mathrm{d}v,$$

and we recover the formula for changing variables in $\mathbb{R}^2$.

Change of variable formula in $\mathbb{R}^3$

In $\mathbb{R}^3$, suppose we have a volume parametrised by $\mathbf{r}(u, v, w)$. Then

$$\delta\mathbf{r} = \frac{\partial\mathbf{r}}{\partial u}\delta u + \frac{\partial\mathbf{r}}{\partial v}\delta v + \frac{\partial\mathbf{r}}{\partial w}\delta w + o(\delta u, \delta v, \delta w).$$

Then the cuboid $\delta u, \delta v, \delta w$ in u, v, w space is mapped to a parallelopiped of volume

$$\delta V = \left|\frac{\partial\mathbf{r}}{\partial u}\delta u \cdot \left(\frac{\partial\mathbf{r}}{\partial v}\delta v \times \frac{\partial\mathbf{r}}{\partial w}\delta w\right)\right| = |J|\ \delta u\ \delta v\ \delta w.$$

So $\mathrm{d}V = |J|\ \mathrm{d}u\ \mathrm{d}v\ \mathrm{d}w$.

5 Geometry of curves and surfaces

Let $\mathbf{r}(s)$ be a curve parametrized by arclength s. Since $\mathbf{t}(s) = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}$ is a unit vector, $\mathbf{t}\cdot\mathbf{t} = 1$. Differentiating yields $\mathbf{t}\cdot\mathbf{t}' = 0$. So $\mathbf{t}'$ is a normal to the curve if $\mathbf{t}' \neq 0$.

Definition. Write $\mathbf{t}' = \kappa\mathbf{n}$, where $\mathbf{n}$ is a unit vector. Then $\mathbf{n}(s)$ is called the *principal normal* and $\kappa(s)$ is called the *curvature*. We make $\kappa > 0$ by choosing an appropriate direction for $\mathbf{n}$.

How can we interpret κ as the curvature? Consider the vector equation for a circle passing through $\mathbf{r}(0)$ with radius a in the plane defined by $\mathbf{t}$ and $\mathbf{n}$.

Then the equation of the circle is

$$\mathbf{r} = \mathbf{r}(0) + a(1 - \cos\theta)\mathbf{n} + a\sin\theta\mathbf{t}.$$

We can expand this to obtain

$$\mathbf{r} = \mathbf{r}(0) + a\theta\mathbf{t} + \frac{1}{2}\theta^2 a\mathbf{n} + o(\theta^3).$$

Since the arclength $s = a\theta$, we obtain

$$\mathbf{r} = \mathbf{r}(0) + s\mathbf{t} + \frac{1}{2}\frac{1}{a}s^2\mathbf{n} + O(s^3).$$

By Taylor's theorem we have:

$$\mathbf{r}(s) = \mathbf{r}(0) + s\mathbf{t}(0) + \frac{1}{2}\kappa(0)s^2\mathbf{n} + O(s^3).$$

So $\kappa = 1/a$, for a the radius of the circle of best fit.

Definition (Radius of curvature)**.** The *radius of curvature* of a curve at a point $\mathbf{r}(s)$ is $1/\kappa(s)$.

Since we are in 3D, given $\mathbf{t}(s)$ and $\mathbf{n}(s)$, there is another normal to the curve. We can add a third normal to generate an orthonormal basis.

Definition (Binormal)**.** The *binormal* of a curve is $\mathbf{b} = \mathbf{t}\times\mathbf{n}$.

Definition (Torsion)**.** Let $\mathbf{b}' = -\tau\mathbf{n}$. Then τ is the *torsion*.

Note that this makes sense, since $\mathbf{b}'$ is both perpendicular to $\mathbf{t}$ and $\mathbf{b}$, and hence must be in the same direction as $\mathbf{n}$. ($\mathbf{b}' = \mathbf{t}'\times\mathbf{n} + \mathbf{t}\times\mathbf{n}' = \mathbf{t}\times\mathbf{n}'$, so $\mathbf{b}'$ is perpendicular to $\mathbf{t}$; and $\mathbf{b}\cdot\mathbf{b} = 1 \Rightarrow \mathbf{b}\cdot\mathbf{b}' = 0$. So $\mathbf{b}'$ is perpendicular to $\mathbf{b}$).

6 Div, Grad, Curl and ∇

We regard the gradient ∇f as obtained from the scalar field f by applying

$$\nabla = \mathbf{e}_i \frac{\partial}{\partial x_i} \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right).$$

for cartesian coordinates x_i and orthonormal righthanded basis $\mathbf{e}_i$. ∇ (*nabla* or *del*) is both an operator and a vector. We can apply it to a vector field $\mathbf{F}(\mathbf{r}) = \mathbf{F}_i(\mathbf{r})\mathbf{e}_i$ using the scalar or vector product.

Definition (Divergence)**.** The *divergence* or *div* of $\mathbf{F}$ is

$$\nabla \cdot \mathbf{F} = \frac{\partial F_i}{\partial x_i} = \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3}.$$

Definition (Curl)**.** The *curl* of $\mathbf{F}$ is

$$\nabla \times \mathbf{F} = \varepsilon_{ijk} \frac{\partial F_k}{\partial x_j} \mathbf{e}_i = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix}$$

Note that ∇ is an operator, so ordering is important. For example,

$$\mathbf{F} \cdot \nabla = F_i \frac{\partial}{\partial x_i} \quad \& \quad \mathbf{F} \times \nabla = \mathbf{e}_k \varepsilon_{ijk} F_i \frac{\partial}{\partial x_j}$$

are differential operators.

Proposition. Let f, g be scalar functions, $\mathbf{F}, \mathbf{G}$ be vector functions, and μ, λ be constants. Then

$$\begin{aligned} \nabla(\lambda f + \mu g) &= \lambda \nabla f + \mu \nabla g \\ \nabla \cdot (\lambda \mathbf{F} + \mu \mathbf{G}) &= \lambda \nabla \cdot \mathbf{F} + \mu \nabla \cdot \mathbf{G} \\ \nabla \times (\lambda \mathbf{F} + \mu \mathbf{G}) &= \lambda \nabla \times \mathbf{F} + \mu \nabla \times \mathbf{G}. \end{aligned}$$

Example. Consider r^α with $r = |\mathbf{r}|$. We know that $\mathbf{r} = x_i \mathbf{e}_i$. So $r^2 = x_i x_i$. Therefore

$$2r \frac{\partial r}{\partial x_j} = 2x_j, \Rightarrow \frac{\partial r}{\partial x_i} = \frac{x_i}{r}.$$

So

$$\nabla r^\alpha = \mathbf{e}_i \frac{\partial}{\partial x_i}(r^\alpha) = \mathbf{e}_i \alpha r^{\alpha - 1} \frac{\partial r}{\partial x_i} = \alpha r^{\alpha - 2} \mathbf{r}.$$

Also,

$$\nabla \cdot \mathbf{r} = \frac{\partial x_i}{\partial x_i} = 3. \quad \& \quad \nabla \times \mathbf{r} = \mathbf{e}_k \varepsilon_{ijk} \frac{\partial x_j}{\partial x_i} = 0.$$

Proposition. We have the following Leibnitz properties:

$$\begin{aligned} \nabla(fg) &= (\nabla f)g + f(\nabla g) \\ \nabla \cdot (f\mathbf{F}) &= (\nabla f) \cdot \mathbf{F} + f(\nabla \cdot \mathbf{F}) \\ \nabla \times (f\mathbf{F}) &= (\nabla f) \times \mathbf{F} + f(\nabla \times \mathbf{F}) \\ \nabla(\mathbf{F} \cdot \mathbf{G}) &= \mathbf{F} \times (\nabla \times \mathbf{G}) + \mathbf{G} \times (\nabla \times \mathbf{F}) + (\mathbf{F} \cdot \nabla)\mathbf{G} + (\mathbf{G} \cdot \nabla)\mathbf{F} \\ \nabla \times (\mathbf{F} \times \mathbf{G}) &= \mathbf{F}(\nabla \cdot \mathbf{G}) - \mathbf{G}(\nabla \cdot \mathbf{F}) + (\mathbf{G} \cdot \nabla)\mathbf{F} - (\mathbf{F} \cdot \nabla)\mathbf{G} \\ \nabla \cdot (\mathbf{F} \times \mathbf{G}) &= (\nabla \times \mathbf{F}) \cdot \mathbf{G} - \mathbf{F} \cdot (\nabla \times \mathbf{G}) \end{aligned}$$

which can be proven by brute-forcing with suffix notation and summation convention.

6.1 Second-order derivatives

We have

Proposition.

$$\nabla \times (\nabla f) = 0$$
$$\nabla \cdot (\nabla \times \mathbf{F}) = 0$$

The converse of each result holds for fields defined in all of $\mathbb{R}^3$:

Proposition. If $\mathbf{F}$ is defined in all of $\mathbb{R}^3$, then

$$\nabla \times \mathbf{F} = 0 \Rightarrow \mathbf{F} = \nabla f$$

for some f.

Definition. If $\mathbf{F} = \nabla f$, then f is the *scalar potential*. We say $\mathbf{F}$ is *conservative* or *irrotational*.

Proposition. If $\mathbf{H}$ is defined over all of $\mathbb{R}^3$ and $\nabla \cdot \mathbf{H} = 0$, then $\mathbf{H} = \nabla \times \mathbf{A}$ for some $\mathbf{A}$.

Definition. If $\mathbf{H} = \nabla \times \mathbf{A}$, $\mathbf{A}$ is the *vector potential* and $\mathbf{H}$ is said to be *solenoidal*.

Note that is is true ONLY IF $\mathbf{F}$ or $\mathbf{H}$ is defined on all of $\mathbb{R}^3$.

Definition (Laplacian operator)**.** The *Laplacian operator* is defined by

$$\nabla^2 = \nabla \cdot \nabla = \frac{\partial^2}{\partial x_i \partial x_i} = \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^3} \right).$$

This operation is defined on both scalar and vector fields:

$$\nabla^2 f = \nabla \cdot (\nabla f), \quad \& \quad \nabla^2 \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla \times (\nabla \times \mathbf{A}).$$

7 Integral theorems

7.1 Statement and examples

7.1.1 Green's theorem (in the plane)

Theorem (Green's theorem)**.** For smooth functions $P(x, y)$, $Q(x, y)$ and A a bounded region in the (x, y) plane with boundary $\partial A = C$,

$$\int_A \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \, \mathrm{d}A = \int_C (P \, \mathrm{d}x + Q \, \mathrm{d}y).$$

Given that C is a piecewise smooth, non-intersecting closed curve, traversed anti-clockwise.

Example. Let $Q = xy^2$ and $P = x^2 y$. If C is the parabola $y^2 = 4ax$ and the line $x = a$, both with $-2a \leq y \leq 2a$, then Green's theorem says

$$\int_A (y^2 - x^2) \, \mathrm{d}A = \int_C x^2 \, \mathrm{d}x + xy^2 \, \mathrm{d}y.$$

From Example Sheet 1, each side gives $\frac{104}{105} a^4$.

