Phys 4900 (Special Topics) Notes

David Murdock, TTU

March 19, 2013

Contents

Chapter 1

Special Functions, Mostly Those of Mr. Bessel

1.1 Introduction

There are a variety of topics in field of math–physics known as "special functions" that we could cover here but I'll focus on the Bessel functions. The reason for specializing is the time limitation and the fact that you have seen quite a bit about special functions already in the EM and QM courses. But it has occurred to me that the standard Bessel functions $J_n(x)$ and $Y_n(x)$ have not come up yet in these courses and if we're going this far in mathematical physics we should say something about them because you will see them later on in advanced physics. And they serve as a good example for other "special functions".

1.2 Bessel's Equation

Bessel's differential equation has the form

$$x^2y'' + xy' + (x^2 - \nu^2)y = 0 \tag{1.1}$$

and we arrive at this differential equation when we solve certain physics problems which have circular or cylindrical symmetry.

In the following we will use the convention of Griffiths' EM book and use (s, ϕ, z) for the cylindrical coordinates of a point. (Many books use ρ in place of s, Griffiths' choice came from the fact that ρ already stands for an important quantity in EM.)

Recall that in these coordinates the Laplacian is given by

$$\nabla^2 t = \frac{1}{s}\frac{\partial}{\partial s}\left(s\frac{\partial t}{\partial s}\right) + \frac{1}{s^2}\frac{\partial^2 t}{\partial \phi^2} + \frac{\partial^2 t}{\partial z^2}$$

where we might sometimes want to expand the $\partial/\partial s$ term.

I will show how Bessel's equation results from four different problems in physics (where we intend to exploit cylindrical symmetry).

1.2.1 Electromagnetism

We'll discuss two problems. The first is the case of electrostatics where we want to solve Laplace's equation with boundary conditions given on a cyliner, say, a cylinder of radius R whose axis lies along the z axis and which extends from $z = 0$ to $z = L$.

With V being the electric potential, Laplace's equation in cylindrical coordinates is

$$\nabla^2 V = \frac{1}{s}\frac{\partial}{\partial s}\left(s\frac{\partial V}{\partial s}\right) + \frac{1}{s^2}\frac{\partial^2 V}{\partial \phi^2} + \frac{\partial^2 V}{\partial z^2} = 0 \tag{1.2}$$

where $V = V(s, \phi, z)$. Note, here the potential does have z dependence, which generalizes similar problems with circular symmetry done at the level of Griffiths' book.

We follow our usual procedure of separation of variables and see what comes out. We look for solutions of the form

$$V(s, \phi, z) = S(s)\,\Phi(\phi)\,Z(z)$$

in the hopes that a general solution would be a linear combination of these that wouldn't be too hard to find. Put this into 1.2 and divide by V. We get

$$\frac{1}{S}\frac{1}{s}\frac{\partial}{\partial s}\left(s\frac{\partial S}{\partial s}\right) + \frac{1}{\Phi}\frac{1}{s^2}\frac{\partial \Phi}{\partial^2 \phi^2} + \frac{1}{Z}\frac{\partial Z}{\partial z^2} = 0 \tag{1.3}$$

We recall our usual argument (which I won't repeat here) that when we have three independent variables and an equation of the type we just got where a term depends only on *one* of those variables, then any such term must be a *constant*. Thus in 1.3, the third term depends only on z and with a little foresight we define that to be the positive quantity k^2.

That gives the fairly simple equation with solutions:

$$\frac{1}{Z}\frac{d^2 Z}{dz^2} = k^2 \qquad \Longrightarrow \qquad \frac{d^2 Z}{dz^2} = k^2 Z \qquad \Longrightarrow \qquad Z(z) = A\sinh(kz) + B\cosh(kz)$$

Now if we proceed with Eq. 1.3 we get

$$\frac{1}{S}s\frac{\partial}{\partial s}\left(s\frac{\partial S}{\partial s}\right) + s^2 k^2 + \frac{1}{\Phi}\frac{\partial^2 \Phi}{\partial \phi^2} = 0 \tag{1.4}$$

and again we get terms which only depend on one of the variables; the first two depend only on s while the last one depends only on ϕ, so it is a constant which we choose to write in the form $-m^2$. That gives the ordinary differential equation and solution

$$\frac{d^2 \Phi}{d\phi^2} = -m^2 \Phi \qquad \Longrightarrow \qquad \Phi = Ce^{\pm im\phi}$$

and we see that for Φ to be single valued, m must be an *integer*. In place of of this solution we can also write

$$\Phi(\phi) = C_1 \sin m\phi + C_2 \cos m\phi$$

Now make the substitution for ϕ in the last version of the DE and after multiplying by s^2 and expanding the terms in s we arrive at

$$s^2\frac{d^2 S}{ds^2} + s\frac{dS}{ds} + (k^2 s^2 - m^2)S = 0 \tag{1.5}$$

We make a change of variable $x \equiv ks$ (so that x is dimensionless) and redefine the solution as

$$y(x) \equiv S(s)$$

and this produces

$$x^2\frac{d^2 y}{dx^2} + x\frac{dy}{dx} + (x^2 - m^2)y = 0$$

which is Bessel's DE of order m.

The next problem I want to solve is that of a very flat conducting cavity which has the shape of a very flat circular cylinder; it has radius R and height a with $a \ll R$. We want to consider a standing electromagnetic wave in this cavity, particularly the ones of lowest frequency.

The z axis points along the axis of the cylinder.

Now I'll be rather loose in getting to the desired equation, but for allwe know at first, the E and B could have any sorts of components along z or in the xy plane. But if we consider the field boundary conditions at the surface of a conductor:

$$\mathbf{E}^{\parallel} = 0 \qquad B^{\perp} = 0$$

then we see that if the B field inside does have a z component B_z, then B_z must be zero at the top and bottom surfaces but non-zero in between. The curvature of the B field would contribute to making the resonant frequency higher and we want to look for the lowest modes. Now E_z can be nonzero and independent of z (it just vanishes abruptly at the surface) so that is what we will assume, and also that there is no xy–plane part of E (though B must have some component in the xy plane). So these assumptions (justifiable with more work!) give us the form:

$$\mathbf{E} = E_z(x,y)e^{-i\omega t}\hat{\mathbf{k}} \qquad \mathbf{B} = \mathbf{B}(x,y)e^{-i\omega t}$$

where we mean the *real part* of these functions and we won't need to know more about $\mathbf{B}$, as we will be solving for E_z.

The Maxwell equations give

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = i\omega \mathbf{B}$$

$$\nabla \times \mathbf{B} = -\frac{1}{i\omega}\nabla^2 \mathbf{E} = -i\omega \mathbf{E}$$

These give

$$\nabla^2 \mathbf{E} = -\frac{\omega^2}{c^2} = -k^2 \mathbf{E} \qquad \text{where} \qquad k = \frac{\omega}{c}$$

But $\mathbf{E}$ has only a z component with (x,y) dependence and this leaves

$$(\nabla^2 + k^2)E_z = 0 \qquad \text{or} \qquad \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + k^2\right) E_z = 0$$

This is referred to as the Helmholtz equation in two-dimensions; note, this time the k^2 term came from *time–dependence* and not a dependence on z.

Clearly, we *really* want to use cylindrical coordinates (s,ϕ) instead of x and y. To solve for $E_z(s,\phi)$. Again we use the strategy of separation of variables and write

$$E_z(s,\phi) = S(s)\Phi(\phi)$$

Substitute, divide by E_z and multiply by s^2 and get

$$\frac{1}{S} s \frac{\partial}{\partial s}\left(s \frac{\partial S}{\partial s}\right) + \frac{1}{\Phi}\frac{\partial^2 \Phi}{\partial \phi^2} + s^2 k^2 = 0$$

As in the last example the middle term is a constant and is set equal to $-m^2$, so that we have the separated DE

$$\frac{1}{\Phi}\frac{d^2\Phi}{d\phi^2} = -m^2 \qquad \Longrightarrow \qquad \frac{d^2\Phi}{d\phi^2} = -m^2\Phi \qquad \Longrightarrow \qquad \Phi(\phi) = Ce^{\pm im\phi}$$

and as before m must be an integer for a sensible answer. The remaining equation is a DE for $S(s)$; splitting up the first term we get

$$s^2\frac{d^2}{ds^2} + s\frac{dS}{ds} + (s^2k^2 - m^2)S = 0$$

and as before define $x = ks$ and $y(x) = S(s)$ and get

$$x^2y'' + xy' + (x^2 - m^2)y = 0\ ,$$

Bessel's DE.

