



Next: [About this document ...](#)

Physics 231

Mathematical Physics and Classical Electrodynamics Part II

Spring Semester, 1997

Robert G. Brown, Instructor

Duke University Physics Department^{[1](#)}

Physics 231 Mathematical Physics and Classical Electrodynamics Part II

Instructor: Robert G. Brown Room 206

Texts: Arfken, Jackson (and Wyld, optional). Phone: 660-2567

Course Description In this course we will cover the following basic topics:

- ODE's (starting at 2nd order - assuming that 1st order was completed last semester).
- Curvilinear coordinates and tensors.
- Orthogonal functions and the idea of functional analysis.
- PDE's of physics: Laplace, Poisson, Helmholtz, Schrödinger, Wave (and maybe even Diffusion/Heat) equations, mostly in three dimensions and spherical or cylindrical coordinates.
- Focus on PDE's of Electrodynamics: Laplace, Poisson, Helmholtz equations, Maxwell's equations, boundary value problems of electro- and magnetostatics.
- Green's Function Methods! Integral Equation Solutions to PDE's of Physics! Multipoles!

There will be, as you may have guessed, *lots* of problems. Actually, there won't as many as last year, but they'll seem like a lot as they'll be considerably harder. The structure and organization of the course will be (approximately!):

30% of grade Homework.
 30% Take-Home Midterm Exam.
 30% Take-Home Final Exam.
 10% Research/Computing project.

In more detail, Homework is Homework, the Exams are Homework (fancied up a bit and with more stringent rules) and the Research Project is described below. These figures are only approximate. I may make homework worth a little more or less, but this is about right.

The way my grading scheme typically works is that if you get below a 50 and have not religiously done (well!) and handed in your homework, you fail (U). If you get less than a 60 and have not religiously handed in your homework, you get an (S). If you get 60 or more you get a G or E of some sort and ``pass". If you have religiously done your homework, but have somehow managed to end up less than a 60 or (worse) 50 you may make sorrowful and wounded noises and perhaps get a G- or S, respectively. If you have not done and handed in your homework on time or have not followed the rules with respect to your homework, don't bother me about your grade - it will likely be bad and you will deserve it. Note that if you get as little as 80 per cent of your homework credit, you will only need 40 percent of your exam credit to get at least a G-.

The Rules

The **RULES** are very serious. In previous years, certain students have reportedly betrayed the trust inherent in the rules. This has led to calls that the rules for this course (and all graduate courses here) be stringently tightened. I would prefer NOT to see this happen, as I think that the rules optimize the learning process as the stand and minimize Mickey Mouse interactions as well, but IF I get any hint of misbehavior (verified or not) I'll tighten things up very, very quickly and we'll all have to work harder and be more frustrated while working. So, PLEASE! Follow the spirit as well as the letter of the rules. You are here to learn your chosen profession, and it has never been truer that choosing an easy path is ultimately cheating yourself. Besides, it will show up on prelims!

Rules:

- You **may** collaborate with your classmates in any permutation on the **homework**. In fact, I **encourage** you to work in groups, as you will probably all learn more that way. However, you must **each** work out and individually write up *all* the solutions even if your ``figure them out" (and they are thus all nearly the same) within a group. Working them out and writing them up from scratch without looking (once you get the idea of the solution) provides invaluable learning reinforcement. Unless the assignment is a computer or mathematica project, I want **handwritten** solutions **neatly** done on **white (laserprinter) paper**.
- You may **not** get worked out solutions to specific problems from more advanced graduate students, the (any) solution manual (if you can find it) or anyplace else. It obviously removes the whole point of the homework in the first place.
- You **may** ask me, your grader, more advanced students, other faculty, personal friends, or your household pets for help or tutoring on particular problems, as long as no worked-out solutions to the assigned problems are present.
- You may (indeed *must*) use the library and all available non-human resources to help solve the problems. I don't even care if you find the solution somewhere and **copy it verbatim** provided that you understand it afterwards (which is the goal), **cite your source**, and provided that you do **not** use the solution manual for Jackson problems (which exists, floating around somewhere), or Arfken (which may or may not exist) see second item above.
- You may **NOT** collaborate with each other or get outside human help on the take home exam problems. They are to be done alone. There will be a time limit (typically 24 hours total working time) on the take home exams, spread out over

four days or so.

- You **may** still use the library and non-human resources on the take home exam problems, provided that (as usual) you don't look up the answer to e.g. - a Jackson problem in a Jackson solution set of one sort or another. Since solutions to difficult problems may well be in the literature, this is an invaluable aid! Last semester, the answer to 6.19c (for example) was the verbatim contents of a paper by Brill and Goodman (as noted in both text and the lectures at the appropriate time). I recommend immediately obtaining a copy of any paper I refer to in lecture, as it may well figure prominently in an exam question...