Green's theorem also holds for a bounded region A, where the boundary ∂A consists of *disconnected* components (each piecewise smooth, non-intersecting and closed) with anti-clockwise orientation on the exterior, and clockwise on the interior boundary.

7.1.2 Stokes' theorem

Theorem (Stokes' theorem)**.** For a smooth vector field $\mathbf{F}(\mathbf{r})$,

$$\int_S \nabla \times \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \int_{\partial S} \mathbf{F} \cdot \mathrm{d}\mathbf{r},$$

S is a smooth, bounded surface and ∂S is a piecewise smooth boundary of S.

The direction of the line integral is as follows: If we walk along C with $\mathbf{n}$ facing up, then the surface is on your left.

It also holds if ∂S disconnected piecewise smooth closed curves, with the orientation determined in the same way as Green's theorem.

7.1.3 Divergence/Gauss theorem

Theorem (Divergence/Gauss theorem)**.** For a smooth vector field $\mathbf{F}(\mathbf{r})$,

$$\int_V \nabla \cdot \mathbf{F} \, \mathrm{d}V = \int_{S=\partial V} \mathbf{F} \cdot \mathrm{d}\mathbf{S},$$

where V is a bounded volume with boundary $\partial V = S$, a piecewise smooth, closed surface, with outward normal $\mathbf{n}$.

Example. Consider a hemisphere.

V is a solid hemisphere

$$x^2 + y^2 + z^2 \leq a^2, \quad z \geq 0,$$

and $\partial V = S_1 + S_2$, the hemisphere and the disc at the bottom.

Take $\mathbf{F} = (0, 0, z + a)$ and $\nabla \cdot \mathbf{F} = 1$. Then

$$\int_V \nabla \cdot \mathbf{F} \, \mathrm{d}V = \frac{2}{3}\pi a^3,$$

the volume of the hemisphere.

On S_1,

$$\mathrm{d}\mathbf{S} = \mathbf{n} \, \mathrm{d}S = \frac{1}{a}(x, y, z) \, \mathrm{d}S.$$

Then

$$\mathbf{F} \cdot \mathrm{d}\mathbf{S} = \frac{1}{a} z(z + a) \, \mathrm{d}S = \cos\theta a(\cos\theta + 1) \underbrace{a^2 \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi}_{\mathrm{d}S}.$$

Then

$$\begin{aligned} \int_{S_1} \mathbf{F} \cdot \mathrm{d}\mathbf{S} &= a^3 \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi/2} \sin\theta(\cos^2\theta + \cos\theta) \, \mathrm{d}\theta \\ &= 2\pi a^3 \left[\frac{-1}{3}\cos^3\theta - \frac{1}{2}\cos^2\theta\right]_0^{\pi/2} = \frac{5}{3}\pi a^3. \end{aligned}$$

On S_2, $\mathrm{d}\mathbf{S} = \mathbf{n}\ \mathrm{d}S = -(0,0,1)\ \mathrm{d}S$. Then $\mathbf{F}\cdot\mathrm{d}\mathbf{S} = -a\ \mathrm{d}S$. So

$$\int_{S_2} \mathbf{F}\cdot\mathrm{d}\mathbf{S} = -\pi a^3.$$

So

$$\int_{S_1} \mathbf{F}\cdot\mathrm{d}\mathbf{S} + \int_{S_2} \mathbf{F}\cdot\mathrm{d}\mathbf{S} = \left(\frac{5}{3} - 1\right)\pi a^3 = \frac{2}{3}\pi a^3,$$

in accordance with Gauss' theorem.

7.2 Relating and proving integral theorems

We will first show the following two equivalences:

- Stokes' theorem $\Leftrightarrow$ Green's theorem
- 2D divergence theorem $\Leftrightarrow$ Green's theorem

Then we prove the 2D version of divergence theorem directly to show that all of the above hold.

Proposition. Stokes' theorem $\Rightarrow$ Green's theorem

Proof. Let A be a region in the (x,y) plane with boundary $C = \partial A$, parametrised by arc length, $(x(s), y(s), 0)$. Then the tangent to C is

$$\mathbf{t} = \left(\frac{\mathrm{d}x}{\mathrm{d}s}, \frac{\mathrm{d}y}{\mathrm{d}s}, 0\right).$$

Given any $P(x,y)$ and $Q(x,y)$, we can have:

$$\mathbf{F} = (P, Q, 0) \Rightarrow \nabla\times\mathbf{F} = \left(0, 0, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right).$$

Then the left hand side of Stokes is

$$\int_C \mathbf{F}\cdot\mathrm{d}\mathbf{r} = \int \mathbf{F}\cdot\mathbf{t}\ \mathrm{d}s = \int_C P\ \mathrm{d}x + Q\ \mathrm{d}y,$$

and the right hand side is

$$\int_A (\nabla\times\mathbf{F})\cdot\hat{\mathbf{k}}\ \mathrm{d}A = \int_A \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\ \mathrm{d}A.$$

□

Proposition. Green's theorem $\Rightarrow$ Stokes' theorem.

Proof. Consider a parametrised surface $S = \mathbf{r}(u,v)$ corresponding to the region A in the u,v plane. Write the boundary as $\partial A = (u(t), v(t))$. Then $\partial S = \mathbf{r}(u(t), v(t))$.

We want to prove

$$\int_{\partial S} \mathbf{F}\cdot\mathrm{d}\mathbf{r} = \int_S (\nabla\times\mathbf{F})\cdot\mathrm{d}\mathbf{S}$$

given

$$\int_{\partial A} F_u\ \mathrm{d}u + F_v\ \mathrm{d}v = \int_A \left(\frac{\partial F_v}{\partial u} - \frac{\partial F_u}{\partial v}\right)\ \mathrm{d}A.$$

Doing some pattern-matching, we want

$$\mathbf{F} \cdot \mathrm{d}\mathbf{r} = F_u \,\mathrm{d}u + F_v \,\mathrm{d}v$$

By the chain rule, we know that $\mathrm{d}\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u}\mathrm{d}u + \frac{\partial \mathbf{r}}{\partial v}\mathrm{d}v$. So we choose

$$F_u = \mathbf{F} \cdot \frac{\partial \mathbf{r}}{\partial u}, \quad F_v = \mathbf{F} \cdot \frac{\partial \mathbf{r}}{\partial v}.$$

This choice matches the left hand sides of the two equations.

To match the right, recall that

$$(\nabla \times \mathbf{F}) \cdot \mathrm{d}\mathbf{S} = (\nabla \times \mathbf{F}) \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right).$$

Therefore, for the right hand sides to match, we want

$$\frac{\partial F_v}{\partial u} - \frac{\partial F_u}{\partial v} = (\nabla \times \mathbf{F}) \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right). \tag{*}$$

Fortunately, this is true. Unfortunately, the proof involves complicated suffix notation and summation convention:

$$\frac{\partial F_v}{\partial u} = \frac{\partial}{\partial u}\left(\mathbf{F} \cdot \frac{\partial \mathbf{r}}{\partial v}\right) = \frac{\partial}{\partial u}\left(F_i \frac{\partial x_i}{\partial v}\right) = \left(\frac{\partial F_i}{\partial x_j}\frac{\partial x_j}{\partial u}\right)\frac{\partial x_i}{\partial v} + F_i \frac{\partial x_i}{\partial u \partial v}.$$

Similarly,

$$\frac{\partial F_u}{\partial u} = \frac{\partial}{\partial u}\left(\mathbf{F} \cdot \frac{\partial \mathbf{r}}{\partial u}\right) = \frac{\partial}{\partial u}\left(F_j \frac{\partial x_j}{\partial u}\right) = \left(\frac{\partial F_j}{\partial x_i}\frac{\partial x_i}{\partial v}\right)\frac{\partial x_j}{\partial u} + F_i \frac{\partial x_i}{\partial u \partial v}.$$

So

$$\frac{\partial F_v}{\partial u} - \frac{\partial F_u}{\partial v} = \frac{\partial x_j}{\partial u}\frac{\partial x_i}{\partial v}\left(\frac{\partial F_i}{\partial x_j} - \frac{\partial F_j}{\partial x_i}\right).$$

This is the left hand side of $(*)$.

The right hand side of $(*)$ is

$$(\nabla \times \mathbf{F}) \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right) = \varepsilon_{ijk}\frac{\partial F_j}{\partial x_i}\varepsilon_{kpq}\frac{\partial x_p}{\partial u}\frac{\partial x_q}{\partial v} = (\delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp})\frac{\partial F_j}{\partial x_i}\frac{\partial x_p}{\partial u}\frac{\partial x_q}{\partial v} = \left(\frac{\partial F_j}{\partial x_i} - \frac{\partial F_i}{\partial x_j}\right)\frac{\partial x_i}{\partial u}\frac{\partial x_j}{\partial v}.$$

So they match. Therefore, given our choice of F_u and F_v, Green's theorem translates to Stokes' theorem. □

Proposition. Greens theorem ⇔ 2D divergence theorem.

Proof. The 2D divergence theorem states that

$$\int_A (\nabla \cdot \mathbf{G}) \,\mathrm{d}A = \int_{C=\partial A} \mathbf{G} \cdot \mathbf{n} \,\mathrm{d}s.$$

with an outward normal $\mathbf{n}$.

Write $\mathbf{G}$ as $(Q, -P)$. Then

$$\nabla \cdot \mathbf{G} = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}.$$

Around the curve $\mathbf{r}(s) = (x(s), y(s))$, $\mathbf{t}(s) = (x'(s), y'(s))$. Then the normal, being tangent to $\mathbf{t}$, is $\mathbf{n}(s) = (y'(s), -x'(s))$ (check that it points outwards!). So

$$\mathbf{G}\cdot\mathbf{n} = P\frac{\mathrm{d}x}{\mathrm{d}s} + Q\frac{\mathrm{d}y}{\mathrm{d}s}.$$

Then we can expand out the integrals to obtain

$$\int_C \mathbf{G}\cdot\mathbf{n}\;\mathrm{d}s = \int_C P\;\mathrm{d}x + Q\;\mathrm{d}y,$$

and

$$\int_A (\nabla\cdot\mathbf{G})\;\mathrm{d}A = \int_A \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial Y}\right)\;\mathrm{d}A.$$

Now 2D version of Gauss' theorem says the two LHS are the equal, and Green's theorem says the two RHS are equal. So the result follows. □

Proposition. 2D divergence theorem.

$$\int_A (\nabla\cdot\mathbf{G})\;\mathrm{d}A = \int_{C=\partial A} \mathbf{G}\cdot\mathbf{n}\;\mathrm{d}s.$$

Proof. For the sake of simplicity, we assume that $\mathbf{G}$ only has a vertical component,and that A is a simple, convex shape. A more complicated shape can be cut into smaller simple regions, and we can apply the simple case to each of the small regions.