1.2.2 Classical Mechanics: Vibration of Drumhead

Recall that the wave equation for wave on a string is

$$\frac{\partial^2 y}{\partial t^2} = v^2\frac{\partial^2 y}{\partial x^2} \tag{1.6}$$

When we impose boundary conditions

$$y(0,t) = 0 \qquad y(L,t) = 0$$

we get the solutions for the standing waves,

$$y(x,t) = A\sin\left(\frac{n\pi x}{L}\right)\cos(\omega_n t) \qquad 0 \le x \le L \tag{1.7}$$

where

$$\lambda_n = \frac{2L}{n} \qquad \lambda_n f_n = v \qquad \omega_n = 2\pi f_n$$

The next step is to solve the problem of a *membrane* vibrating in response to some disturbance but having a boundary. We consider a circular membrane of radius R; an (ideal) drumhead! Waves will also travel on a membrane and they satisfy a similar wave equation. With $z(x,y,t)$ being the displacement of the material at (x,y) at time t, z can be shown to satisfy a generalized version of 1.6:

$$\frac{\partial^2 z}{\partial t^2} = v^2\nabla^2 z \tag{1.8}$$

where ∇^2 is the two–dimensional laplacian. If we search for separated solutions of the form $z(x,y,t) = Z(x,y)T(t)$, we find that these solutions are

$$z(x,y,t) = Z(x,y)e^{-i\omega t} \tag{1.9}$$

Define $k^2 \equiv \frac{\omega^2}{v^2}$, then this is

$$\nabla^2 Z + k^2 Z = 0 \tag{1.10}$$

which is the Helmholtz equation in two dimensions. We want to use this for a circular membrane constrained to have no motion at $r = R$ so we work in cylindrical coordinates, for which we'll use (s,ϕ). Do a further separation of variables:

$$Z(s,\phi) = S(s)\Phi(\phi)$$

then substitution and a little math gives

$$\frac{1}{S}s\frac{\partial}{\partial s}\left(s\frac{\partial S}{\partial s}\right)+\frac{1}{\Phi}\frac{\partial^2\Phi}{\partial\phi^2}+k^2s^2=0$$

We appeal to the usual argument that when we have a sum of terms where each terms depends only *one* of those independent variable, then each of the terms is equal to a constant. The middle term here is set equal to the constant written as $-m^2$, and will have solutions

$$\Phi(\phi)=e^{\pm im\phi}$$

which leaves

$$s\frac{\partial}{\partial s}\left(s\frac{\partial S}{\partial s}\right)+(k^2s^2-m^2)S=0$$

Change variables to $x=ks$ and rename the function $S(s)\to y(x)$ (these have nothing to do with our original Cartesian coordinates x and y) and we get the differential equation

$$x^2y''+xy'+(x^2-m^2)y=0\ ,$$

which is (again) Bessel's equation.

1.2.3 Quantum Mechanics; Trapped Particle in Circular Geometry

Lastly, we solve a problem where a particle of mass m moves in two dimensions and is trapped in a circular region (radius R) where the potential is $V=0$ while outside the circle the potential is infinite. We want to find the energy eigenvalues.

Inside the region the wave function satisfies

$$-\frac{\hbar^2}{2m}\nabla^2\psi=E\psi$$

and doing our usual pre-math clean-up work we have

$$\nabla^2\psi=-\frac{2mE}{\hbar^2}\psi\equiv-k^2\psi$$

giving the 2-D Helmholtz equation,

$$(\nabla^2+k^2)\psi=0$$

As before we will want to use cylindrical (well, plane polar) coordinates and separate variables; thus we look for solutions of the form

$$\psi(s,\phi)=S(s)\Phi(\phi)$$

and then as before substitute and find the possibilities for $S(s)$ and $\Phi(\phi)$. Again we get angular solutions of the form $\Phi(\phi)=e^{\pm im\phi}$ and we get Bessel's differential equation for $y(x)=S(s)$, with $x=ks$.

These four examples should convince you that the conventional (integer-order) Bessel functions are important in mathematical physics! As important as circles and cylinders at any rate.

1.3 Bessel Functions

1.3.1 A Sketch of How to Solve the DE

We will solve Bessel's DE by the series method, and the series solutions will be *the* solutions (though in the problems and examples we'll see other exotic ways to represent them). If this seems awkward recall that when you really get down it, a series is basically all we have for the "closed–form" functions like $\sin(x)$ and e^x as well, so it's really a bit arbitrary what we call "closed –form" or a "standard" function.

Rewrite Bessel's DE of order ν as

$$y'' + \frac{1}{x}y' + \left(1 - \frac{\nu^2}{x^2}\right) y = 0$$

and try a solution of the form

$$y = x^\sigma \sum_{n=0}^{\infty} a_n x^n$$

which because of extra power of x out in front is often called a Froebenius series. Substituting, the DE can be shown to give

$$\sum_{n=0}^{\infty}[(\sigma+n)^2 - \nu^2]a_n x^n + \sum_{n=0}^{\infty} a_n x^{n+2} = 0$$

and matching the coefficient of x^0 gives the "indicial equation",

$$\sigma^2 - \nu^2 = 0 \qquad \Longrightarrow \qquad \sigma = \pm\nu$$

Matching coefficients of higher powers of x gives

$$[(\sigma+1)^2 - \nu^2]a_1 = 0$$

and for $n \geq 2$ the recursion relation

$$[(\sigma+n)^2 - \nu^2]a_n + a_{n-2} = 0$$

which result in

$$(1 \pm 2\nu)a_1 = 0 \qquad \text{and} \tag{1.11}$$

$$n(n \pm 2\nu)a_n + a_{n-2} = 0 \qquad \text{for} \quad n \geq 2 \tag{1.12}$$

From 1.11 we deduce that a_1=0. (Actually that doesn't follow if $\nu = \pm\frac{1}{2}$, but we will also made that choice for that case and we will still obtain two independent solutions.)

We will make some standard choice for a_0 and then the recursion relation Eq. 1.11 will give all the even a_n's while setting all the odd a_n's equal to zero. We note that **??** will give us trouble if ν is an integer because then for the "$-$" case (with positive ν) n eventually takes on the value 2ν so that we are dividing by zero! So we divide up our work into the cases where ν is not an integer and where it is an integer.

Two cases to consider:

1) Non-integer ν

Then (except for the half–odd–integer cases which need to be checked separately) the two possibilities for σ, $\sigma = \nu$ and $\sigma = -\nu$ don't differ by an integer. Then the steps given above two *linearly independent* solutions, with recursion relations

$$a_n = \begin{cases} -\frac{1}{n(n\pm 2\nu)} a_{n-2} & n = 2, 4, 6, \ldots \\ 0 & n = 1, 3, 5, \ldots \end{cases}$$

By convention we set

$$a_0 = \frac{1}{2^{\pm\nu}\Gamma(1 \pm \nu)}$$

which give the **Bessel functions** $J_\nu(x)$ and $J_{-\nu}(x)$, and the general solution to Bessel's DE is

$$y(x) = c_1 J_\nu(x) + c_2 J_{-\nu}(x)$$

Example: Find the solution to

$$x^2 y'' + xy' + (x^2 - \tfrac{1}{4})y = 0$$

Here, $\nu = \frac{1}{2}$ so the general solution is

$$y(x) = c_1 J_{1/2}(x) + c_2 J_{-1/2}(x)$$

In addition, one can show that

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}} \sin(x) \qquad \text{and} \qquad J_{-1/2}(x) = \sqrt{\frac{2}{\pi x}} \cos(x)$$

2) Integer ν

We will *not* have a general solution of the form

$$y(x) = c_1 J_\nu(x) + c_2 J_{-\nu}(x) \; .$$

for this case as the two J's are not independent functions.

Of course, the case $\nu = 0$ gives the same thing for $J_{\pm\nu}(x)$. Its Frobenius series is

$$J_0(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{2^n \, n! \, \Gamma(n+1)} = 1 - \frac{x^2}{2^2} + \frac{x^4}{2^2 4^2} - \frac{x^6}{2^2 4^2 6^2} + \cdots$$

where as we'll discuss later, $\Gamma(n+1) = n!$.

See a plot of $J_0(x)$ in the books!