Research Project: I'm offering you instead the following assignment, with several choices. You may prepare any of:

1. A set of lecture notes on a topic, relevant to the material we will cover, that interests you. If you select this option you may be asked to present the lecture for that topic in my place, time permitting.
2. A review paper on a mathematical physics topic, relevant to the material we will cover (or not cover!), that interests you. Typically, in the past, students going into (e.g.) FEL have prepared review papers on the electromechanism of the FEL. That is, relevance to your future research is indicated but not mandated.
3. A computer demonstration or simulation of some important electrodynamical principle or system. Possible projects here include solving the Poisson and inhomogeneous Helmholtz equation numerically, evaluating and plotting radiation patterns and cross-sections for complicated but interesting time dependent charge density distributions, etc. Resources here include Mathematica, maple, SuperMongo, Numerical Recipes, and more. Obviously now is not the time to learn to program; presumably you are all competent in f77 or c or both. Better late than never, if not, and this is a possible learning project.

If you choose to do a project, it is due **TWO WEEKS** before the last class² so don't blow them off until the end. It is strongly recommended that you clear the topic with me, too.

I will grade you on: doing a decent job (good algebra), picking an interesting topic (somewhat subjective, but I can't help it and that's why I want to talk to you about it ahead of time), adequate preparation (enough algebra), adequate documentation (where did you find the algebra), organization, and Visual Aids (pictures or interactive demos are sometimes worth a thousand equations). Those of you who do numerical calculations (applying the algebra) must also write it up and (ideally) submit some nifty graphics, if possible.

I'm not going to grade you particularly brutally on this -- it is supposed to be fun as well as educational. However, if you do a miserable job on the project, it doesn't count. If you do a decent job (evidence of more than 20 hours of work) you get your ten percent of your total grade (which works out to maybe a third-of-a-grade credit and may be promoted from, say, a G+ to a E-).

I will usually be available for questions after class. It is best to make appointments to see me via e-mail. My third department job is managing the computer network (teaching this is my second and doing research is my first) so I'm usually on the computer and always insanely busy. However, I will nearly *always* try to answer questions if/when you catch me. That doesn't mean that I will know the answers, of course ...

Our grader is Matt Sexton, Room 203A, 660-2566, and is also generally available for help in the coursework. He has tentatively set "office hours" at 2:30 to 3:30 Tuesday afternoons (plus runover allowance). He will answer questions too, when He can. Otherwise He'll look busy and say that He needs to think about it and come bug me. I, in turn, will look puzzled and say I'll think about it and spend all night in the library trying to figure it out so I can tell her and she can tell you. I find (e.g.) Jackson problems just as hard as you do -- and I've done a bunch of them a bunch of times.

I welcome feedback and suggestions at any time during the year. I would prefer to hear constructive suggestions early so that I have time to implement them this semester.

I will TRY to put a complete set of lecture notes, printed out like this up on the Web in both PS and html/gif form. Exemplary problems from other sources may also be included.

Lecture 1

General, Nth Order Linear Homogeneous ODE's

$$a_N(x) \frac{d^N y}{dx^N} + a_{N-1}(x) \frac{d^{N-1} y}{dx^{N-1}} + \dots + a_0(x) y = 0 \quad (1)$$

If we divide out the $a_N(x) \rightarrow A_k(x) = \frac{a_k(x)}{a_N(x)}$, this becomes:

$$\frac{d^N y}{dx^N} + A_{N-1}(x) \frac{d^{N-1} y}{dx^{N-1}} + \dots + A_0(x) y = 0 \quad (2)$$

(making the leading order linearity obvious). Let x be complex - most general abelian algebra. Then a point x_0 is:

- **ordinary** if $A_k(x)$'s are all analytic at $x = x_0$.
- **singular** if not:
 1. **regular singular** if $(x - x_0)^{N-k} A_k(x)$ are all analytic, so that $A_{N-m}(x)$ has no more than an m th order pole.
 2. **irregular point** or **essential singularity** (otherwise).

Fuch's Theorem

- At an ordinary point x_0 there are N linearly independent solutions.
- The Taylor series for these, $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ converges at least up to the nearest singular point.
- At a **regular singular point**, there is at least one solution of the form:

$$(x - x_0)^k \sum_{n=0}^{\infty} a_n(x - x_0)^n \quad (3)$$

where k can be any complex number. k is called the **indicial exponent**. This expansion, too, converges at least to the nearest singular point.

Example: Bessel (Differential) Equation

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + (x^2 - m^2) y = 0 \quad (4)$$

or

$$\frac{d^2 y}{dx^2} + \frac{1}{x} \frac{dy}{dx} + \left(1 - \frac{m^2}{x^2}\right) y = 0 \quad (5)$$

so (obviously) $x = 0$ is a *regular singular point*.