Suppose $\mathbf{G} = G(x, y)\hat{\mathbf{j}}$. Then

$$\nabla\cdot\mathbf{G} = \frac{\partial G}{\partial y}.$$

Then

$$\int_A \nabla\cdot\mathbf{G}\;\mathrm{d}A = \int_X \left(\int_{Y_x} \frac{\partial G}{\partial y}\;\mathrm{d}y\right)\;\mathrm{d}x.$$

Now we divide A into an upper and lower part, with boundaries $C_+ = y_+(x)$ and $C_- = y_-(x)$ respectively.

We see that the boundary of Y_x at any specific x is given by $y_-(x)$ and $y_+(x)$. Hence by the Fundamental theorem of Calculus,

$$\int_{Y_x} \frac{\partial G}{\partial y}\;\mathrm{d}y = \int_{y_-(x)}^{y_+(x)} \frac{\partial G}{\partial y}\;\mathrm{d}y = G(x, y_+(x)) - G(x, y_-(x)).$$

To compute the full area integral, we want to integrate over all x. However, the divergence theorem talks in terms of $\mathrm{d}s$, not $\mathrm{d}x$. If we move a distance δs, the change in x is $\delta s\cos\theta$, where θ is the angle between the tangent and the horizontal. But θ is also the angle between the normal and the vertical. So $\cos\theta = \mathbf{n}\cdot\hat{\mathbf{j}}$. Therefore $\mathrm{d}x = \hat{\mathbf{j}}\cdot\mathbf{n}\;\mathrm{d}s$.

In particular, $G\;\mathrm{d}x = G\,\hat{\mathbf{j}}\cdot\mathbf{n}\;\mathrm{d}s = \mathbf{G}\cdot\mathbf{n}\;\mathrm{d}s$, since $\mathbf{G} = G\,\hat{\mathbf{j}}$.

However, at C_-, $\mathbf{n}$ points downwards, so $\mathbf{n}\cdot\hat{\mathbf{j}}$ happens to be negative. So, actually, at C_-, $\mathrm{d}x = -\mathbf{G}\cdot\mathbf{n}\;\mathrm{d}s$.

Therefore, our full integral is

$$\begin{aligned}\int_A \nabla\cdot\mathbf{G}\;\mathrm{d}A = \int_X \left(\int_{y_x} \frac{\partial G}{\partial y}\;\mathrm{d}Y\right)\;\mathrm{d}x &= \int_X G(x, y_+(x)) - G(x, y_-(x))\;\mathrm{d}x\\ &= \int_{C_+} \mathbf{G}\cdot\mathbf{n}\;\mathrm{d}s + \int_{C_-} \mathbf{G}\cdot\mathbf{n}\;\mathrm{d}s = \int_C \mathbf{G}\cdot\mathbf{n}\;\mathrm{d}s.\end{aligned}$$

□

To prove the 3D version, we again consider $\mathbf{F} = F(x, y, z)\hat{\mathbf{k}}$, a purely vertical vector field. Then

$$\int_V \nabla \cdot \mathbf{F} \, \mathrm{d}V = \int_D \left(\int_{Z_{xy}} \frac{\partial F}{\partial z} \, \mathrm{d}z \right) \, \mathrm{d}A.$$

Again, split $S = \partial V$ into the top and bottom parts S_+ and S_- (ie the parts with $\hat{\mathbf{k}} \cdot \mathbf{n} \geq 0$ and $\hat{\mathbf{k}} \cdot \mathbf{n} < 0$, and parametrize by $z_+(x, y)$ and $z_-(x, y)$. Then the integral becomes

$$\int_V \nabla \cdot \mathbf{F} \, \mathrm{d}V = \int_D (F(x, y, z_+) - F(x, y, z_-)) \, \mathrm{d}A = \int_S \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}S.$$

8 Some applications of integral theorems

8.1 Integral expressions for div and curl

From Gauss' theorem we can have the following:

Proposition.

$$\nabla \cdot \mathbf{F} = \lim_{V \to 0} \frac{1}{V} \int_{\partial V} \mathbf{F} \cdot \mathrm{d}\mathbf{S}.$$

Similarly, Stokes' gives

Proposition.

$$\mathbf{n} \cdot \nabla \times \mathbf{F} = \lim_{A \to 0} \frac{1}{A} \int_{\partial A} \mathbf{F} \cdot \mathrm{d}\mathbf{r}.$$

These are coordinate-independent definitions of div and curl.

Example. Suppose $\mathbf{u}$ is a velocity field of fluid flow. Then

$$\int_S \mathbf{u} \cdot \mathrm{d}\mathbf{S}$$

is the rate of which fluid crosses S. Taking V to be the volume occupied by a fixed quantity of fluid material, we have

$$\dot{V} = \int_{\partial V} \mathbf{u} \cdot \mathrm{d}\mathbf{S}$$

Then, at $\mathbf{r}_0$,

$$\nabla \cdot \mathbf{u} = \lim_{V \to 0} \frac{\dot{V}}{V},$$

the relative rate of change of volume. For example, if $\mathbf{u}(\mathbf{r}) = \alpha \mathbf{r}$ (ie fluid flowing out of origin), then $\nabla \cdot \mathbf{u} = 3\alpha$, which increases at a constant rate everywhere.

8.2 Conservative fields and scalar products

Definition (Conservative field). A vector field f is *conservative* if

(i) $\mathbf{F} = \nabla f$ for some scalar field f; or

(ii) $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ is independent of C, for fixed end points and orientation; or

(iii) $\nabla \times \mathbf{F} = 0$.

We have previously shown (i) $\Rightarrow$ (ii) and (i) $\Rightarrow$ (iii) since

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = f(\mathbf{b}) - f(\mathbf{a}). \quad \& \quad \nabla \times (\nabla f) = 0.$$

So we want to show that (iii) $\Rightarrow$ (ii) and (ii) $\Rightarrow$ (i)

Proposition. If (iii) $\nabla \times \mathbf{F} = 0$, then (ii) $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ is independent of C.

Proof. Given $\mathbf{F}(\mathbf{r})$ satisfying $\nabla \times \mathbf{F} = 0$, let C and $\tilde{C}$ be any two curves from $\mathbf{a}$ to $\mathbf{b}$. If S is any surface with boundary $\partial S = C - \tilde{C}$, By Stokes' theorem,

$$\int_S \nabla \times \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \int_{\partial S} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} - \int_{\tilde{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}.$$

But $\nabla \times \mathbf{F} = 0$. So

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_{\tilde{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}.$$

□

Proposition. If (ii) $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ is independent of C for fixed end points and orientation, then (i) $\mathbf{F} = \nabla f$ for some scalar field f.

Proof. We fix $\mathbf{a}$ and define $f(\mathbf{r}) = \int_C \mathbf{F}(\mathbf{r}') \cdot \mathrm{d}\mathbf{r}'$ for any curve from $\mathbf{a}$ to $\mathbf{r}$. Assuming (ii), f is well-defined. For small changes $\mathbf{r}$ to

$$f(\mathbf{r} + \delta\mathbf{r}) = \int_{C+\delta C} \mathbf{F}(\mathbf{r}') \cdot \mathrm{d}\mathbf{r}' = \int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}' + \int_{\delta C} \mathbf{F} \cdot \mathrm{d}\mathbf{r}' = f(\mathbf{r}) + \mathbf{F}(\mathbf{r}) \cdot \delta\mathbf{r} + o(\delta\mathbf{r}).$$

So

$$\delta f = f(\mathbf{r} + \delta\mathbf{r}) - f(\mathbf{r}) = \mathbf{F}(\mathbf{r}) \cdot \delta\mathbf{r} + o(\delta\mathbf{r}).$$

But the definition of grad is exactly

$$\delta f = \nabla f \cdot \delta\mathbf{r} + o(\delta\mathbf{r}).$$

So we have $\mathbf{F} = \nabla f$. □

Note that these results assume "niceness" (simply connected) of the domain. If it isn't, the laws do NOT hold.

8.3 Conservation laws

Definition (Conservation equation)**.** Suppose we are interested in a quantity Q. Let $\rho(\mathbf{r}, t)$ be the amount of stuff per unit volume and $\mathbf{j}(\mathbf{r}, t)$ be the flow rate of the quantity (eg if Q is charge, j is the current density).

The conservation equation is

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0.$$

It says that Q cannot just disappear here and appear elsewhere. It must continuously flow out. In particular, let V be a fixed time-independent volume with boundary $S = \partial V$. Then

$$Q(t) = \int_V \rho(\mathbf{r}, t)\ \mathrm{d}V$$

Then the rate of change of amount of Q in V is

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \int_V \frac{\partial \rho}{\partial t}\ \mathrm{d}V = -\int_V \nabla \cdot \mathbf{j}\ \mathrm{d}V = -\int_S \mathbf{j} \cdot \mathrm{d}\mathbf{s}.$$

by divergence theorem. So this states that the rate of change of the quantity Q in V is the flux of the stuff flowing out of the surface. ie Q cannot just disappear but must smoothly flow out.

In particular, if V is the whole universe (ie $\mathbb{R}^3$), and $\mathbf{j} \to 0$ sufficiently rapidly as $|\mathbf{r}| \to \infty$, then we have:

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = 0,$$

Example. If $\rho(\mathbf{r}, t)$ is the charge density (ie. $\rho \delta V$ is the amount of charge in a small volume δV), then $Q(t)$ is the total charge in V. $\mathbf{j}(\mathbf{r}, t)$ is the electric current density. So $\mathbf{j} \cdot \mathrm{d}\mathbf{S}$ is the charge flowing through δS per unit time.

9 Orthogonal curvilinear coordinates

9.1 Line, area and volume elements

In this section, we study funny coordinate systems. We can think of a coordinate system as a function $\mathbf{r}(u, v, w)$.

By the chain rule, for (good) parametrizations, we have

$$\mathrm{d}\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u}\mathrm{d}u + \frac{\partial \mathbf{r}}{\partial v}\mathrm{d}v + \frac{\partial \mathbf{r}}{\partial w}\mathrm{d}w \text{ with } \frac{\partial \mathbf{r}}{\partial u} \cdot \left(\frac{\partial \mathbf{r}}{\partial v} \times \frac{\partial \mathbf{r}}{\partial w}\right) \neq 0,$$

ie. $\frac{\partial \mathbf{r}}{\partial u}$, $\frac{\partial \mathbf{r}}{\partial v}$ and $\frac{\partial \mathbf{r}}{\partial w}$ are linearly independent. These vectors are tangent to the curves parametrized by u, v, w respectively when the other two are being fixed.

Even better, they should be orthogonal:

Definition (Orthogonal curvilinear coordinates)**.** u, v, w are *orthogonal curvilinear* if the tangent vectors are orthogonal.