But if we use the series to *define* $J_\nu(x)$ we find that

$$J_{-\nu}(x) = (-1)^\nu J_\nu(x)$$

no they are not independent.

We can find a second solution to the equation of order (integer) ν by using the definition

$$Y_\nu(x) \equiv \frac{J_\nu(x)\cos\nu\pi - J_{-\nu}(x)}{\sin\nu\pi}$$

which is OK for non-integer ν, and its *limit* is also OK as ν approaches an integer value. Thus for integer ν we use

$$Y_\nu(x) \equiv \lim_{\mu\to\nu}\left[\frac{J_\mu(x)\cos\mu\pi - J_{-\mu}(x)}{\sin\mu\pi}\right]$$

which are called the **Neumann functions**, or sometimes **Bessel functions of the second kind**.

Note: Many books use $N_\nu(x)$ for the Neumann functions (though everybody uses $J_\nu(x)$ for the Bessel functions).

There is a series representation for the Neumann functions of integer order, and it's complicated:

$$\begin{aligned} Y_n(x) = \frac{2}{\pi}\{\ln(x/2)+\gamma\}J_n(x) \quad &- \quad \frac{1}{\pi}\sum_{k=0}^{n-1}\frac{(n-k-1)!(x/2)^{2k-n}}{k!} \\ &- \quad \frac{1}{\pi}\sum_{k=0}^{\infty}(-1)^k\{\Phi(k)+\Phi(n+k)\}\frac{(x/2)^{2k+n}}{k!(n+k)!} \end{aligned}$$

where γ is the Euler constant, $\gamma \approx 0.5772156$ and

$$\Phi(p) \equiv 1 + \frac{1}{2} + \frac{1}{3} + \cdots \frac{1}{p}$$

You may have a reason to use the Y_n's some day, but clearly you will want a computer to do the work.

Thus for integer ν the general solution to Bessel's differential equation is

$$y(x) = c_1 J_\nu(x) + c_2 Y_\nu(x)$$

1.4 Properties of the Bessel Functions

1.4.1 Imposing $y = 0$ at $s = a$

Many of the properties of the Bessel functions follow from the fact that Bessel's DE is of a general type of DE known as the **Sturm–Liouville** type of DE's. These are of the general form

$$(py')' + qy + \lambda\rho y = 0 \tag{1.13}$$

or as

$$\mathcal{L}y = \lambda\rho y \qquad \text{with} \quad \mathcal{L} \equiv -\left[\frac{d}{dx}\left(p(x)\frac{d}{dx}\right) + q(x)\right] \tag{1.14}$$

where p and q are functions of x and λ is a constant.

The Bessel DE is the case where

$$p = x \qquad q = -\frac{\nu^2}{x} \qquad \lambda = \alpha^2 \qquad \rho = x$$

where the *function* y in these equation is $y(\alpha x)$.

Before stating the orthogonality properties of the J_ν's, we need a little nomenclature.

First off, recall that recall that our solution to the radial DE (of the *physical problem*) was originally called $S(s) = y(x)$, with $x = ks$. We need our solutions in terms of the coordinate s, so for example we will be writing

$$S(s) = y(x) = J_\nu(x) = J_\nu(ks)$$

so if we want to impose the boundary condition $S(s) = 0$ at $s = a$ (as we did in all the physics examples) then we are demanding that ka be a zero of $J_\nu(x)$ and this sets the value of k. Thus it is quite important to know the values of the zeroes of the J_ν to get the full solutions.

Let α_{mn} stand for the n^{th} (nontrivial) zero of J_m. Examples:

$$\alpha_{01} = 2.4018 \qquad \alpha_{02} = 5.5201 \qquad \alpha_{03} = 8.65377$$

Big tables of these exist.

For problems where where boundary condition at the circle's radius is that the *slope* of the function is zero. Thus it also important to know where the *derivatives* of the Bessel functions are zero, and those points are given by β_{mn}; this is the n^{th} value of x such that $J'_m(x) = 0$.

So, invoking our Dirichlet boundary condition (the value of the function is zero) at $r = a$, for each m we have

$$ka = \alpha_{mn} \quad \text{for} \qquad n = 1, 2, 3, \ldots \qquad \Longrightarrow \qquad k = \frac{\alpha_{mn}}{a}$$

so the radial solution is

$$S(s) = J_m\left(\frac{\alpha_{mn}}{a}s\right)$$

There is an important orthogonality relation between the Bessel functions when we consider a restricted range of the radial coordinate; we will just consider s between 0 and a particular radius a. use Sturm-Liouville theory one can show that Bessel functions of the *same* order are orthogonal in the following sense:

$$\int_0^1 J_m\left(\alpha_{mn}\frac{s}{a}\right) J_m\left(\alpha_{mn'}\frac{s}{a}\right)\, s\, ds = 0 \qquad \text{for} \quad \neq n' \tag{1.15}$$

A more general orthogonality condition is given in RHB pp 608-611.

One can also show the continuum version of relation,

$$\int_0^\infty J_\nu(\alpha s) J_\nu(\alpha' s)\, s\, ds = \frac{1}{\alpha}\delta(\alpha - \alpha') \tag{1.16}$$

From the orthogonality condition we can functions of the radial coordinate s, restricted to $0 \le s \le a$, as a linear combination of Bessel functions. Thus, for any reasonably well-behaved function $f(s)$ defined on $0 \le s \le a$ we have

$$f(s) = \sum_{m=1}^{\infty} c_{\nu m} J_\nu\left(\alpha_{\nu m}\frac{s}{a}\right) \qquad 0 \le sa$$

where

$$c_{\nu m} = \frac{2}{a^2[J_{\nu+1}(\alpha_{\nu m})]^2}\int_0^a f(s) J_\nu\left(\alpha_{\nu m}\frac{s}{a}\right)\, s\, ds$$

1.4.2 Recurrence Relations

The Bessel functions can be shown to satisfy certain relations between J_ν and $J_{\nu\pm1}$. One can show that for integer n we have

$$J_{n-1}(x) + J_{n+1}(x) = \frac{2n}{x}J_n(x)$$

$$J'_n(x) = \tfrac{1}{2}[J_{n-1}(x) - J_{n+1}(x)]$$

$$xJ'_n(x) = nJ_n(x) - xJ_{n+1}(x)$$

$$xJ_n'(x) = xJ_{n-1}(x) - nJ_n(x)$$

and two more relations involving derivatives,

$$\frac{d}{dx}[x^\nu J_\nu(x)] = x^\nu J_{\nu-1}(x)$$

$$\frac{d}{dx}[x^{-\nu} J_\nu(x)] = -x^{-\nu} J_{\nu+1}(x)$$

1.4.3 Generating Functions

One can produce the special functions by picking off the coefficients of a power series. Before showing how this works for the Bessel functions, we show how this works for the more familiar functions.

Consider the function

$$G(x,h) = (1 - 2xh + h^2)^{-1/2} \tag{1.17}$$

This can be expanded as a power series in h, and then the coefficients will depend on x:

$$G(x,h) = \sum_{n=0}^{\infty} g_n(x) h^n$$

and in fact when this is done for the function in 1.17, the coefficients $g(x)$ are the Legendre polynomials:

$$(1 - 2xh + h^2)^{-1/2} = \sum_{n=0}^{\infty} P_n(x) h^n \tag{1.18}$$

Now, you'll note that in *this* case the generating function is an important one. It shows up as a factor in all sorts of physics problems, since for two vectors $\mathbf{r}$ and $\mathbf{r}'$ with $r > r'$ we have

$$|\mathbf{r} - \mathbf{r}'| = \sqrt{r^2 + r'^2 - 2rr'\cos\theta} = r\sqrt{1 - 2xh + h^2}$$

where $x = \cos\theta$ and $h = \frac{r'}{r} < 1$. And this functions now has an expansion involving the P_n's.