Solutions: Frobenius Method

...follows directly from Fuch's Theorem:

1. If you want to expand around $x_0 \neq 0$, change variables so that you are expanding around 0: $x' = x - x_0$ or (if $x_0 = \infty$) $x' = 1/x$. This set of coordinates defines the **radius of convergence** of the solution derived.
2. Decide which case/kind of point x_0 is:
 - Ordinary: Use $y(x) = \sum_n a_n x^n$.
 - Regular singular: Use $y(x) = x^k \sum_n a_n x^n$ with $a_0 \neq 0$.
 - Irregular singular, essential: Use Famous Mathematician Method (i.e. - look it up if you can find it) or Give Up.
3. Substitute into ODE, also expanding any $A_k(x)$ coefficients (e.g. - $\sin x \rightarrow x - \frac{x^3}{3!} + \dots$. Match up coefficients of each power of x .
 - Take especial care with the first N coefficients (usually 2, for our purposes).
 - After that, write a recurrence relation and hence the whole solution. Make pithy observations, if any exist to be made.
 - For the regular singular point case, the lowest remaining power of x determines k via the **indicial equation**.

Examples

1. Airy's Equation

$$\frac{d^2 y}{dx^2} - xy = 0 \quad (6)$$

Note that $x = 0$ is ordinary, no singularities but $x \rightarrow \infty$.

$$y(x) = \sum_{n=0}^{\infty} a_n x^n \quad (7)$$

$$\frac{d^1 y}{dx^1}(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} \quad (8)$$

$$\frac{d^2 y}{dx^2}(x) = \sum_{n=2}^{\infty} n(n-1) a_n x^{n-2} \quad (9)$$

or

$$\left(\sum_{n=2}^{\infty} n(n-1) a_n x^{n-2} \right) = \left(\sum_{n=0}^{\infty} a_n x^{n+1} \right) = 0 \quad (10)$$

Examine the coefficients of:

- $x^0: (2) \times (1)a_2 = 0$ or $a_2 = 0$.
- $x^1: (3) \times (2)a_3 - a_0 = 0$
- $x^{m+1}: (m+3) \times (m+2)a_{m+3} - a_m = 0$

or

$$\frac{a_{m+3}}{a_m} = \frac{1}{(m+3)(m+2)} \quad (11)$$

We can now reconstruct the entire solution:

$$y(x) = a_0 \left(1 + \frac{x^3}{6} + \frac{x^6}{360} + \dots \right) + a_1 \left(x + \frac{x^4}{12} + \frac{x^7}{504} + \dots \right) + (a_2 = 0 \dots) \quad (12)$$

Remarks:

- **Two** linearly independent solutions with coefficients a_0 and a_1 .
- Converges everywhere (ratio test for x^3 or Fuch's Theorem).
- $Ai(x)$ and $Bi(x)$ (Airy functions/Airy integrals) correspond to two particular choices for a_0, a_1 . These are not too important in E&M, but are important in quantum theory.

2. Legendre Equation:

$$(1-x^2)\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + \ell(\ell+1)y = 0 \quad (13)$$

$x = 0$ is ordinary. $x = \pm 1$ are regular singular points. Start with $x = 0$. Using same expansions for $y(x)$ as above:

$$(1-x^2) \sum_{n=2}^{\infty} n(n-1)a_n x^{n-2} - 2x \sum_{n=1}^{\infty} n a_n x^{n-1} + \ell(\ell+1) \sum_{n=0}^{\infty} a_n x^n = 0 \quad (14)$$

With a little work (and treating the first two carefully) we find that

=

$$\begin{aligned} \frac{a_{n+2}}{a_n} &= \frac{-\ell^2 - \ell + n^2 + n}{(n+2)(n+1)} \\ &= -\frac{(\ell-n)(\ell+n+1)}{(n+1)(n+2)} \end{aligned} \quad (15)$$

or

$$\begin{aligned} y(x) &= a_0 \left(1 - \ell(\ell+1) \frac{x^2}{2!} + \ell(\ell+1)(\ell-2)(\ell+3) \frac{x^4}{4!} + \dots \right) \\ &+ a_1 \left(x - (\ell-1)(\ell+2) \frac{x^3}{3!} + (\ell-1)(\ell+2)(\ell-3)(\ell+4) \frac{x^5}{5!} + \dots \right) \end{aligned} \quad (16)$$

Remarks:

The solution is supposed to exist on the interval $[-1, 1]$ (right up to the nearest singular points). If one examines the limit of the ratio of two successive terms (the ratio test) one finds that:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{a_{n+2} x^{n+2}}{a_n x^n} &= \frac{(-\ell^2 - \ell + n^2 + n) x^2}{(n+2)(n+1)} \\ &= \lim_{n \rightarrow \infty} \frac{a_{n+2} x^{n+2}}{a_n x^n} - \frac{(\ell-n)(\ell+n+1) x^2}{(n+1)(n+2)} \\ &= x^2 \end{aligned} \quad (17)$$

This indeed does converge for $|x| < 1$, but the end point behavior is ambiguous.