We can then set

$$\frac{\partial \mathbf{r}}{\partial u} = h_u \mathbf{e}_u,\ \frac{\partial \mathbf{r}}{\partial v} = h_v \mathbf{e}_v,\ \frac{\partial \mathbf{r}}{\partial w} = h_w \mathbf{e}_w, \Rightarrow \mathrm{d}\mathbf{r} = h_u \mathbf{e}_u\ \mathrm{d}u + h_v \mathbf{e}_v\ \mathrm{d}v + h_w \mathbf{e}_w\ \mathrm{d}w,$$

with the unit vectors forming an orthonormal righthanded basis and h_u, h_v, h_w determining the changes in length along each orthogonal direction resulting from changes in u, v, w. Equivalently, we have:

$$|\delta \mathbf{r}|^2 = h_u^2 \delta u^2 + h_v^2 \delta v^2 + h_w^2 \delta w^2 + o\text{-terms}.$$

We do not have the cross terms because the basis is orthonormal. Then

$$\left|\frac{\partial \mathbf{r}}{\partial u}\right|^2 = h_u^2, \text{ or } h_u = \left|\frac{\partial \mathbf{r}}{\partial u}\right|.$$

Example. In cylindrical polars, $\mathbf{r}(\rho, \varphi, z) = \rho[\cos\varphi\hat{\mathbf{i}} + \sin\varphi\hat{\mathbf{j}}] + z\hat{\mathbf{k}}$. Then $h_\rho = h_z = 1$:

$$h_\varphi = \left|\frac{\partial \mathbf{r}}{\partial \varphi}\right| = |(-\rho\sin\varphi, \rho\sin\varphi, 0)| = \rho.$$

The basis vectors $\mathbf{e}_\rho, \mathbf{e}_\varphi, \mathbf{e}_z$ are as in section 1.

Consider a surface with w constant and parametrised by u and v. The vector area element is

$$\mathrm{d}\mathbf{S} = \frac{\mathrm{d}\mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\, \mathrm{d}u\, \mathrm{d}v = h_u \mathbf{e}_u \times h_v \mathbf{e}_v\, \mathrm{d}u\, \mathrm{d}v = h_u h_v \mathbf{e}_w\, \mathrm{d}u\, \mathrm{d}v.$$

We interpret this as δS having a small rectangle with sides approximately $h_u \delta u$ and $h_v \delta v$. The volume element is

$$\mathrm{d}V = \frac{\partial \mathbf{r}}{\partial u} \cdot \left(\frac{\partial \mathbf{r}}{\partial v} \times \frac{\partial \mathbf{r}}{\partial w}\right) \mathrm{d}u\, \mathrm{d}v\, \mathrm{d}w = h_u h_v h_w\, \mathrm{d}u\, \mathrm{d}v\, \mathrm{d}w,$$

ie. a small cuboid with sides $h_u\delta_u$, $h_v\delta_v$ and $h_w\delta_w$ respectively.

9.2 Grad, Div and Curl

We have the following:

$$\mathrm{d}f = \frac{\partial f}{\partial u}\, \mathrm{d}u + \frac{\partial f}{\partial v}\, \mathrm{d}v + \frac{\partial f}{\partial w}\, \mathrm{d}w, \quad \& \quad \mathrm{d}\mathbf{r} = h_u \mathbf{e}_u\, \mathrm{d}u + h_v \mathbf{e}_v\, \mathrm{d}v + h_w \mathbf{e}_w\, \mathrm{d}v,$$

we can use $\mathrm{d}f = (\nabla f) \cdot \mathrm{d}\mathbf{r}$ and compare the terms to know that

Proposition.

$$\nabla f = \frac{1}{h_u}\mathbf{e}_u \frac{\partial f}{\partial u} + \frac{1}{h_v}\mathbf{e}_v \frac{\partial f}{\partial v} + \frac{1}{h_w}\mathbf{e}_w \frac{\partial f}{\partial w}.$$

Or, in differential operator form:

Proposition.

$$\nabla = \frac{1}{h_u}\mathbf{e}_u \frac{\partial}{\partial u} + \frac{1}{h_v}\mathbf{e}_v \frac{\partial}{\partial v} + \frac{1}{h_w}\mathbf{e}_w \frac{\partial}{\partial w}.$$

We can apply this to a vector field

$$\mathbf{F} = F_u \mathbf{e}_u + F_v \mathbf{e}_v + F_w \mathbf{e}_w$$

using scalar or vector products to obtain

Proposition.

$$\nabla \times \mathbf{F} = \frac{1}{h_u h_v h_w}\begin{vmatrix} h_u \mathbf{e}_u & h_v \mathbf{e}_v & h_w \mathbf{e}_w \\ \frac{\partial}{\partial u} & \frac{\partial}{\partial v} & \frac{\partial}{\partial w} \\ h_u f_u & h_v F_v & h_w F_w \end{vmatrix}$$

and

$$\nabla \cdot \mathbf{F} = \frac{1}{h_u h_v h_w}\left[\frac{\partial}{\partial u}(h_v h_w F_u) + \text{ two similar terms}\right].$$

There are many ways to get this but we don't need to know about them. :)

Example. Let $\mathbf{A} = \frac{1}{r}\tan\frac{\theta}{2}\mathbf{e}_\varphi$ in spherical polars. Then

$$\nabla \times \mathbf{A} = \frac{1}{r^2 \sin\theta}\begin{vmatrix} \mathbf{e}_r & r\mathbf{e}_\theta & r\sin\theta\mathbf{e}_\varphi \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \varphi} \\ 0 & 0 & r\sin\theta \cdot \frac{1}{r}\tan\frac{\theta}{2} \end{vmatrix} = \frac{\mathbf{e}_r}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left[\sin\theta\tan\frac{\theta}{2}\right] = \frac{1}{r^2}\mathbf{e}_r.$$

10 Gauss' Law and Poisson's equation

10.1 Laws of gravitation

Consider a distribution of mass producing a gravitational force $\mathbf{F}$ on a point mass m at $\mathbf{r}$. The total force is a sum of contributions from each part of the mass distribution, and is proportional to m. Write

$$\mathbf{F} = m\mathbf{g}(\mathbf{r}),$$

Definition (Gravitational field)**.** $\mathbf{g}(\mathbf{r})$ is the *gravitational field*, *acceleration due to gravity*, or *force per unit mass.*

The gravitational field is conservative,so if you walk around the place and return to the same position, the total work done is 0 and you did not gain energy, ie. gravitational potential energy is conserved.

Gauss' law tells us:

Law. Given any volume V bounded by closed surface S,

$$\int_S \mathbf{g} \cdot \mathrm{d}\mathbf{S} = -4\pi GM,$$

where G is the gravitational constant, and M is the mass contained in V.

Example. Consider a total mass M distributed with a spherical symmetry about the origin $\mathbf{O}$, with all the mass contained within some radius $r = a$. By spherical symmetry, we have $\mathbf{g}(\mathbf{r}) = g(\mathbf{r})\hat{\mathbf{r}}$.

Consider Gauss' law with S being a sphere of radius $r = R > a$. Then $\hat{\mathbf{n}} = \hat{\mathbf{r}}$. So

$$\int_S \mathbf{g} \cdot \mathrm{d}\mathbf{S} = \int_S g(R)\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}\ \mathrm{d}S = \int g(R)\mathrm{d}S = 4\pi R^2 g(R).$$

By Gauss' law, we obtain

$$4\pi R^2 g(R) = -4\pi GM. \Rightarrow g(R) = -\frac{GM}{R^2} \Rightarrow \mathbf{F}(\mathbf{r}) = -\frac{GMm}{r^2}\hat{\mathbf{r}}.$$

If we take the limit as $a \to 0$ (point mass), we recover Newton's law of gravitation.

Since gravity is conservative, we have, from Stokes' theorem (and noticing this is true for arbitrary S):

$$\int_C \mathbf{g} \cdot \mathrm{d}\mathbf{r} = 0 \Rightarrow \int_S \nabla \times \mathbf{g} \cdot \mathrm{d}\mathbf{S} = 0, \Rightarrow \nabla \times \mathbf{g} = 0.$$

Note that we exploited symmetry to solve Gauss' law. However, if the mass distribution is not sufficiently symmetrical, we use the differential form, assuming the density only depends on $\mathbf{r}$:

$$\int_S \mathbf{g} \cdot \mathrm{d}\mathbf{S} = -4\pi GM \Rightarrow \int_V \nabla \cdot \mathbf{g}\ \mathrm{d}V = \int_V -4\pi G\rho\ \mathrm{d}V.$$

Since this is true for all V, we must have

Law.

$$\nabla \cdot \mathbf{g} = -4\pi G\rho.$$

Since $\nabla \times \mathbf{g} = 0$, we can introduce a gravitational potential $\varphi(\mathbf{r})$ with $\mathbf{g} = -\nabla\varphi$. Then Gauss' Law becomes

$$\nabla^2\varphi = 4\pi G\rho.$$

10.2 Laws of electrostatics

Definition (Electric field)**.** The force produced by electric charges on another charge q is $\mathbf{F} = q\mathbf{E}(\mathbf{r})$, where $\mathbf{E}(\mathbf{r})$ is the *electric field*, or force per unit charge.

Again, this is conservative. So it obeys

Law (Gauss' law for electrostatic forces)**.**

$$\int_S \mathbf{E} \cdot \mathrm{d}\mathbf{S} = \frac{Q}{\varepsilon_0},$$

where ε_0 is the *permittivity of free space*, or *electric constant.*

Then we can write it in differential form, as in the gravitational case.

Law (Gauss' law for electrostatic forces in differential form)**.**

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}.$$

If we only have a constant magnetic field, we can have $\mathbf{E} = -\nabla\varphi$.

Definition (Electrostatic potential)**.** If we write $\mathbf{E} = -\nabla\varphi$, then φ is the *electrostatic potential*, and

$$\nabla^2\varphi = \frac{\rho}{\varepsilon_0}.$$

Example. Take a spherically symmetric charge distribution about O with total charge Q. Suppose all charge is contained within a radius $r = a$. Then similar to the gravitational case, we have

$$\mathbf{E}(\mathbf{r}) = \frac{Q\hat{\mathbf{r}}}{4\pi\varepsilon_0 r^2}, \quad \& \quad \varphi(\mathbf{r}) = \frac{-Q}{4\pi\varepsilon_0 r}.$$

As $a \to 0$, we get *point charges.* We can recover coulomb's law:

$$\mathbf{F} = q\mathbf{E} = \frac{qQ\hat{\mathbf{r}}}{4\pi\varepsilon_0 r^2}.$$

Example (Line charge)**.** Consider an infinite line with uniform charge density *per unit length* σ. Using cylindrical polar coordinates(the field is radial):

$$\mathbf{E}(r) = E(r)\hat{\mathbf{r}}.$$

Pick S to be a cylinder of length L and radius r. We know that the end caps do not contribute to the flux since the field lines are perpendicular to the normal. Also, the curved surface has area $2\pi rL$. Then by Gauss' law in integral form,

$$\int_S \mathbf{E} \cdot \mathrm{d}\mathbf{S} = E(r)2\pi rL = \frac{\sigma L}{\varepsilon_0}. \ \& \ \mathbf{E}(r) = \frac{\sigma}{2\pi\varepsilon_0 r}\hat{\mathbf{r}}.$$

10.3 Poisson's Equation and Laplace's equation

Definition (Poisson's equation)**.** The *Poisson's equation* is

$$\nabla^2\varphi = -\rho,$$

where ρ is given and $\varphi(\mathbf{r})$ is to be solved.