The *associated* Legendre polynomials (a variant of the regular Legendre polynomials, which you'll recall went into making up the spherical harmonics) which we deonted by $P_n^m(x)$ have their own generating function, which is bit more obscure, but its expansion produces the P_n^m:

$$G(x,h) = \frac{(2m!)(1-x^2)^{m/2}}{2^m m!(1 - 2xh + h^2)^{m+\frac{1}{2}}} = \sum_{n=0}^{\infty} P_{n+m}^m(x) h^n$$

And in case you'd like to collect these things, here are the generating functions for the Laguerre and Associated Laguerre polynomials where were encountered in quantum mechanics:

$$G(x,h) = \frac{e^{-xh/(1-h)}}{1-h} = \sum_{n=0}^{\infty} L_n(x) h^n$$

$$G(x,h) = \frac{e^{-xh/(1-h)}}{(1-h)^{m+1}} \sum_{n=0}^{\infty} L_n^m(x) h^n$$

Now in each of the above cases the coefficient of h^n was a *polynomial*. But the coefficient of h^n could be in itself *another* infinite series (in powers of x) and that is what happens with the Bessel functions. Its generating function is

$$G(x,h) = \exp\left[\frac{x}{2}\left(h - \frac{1}{h}\right)\right]$$

and then the Bessel functions $J_n(x)$ can be found from the series for $G(x,h)$:

$$G(x,h) = \exp\left[\frac{x}{2}\left(h - \frac{1}{h}\right)\right] = \sum_{n=-\infty}^{\infty} J_n(x)h^n \tag{1.19}$$

Note here also that this is a sum on positive and negative powers of h. While we have a simple relation between $J_n(x)$ and $J_{-n}(x)$, the expansion of the generating function requires all the n's.

Finally, one can derive lots of curious representations of the J_n's and these are given in many old fat math books written by old fat 19$^{\text{th}}$–century mathematicians. An example is

$$J_n(x) = \frac{1}{\pi}\int_0^{\pi} \cos(n\theta - x\sin\theta)\, d\theta$$

This is a little surprising; the Bessel functions are related to the trig functions, but notice that the integrand is the cosine *of* another trig function.

One can also show

$$J_n(x) = \frac{1}{\pi}\int_0^{\pi} \cos(x\sin\theta)\cos n\theta\, d\theta \qquad \text{for} \quad n \text{ even.}$$

$$J_n(x) = \frac{1}{\pi}\int_0^{\pi} \sin(x\sin\theta)\sin n\theta\, d\theta \qquad \text{for} \quad n \text{ odd.}$$

1.4.4 DE's Which Can Be Made Into Bessel's DE

A differential equation of the form

$$x^2y'' + (2k+1)xy' + (\alpha^2x^{2r} + \beta^2)y = 0 \tag{1.20}$$

where k, α, r and β are constants can be made into Bessel's DE with a change of variables. The result is that 1.20 has the general solution

$$y = x^{-k}[c_1 J_{\kappa/r}(\alpha x^r/r) + c_2 Y_{\kappa/r}(\alpha x^r/r)] \tag{1.21}$$

where $\kappa = \sqrt{k^2 - \beta^2}$.

The case where $\alpha = 0$ is special in the DE of 1.20; in that case the solution is

$$y = x^{-k}(c_3x^{\kappa} + c_4x^{-\kappa})$$

1.5 Close Relatives of the Bessel Functions

1.5.1 Hankel Functions

Hankel functions are just simple combinations of the J_ν's and Y_ν's which are useful for problems involving waves. They are defined as

$$H_\nu^{(1)}(x) \equiv J_\nu(x) + iY_\nu(x) \qquad H_\nu^{(2)}(x) \equiv J_\nu(x) - iY_\nu(x) \tag{1.22}$$

which are called, respectively, Hankel functions of the **first and second kind**. The H_ν's satisfy the same recurrence relations as the J_ν and Y_ν.

Then we have...

1.5.2 Modified Bessel Functions

Now suppose we started this whole game with a DE similar to Bessel's, DE:

$$x^2\frac{d^2}{dx^2}y(x) + x\frac{d}{dx}y(x) - (x^2+\nu^2)y(x) = 0 \ , \tag{1.23}$$

the difference from Bessel's DE being the sign on the x^2 term.

1.23 can be put in the form of Bessel's DE with the replacement $x = -it$. Then we get

$$t^2\frac{d^2}{dt^2}y(-it) + t\frac{d}{dt}y(-it) + (t^2-\nu^2)y(-it) = 0 \tag{1.24}$$

One solution to this equation is then

$$y(x) = I_\nu(x) \equiv i^{-\nu}J_\nu(ix) \qquad \text{or} \qquad I_\nu(x) = e^{-\nu\pi i/2}J_\nu\left(xe^{i\pi/2}\right)$$

The second (independent) solution to 1.24 is chosen in various ways, but one popular choice is

$$K_\nu(x) \equiv \frac{\pi}{2}i^{\nu+1}H_\nu^{(1)}(ix) = \frac{\pi}{2}i^{\nu+1}[J_\nu(ix) + iY_\nu(ix)]$$

but this K_ν doesn't satisfy the recurrence relations. $I_\nu(x)$ and $K_\nu(x)$ are known as the **modified Bessel functions**.

Always be careful of which definition you're using when working with special functions!

1.5.3 Connection With the Spherical Bessel Functions

Recall the spherical Bessel function which we encountered when studying the Schrödinger equation in spherical coordinates for the case of zero potential. The equation

$$(\nabla^2 + k^2)\psi(\mathbf{r}) = 0$$

in spherical coordinates led to the radial equation for an l angular solution:

$$r^2\frac{d^2R}{dr^2} + 2r\frac{dR}{dr} + [k^2r^2 - l(l+1)]R = 0$$

Define $R(r) = r^{-1/2}u(r)$ and let

$$x = kr, \qquad \text{and} \qquad y(x) = u(kr)$$

Then we get

$$x^2y'' + xy' + [x^2 - (l+\tfrac{1}{2})^2]y = 0$$

which is Bessel's equation of order $l+\frac{1}{2}$, with solutions $J_{l+1/2}(x)$ and $Y_{l+1/2}(x)$, so that the general solution is

$$R(r) = r^{-1/2}[c_1J_{l+\frac{1}{2}}(x) + c_2Y_{l+\frac{1}{2}}(x)]$$

The functions

$$\frac{1}{\sqrt{x}}J_{l+\frac{1}{2}}(x) \qquad \text{and} \qquad \frac{1}{\sqrt{x}}Y_{l+\frac{1}{2}}(x)$$

when normalized in the conventional way are the spherical Bessel functions. Specifically,

$$j_l(x) = \sqrt{\frac{\pi}{2x}}J_{l+\frac{1}{2}}(x) \qquad n_l(x) = \sqrt{\frac{\pi}{2x}}Y_{l+\frac{1}{2}}(x)$$

Particular expressions for the j_l and n_l were given in the quantum mechanics textbooks; the first two are

$$j_0(x) = \frac{\sin x}{x} \qquad n_0(x) = -\frac{\cos x}{x} \ .$$

Recall that in scattering theory we had need for the spherical Hankel functions,

$$h_l^{(1)}(x) \equiv j_l(x) + in_l(x) \qquad \text{and} \qquad h_l^{(2)}(x) \equiv j_l(x) - in_l(x)$$

1.6 The Gamma Function and More

The **gamma function** is defined as

$$\Gamma(x) \equiv \int_0^\infty x^{n-1} e^{-x}\, dx \qquad \text{for} \quad n > 0 \ . \tag{1.25}$$

This holds for real $n > 0$; for integer n it is related to $n!$ (see below).

$\Gamma(x)$ satisfies a recurrence relation. We evaluate $\Gamma(n+1)$:

$$\begin{aligned} \Gamma(n+1) = \int_0^\infty x^n e^{-x}\, dx &= \left[-x^n e^{-x}\right]\Big|_0^\infty + \int_0^\infty n x^{n-1} e^{-x}\, dx \\ &= 0 + n\int_0^\infty x^{n-1} e^{-x}\, dx \\ &= n\Gamma(n) \end{aligned}$$

Since $\Gamma(1) = 1$, the recursion relation gives

$$\Gamma(n+1) = n! \qquad \text{for} \quad n = 1, 2, 3, \ldots \tag{1.26}$$

(It also holds for $n = 0$ if we use the common definition $0! = 1$.)

Here we note that with the recursion relation $\Gamma(n+1) = \Gamma(n)$ and the convention $0! = 1$ it is possible to define a factorial for *negative* n. We would get

$$n! = \frac{(n+m)!}{(n+m)(n+m-1)\cdots(n+1)} \qquad \text{for any } m \text{ with } n+m > 0.$$

but this definition is not useful in my experience, and it conflicts with the basic relation between the factorial and the gamma function, as $\Gamma(x)$ blows up for negative x.

See a plot of $\Gamma(x)$ in the books.