If ℓ is an integer, **one** of the two independent series terminates, giving a finite polynomial solution that is defined at the end points. The other (as it turns out, see e.g. - Courant and Hilbert for discussion) diverges and must be rejected. We can tabulate these solutions for each ℓ (including some optional normalization so that $P_\ell(x) = 1$):

$$1P_0(x) = 1 \quad (18)$$

$$1P_1(x) = x \quad (19)$$

$$-2P_2(x) = 1 - 3x^2 \quad (20)$$

$$-\frac{2}{3}P_3(x) = x - \frac{5}{3}x^3 \quad (21)$$

$$(22)$$

...

or in more familiar form we have the **Legendre Polynomials**:

$$P_0(x) = 1 \quad (23)$$

$$P_1(x) = x \quad (24)$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1) \quad (25)$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x) \quad (26)$$

$$\dots \quad (27)$$

Another solution exists around the $x = \pm 1$ points with a radius of converge of 1, $\lim \left| \frac{a_{n+2}}{a_n} \right| = 1$. We'll look at this later, maybe.

3. Bessel's Equation (with $m = 0$):

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + x^2 y = 0 \quad (28)$$

$x = 0$ is a regular singular point, so we try:

$$y(x) = x^k \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n x^{n+k} \quad (29)$$

$$y'(x) = x^k \sum_{n=0}^{\infty} n a_n x^{n-1} = \sum_{n=0}^{\infty} n a_n x^{n+k-1} \quad (30)$$

$$y''(x) = x^k \sum_{n=0}^{\infty} n(n-1) a_n x^{n-2} = \sum_{n=0}^{\infty} n(n-1) a_n x^{n+k-2} \quad (31)$$

Substituting these into the ODE we obtain:

$$\sum_{n=0}^{\infty} \left\{ (n+k)(n+k-1) a_n x^{n+k} + (n+k) a_n x^{n+k} \right\} + \sum_{n=0}^{\infty} a_n x^{n+k+2} = 0 \quad (32)$$

The lowest power of x in these sums is x^k (or $n = 0$):

$$x^k : \quad k(k-1)a_0 + k a_0 = 0 \rightarrow k^2 a_0 = 0 \rightarrow (k = 0, a_0 \neq 0) \quad (33)$$

where the latter follows because we insist that $a_0 \neq 0$ (or we would be starting the series somewhere else). With $k = 0$, we look at the power of $x^{k+1} = x$:

$$x^{k+1} : \quad [(k+1)k + k] a_1 = 0 \rightarrow a_1 = 0. \quad (34)$$

Finally, for general $n, k = 0$:

$$x^n : \quad [n(n-1) + n] a_n + a_{n-2} = 0 \quad (35)$$

which is the recursion relation:

$$\frac{a_n}{a_{n-2}} = -\frac{1}{n^2} \quad (36)$$

and we see that all odd x terms vanish.

Reconstructing the series is now easy:

$$y(x) = a_0 \left(1 - \frac{x^2}{2^2} + \frac{x^4}{2^2 4^2} - \frac{x^6}{2^2 4^2 6^2} + \dots \right) \quad (37)$$

and we identify the part in parentheses with:

$$J_0(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(n!)^2} \left(\frac{x}{2} \right)^{2n} \quad (38)$$

We will shortly solve the general case for $m \neq 0$. Or, I should say, you will for homework. The general theory of solutions around a regular singular point suggests that we should try a second solution of the form:

$$N_0 = J_0(x) \ln(x) + \sum_{n=0}^{\infty} a_n x^n \quad (39)$$

(which has a logarithmic singularity at $x = 0$!) and substitute to try to solve for the a_n or proceed with the Wronskian method or proceed according to Wyld, Section 4.3. Eventually we'll probably do one of these things, but if we don't you can always look it up.