This is the form of the equations for gravity and electrostatics. When $\rho = 0$, we get *Laplace's Equation.*

We're concerned here mainly with cases exhibiting spherical or cylindrical symmetry, ie. when $\varphi(\mathbf{r})$ has spherical or cylindrical symmetry. Write $\varphi = \varphi(r)$:

$$\nabla\varphi = \varphi'(r)\hat{\mathbf{r}}.$$

Then Laplace's equation $\nabla^2\varphi = 0$ becomes an ordinary differential equation.

- For spherical symmetry, using the chain rule, we have

$$\nabla^2\varphi = \varphi'' + \frac{2}{r}\varphi' = \frac{1}{r^2}(r^2\varphi')' = 0. \Rightarrow \varphi = \frac{A}{r} + B.$$

- For cylindrical symmetry, with $r^2 = x_1^2 + x_2^2$, we have

$$\nabla^2\varphi = \varphi'' + \frac{1}{r}\varphi' = \frac{1}{r}(r\varphi')' = 0. \Rightarrow \varphi = A\ln r + B.$$

Then solutions to Poisson's equations can be obtained in a similar way, ie. by integrating the differential equations directly, or by adding particular integrals to the solutions above.

For example, for a spherically symmetric solution of $\nabla^2\varphi = -\rho_0$, with ρ_0 constant, recall that $\nabla^2 r^\alpha = \alpha(\alpha+1)r^{\alpha-2}$. Taking $\alpha = 2$, we find the particular integral

$$\varphi = -\frac{\rho_0}{6}r^2,$$

To determine A, B, we must specify boundary conditions. If φ is defined on all of $\mathbb{R}^3$, we often require $\varphi \to 0$ as $|\mathbf{r}| \to \infty$. If φ is defined on a bounded volume V, then there are two kinds of common boundary conditions on ∂V:

- Specify φ on ∂V - a *Dirichlet* condition
- Specify $\mathbf{n}\cdot\nabla\varphi$ (sometimes written as $\frac{\partial\varphi}{\partial\mathbf{n}}$): a *Neumann* condition. ($\mathbf{n}$ is the outward normal on ∂V).

We can also specify different boundary conditions on different boundary components.

Example. We might have a spherically symmetric distribution with constant ρ_0, defined in $a \leq r \leq b$, with $\varphi(a) = 0$ and $\frac{\partial\varphi}{\partial n}(b) = 0$.

Then the general solution is

$$\varphi(r) = \frac{A}{r} + B - \frac{1}{6}\rho_0 r^2.$$

We apply the boundary conditions to obtain

$$\frac{A}{a} + B - \frac{1}{6}\rho_0 a^2 = 0. \quad \& \quad \mathbf{n}\cdot\nabla\varphi = -\frac{A}{b^2} - \frac{1}{3}\rho_0 b = 0.$$

These conditions give

$$A = -\frac{1}{3}\rho_0 b^3, \quad B = \frac{1}{5}\rho_0 a^2 + \frac{1}{3}\rho_0 \frac{b^3}{a}.$$

In general, even if the problem has nothing to do with gravitation or electrostatics, if we want to solve $\nabla^2\varphi = -\rho$ with ρ and φ sufficiently symmetric, we can consider the flux of $\nabla\varphi$ out of a surface $S = \partial V$:

$$\int_S \nabla\varphi \cdot \mathrm{d}\mathbf{S} = -\int_V \rho \, \mathrm{d}V,$$

by divergence theorem. This is called the Gauss Flux method.

11 Laplace's and Poisson's equations

11.1 Uniqueness theorems

Theorem. Consider $\nabla^2\varphi = -\rho$ for some $\rho(\mathbf{r})$ on a bounded volume V with $S = \partial V$ being a closed surface, with an outward normal $\mathbf{n}$. Then:

(i) If φ satisfies the Dirichlet condition on the surface, then $\varphi(\mathbf{r})$ is unique.

(ii) If φ satisfies the Neumann condition on the surface, then $\varphi(\mathbf{r})$ is unique up to a constant.

This theorem is practically important - if you find a solution by any magical means, you know it is the *only* solution (up to a constant).

Proof. Let $\varphi_1(\mathbf{r})$ and $\varphi_2(\mathbf{r})$ satisfy Poisson's equation, each obeying the boundary conditions (N) or (D). Then $\Psi(\mathbf{r}) = \varphi_2(\mathbf{r}) - \varphi_1(\mathbf{r})$ satisfies $\nabla^2\Phi = 0$ on V by linearity, and

$$(i)\ \Psi = 0 \text{ on } S; \text{ or } (ii)\ \tfrac{\partial\Psi}{\partial\mathbf{n}} = 0 \text{ on } S.$$

Combining these two together, we know that $\Psi\frac{\partial\Psi}{\partial\mathbf{n}} = 0$ on the surface. So using the divergence theorem,

$$\int_V \nabla \cdot (\Psi\nabla\Psi) = \int_S (\Psi\nabla\Psi) \cdot \mathrm{d}\mathbf{S} = 0.$$

But

$$\nabla \cdot (\Psi\nabla\Psi) = (\nabla\Psi) \cdot (\nabla\Psi) + \Psi \underbrace{\nabla^2\Psi}_{=0} = |(\nabla\Psi)|^2.$$

So

$$\int_V |\nabla\Psi|^2 \, \mathrm{d}V = 0.$$

Since $|\nabla\Psi|^2 \geq 0$, the integral can only vanish if $|\nabla\Psi| = 0$. So $\nabla\Psi = 0$. So $\Psi = c$, a constant on V. So

(i) $\Psi = 0$ on $S \Rightarrow c = 0$. So $\varphi_1 = \varphi_2$ on V.

(ii) $\varphi_2(\mathbf{r}) = \varphi_1(\mathbf{r}) + C$, as claimed.

□

We've proven uniqueness. How about existence? It turns out that existence is not guaranteed at all.

For example, if we have $\nabla^2\varphi = -\rho$ on V with the condition $\frac{\partial \varphi}{\partial \mathbf{n}} = g$, then by the divergence theorem,

$$\int_V \nabla^2\varphi \;\mathrm{d}V = \int_{\partial S} \frac{\partial \varphi}{\partial \mathbf{n}} \;\mathrm{d}S.$$

Using Poisson's equation and the boundary conditions, we have

$$\int_V \rho \;\mathrm{d}V + \int_{\partial V} g \;\mathrm{d}S = 0$$

So if ρ and g don't satisfy this equation, then we can't have any solutions.

Note. The theorem can be similarly proved and stated for regions in $\mathbb{R}^2, \mathbb{R}^3, \cdots$, by using the definitions of grad, div and the divergence theorem. The result also extends to unbounded domains. To prove it, we can take a sphere of radius R and impose the boundary conditions $|\Psi(\mathbf{r})| = O(1/R)$ or $|\Psi(\frac{\partial \Psi}{\partial \mathbf{n}})| = O(1/R^2)$ as $R \to \infty$. Then we just take the relevant limits to complete the proof.

The proof uses a special case of the result

Proposition (Green's first identity)**.**

$$\int_S (u\nabla v)\cdot \mathrm{d}S = \int_V (\nabla u)\cdot(\nabla v)\;\mathrm{d}V + \int_V u\nabla^2 v\;\mathrm{d}V,$$

By swapping u and v around and subtracting the equations, we have

Proposition (Green's second identity)**.**

$$\int_S (u\nabla v - v\nabla u)\cdot \mathrm{d}\mathbf{S} = \int_V (u\nabla^2 v - v\nabla^2 u)\;\mathrm{d}V.$$

11.2 Laplace's equation and harmonic functions

Definition. A *harmonic function* is a solution to Laplace's equation.

11.2.1 The mean value property

Proposition (Mean value property)**.** Suppose $\varphi(\mathbf{r})$ is harmonic on region V containing a solid sphere defined by $|\mathbf{r}-\mathbf{a}| \leq R$, with boundary $S_R = |\mathbf{r}-\mathbf{a}| = R$, for some R. Define

$$\bar{\varphi}(R) = \frac{1}{4\pi R^2}\int_{S_R} \varphi(\mathbf{r})\;\mathrm{d}S.$$

Then $\varphi(\mathbf{a}) = \bar{\varphi}(R)$.

In words, this says that the value at the center of a sphere is the average of the values on the surface on the sphere.

Proof. Note that $\bar{\varphi}(R) \to \varphi(\mathbf{a})$ as $R \to 0$. We take spherical coordinates (u, θ, χ) centered on $\mathbf{r} = \mathbf{a}$. The scalar element (when $u = R$) on S_R is

$$\mathrm{d}S = R^2 \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\chi.$$

So $\frac{\mathrm{d}S}{R^2}$ is independent of R. So

$$\bar{\varphi}(R) = \frac{1}{4\pi} \int \varphi(u) \, \frac{\mathrm{d}S}{R^2}.$$

Differentiate this with respect to R, noting that $\mathrm{d}S/R^2$ is independent of R. Then we obtain

$$\frac{\mathrm{d}}{\mathrm{d}R}\bar{\varphi}(R) = \frac{1}{4\pi R^2} \int \left.\frac{\partial \varphi}{\partial u}\right|_{u=R} \mathrm{d}S$$

But

$$\frac{\partial \varphi}{\partial u} = \mathbf{e}_u \cdot \nabla\varphi = \mathbf{n} \cdot \nabla\varphi = \frac{\partial \varphi}{\partial \mathbf{n}}$$

on S_R. So

$$\frac{\mathrm{d}}{\mathrm{d}R}\bar{\varphi}(R) = \frac{1}{4\pi R^2} \int \left.\frac{\partial \varphi}{\partial u}\right|_{u=R} \mathrm{d}S = \frac{1}{4\pi R^2} \int_{V_R} \nabla^2 \varphi \, \mathrm{d}V = 0$$

by divergence theorem. So $\bar{\varphi}(R)$ does not depend on R, and the result follows. □

11.2.2 The maximum (or minimum) principle

The results in this section also hold for minima.

Definition (Local maximum)**.** We say that $\varphi(\mathbf{r})$ has a *local maximum* at $\mathbf{a}$ if for some $\varepsilon > 0$, $\varphi(\mathbf{r}) < \varphi(\mathbf{a})$ when $0 < |\mathbf{r} - \mathbf{a}| < \varepsilon$.

Proposition (Maximum principle)**.** If a function φ is harmonic on a region V, then φ cannot have a maximum at an interior point of $\mathbf{a}$ of V.

Proof. Suppose that φ had a local maximum at $\mathbf{a}$ in the interior. Then there is an ε such that for any $\mathbf{r}$ such that $0 < |\mathbf{r} - \mathbf{a}| < \varepsilon$, we have $\varphi(\mathbf{r}) < \varphi(\mathbf{a})$. Pick an ε such that the region $|\mathbf{r} - \mathbf{a}| < \varepsilon$ lies within V.