We can get another closed for Γ for "half-integer" values, since

$$\Gamma(\tfrac{1}{2}) = \sqrt{\pi}$$

so that the recursion relation gives

$$\tfrac{3}{2}! = \tfrac{3}{4}\sqrt{\pi} \qquad \tfrac{1}{2}! = \frac{\sqrt{\pi}}{2} \qquad (-\tfrac{1}{2})! = \sqrt{\pi} \qquad \text{etc.}$$

1.6.1 More "Interesting" Results

The old fat 19th–century mathematicians have shown more interesting results for the gamma function. All right, you *might* find them interesting.

1)

$$\Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin x\pi}$$

2)

$$2^{2x-1}\Gamma(x)\Gamma(x+\tfrac{1}{2}) = \sqrt{\pi}\Gamma(2x)$$

3)

$$\Gamma(x)\Gamma\left(1+\tfrac{1}{m}\right)\Gamma\left(1+\tfrac{2}{m}\right)\cdots\Gamma\left(1+\tfrac{m-1}{m}\right) = m^{\frac{1}{2}-mx}(2\pi)^{(m-1)/2}\Gamma(mx)$$

Yes, I have to admit I don't find this one so interesting.

4)

$$\Gamma(x+1) = \sqrt{2\pi x}x^x e^{-x}\left\{1+\frac{1}{12x}+\frac{1}{288x^2}-\frac{139}{51840x^3}+\cdots\right\}$$

which is called Stirling's asymptotic series and might look pretty obscure, but in fact it is quite useful in statistical mechanics where we need analytic approximations for the factorial function. There we find that it is adequate to just include the first couple terms inside the curly braces.

5)

$$\Gamma'(1) = \int_0^\infty e^{-x}\ln x\,dx = -\gamma$$

where γ is the Euler–Mascheroni constant, which is also given by

$$\gamma = \lim_{m\to\infty}\left(1+\tfrac{1}{2}+\tfrac{1}{3}+\cdots+\tfrac{1}{M}-\ln M\right) \approx 0.5772156$$

6)

$$\frac{\Gamma'(p+1)}{\Gamma(p+1)} = \tfrac{1}{1}+\tfrac{1}{2}+\tfrac{1}{3}+\cdots+\tfrac{1}{p}-\gamma$$

Yes, a fine set of collectibles from 19th–century mathematics. But the Stirling expansion is useful.

One can also define a version of the gamma function where the integral is only taken over part of the real axis beginning at x. The **incomplete gamma function** thus has two arguments and is given by

$$\gamma(n,x) \equiv \int_x^\infty u^{n-1}e^{-u}\,du$$

1.6.2 The Beta Function

The gamma function is related to the **beta function**, which is defined by

$$B(m,n) \equiv \int_0^1 x^{m-1}(1-x)^{n-1}\,dx \tag{1.27}$$

which does converge for all $m > 0$ and $n > 0$. It is related to the gamma function by

$$B(m,n) = \frac{\Gamma m \Gamma(n)}{\Gamma(m+n)}$$

and it can crop up when we have to evaluate certain integrals. For example,

$$\int_0^{\pi/2} \sin^{2m-1}\theta\,\cos^{2n-1}\theta\,d\theta = \tfrac{1}{2}B(m,n) = \frac{\Gamma m\Gamma(n)}{2\Gamma(m+n)}$$

1.6.3 Error Function

Though you've already come across this one and it has a pretty simple definition, you should be aware that the following function exists and has a name; the **error function** is

$$\mathrm{erf}(x) \equiv \frac{2}{\sqrt{\pi}}\int_0^x e^{-u^2}\,du = 1 - \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-u^2}\,du$$

and the **complementary error function** is

$$\mathrm{erfc}(x) \equiv 1 - \mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-u^2}\,du = \frac{\Gamma(\frac{1}{2}, x^2)}{\sqrt{\pi}}$$

1.6.4 References

Probably the one reference to use nowadays on the special functions is the overhaul of the venerable Abramowitz and Stegun, namely The NIST Handbook of Mathematical Functions, Frank Olver et al, eds., Cambr Univ Press

Being funded by your tax dollars, much or all of the book information has been put on a web site,

`dlmf.nist.gov`

Chapter 2

Tensors

2.1 Introduction

The story of theoretical physics in the 20^{th} century is largely the search for (successful) **symmetries** in the basic equations of physics. A symmetry in an equations mean we make a change in some quantity or a function and we somehow arrive at an equation that has the same form as the original. We insist on equations that have certain symmetries.

Symmetry is also used in physics when we already know the correct equation but we are solving a very complicated problem, such as the dynamics of a molecule. Then one can use the required symmetries of the solutions to simplify the calculations.

The emphasis on symmetries has led to the importance of quantities known as tensors and also of the mathematical theory of groups. So it's my intention for the rest of this class to give an introduction to these mathematical ideas and their use in physics.

I hope it works.

2.2 Change of Coordinates

Physics should not depend on the arbitrary choices made by humans in measuring the phenomena, so the choice of a coordinate system can't affect the answers.

Throughout the following we will be dealing with a change in our description of the coordinates of space from a set of numbers $(x^1, x^2, \ldots, x^N)$ to one given by $(\bar{x}^1, \bar{x}^2, \ldots, \bar{x}^N)$. For both sets we put the indices in the "up" positions. There's a reason for this and though there is a possible confusion with *taking a power* of a certain coordinate, usually the context will determine what is going on.

We will assume from the start that we know there relation between the x^i and $\bar{x}^i$ coordinates and that there is mathematically well-behaved relation between them. Thus to change the coordinates one way, we use the functions

$$\begin{aligned} \bar{x}^1 &= \bar{x}^1(x^1, x^2, \ldots x^N) \\ \bar{x}^2 &= \bar{x}^2(x^1, x^2, \ldots x^N)\,, \qquad \text{etc.} \end{aligned}$$

and we also have the reverse relations,

$$\begin{aligned} x^1 &= x^1(\bar{x}^1, \bar{x}^2, \ldots \bar{x}^N) \\ x^2 &= x^2(\bar{x}^1, \bar{x}^2, \ldots \bar{x}^N)\,, \qquad \text{etc.} \end{aligned}$$

Now it may be the case that *in practice* these relations will be messy and will require some steps to

2.3 First-Order Tensors

Also known as vectors, but here we give a very precise definition of what we mean by a vector. We then then the definitions to general tensors.

A vector (in N spatial dimensions) is a set of N numbers (defined at each point in space) for which we have instructions for calculating their values from the coordinate system we're currently using the describe the points. If we get the numbers A^p in one coordinate system, we will definitely get a different set of numbers $\bar{A}^p$ when we use different coordinates even if the set of A's is associated with the *same* point in space. This is not a great surprise as the Cartesian components of a regular vector are different if we use a rotated set of coordinates.

We say that the set of numbers A^p are the components of a **contravariant vector** if in the new coordinate system the values of $\bar{A}^p$ are given by

$$\bar{A}^p = \sum_{q=1}^{N} \frac{\partial \bar{x}^p}{\partial x^q} A^q = \frac{\partial \bar{x}^p}{\partial x^q} A^q \tag{2.1}$$

where the second equality uses the assumption of the summation convention.

We say that the set of number A^p are the components of a **covariant vector** if in the new coordinate system the values of $\bar{A}^p$ are given by

$$\bar{A}^p = \sum_{q=1}^{N} \frac{\partial x^p}{\partial \bar{x}^q} A^q = \frac{\partial x^p}{\partial \bar{x}^q} A^q \tag{2.2}$$

and again the second equality is from the summation convention which we will use from now on!

2.3.1 Gradient of a Scalar

If ϕ is a scalar field then with we can obtain a *covariant* vector field A_i from

$$A_i \equiv \frac{\partial \phi}{\partial x^i} \tag{2.3}$$

To *prove* that A_i a covariant vector we write down what $\bar{A}_i$ would be if we used the coordinates $\bar{x}^i$:

$$\bar{A}_i = \frac{\partial \bar{\phi}}{\partial \bar{x}^i}$$

Now use the fact that since ϕ is a scalar field, $\bar{\phi} = \phi$ and expand the partial derivative on the rhs using the chain rule for partial derivatives:

$$\bar{A}_i = \frac{\partial x^j}{\partial \bar{x}^i} \frac{\partial \phi}{\partial x^j} = \frac{\partial x^j}{\partial \bar{x}^i} A_j \tag{2.4}$$

Comparing the rhs of 2.4 with Eq. 2.2 we can say the the components (numbers) A_i transform in the proper manner for a covariant vector.