Orthogonal Functions, Representations

Suppose we are given (by God, if you like) an interval (a, b) , a "weight function" (or density function) $w(x) > 0$, and a set of reasonably smooth *linearly independent* functions $u_n(x)$ defined on the entire interval. These functions are **orthogonal** if:

$$\langle u_m | u_n \rangle = \int_a^b u_m^*(x) u_n(x) w(x) dx = 0 \quad (m \neq n) \quad (40)$$

Let:

$$\langle u_n | u_n \rangle = \int_a^b u_n^*(x) u_n(x) w(x) dx = N_n > 0 \quad (41)$$

then

$$\langle u_n | u_n \rangle = N_n \delta_{nn} \quad (42)$$

and we can *always* construct an **orthonormal** set:

$$\phi_n(x) = \frac{1}{\sqrt{N_n}} u_n(x) \quad (43)$$

such that

$$\langle \phi_n | \phi_n \rangle = \delta_{nn} \quad (44)$$

We could also (if we wished) absorb $w(x)$ into the orthogonal representation, but it is not always useful to do so. In fact, it is not always useful to normalize to unity - sometimes we will use an n -dependent normalization to help us cancel or improve a factor that appears elsewhere in our algebraic travails.

Note that orthogonal (or orthonormal) functions are *very useful in physics!* They are the basis of functional analysis, both in quantum mechanics and in the DE's of the rest of physics as well. Hence our formalization of the process:

Notation: Let $|f\rangle = f(x)$, $|\phi_n\rangle = \phi_n(x)$, $\langle \phi_n| = \phi_n^*(x)$, etc. Then we can notationally formalize the relationship between **functional orthogonality** and the underlying **vector space** with its suitably defined **norm**.

That is, suppose $f(x)$ is a piecewise continuous function on (a, b) and suppose further that the $u_n(x) \rightarrow \phi_n(x)$ are orthogonal and **complete**. Then:

$$f(x) = |f\rangle = \sum_{n=0}^{\infty} a_n |\phi_n\rangle \quad (45)$$

where

$$a_n = \langle \phi_n | f \rangle = \int_a^b \phi_n^*(x) f(x) w(x) dx \quad (46)$$

so that

$$\begin{aligned} |f\rangle &= \sum_{n=0}^{\infty} a_n |\phi_n\rangle \\ &= \sum_{n=0}^{\infty} \langle \phi_n | f \rangle |\phi_n\rangle \\ &= \end{aligned} \quad (47)$$

$$\sum_{n=0}^{\infty} |\phi_n\rangle \langle \phi_n| |f\rangle$$

is a consistent statement for general $f(x)$ only if

$$\sum_{n=0}^{\infty} |\phi_n\rangle \langle \phi_n| \equiv 1 \quad (48)$$

In these equations we repeatedly write down series sums over a (possibly and indeed usually infinite) set of functions. We must, therefore, address the issue of whether, and how, these sums **converge**. Noting that we were pretty slack in our requirements on $f(x)$ - it only had to be piecewise continuous, for example, and we said nothing about how it behaved at the points a and b themselves.

Clearly we cannot converge uniformly (at each and every point) since at certain points the "function" really isn't - whether or not you like to think that it has two values at the discontinuities, it clearly approaches different values from the two sides in a limiting sense that makes its value at the limit point hard to uniquely define. It turns out that the kind of convergence that we can expect is at least:

$$\lim_{N \rightarrow \infty} \int_a^b \left| f(x) - \sum_{n=0}^N a_n \phi_n(x) \right|^2 w(x) dx \rightarrow 0. \quad (49)$$

This is called **convergence in the mean** since the **mean square error** measured in this way goes to zero. This is a wee bit weaker than plain old "convergence" or "uniform convergence", but for a continuous, smooth, well-behaved function $f(x)$ the convergence is effectively uniform.

Note the ϵ, δ (Cauchy criterion for convergence) inherent in this statement. This **must** exist for us to be able to talk about the sum with a straight face instead of a smirk. This is because in the real world we cannot do infinite sums. If we make a **finite approximation**:

$$f(x) \approx \sum_{n=0}^N a_n \phi_n(x) \quad (50)$$

then a very important conclusion that can be shown (we won't show it) is that the a_n 's that minimize the least squares error (the integral above) **do not depend on N** and are in fact

$$a_n = \langle \phi_n | f \rangle. \quad (51)$$

This is **not true** if the $\phi_n(x)$ are not orthogonal, for example, if they are x^n so the sum is a power series, adding each term requires that one rearrange all the a_n to obtain the new "best fit". You can check this for homework, too.

In the meantime, we can understand the fundamental idea by multiplying out the convergence relation and noting that (since $w(x) > 0$):

$$\begin{aligned} \int_a^b \left| f(x) - \sum_{n=0}^N a_n \phi_n(x) \right|^2 w(x) dx &= \int |f(x)|^2 w(x) dx \\ &- \sum_n \int f(x)^* a_n \phi_n(x) w(x) dx - \sum_n \int f(x) a_n^* \phi_n^*(x) w(x) dx \\ &+ \sum_n \sum_m \int a_m^* a_n \phi_m^*(x) \phi_n^*(x) w(x) dx. \geq 0 \end{aligned} \quad (52)$$

Rearranging, we get **Bessel's Inequality**:

$$\int |f(x)|^2 w(x) dx \geq - \sum_n |a_n|^2 - \sum_n |a_n|^2 + \sum_n |a_n|^2 \geq \sum_n |a_n|^2 \quad (53)$$

If and only if the representation is **complete**, we obtain

$$\int |f(x)|^2 w(x) dx = \sum_n |a_n|^2. \quad (54)$$

This is known as **Parseval's Theorem**.