Then for any $\mathbf{r}$ such that $|\mathbf{r} - \mathbf{a}| = \varepsilon$, we have $\varphi(\mathbf{r}) < \varphi(\mathbf{a})$.

$$\bar{\varphi}(\varepsilon) = \frac{1}{4\pi R^2} \int_{S_R} \varphi(\mathbf{r}) \, \mathrm{d}S < \varphi(\mathbf{a}),$$

which contradicts the mean value property. □

11.3 Integral solutions of Poisson's equations

11.3.1 Statement and informal derivation

We want to find a solution to Poisson's equations. We start with a discrete case, and try to generalize it to a continuous case.

If there is a single point source of strength λ_α, the potential φ is

$$\varphi = \frac{\lambda}{4\pi} \frac{1}{|\mathbf{r} - \mathbf{a}|}.$$

(we have $\lambda = -4\pi GM$ for gravitation and Q/ε_0 for electrostatics)

If we have many sources λ_α at positions $\mathbf{r}_\alpha$, having a distribution of $\rho(\mathbf{r})$ with $\rho(\mathbf{r}')\ \mathrm{d}V'$ being the contribution from a small volume at position $\mathbf{r}'$. It would be reasonable to guess that the solution is what we obtain by replacing a discrete sum with an integral:

Proposition. The solution to Poisson's equation $\nabla^2\varphi = -\rho$, with boundary conditions $|\varphi(\mathbf{r})| = O(1/|\mathbf{r}|)$ and $|\nabla\varphi(\mathbf{r})| = O(1/|\mathbf{r}|^2)$, is

$$\varphi(\mathbf{r}) = \frac{1}{4\pi}\int_{V'} \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\ \mathrm{d}V'$$

For $\rho(\mathbf{r}')$ non-zero everywhere, but suitably well-behaved as $|\mathbf{r}'| \to \infty$, we can also take $V' = \mathbb{R}^3$.

Example. Suppose

$$\nabla^2\varphi = \begin{cases} -\rho_0 & |\mathbf{r}| \le a \\ 0 & |\mathbf{r}| > a. \end{cases}$$

Fix $\mathbf{r}$ and introduce polar coordinates r', θ, χ for $\mathbf{r}'$. We take the $\theta = 0$ direction to be the direction along the line from $\mathbf{r}'$ to $\mathbf{r}$.

Then

$$\varphi(\mathbf{r}) = \frac{1}{4\pi}\int_{V'} \frac{\rho_0}{|\mathbf{r}-\mathbf{r}'|}\ \mathrm{d}V'.$$

We have

$$\mathrm{d}V' = r'^2 \sin\theta\ \mathrm{d}r'\ \mathrm{d}\theta\ \mathrm{d}\chi.$$

We also have

$$|\mathbf{r}-\mathbf{r}'| = \sqrt{r^2 + r'^2 - 2rr'\cos\theta}$$

by the cosine rule ($c^2 = a^2 + b^2 - 2ab\cos C$). So

$$\begin{aligned}
\varphi(\mathbf{r}) &= \frac{1}{4\pi}\int_0^a \mathrm{d}r' \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\chi \frac{\rho_0 r'^2 \sin\theta}{\sqrt{r^2 + r'^2 - 2rr'\cos\theta}} \\
&= \frac{\rho_0}{2}\int_0^a \mathrm{d}r' \frac{r'^2}{rr'}\left[\sqrt{r^2 + r'^2 - rr'\cos\theta}\right]_{\theta=0}^{\theta=\pi} \\
&= \frac{\rho_0}{2}\int_0^a \mathrm{d}r' \frac{r'}{r}(|\mathbf{r}+\mathbf{r}'| + |\mathbf{r}-\mathbf{r}'|) \\
&= \frac{\rho_0}{2}\int_0^a \left[\mathrm{d}r' \frac{r'}{r}\left(\begin{cases} 2r' & r > r' \\ 2r & r < r' \end{cases}\right)\right]
\end{aligned}$$

If $r > a$, then $r > r'$ always. So

$$\varphi(\mathbf{r}) = \rho_0 \int_0^a \frac{r'^2}{r} = \frac{r_0 a^3}{3r}.$$

If $r < a$, then the integral splits into two parts:

$$\varphi(\mathbf{r}) = \rho_0\left(\int_0^r \mathrm{d}r' \frac{r'^2}{r} + \int_r^a \mathrm{d}r' r'\right) = \rho_0\left[-\frac{1}{6}r^2 + \frac{a^2}{2}\right].$$

12 Maxwell's equations

12.1 Laws of electromagnetism

Definition (Electric and magnetic fields)**.** *Moving* electric charges interact in a way that can be described by *electric* and *magnetic* fields, $\mathbf{E}(r,t)$ and $\mathbf{B}(r,t)$ respectively with $\rho(\mathbf{r},t)$ is the *charge density* and $\mathbf{j}(\mathbf{r},t)$ is the *current density.*

Then Maxwell's equations say

Law (Maxwell's equations)**.**

$$\nabla\cdot\mathbf{E}=\frac{\rho}{\varepsilon_0}\qquad\&\qquad\nabla\cdot\mathbf{B}=0$$

$$\nabla\times\mathbf{E}+\frac{\partial\mathbf{B}}{\partial t}=0\qquad\&\qquad\nabla\times\mathbf{B}-\mu_0\varepsilon_0\frac{\partial\mathbf{E}}{\partial t}=\mu_0\mathbf{j},$$

where ε_0 is the electric constant (permittivity of free space) and μ_0 is the magnetic constant (permeability of free space), which are constants determined experimentally.

Law (Lorentz force law)**.** A point charge q experiences a force

$$\mathbf{F}=q(\mathbf{E}+\dot{\mathbf{r}}\times\mathbf{B}).$$

We can derive the conservation of electric charge from Maxwell's equations. Take the divergence of the last equation to obtain

$$\underbrace{\nabla\cdot(\nabla\times\mathbf{B})}_{=0}-\mu_0\varepsilon_0\frac{\partial}{\partial t}\underbrace{(\nabla\cdot\mathbf{E})}_{=\rho/\varepsilon_0}=\mu_0\nabla\cdot\mathbf{j}.\quad\Rightarrow\quad\frac{\partial\rho}{\partial t}+\nabla\cdot\mathbf{j}=0.$$

We can also take the volume integral of the first one and use the divergent theorem:

$$\int_V\nabla\cdot\mathbf{E}\,\mathrm{d}V=\frac{1}{\varepsilon_0}\int_V\rho\,\mathrm{d}V=\frac{Q}{\varepsilon_0}.\quad\Rightarrow\quad\int_S\mathbf{E}\cdot\mathrm{d}\mathbf{S}=\frac{Q}{\varepsilon_0},$$

which is Gauss' law for electric fields. Similarly, we can integrate the other equations to obtain

$$\int_S\mathbf{B}\cdot\mathrm{d}\mathbf{S}=0.\ \ \&\ \ \int_{C=\partial S}\mathbf{E}\cdot\mathrm{d}\mathbf{r}=\int_S\nabla\times\mathbf{E}\,\mathrm{d}S=-\frac{\mathrm{d}}{\mathrm{d}t}\int_S\mathbf{B}\cdot\mathrm{d}\mathbf{S}$$

The left side roughly states that there are no "magnetic charges". The right side (using Stokes' Theorem) says that a changing magnetic field produces a current.

12.2 Static charges and steady currents

If $\rho,\mathbf{j},\mathbf{E},\mathbf{B}$ are all independent of time, $\mathbf{E}$ and $\mathbf{B}$ are no longer linked.

We can solve the equations for electric fields:

$$\nabla\cdot\mathbf{E}=\rho/\varepsilon_0\quad\&\quad\nabla\times\mathbf{E}=\mathbf{0}$$

Second equation gives $\mathbf{E}=-\nabla\varphi$. Substituting into first gives $\nabla^2\varphi=-\rho/\varepsilon_0$.

The equations for the magnetic field are

$$\nabla\cdot\mathbf{B}=0\quad\&\quad\nabla\times\mathbf{B}=\mu_0\mathbf{j}$$

First equation gives $B = \nabla \times \mathbf{A}$ for some *vector potential* $\mathbf{A}$. But the vector potential is ambiguous. Making the transformation $\mathbf{A} \mapsto \mathbf{A} + \nabla\chi(\mathbf{x})$ produces the same $\mathbf{B}$, since $\nabla \times (\nabla\chi) = 0$. So choose χ such that $\nabla \cdot \mathbf{A} = 0$. Then

$$\nabla^2 \mathbf{A} = \nabla(\underbrace{\nabla \cdot \mathbf{A}}_{=0}) - \nabla \times (\underbrace{\nabla \times \mathbf{A}}_{\mathbf{B}}) = -\mu_0 \mathbf{j}.$$

12.3 Electromagnetic waves

Consider Maxwell's equations in empty space, ie. $\rho = 0$, $\mathbf{j} = \mathbf{0}$. Then Maxwell's equations give

$$\nabla^2 \mathbf{E} = \nabla(\nabla \cdot \mathbf{E}) - \nabla \times (\nabla \times \mathbf{E}) = \nabla \times \frac{\partial \mathbf{B}}{\partial t} = \frac{\partial}{\partial t}(\nabla \times \mathbf{B}) = \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial^2 t}.$$

Define $c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$. Then the equation (similarly with magnetic field) gives

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\mathbf{E} = 0. \quad \& \quad \left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\mathbf{B} = 0.$$

So Maxwell's equations predict that there exists electromagnetic waves in free space, which move with speed $C = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} \approx 2.00 \times 10^8 m\ s^{-1}$, which is the speed of light! Maxwell then concluded that light is electromagnetic waves!

13 Tensors and tensor fields

13.1 Definition

Recall the bases transformation rule of vectors from Vectors and Matrices. Suppose two orthonormal right-handed bases $\{\mathbf{e}_i\}$, $\{\mathbf{e}'_i\}$ (of $\mathbb{R}^3$) are related by

$$\mathbf{e}'_i = R_{ip}\mathbf{e}_p,$$

where R_{ip} is a rotation. To be precise, R_{ip} satisfies

(i) $R_{ip}R_{jp} = R_{qi}R_{qj} = \delta_{ij}$, ie. $RR^T = R^T R = I$, or R is *orthogonal*.

(ii) $\det R = 1$, ie. R is *special orthogonal*.

Then given a vector $\mathbf{v}$, its coordinates in these bases transform as

$$v'_i = R_{ip}v_p.$$

We can generalize these to *tensors*, which are multi-dimensional arrays that obey this transformation rule:

Definition (Tensor). A *tensor* of rank n has components $T_{ij\cdots k}$ (with n indices) with respect to each basis $\{\mathbf{e}_i\}$ or coordinate system $\{x_i\}$, and satisfies the following rule of change of basis:

$$T'_{ij\cdots k} = R_{ip}R_{jq}\cdots R_{kr}T_{pq\cdots r}.$$

So a tensor is a multi-dimensional array that transforms *nicely*.