2.4 Second–Order and Higher Tensors

Consider the set of N^2 numbers A^{pr} which, again have some prescription for being calculated when one uses one set of coordinates or another. We say that A^{pr} is a contravariant tensor of the second rank if in the new coordinate system the components are

$$\bar{A}^{pr} = \frac{\partial \bar{x}^p}{\partial x^q}\frac{\partial \bar{x}^r}{\partial x^s}A^{qs} \tag{2.5}$$

and we can generalize the classification to multiple ranks of the contravariant and covariant type. As an example bordering on the silly, the set of numbers A^{qst}_{kl} is contravariant of rank three and covariant two if the elements in the new coordinate system are given by

$$\bar{A}^{prm}_{ij} = \frac{\partial \bar{x}^p}{\partial x^q}\frac{\partial \bar{x}^r}{\partial x^s}\frac{\partial \bar{x}^m}{\partial x^t}\frac{\partial x^k}{\partial \bar{x}^i}\frac{\partial x^l}{\partial \bar{x}^j}A^{qst}_{kl}$$

If the tensor is defined at each point in space then we are discussing a **tensor field**. When working with these we must watch out for the distinction between functions and variables, which is sometimes obscured by physicists. All of the transformation relations refer to the *same point in space* which of course is given by different expressions depending on whether we use the x^i or $\bar{x}^i$ coordinates.

For the special case of Cartesian coordinates related by a rotation, with $\bar{x}^i = L_{ij}x^j$ the transformation equations for a second–order tensor are

$$\bar{T}_{ij} = L_{ik}L_{jl}T_{kl}$$

$$T_{ij} = L_{ki}L_{lj}\bar{T}_{kl}$$

which follows from the fact that L is an orthogonal matrix, with $L^{-1} = L^T$.

2.5 How To Get Tensors

As noted above, if we take the generalized gradient of a scalar field we do get a covariant tensor of the first rank. But that trick doesn't work in general! For example taking the gradient of components of a general vector field:

$$A_{ij} \equiv \frac{\partial A_i}{\partial x^j}$$

does *not* produce a tensor field in general (as we'll show later). Now, for linear transformations such as rotations of axes in fact it *does* give a tensor, but when a transformation to curvilinear coordinates is involved some extra machinery is needed to produce a tensor from such an operation.

One *can* get a tensor by taking the **outer product** of two given tensors. For example, if we are given the (covariant) vector fields u_i and v_j we can form a second-rank (covariant) tensor from

$$T_{ij} \equiv u_i v_j$$

or for a more exotic example, if we are given the tensor fields A^{pr}_q and B^m_s we can get a tensor of fifth rank from

$$C^{prm}_{qs} = A^{pr}_q B^m_s$$

We can also form tensors of *smaller* rank from ones of larger rank. We can perform a **contraction** by setting one covariant and one contravariant index equal and summing over that index. So

if we are given the tensor A^{mpr}_{qs} we can set $r = s$ and sum (no "Σ" is needed as this falls under the Einstein summation convention):

$$B^{mp}_{q} \equiv A^{mpr}_{qr}$$

In a similar way we can get a smaller tensor from multiplication of elements of two tensors as described above for the outer product but then setting a contravariant index of one tensor and a covariant index of the other tensor equal and then summing. (This operation is actually the generalization of the familiar dot product of ordinary 3D vectors.)

For example, given the tensors A^{mp}_{q} and B^{r}_{st}, if we multiply, set $q = r$ and $p = s$ and sum, we get the tensor C^{m}_{t}:

$$C^{m}_{t} = A^{mp}_{r}\, B^{r}_{pt}$$

A frequent operation on an arbitrary tensor is to write it as a sum of symmetric and antisymmetric tensors. For example, given the tensor A_{ij} we can form the symmetric and antisymmetric tensors via

$$\begin{aligned} T_{ij} &= \tfrac{1}{2}(T_{ij} + T_{ji}) + \tfrac{1}{2}(T_{ij} - T_{ji}) \\ &\equiv S_{ij} + A_{ij} \end{aligned}$$

2.5.1 Pseudotensors: More Confusing Than They Ought To Be

Still considering linear transformations of coordinates, we can consider a transformation where the transformed coordinates have a different relation to original ones. Instead of the usual rule $\bar{x}^i = L_{ij}x^j$ we could consider a rule which works like

$$\bar{x}^i = -L_{ij}x^j \tag{2.6}$$

and we don't have to go far to find such a transformation: The inversion of coordinates in 3D given by $\bar{x}^i = -x^i$ (so that $L_{ij} = -\delta_{ij}$) is such a transformation, but it is not equivalent to any rotation. More generally we can incorporate the -1 factor in 2.6 as the determinant of the matrix L and write the rule as

$$\bar{x}^i = |L|\, L_{ij}x^j \tag{2.7}$$

A vector which transforms in this way is called a **pseudovector**.

This topic is of great interest in sorting out the basic laws of physics, and as we'll discuss, most authors will classify angular momentum and the magnetic field as pseudovectors. Alas, our textbook authors RHB seem to be on a misguided religious mission about this and insist that all quantities used in physics equations are *genuine* vectors or scalars and much confusion can come from this.

The basic point of contention is what *prescription* to use for the cross product of two vectors in the new (inverted) coordinates, that is, the meaning of $\bar{\mathbf{a}} = \bar{\mathbf{b}} \times \bar{\mathbf{c}}$. Most authors take the view that we should respect the cyclic order of the coordinate axes and use

$$\bar{a}^i = \epsilon_{ijk}\, \bar{b}^j\, \bar{c}^k$$

which has the curious feature that we now have a *left-hand rule* for the cross product in the inverted coordinates.

RHB insist that the cross product is an inherently *right-handed* operation and gives one definite result regardless of the coordinates used to describe the points of space. But to bring this about

they must now include a minus sign in the "ϵ" formula for the cross product for the $\bar{x}$ system and thus they say the prescription is

$$\bar{a}^i = -\epsilon_{ijk}\,\bar{b}^j\,\bar{c}^k$$

The two prescriptions differ in that most authors say that the components of (say) **B** are *the same* are the are in the unprimed system, which is not the way regular vectors behave... normally, the components change sign. RHB insist that for all of these physical objects the components change sign for the new system, and thus they are all true vectors.

2.6 Physical Applications

Here we list a few places where we encounter *Cartesian* tensors in physics

In classical mechanics, we have reason to consider the total angular momentum of a system of particles. For a set of mass points indexed by (α), it is

$$\mathbf{J} = \sum_\alpha (\mathbf{r}^{(\alpha)} \times \mathbf{p}^{(\alpha)}$$

where for a rotating rigid body

$$\mathbf{p}^{(\alpha)} = m^{(\alpha)}\dot{\mathbf{r}}^{(\alpha)} \quad \text{and} \quad \dot{\mathbf{r}}^{(\alpha)} = \boldsymbol{\omega} \times \mathbf{r}^{(\alpha)}$$

with $\boldsymbol{\omega}$ being the vector which gives the (instantaneous) angular velocity of the object. One can who that this gives

$$\sum_\alpha m^{(\alpha)}[{r^{(\alpha)}}^2\delta^{il} - x^{(\alpha)\,i}x^{(\alpha)\,l}]\omega^l \equiv I^{il}\omega^l$$

For a continuum mass distribution, replace the sum by an integral with a mass density factor. We can write I as a matrix as:

$$I = \begin{pmatrix} \int(y^2+z^2)\rho\,dV & -\int xy\rho\,dV & -\int xz\rho\,dV \\ -\int xy\rho\,dV & \int(x^2+z^2)\rho\,dV & -\int yz\rho\,dV \\ -\int xz\rho\,dV & -\int yz\rho\,dV & \int(x^2+y^2)\rho\,dV \end{pmatrix}$$

One can show that the kinetic energy of the rotating object is

$$T = \tfrac{1}{2}I^{il}\omega^j\omega^l$$

We also find that a particular choice of axes attached to the rigid body make the motion especially simple. These are the **principal axes** of the rigid body and in fact are the eigenvectors of the moment of inertial matrix.

The magnetic susceptibility of a substance (the proportionality factor between the H field and the magnetization M) is in general a *tensor*:

$$M^i = \chi^{ij}H^i$$

which covers the materials where **M** is not parallel to the field **H**.