Let us examine all this in the context of a well-known example. Consider a **Fourier Series** for $f(x)$ on the interval $0 \leq x \leq 2\pi$:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx) \quad (55)$$

Thus the $u_n(x)$ functions are $\frac{1}{2}, \cos(nx), \sin(nx)$, with $w(x) = 1$. We can normalize these functions easily enough:

$$\phi_0(x) = \sqrt{\frac{2}{\pi}} \times \frac{1}{2} \quad (n=0) \quad (56)$$

$$\phi_n(x) = \sqrt{\frac{1}{\pi}} \times u_n(x) \quad (n \neq 0). \quad (57)$$

With these definitions,

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx) dx \quad (58)$$

$$(59)$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) dx$$

$$\int_0^{2\pi} |f(x)|^2 = \left[\frac{a_0^2}{2} + \sum_{n=1}^{\infty} (|a_n|^2 + |b_n|^2) \right] \pi \quad (60)$$

where the first relation works for a_0 too.

If we use only the $\sin(nx)$ functions, they are orthogonal but not (by themselves) complete. The expansion may only poorly approximate the function and

$$\int_0^{2\pi} |f(x)|^2 \geq \sum_{n=1}^{\infty} |b_n|^2 \quad (61)$$

Closure and Completeness

Now we can state the **closure relation** that is directly and intimately connected to completeness. Assume that an orthonormal set has $w(x) = 1$ (or that we have absorbed the weight factor into the normalized functions). Then:

$$a_n = \int_a^b \phi_n^*(y) f(y) dy \quad (62)$$

$$\begin{aligned} f(x) &= \sum_n a_n \phi_n(x) \\ &= \sum_n \int_a^b \phi_n^*(y) f(y) dy \phi_n(x) \\ &= \int_a^b \left(\sum_n \phi_n^*(y) \phi_n(x) \right) f(y) dy. \end{aligned} \quad (63)$$

But:

$$f(x) = \int_a^b \delta(x - y) f(y) dy \quad (64)$$

(by **definition** of the Dirac delta function!) Thus:

$$\left(\sum_n \phi_n^*(y) \phi_n(x) \right) = \delta(x - y) \quad (65)$$

$$\int_a^b \phi_m^*(x) \phi_n(x) dx = \delta_{mn} \quad (66)$$

together make the closure relation. It is in this sense that we can write the identity as the sum of unit vector projections:

$$\sum_n |\phi_n\rangle \langle \phi_n| = 1 \quad (67)$$

To continue using Fourier representations as an example, consider the normalized function:

$$\phi_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx} \quad (68)$$

on the interval $(0, 2\pi)$ with $w(x) = 1$. We can read off of the above that:

$$\frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{-iny} e^{inx} = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{in(x-y)} \quad (69)$$

$$= \delta(x - y) = \delta(y - x) \quad (70)$$

which is a *crucial* relation in physics. At some point in the near future I will probably lecture you to tears on how $\delta(x - y)$ is *not* a function but a distribution, or an integral operator, and how you can get into real trouble treating it as a function. However, you have probably heard this lecture four or five times already, so we'll pass it by for now.

Gram-Schmidt Orthogonalization

In many cases one can obtain a linearly independent set of functions $u_n(x)$, but find upon examination that they are (alas!) not orthogonal. For example, $u_n(x) = x^n$. Fortunately, it is always possible to **orthogonalize** the set of functions via a simple procedure. Let us define $\{u_n(x)\}$ to be the set of non-orthogonal but linearly independent functions. We will take suitable linear combinations of them to generate an orthogonal set which we will call $\{\psi_n(x)\}$. Finally, we can normalize this orthogonal set any way we like to form an orthonormal representation $\{\phi_n(x)\}$.