Example.

(i) (Physical example) In some substances, an applied electric field $\mathbf{E}$ gives rise to a non-parallel current density $\mathbf{j}$, according to the linear relation $j_i = \epsilon_{ij}E_j$, where ϵ_{ij} is the *conductivity tensor*.

However, if the substance is *isotropic*, we have $\epsilon_{ij} = \sigma\delta_{ij}$.

13.2 Tensor algebra

Definition (Tensor addition)**.** Tensors T and S of the same rank can be *added*; $T+S$ is also a tensor of the same rank, defined as

$$(T+S)_{ij\cdots k} = T_{ij\cdots k} + S_{ij\cdots k}.$$

in any coordinate system.

To check that this is a tensor, we check the transformation rule. Again, we only show for $n=2$:

$$(T+S)'_{ij} = T'_{ij} + S'_{ij} = R_{ip}R_{jq}T_{pq} + R_{ip}R_{jq}S_{pq} = (R_{ip}R_{jq})(T_{pq} + S_{pq}).$$

Definition (Scalar multiplication)**.** A tensor T of rank n can be multiplied by a scalar α. αT is a tensor of the same rank, defined by

$$(\alpha T)_{ij} = \alpha T_{ij}.$$

Definition (Tensor product)**.** Let T be a tensor of rank n and S be a tensor of rank m. The *tensor product* $T \otimes S$ is a tensor of rank $n+m$ defined by

$$T_{x_1 x_2 \cdots x_n y_1 y_2 \cdots y_m} = T_{x_1 x_2 \cdots x_n} S_{y_1 y_2 \cdots y_n}.$$

Definition (Tensor contraction)**.** For a tensor T of rank n with components $T_{ijp\cdots q}$, we can *contract on* the indices i, j to obtain a new tensor of rank $n-2$:

$$S_{p\cdots q} = \delta_{ij} T_{ijp\cdots q} = T_{iip\cdots q}$$

Note that we can contract any pair we like.

If we view T_{ij} as a matrix, then the contraction is simply the trace of the matrix. So our result above says that the trace is invariant basis transformations - as we already know in Vectors and Matrices.

Note that our usual matrix product can be formed by first applying a tensor product to obtain $M_{ij}N_{pq}$, then contract with δ_{jp} to obtain $M_{ij}N_{jq}$.

13.3 Symmetric and antisymmetric tensors

Definition (Symmetric and anti-symmetric tensors)**.** A tensor T of rank n is *symmetric* in the indices i, j if it obeys

$$T_{ijp\cdots q} = T_{jip\cdots q}.$$

It is *anti-symmetric* if

$$T_{ijp\cdots q} = -T_{jip\cdots q}.$$

This is a property that holds in any coordinate systems, if it holds in one, since

$$T'_{k\ell r\ldots s} = R_{ki}R_{\ell j}R_{rp}\cdots R_{sq}T_{ijp\cdots q} = \pm R_{ki}R_{\ell j}R_{rp}\cdots R_{sq}T_{jip\cdots q} = \pm T_{\ell kr\cdots s}$$

as required.

Definition (Totally symmetric and anti-symmetric tensors)**.** A tensor is *totally (anti-)symmetric* if it is (anti-)symmetric in every pair of indices.

Example. In $\mathbb{R}^3$,

- Any totally antisymmetric tensor of rank 3 is $\lambda\varepsilon_{ijk}$ for some scalar λ.
- There are no totally antisymmetric tensors of rank greater than 3, except for the trivial tensor with all components 0.

13.4 Tensors, multi-linear maps and the quotient rule

13.4.1 Tensors as multi-linear maps

Definition (Multilinear map)**.** A map T that maps n vectors $\mathbf{a}, \mathbf{b}, \cdots, \mathbf{c}$ to $\mathbb{R}$ is multi-linear if it is linear in each of the vectors $\mathbf{a}, \mathbf{b}, \cdots, \mathbf{c}$ individually.

We will show that a tensor T or rank n is a equivalent to a multi-linear map from n vectors $\mathbf{a}, \mathbf{b}, \cdots, \mathbf{c}$ to $\mathbb{R}$ defined by

$$T(\mathbf{a}, \mathbf{b}, \cdots, \mathbf{c}) = T_{ij\cdots k} a_i b_j \cdots c_k.$$

To show that tensors are *equivalent* to multi-linear maps, we have to show the following:

(i) Defining a map with a tensor makes sense, ie. the expression $T_{ij\cdots k} a_i b_j \cdots c_k$ is the same regardless of the basis chosen;

(ii) While it is always possible to write a multi-linear map as $T_{ij\cdots k} a_i b_j \cdots c_k$, we have to show that $T_{ij\cdots k}$ is indeed a tensor, ie. transform according to the tensor transformation rules.

To show the first property, just note that the $T_{ij\cdots k} a_i b_j \cdots c_k$ is a tensor product (followed by contraction), which retains tensor-ness. So it is also a tensor. In particular, it is a rank 0 tensor, ie. a scalar, which is independent of the basis.

To show the second property, assuming that T is a multi-linear map, it must be independent of the basis, so

$$T_{ij\cdots k} a_i b_j \cdots c_k = T'_{ij\cdots k} a'_i b'_j \cdots c'_k.$$

Since $v'_p = R_{pi} v_i$ by tensor transformation rules, multiplying both sides by R_{pi} gives $v_i = R_{pi} v'_p$. Substituting in gives

$$T_{ij\cdots k}(R_{pi} a'_p)(R_{qj} b'_q) \cdots (R_{kr} c'_r) = T'_{pq\cdots r} a'_p b'_q \cdots c'_r.$$

Since this is true for all $\mathbf{a}, \mathbf{b}, \cdots \mathbf{c}$, we must have

$$T_{ij\cdots k} R_{pi} R_{qj} \cdots R_{rk} = T'_{pq\cdots r}$$

Hence $T_{ij\cdots k}$ obeys the tensor transformation rule, and is a tensor.

This shows that there is a one-to-one correspondence between tensors of rank n and multi-linear maps so we can think about tensors without coordinate systems.

13.4.2 The quotient rule

Proposition (Quotient rule)**.** Suppose that $T_{i\cdots jp\cdots q}$ is an array defined in each coordinate system, and that $v_{i\cdots j} = T_{i\cdots jp\cdots q} u_{p\cdots q}$ is also a tensor for any tensor $u_{p,\cdots q}$. Then $T_{i\cdots jp\cdots q}$ is also a tensor.

Proof. We can check the tensor transformation rule directly. However, we can reuse the result above to save some writing.

Consider the special form $u_{p\cdots q} = c_p \cdots d_q$ for any vectors $\mathbf{c}, \cdots \mathbf{d}$. By assumption,

$$v_{i\cdots j} = T_{i\cdots jp\cdots q} c_p \cdots d_q$$

is a tensor. Then

$$v_{i\cdots j}a_i\cdots b_j = T_{i\cdots jp\cdots q}a_i\cdots b_j c_p\cdots d_q$$

is a scalar for any vectors $\mathbf{a},\cdots,\mathbf{b},\mathbf{c},\cdots,\mathbf{d}$. Since $T_{i\cdots jp\cdots q}a_i\cdots b_j c_p\cdots d_q$ is a scalar and hence gives the same result in every coordinate system, $T_{i\cdots jp\cdots q}$ is a multilinear map. So $T_{i\cdots jp\cdots q}$ is a tensor. □

13.5 Tensor calculus

13.5.1 Tensor fields and derivatives

Just as with scalars or vectors, we can define tensor fields:

Definition (Tensor field)**.** A *tensor field* is a tensor, which we assume to be *smooth*, at each point $T_{ij\cdots k}(\mathbf{x})$,(other than obvious points where it goes wrong), which can also be written as $T_{ij\cdots k}(x_\ell)$.

We claim:

Proposition.

$$\underbrace{\frac{\partial}{\partial x_p}\cdots\frac{\partial}{\partial x_q}}_{m}T_{\underbrace{ij\cdots k}_{n}}, \qquad (*)$$

is a tensor of rank $n+m$.

Proof. Since $x_i' = R_{iq}x_q$, we have

$$\frac{\partial x_i'}{\partial x_p} = R_{ip}. \;\;\&\;\; \frac{\partial x_q}{\partial x_i'} = R_{iq}.$$

Note. Note that $\frac{\partial x_p}{\partial x_q} = \delta_{pq}$). R_{ip}, R_{iq} are constant matrices.

Hence by the chain rule,

$$\frac{\partial}{\partial x_i} = \left(\frac{\partial x_q}{\partial x_i'}\right)\frac{\partial}{\partial x_q} = R_{iq}\frac{\partial}{\partial x_q}.$$

So $\frac{\partial}{\partial x_p}$ obeys the vector transformation rule.

Note that this is expected for components of a vector, as we previously showed that $\nabla = \mathbf{e}_i\frac{\partial}{\partial x_i} = \mathbf{e}_i'\frac{\partial}{\partial x_i'}$.

□

13.5.2 Integrals and the tensor divergence theorem

We can sum tensors and (by assumption) take limits of the definition of a tensor-valued integral. For $\int_V T_{ij\cdots k}(\mathbf{x})\;\mathrm{d}V$ and the result is the tensor (think of the integral as the limit of a sum).

For a physical example, for a fluid with velocity $\mathbf{u}(\mathbf{x})$ through a surface element - assume a uniform density ρ. The flux of volume is $\mathbf{u}\cdot\mathbf{n}\delta s = u_j n_j\delta S$. So the flux of mass is $\rho u_j n_j\delta S$. Then the (ith component of momentum) is $\rho u_i u_j n_j\delta S = T_{ij}n_j k\delta S$ (mass times velocity), where $T_{ij} = \rho u_i u_j$. Then the flux through the surface S is $\int_S T_{ij}n_j\;\mathrm{d}S$. Let V be a volume bounded by a surface $S = \nabla V$ and $T_{ij\cdots k\ell}$ be a smooth tensor field. Then

Theorem (Divergence theorem for tensors).

$$\int_S T_{ij\cdots k\ell} n_\ell \, \mathrm{d}S = \int_V \frac{\partial}{\partial x_\ell}(T_{ij\cdots k\ell}) \, \mathrm{d}V,$$

with $\mathbf{n}$ being an outward pointing normal.

Proof. Apply the usual divergence theorem to the vector field $\mathbf{v}$ defined by $v_\ell = a_i b_j \cdots c_k T_{k\ell}$, where $\mathbf{a}, \mathbf{b}, \cdots, \mathbf{c}$ are fixed constant vectors.