Likewise, although the current density **J** is usually parallel to the electric field **E** in general it may differ in direction (while still being proportioinal to **E** so that we have a conductivity *tensor*:

$$J^i = \sigma^{ij}E^j$$

We also have to electromagnetic stress–energy tensor, which gives the momentum and energy content of the electromagnetic field:

$$T^{ij} = \epsilon_0(E^iE^j - \tfrac{1}{2}\delta^{ij}E^2) + \frac{1}{\mu_0}(B^iB^j - \delta^{ij}B^2)$$

2.7 Non–Cartesian Coordinates

We now focus strictly on coordinate systems which are *not* related to the standard Cartesian set by a rotation. They are *curved* with respect to the Cartesian set, the most popular examples being cylindrical and spherical polar coordinates, but we will let be as general as possible, as long as the points is space are given by *some* function of (say) three coordinates, (u^1, u^2, u^3).

We will follow the treatment of RHB, who want to work with basis vectors, much in the manner of our vector calculus material. While some books (especially those on general relativity) also do this, many books of a more mathematical bent, focus on the coordinates alone. In the end we will get equations which *do* deal only with the coordinates but along the way we may get some insight using the vectors, thus I will go with the RHB approach.

First, we assume that the displacements from the origin (vectors!) can be written in terms of the curvilinear coordinates by some relation given by

$$\mathbf{r} = \mathbf{r}(u^1, u^2, u^3)$$

We first define the basis vectors $\mathbf{e}_i$ and $\boldsymbol{\epsilon}_i$ by

$$\mathbf{e}_i \equiv \frac{\partial \mathbf{r}}{\partial u^i} \qquad \boldsymbol{\epsilon}_i \equiv \nabla u^i \qquad \text{for } i = 1, 2, 3$$

One can show that while each set *in itself* is not orthogonal, the sets have a *mutual* orthonormality property:

$$\mathbf{e}_i \cdot \boldsymbol{\epsilon}_j = \delta_{ij}$$

And now, from on we will always pay attention to whether an index is up or down, and we will redefine

$$\mathbf{e}^j \equiv \boldsymbol{\epsilon}_j \qquad \text{so that} \quad \mathbf{e}_i \cdot \mathbf{e}^j = \delta_i^j$$

and when we use the summation convention we will always be summing over an index pair where one of them is up and the other is down.

Any vector (which RHB want us to think of as a real, physical arrow) can be expanded in terms of either of these sets of vectors, and accordingly the components are written with indices in the up or down position opposite that of the basis vector it goes with. Thus for the vector $\mathbf{a}$,

$$\mathbf{a} = a^1\mathbf{e}_1 + a^2\mathbf{e}_2 + a^3\mathbf{e}_3$$

and

$$\mathbf{a} = a_1\mathbf{e}^2 + a_2\mathbf{e}^2 + a_3\mathbf{e}^3$$

and we call the a^i the **contravariant** components of the vector $\mathbf{a}$ and the a_i the **covariant** components of $\mathbf{a}$.

We can extend the basis-vector approach to give tensor of higher rank by their components that go with out products of the basis vectors, $\mathbf{e}_i \otimes \mathbf{e}_j$ (or products with the $\mathbf{e}^i$'s:

$$\mathbf{T} = T^{ij}\,\mathbf{e}_i \otimes \mathbf{e}_j = T^i_j \mathbf{e}_i \otimes \mathbf{e}^j = T_{ij}\mathbf{e}^i \otimes \mathbf{e}^j$$

2.7.1 The Metric Tensor

For a particular coordinate system, we define the numbers g_{ij} by:

$$g_{ij} \equiv \mathbf{e}_i \cdot \mathbf{e}_j$$

Now, an infinitesimal space displacement can be expressed in terms of the corrdinates and basis vectors as

$$d\mathbf{r} = du^i \mathbf{e}_i$$

and one can show that the squared length of this vector is

$$(ds)^{@} = d\mathbf{r} \cdot d\mathbf{r} = du^i \mathbf{e}_i \cdot du^i \mathbf{e}_j = g_{ij} du^i \, du^j$$

One can also show that the volume element resulting from set of increments in the three coordinates is

$$dV = \sqrt{g} \, du^1 \, du^2 \, du^3$$

where g is the determinant of the matrix given by g_{ij}.

As one might expect, the g_{ij} are related to the scale factors h_i that arise when the standard orthogonal coordinate systems are discussed; in these systems,

$$\mathbf{e}_i \cdot \mathbf{e}_j = 0 \quad \text{for} \quad i \neq j \, , \quad \text{and}$$

$$g_{ij} = \begin{cases} h_i^2 & i = j \\ 0 & i \neq j \end{cases} \qquad \text{giving} \quad g = h_1^2 h_2^2 h_3^2$$

It can be shown that the scalar product is

$$\mathbf{a} \cdot \mathbf{b} = g_{ij} a^i b^j = g^{ij} a_i b_j = a_i b^i = a^j b_j$$

from which it follows that the contravariant and covariant can be gotten from one another by means of the metric tensor, which is then thought of as a raising or lowering operator for the indices:

$$g_{ij} b^j = b_i \qquad g^{ij} b_j = b^i \qquad \text{and} \quad g^{pq} g_{rq} = \delta_r^p$$

The basis vectors themselves are mutually related in this way:

$$\mathbf{e}^i = g^{ij} \mathbf{e}_j \qquad \mathbf{e}_i = g_{ij} \mathbf{e}^j$$

One can also show that the sets of 3 basis vectors are related by

$$\mathbf{e}^i = \frac{\mathbf{e}_j \times \mathbf{e}_k}{\mathbf{e}_i \cdot (\mathbf{e}_j \times \mathbf{e}_k)}$$

for combinations $1, 2, 3$ of the subscripts and cyclic permutations thereof. Using terminology that arises in solid state physics, we would say that the $\mathbf{e}^i$ and $\mathbf{e}_i$ are **reciprocal** sets of vectors. We also have

$$\mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{e}_3) = \sqrt{g}$$

2.8 Derivatives of Tensors

We have a big job ahead of us in thinking about taking derivatives of tensors in the curvilinear coordinates we are now considering. First off, we can take $\partial / \partial u^i$ or combinations thereof on objects (i.e. tensors) with possibly a multitude of components... of two types! But to make physical sense of the results we also insist that we take derivatives of tensors in such a way that we *get another tensor as a result.*

Consider the basis vectors $\mathbf{e}_i$. As we change one coordinate, the basis vector will change, as we are well aware from our study of curvilinear coordinates in vector calculus. So the derivative is non-zero, but since the basis vectors span the space, the derivative can be expressed as a linear combination of the basis vectors. The coefficients are how we will define the **Christoffel symbol of the second type**, Γ^k_{ij}:

$$\frac{\partial \mathbf{e}_i}{\partial u^j} \equiv \Gamma^k_{ij}\mathbf{e}_k \tag{2.8}$$

that is to say, Γ^k_{ij} is the k^{th} component of $\frac{\partial \mathbf{e}_i}{\partial u^j}$. Or using the ortogonality properties,

$$\Gamma^k_{ij} = \mathbf{e}^k \cdot \frac{\partial \mathbf{e}_i}{\partial u^j}$$

One can show that for the derivative of a contravariant basis vector we have

$$\frac{\partial \mathbf{e}^i}{\partial u^j} = -\Gamma^i_{kj}\mathbf{e}^k$$

Despite the notation, the Christoffel symbols Γ^k_{ij} don't form the components of a tensor themselves, and they have their own formula for how to transform them (later). One can derive an explicit formula for them in terms of the matric tensor, and that is some books begin the discussion of them. We have:

$$\Gamma^m_{ij} = \tfrac{1}{2}g^{mk}\left(\frac{\partial g_{jk}}{\partial u^i} + \frac{\partial g_{ki}}{\partial u^j} - \frac{\partial g_{ij}}{\partial u^k}\right) \tag{2.9}$$

For completeness we note that many books also give the **Christoffel symbol of the first kind** as

$$[pq, r] = \tfrac{1}{2}\left(\frac{\partial g_{pr}}{\partial u^q} + \frac{\partial g_{qr}}{\partial u^p} - \frac{\partial g_{pq}}{\partial u^r}\right) \tag{2.10}$$

with the symbol of the second kind gotten from

$$\Gamma^s_{pq} \equiv \left\{ \begin{array}{c} s \\ pq \end{array} \right\} = g^{sr}[pq, r] \tag{2.11}$$

In this expression, the curly brace notation for the Christoffel symbol of the second kind is given; many books prefer this notation since as we noted, Γ is not a tensor.