To understand the Gram-Schmidt procedure, it is easiest to consider it for ordinary Cartesian vectors. Suppose \vec{A} and \vec{B} are two non-orthogonal, but linearly independent vectors that span a two-dimensional plane as drawn below:

In this figure, we see that we can systematically construct $\vec{C} \perp \vec{A}$ by projecting \vec{B} onto \vec{A} and subtracting its \vec{A} -directed component from \vec{B} itself. What is left is necessarily orthogonal to \vec{A} . Algebraically:

$$|\vec{A}| |\vec{B}| \geq \vec{A} \cdot \vec{B} > 0 \quad (71)$$

We want to find \vec{C} such that $\vec{C} \cdot \vec{A} = 0$. We try

$$\vec{C} = \vec{B} + \lambda \vec{A}. \quad (72)$$

Putting this into the condition,

$$\vec{A} \cdot \vec{C} = \vec{A} \cdot \vec{B} + \lambda |\vec{A}|^2 = 0 \quad (73)$$

or

$$\lambda = -\frac{\vec{A} \cdot \vec{B}}{|\vec{A}|^2} \quad (74)$$

so that

$$\vec{C} = \vec{B} - \frac{\vec{A}}{|\vec{A}|} \left(\frac{\vec{A}}{|\vec{A}|} \cdot \vec{B} \right) \quad (75)$$

exactly as we expected from the figure.

This procedure works just as well for n sequential operations and with functional "vectors" instead of real space vectors. Pick any function $u_1(x)$. Let

$$|\psi_1\rangle = |u_1\rangle \quad (76)$$

and go ahead and normalize it:

$$|\phi_1\rangle = \frac{|\psi_1\rangle}{\sqrt{\langle \psi_1 | \psi_1 \rangle}} \quad (77)$$

(I won't explicitly include the normalization step in the future, but you'll see how/where it occurs). Assume

$$|\psi_2\rangle = |u_2\rangle - a |\phi_1\rangle \quad (78)$$

The condition $\langle \phi_1 | \psi_2 \rangle = 0$ leads to

$$a = -\langle \phi_1 | \psi_2 \rangle \quad (79)$$

(since $|\phi_1\rangle$ is normalized) and

$$|\psi_2\rangle = |u_2\rangle - |\phi_1\rangle \langle \phi_1 | u_2 \rangle \quad (80)$$

We can now normalize *this* into $|\phi_2\rangle$.

We then try to make

$$|\psi_3\rangle = |u_3\rangle - b|\phi_1\rangle - c|\phi_2\rangle \quad (81)$$

such that

$$\langle \phi_1 | \psi_3 \rangle = 0 \quad (82)$$

$$\langle \phi_2 | \psi_3 \rangle = 0 \quad (83)$$

These are two equations, there are two unknowns, and everything is hunky-dory if $|u_3\rangle$ is linearly independent of the first two (true by hypothesis and revealed even if untrue!). So iterate until you run out of linearly independent functions to orthogonalize or you get bored.

There is a clever example in both Arfken and Wyld where they show that if you take $u_n(x) = x^n$ and apply this procedure on the interval $[-1, 1]$ with $w(x) = 1$, you obtain the (gasp!) Legendre Polynomials! In fact, if one varies the interval and the weight function, one can obtain *all* the known orthogonal polynomials in this manner!

Insert Table

These are quite **useful** for both expansions and numerical integration (quadrature).

The Sturm-Liouville Theorem

Suppose one has a general 2nd order linear homogeneous ODE:

$$p_0(x)y''(x) + p_1(x)y'(x) + [p_2(x) + \lambda p_3(x)]y(x) = 0 \quad (84)$$

where we assume only that the $p_i(x)$'s are analytic on (a, b) (although a regular singularity at a or b is ok). We also insist that $p_0(x) \neq 0$ on the interior - all interior points must be "ordinary". This sort of 2OLHODE occurs very, very frequently in physics, and hence in this course. It can always (by suitable algebraic manipulations) be put in the **self adjoint** or **Sturm-Liouville** form:

$$\frac{d}{dx} \left(p(x) \frac{dy}{dx} \right) + q(x)y(x) + \lambda w(x)y(x) = 0 \quad (85)$$

Once it is in this form, we can easily show that it has a really nifty property. Let us define the **linear operator** \mathcal{L} such that:

$$\mathcal{L}u = \quad (86)$$

$$\mathcal{L}v = \frac{d}{dx} \left[p(x) \frac{dv}{dx} \right] + q(x)v \quad (87)$$

(etc.) where u and v are (almost) *arbitrary twice differentiable functions!* We can **integrate by parts twice**:

$$\int_a^b dx [v\mathcal{L}u - u\mathcal{L}v] = \left[p \left(v \frac{du}{dx} - u \frac{dv}{dx} \right) \right]_a^b \quad (88)$$

(where the q terms cancel on the boundary, as do all the terms with derivatives of both u and v). If we assume only the very weak condition that u and v are well-enough-behaved on the boundaries so that

$$Au(a) + Bu'(a) = 0 \quad (89)$$

$$Cu(b) + Du'(b) = 0 \quad (90)$$

(etc.) the boundary term vanishes and

$$\int_a^b dx [v\mathcal{L}u - u\mathcal{L}v] = 0 \quad (91)$$

Thus

$$\int_a^b dx v\mathcal{L}u = \int_a^b dx u\mathcal{L}v \quad (92)$$

This is called the *Hermitian* property of the differential operator; Hermitian and Self-Adjoint mean almost the same thing.