Then

$$\nabla \cdot \mathbf{v} = \frac{\partial v_\ell}{\partial x_\ell} = a_i b_j \cdots c_k \frac{\partial}{\partial x_\ell} T_{ij\cdots k\ell},$$

and

$$\mathbf{n} \cdot \mathbf{v} = n_\ell v_\ell = a_i b_j \cdots c_k T_{ij\cdots k\ell} n_\ell.$$

Since $\mathbf{a}, \mathbf{b}, \cdots, \mathbf{c}$ are arbitrary, therefore they can be eliminated, and the tensor divergence theorem follows. □

14 Tensors of rank 2

14.1 Decomposition of a second-rank tensor

Any second rank tensor can be written as a sum of its symmetric and anti-symmetric parts

$$T_{ij} = S_{ij} + A_{ij},$$

where

$$S_{ij} = \frac{1}{2}(T_{ij} + T_{ji}), \quad A_{ij} = \frac{1}{2}(T_{ij} - T_{ji}).$$

Here T_{ij} has 9 independent components, whereas S_{ij} and A_{ij} have 6 and 3 independent components, since they must be of the form

$$(S_{ij}) = \begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix}, \quad (A_{ij}) = \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}.$$

The symmetric part can be be further reduced to a *traceless* part plus an *isotropic* (ie. multiple of δ_{ij}) part:

$$S_{ij} = P_{ij} + \frac{1}{3}\delta_{ij} Q,$$

where $Q = S_{ii}$ is the trace of S_{ij} and $P_{ij} = S_{ij} - \frac{1}{3}\delta_{ij} Q$ is traceless. Then P_{ij} has 5 independent components while Q has 1.

Since the antisymmetric part has 3 independent components, just like a usual vector, we should be able to write A_i in terms of a single vector. In fact, we can write the antisymmetric part as

$$A_{ij} = \varepsilon_{ijk} B_k$$

for some vector B. To figure out what this B is, we multiply by $\varepsilon_{ij\ell}$ on both sides and use some magic algebra to obtain

$$B_k = \frac{1}{2}\varepsilon_{ijk} A_{ij} = \frac{1}{2}\varepsilon_{ijk} T_{ij},$$

where the last equality is from the fact that only antisymmetric parts contribute to the sum.

Then

$$(A_{ij}) = \begin{pmatrix} 0 & B_3 & -B_2 \\ -B_3 & 0 & B_1 \\ B_2 & -B_1 & 0 \end{pmatrix}$$

To summarize,

$$T_{ij} = P_{ij} + \varepsilon_{ijk} B_k + \frac{1}{3}\delta_{ij} Q,$$

where $B_k = \frac{1}{2}\varepsilon_{pqj} T_{pq}$ and $Q = T_{kk}$.

14.2 The inertia tensor

Consider masses m_α with positions $\mathbf{r}_\alpha$, all rotating with angular velocity $\boldsymbol{\omega}$ about $\mathbf{0}$. So the velocities are $\mathbf{v}_\alpha = \boldsymbol{\omega} \times \mathbf{r}_\alpha$. The total angular momentum is

$$\mathbf{L} = \sum_\alpha \mathbf{r}_\alpha \times m_\alpha \mathbf{v}_\alpha = \sum_\alpha m_\alpha \mathbf{r}_\alpha \times (\boldsymbol{\omega} \times \mathbf{r}_\alpha) = \sum_\alpha m_\alpha (|\mathbf{r}_\alpha|^2 \boldsymbol{\omega} - (\mathbf{r}_\alpha \cdot \boldsymbol{\omega})\mathbf{r}_\alpha).$$

by vector identities. In components, we have

$$L_i = I_{ij}\omega_j,$$

where

Definition (Inertia tensor)**.** The *inertia tensor* is

$$I_{ij} = \sum_\alpha m_\alpha [|\mathbf{r}_\alpha|^2 \delta_{ij} - (\mathbf{r}_\alpha)_i (\mathbf{r}_\alpha)_j] \text{ or } I_{ij} = \int_V \rho(\mathbf{r})(x_k x_k \delta_{ij} - x_i x_j)\, \mathrm{d}V.$$

For a rigid body occupying volume V with mass density $\rho(\mathbf{r})$.

Note. In general, the inertia tensor is not always parallel to $\boldsymbol{\omega}$.

14.3 Diagonalization of a symmetric second rank tensor

Using matrix notation, the rule $T'_{ij} = R_{ip} R_{jq} T_{pq}$ becomes $T' = RTR^T = RTR^{-1}$..

If T is symmetric, the directions defined by its orthonormal eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are the *principal axes* for T, and the tensor is diagonal in Cartesian coordinates along these axes.

This applies to any symmetric rank-2 tensor. For the special case of the inertia tensor, the eigenvalues are called the *principal moments of inertia.*

As exemplified in the previous example, we can often guess the correct principal axes for I_{ij} based on the symmetries of the body. With the axes we chose, I_{ij} was found to be diagonal by direct calculation.

15 Invariant and isotropic tensors

15.1 Definitions and classification results

Definition (Invariant and isotropic tensor)**.** A tensor T is *invariant* under a particular rotation R if

$$T'_{ij\cdots k} = R_{ip} R_{jq} \cdots R_{kr} T_{pq\cdots r} = T_{ij\cdots k},$$

ie. every component is unchanged under the rotation.

A tensor T which is invariant under every rotation is *isotropic*, ie. the same in every direction.

Isotropic tensors in $\mathbb{R}^3$ can be classified:

Theorem.

(i) There are no isotropic tensors of rank 1, except the zero tensor.

(ii) The most general rank 2 isotropic tensor is $T_{ij} = \alpha\delta_{ij}$ for some scalar α.

(iii) The most general rank 3 isotropic tensor is $T_{ijk} = \beta\varepsilon_{ijk}$ for some scalar β.

(iv) All isotropic tensors of higher rank are obtained by combining δ_{ij} and ε_{ijk} using tensor products, contractions, and linear combinations.

We will provide a sketch of the proof:

Proof. (i) Suppose T_i is rank-1 isotropic. Consider a rotation about x_3 through π:

$$(R_{ij}) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We want $T_1 = R_{ip}T_p = R_{11}T_1 = -T_1$. So $T_1 = 0$. Similarly, $T_2 = 0$. By consider a rotation about, say x_1, we have $T_3 = 0$.

(ii) Suppose T_{ij} is rank-2 isotropic. Consider

$$(R_{ij}) = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which is a rotation through $\pi/2$ about the x_3 axis. Then

$$T_{13} = R_{1p}R_{3q}T_{pq} = R_{12}R_{33}T_{23} = T_{23}$$

and

$$T_{23} = R_{2p}R_{3q}T_{pq} = R_{21}R_{33}T_{13} = -T_{13}$$

So $T_{13} = T_{23} = 0$. Similarly, we have $T_{31} = T_{32} = 0$.

We also have

$$T_{11} = R_{1p}R_{1q}T_{pq} = R_{12}R_{12}T_{22} = T_{22}.$$

So $T_{11} = T_{22}$.

By picking a rotation about a different axis, we have $T_{21} = T_{12}$ and $T_{22} = T_{33}$.

Hence $T_{ij} = \alpha\delta_{ij}$.

(iii) We use the same method in (ii) with one π and one $\frac{\pi}{2}$ rotation by one axis.

□

Example. The most general isotropic tensor of rank 4 is

$$T_{ijk\ell} = \alpha\delta_{ij}\delta_{k\ell} + \beta\delta_{ik}\delta_{j\ell} + \gamma\delta_{i\ell}\delta_{jk}$$

for some scalars α, β, γ. There are no other independent combinations. (we might think we can write a rank-4 isotropic tensor in terms of ε_{ijk}, like $\varepsilon_{ijp}\varepsilon_{k\ell p}$, but this is just $\delta_{ik}\delta_{j\ell} - \delta_{i\ell}\delta_{jk}$. It turns out that anything you write with ε_{ijk} can be written in terms of δ_{ij} instead)

15.2 Application to invariant integrals

We have the following very useful theorem. It might seem a bit odd and arbitrary at first sight - if so, read the example below first (after reading the statement of the theorem), and things will make sense!

Theorem. Let

$$T_{ij\cdots k} = \int_V f(\mathbf{x}) x_i x_j \cdots x_k \ \mathrm{d}V.$$

where $f(\mathbf{x})$ is a scalar function and V is some volume.

Given a rotation R_{ij}, consider an *active* transformation: $\mathbf{x} = x_i \mathbf{e}_i$ is mapped to $\mathbf{x}' = x_i' \mathbf{e}_i$ with $x_i' = R_{ij} x_i$, ie. we map the components but not the basis, and V is mapped to V'.

Suppose that under this active transformation,

(i) $f(\mathbf{x}) = f(\mathbf{x}')$,

(ii) $V' = V$ (eg. if V is all of space).

Then $T_{ij\cdots k}$ is invariant under the rotation.

Proof. First note that the Jacobian of the transformation R is 1, since it is simply the determinant of R ($x_i' = R_{ip} x_p \Rightarrow \frac{\partial x_i'}{\partial x_p} = R_{ip}$), which is by definition 1. So $\mathrm{d}V = \mathrm{d}V'$.

Then we have

$$\begin{aligned} R_{ip} R_{jq} \cdots R_{kr} T_{pq\cdots r} = \int_V f(\mathbf{x}) x_i' x_j' \cdots x_k' \ \mathrm{d}V &= \int_V f(\mathbf{x}') x_i' x_j' \cdots x_k' \ \mathrm{d}V && \text{using (i)} \\ &= \int_{V'} f(\mathbf{x}') x_i' x_j' \cdots x_k' \ \mathrm{d}V' && \text{using (ii)} \\ &= \int_V f(\mathbf{x}) x_i x_j \cdots x_k \ \mathrm{d}V = T_{ij\cdots k} \end{aligned}$$

The second last step follows because the x_i or x_i' are just dummy variables. □

The result is particularly useful if (1) and (2) hold for *any* rotation R, in which case $T_{ij\cdots k}$ is isotropic.

Example. We look at the inertia tensor of a solid sphere of constant density ρ_0, or of mass $M = \frac{4}{3}\pi a^3 \rho_0$.

Recall that

$$I_{ij} = \int_V \rho_0 (x_k x_k \delta_{ij} - x_i x_j) \ \mathrm{d}V.$$

We see that I_{ij} is isotropic (since we have just shown that $\int x_i x_j \ \mathrm{d}V$ is isotropic, and

$x_k x_k \delta_{ij}$ is also isotropic). Let $I_{ij} = \beta \delta_{ij}$. Then

$$\begin{aligned}
I_{ij} &= \int_V \rho_0 (x_k x_k \delta_{ij} - x_i x_j)\ \mathrm{d}V \\
&= \rho_0 \left(\delta_{ij} \int_V x_k x_k\ \mathrm{d}V - \int_V x_i x_j\ \mathrm{d}V \right) \\
&= \rho_0 \left(\delta_{ij} T_{kk} - T_{ij} \right) \\
&= \rho_0 \left(\frac{4}{5} \pi a^5 \delta_{ij} - \frac{4}{15} \pi a^5 \delta_{ij} \right) \\
&= \frac{8}{15} \rho_0 \pi a^5 \delta_{ij} \\
&= \frac{2}{5} M a^2 \delta_{ij}.
\end{aligned}$$