2.8.1 Covariant Differentiation

While the partial derivatives of the individual components of a vector $\mathbf{v}$ do *not* give a tensor, the partial derivatives of the *vector itself* does. Since $\mathbf{v} = v^i\mathbf{e}_i$, the product rule gives us

$$\frac{\partial \mathbf{v}}{\partial u^j} = \frac{\partial v^i}{\partial u^j}\mathbf{e}_i + v^i\frac{\partial \mathbf{e}_i}{\partial u^j}$$

and we now know how the take partial derivatives of basis vectors, using the Christoffel symbols! Thus we have (using some relabeling of the indices)

$$\frac{\partial \mathbf{v}}{\partial u^j} = \frac{\partial v^i}{\partial u^j}\mathbf{e}_i + v^i\Gamma^k_{ij}\mathbf{e}_k = \left(\frac{\partial v^i}{\partial u^j} + v^k\Gamma^i_{kj}\right)\mathbf{e}_i$$

The quantity in the parentheses of the last expression is known as the **covariant derivative** of the vector components v^i. To distinguish from the normal partial derivative with respect to u^j (which we might denote as $v^i{}_{,j}$) we use a semicolon and write

$$v^i{}_{;j} \equiv \frac{\partial v^i}{\partial u^j} + \Gamma^i_{kj}v^k \tag{2.12}$$

The distinction is not necessary in rectangular coordinates as in that case the Christoffel symbols are all zero.

For the covariant derivative of a covariant vector we have

$$v_{i\,;\,j} = \frac{\partial v_i}{\partial u^j} - \Gamma^k_{ij} v_k$$

The Christoffel symbols *can* be transformed to other coordinates, but we have to use the relation

$$\bar{\Gamma}^i_{jk} = \Gamma^r_{st} \frac{\partial \bar{x}^i}{\partial x^r} \frac{\partial x^s}{\partial \bar{x}^j} \frac{\partial x^t}{\partial \bar{x}^k} + \frac{\partial^2 x^r}{\partial \bar{x}^j \, \partial \bar{x}^k} \frac{\partial \bar{x}^i}{\partial x^r}$$

2.8.2 Our Old Vector Operators in Tensor Form

2.8.3 The Absolute Derivative

Consider the space curve $\mathbf{r}(t)$ (set of points parametrized by t). We consider the derivative of the vector $\mathbf{v}$ along the curve. Again, this not the same as the set of (total) derivatives of the components of $\mathbf{v}$ because the basis vectors $\mathbf{e}_i$ also change with t.

One can show:

$$\frac{d\mathbf{v}}{dt} = \left(\frac{dv^i}{dt} + \Gamma^i_{jk} v^j \frac{du^k}{dt} \right) \mathbf{e}_i$$

The term in the parenthesis is known as the **absolute derivative** of the vector field $\mathbf{v}$. In the notation of RHB, it is

$$\frac{\delta v^i}{\delta t} = \frac{dv^i}{dt} + \Gamma^i_{jk} \frac{du^k}{dt} = v^i_{\,;\,k} \frac{du^k}{dt} \tag{2.13}$$

2.8.4 Physical Components of a Vector

The contravariant vector components v^i we've been discussing for curvilinear coordinates, are *associated* with the different coordinates of these systems (for example (r, θ, ϕ) for spherical coordinates) are *not the same* as the vector components as used in physics and in conventional vector analysis! For comparison of our formula with those which we knew before, we must cast them in terms of the **physical components** of these vectors.

Consider an orthogonal coordinates system with metric components g_{ii}, with $i = 1, 2, 3$ which go with the coordinate names u, v, w. The familiar (physical) components of the vector A^i or A_i are given by

$$A_u = \sqrt{g_{11}}\, A^1 \qquad A_v = \sqrt{g_{22}}\, A^2 \qquad A_w = \sqrt{g_{33}}\, A^3$$

2.9 The Equations of General Relativity

While general relativity is probably the most common reason why any physicist would want to learn about tensor analysis, the extension of our ideas to include a time coordinate (along with a relative minus sign in the metric tensor) as well as the development of other tensor tools that one would need is probably not the best use of our time. Nevertheless, as long as we have come this far, it might be worthwhile to show what the central equation of general relativity *looks like*; as we already have the Christoffel symbols, it isn't all that much further.

First off, one needs to construct the **Riemann tensor**, which is a further reflection of the curvature properties of the coordinates in question. It is given by:

$$R^\sigma_{\mu\nu\lambda} \equiv \Gamma^\sigma_{\mu\lambda,\,\nu} - \Gamma^\sigma_{\mu\nu,\,\lambda} - \Gamma^\sigma_{\alpha\nu}\Gamma^\alpha_{\mu\lambda} - \Gamma^\sigma_{\alpha\lambda}\Gamma^\alpha_{\mu\nu} \tag{2.14}$$

Next we make the relatively easy definition of the **Ricci tensor**:

$$R_{\mu\nu} \equiv R^{\sigma}_{\mu\sigma\nu} \tag{2.15}$$

The **Einstein tensor** is defined by

$$G^{\mu\nu} \equiv R^{\mu\nu} - \tfrac{1}{2} g^{\mu\nu} R \tag{2.16}$$

With $G^{\mu\nu}$ we have characterized the curvature aspect of space-time. Space-time is told *how* to curve from the local properties of matter. For a perfect fluid, $T_{\mu\nu}$ takes the following form:

$$T^{\mu\nu} = (\varepsilon + p)\eta^{\mu}\eta^{\nu} - p g^{\mu\nu} \tag{2.17}$$

where ε is the mass–energy density, p is the pressure and η is the four-velocity of the matter in the notation of Griffiths' EM book:

$$\eta^{\mu} = \frac{dx^{\mu}}{d\tau}$$

With these definitions, Einstein's equations are

$$G_{\mu\nu} = \left(\frac{8\pi G}{c^4} \right) T_{\mu\nu} \tag{2.18}$$

where the G inside the parenthesis is Newton's gravitational constant. This is an unfortunate overlap in symbols but in fact most books on general relativity just units where $G = c = 1$ and then the Einstein equation is the famous

$$G = 8\pi T \tag{2.19}$$

Chapter 3

Group Theory

3.1 Introduction

Symmetry of a verby basic sort is commonly used in elementary mathematical operations, such as when we recognize an odd function to see that its integral over $(-\infty, \infty)$ vanishes:

$$f(-x) = -f(x) \qquad \Longrightarrow \qquad \int_{-\infty}^{\infty} f(x)\, dx = 0$$

3.2 Definition of Group

A group $\mathcal{G}$ is a set of elements $\{X, Y, \ldots Z\}$ with a rule for combining them to give another element, written as

$$X \cdot Y = Z$$

which satisfies the following basic properties:

i) If $X \cdot Y = Z$ then Z also belongs to $\mathcal{G}$.

ii) For all triples X, Y, Z, the associative law holds:

$$X \cdot (Y \cdot Z) = (X \cdot Y) \cdot Z$$

iii) There is a unique element I known as the **identity element** in $\mathcal{G}$ with the property

$$I \cdot X = X \cdot I = X$$

iv) For every element X of $\mathcal{G}$ there exists an element X^{-1} in $\mathcal{G}$ such that

$$X^{-1} \cdot X = X \cdot X^{-1} = I$$

X^{-1} is called the **inverse** of X

Simple examples of groups are:

1) All integers under regular addition. Note, this group has an infinite of elements. But that's OK.

2) The numbers 1 and -1 under multiplication. This is a very small group!

3) The rational numbers except for 0 under regular multiplication. Or, all numbers except for 0 under multiplication.

4) The symmetry operations that we want to apply to physical systems. There, the identity element (operation) is the act of doing nothing!

3.2.1 Notation, Nomenclature and Conventions

When we combine symmetry operations X and Y on a system and we write the combination as $Z = X \cdot Y$ we will always mean that we perform the operation Y first and then perform X. (This will be consistent with our later usage of matrices to carry out the symmetry operations.) The order will be important because in general we *don't* have $X \cdot Y = Y \cdot X$ for a comination of two symmetry operations.

3.3 Basic Theorems

3.4 Classes