With this observation in hand, we can easily proceed to prove 2/3 of the **Sturm-Liouville Theorem** for solutions to nearly general (self-adjoint) 2OLHODE's.

We assume (here as above):

- p, q, w real, analytic in (a, b)
- $p \neq 0$ and $w \neq 0$ on (a, b)
- Since $w \neq 0$ it doesn't change sign, and we can always arrange for it to be positive, so $w \geq 0$ on (a, b) .
- That $y(x)$ satisfies "suitable" homogeneous B.C.'s at a and b . We require that if $y(x)$ is a solution, so is $\alpha y(x)$.

Almost any homogeneous BC works, any BC that works is homogeneous. Exceptions will be clear in the context of physics.

Then solutions to S-L 2OLHODE are a **discrete set of eigenfunctions** ($u_n(x)$) and corresponding **eigenvalues** (λ_n) where:

- λ_n are all real! (Self-adjoint property)

- $u_n(x)$'s are all **orthogonal**, $\langle u_m | u_n \rangle = \int_a^b u_m^*(x) u_n(x) w(x) dx = 0$ if $\lambda_n \neq \lambda_m$.

Note that sometimes $\lambda_m = \lambda_n$ for distinct u 's so you might think that the u 's don't have to be orthogonal. However, they do have to be linearly independent or they are distinct solutions with a vanishing Wronskian! So, we can GSO them to orthogonalize the λ -degenerate subspace. So for all practical purposes, after a bit of work, all the $u_n(x)$ are orthogonal or can be made so even if their eigenvalues are the same.

- The set of $\{u_n(x)\}$ are **complete**!

In summary, given a set of $u_n(x)$'s that solve a 2OLHODE, we can always write them as a complete orthonormal set $\{\phi_n(x)\}$.

This is a **really useful theorem**. Immediately implies the orthogonality of all the common ODE solution sets and orthogonal function sets in use in physics, e.g. $y'' + \lambda y = 0$, with its well-known Fourier solutions on the interval $(0, L)$ with Dirichlet boundary conditions $y(0) = 0$, $y(L) = 0$, $y_n(x) = \sin(\frac{n\pi x}{L})$ with $\lambda = \frac{n^2\pi^2}{L^2}$. Whew!

It is easy and instructive to prove the first two properties predicted by the SL theorem. The Hermitian property above implies:

$$\int_a^b u^* \mathcal{L}v dx = \left(\int_a^b v^* \mathcal{L}u dx \right)^* \quad (93)$$

$$= \int_a^b v (\mathcal{L}u)^* dx \quad (94)$$

for any $u(x)$, $v(x)$ that satisfy the B.C.'s (not necessarily the ODE). Note well that there is no $w(x)$ in these equations. So we define:

$$(u|\mathcal{L}v) = \int_a^b u^* \mathcal{L}v dx \quad (95)$$

in analogy with the bracket notation. In this shorthand, Hermitian means:

$$(u|\mathcal{L}v) = (v|\mathcal{L}u)^* = (\mathcal{L}u|v) \quad (96)$$

and we can conclude that $(u|\mathcal{L}v)$ is *real*.

NOW, let u and v be any two solutions $u_1(x)$ and $u_2(x)$ corresponding to λ_1 and λ_2 . Then:

$$\begin{aligned} (u_1|\mathcal{L}u_2) &= (u_1| -\lambda_2 w u_2) = -\lambda_2 (u_1|w u_2) = -\lambda_2 \langle u_1 | u_2 \rangle = \\ (u_2|\mathcal{L}u_1)^* &= (u_2| -\lambda_1 w u_1)^* = -\lambda_1^* \langle u_2 | u_1 \rangle^* = -\lambda_1^* \end{aligned} \quad (97)$$

or (subtracting)

$$(\lambda_2 - \lambda_1^*) \langle u_1 | u_2 \rangle = 0 \quad (98)$$

THUS

$$u_1 = u_2, \quad \langle u_1 | u_1 \rangle \neq 0, \quad \lambda_1 = \lambda_1^* \quad (99)$$

(and we conclude that the λ_i are **real**).

$$\lambda_1 \neq \lambda_2, \quad \langle u_1 | u_2 \rangle = 0 \quad (100)$$

(and we conclude that the $\{u_n(x)\}$ are all **orthogonal**).

It turns out to be quite difficult and involved to prove completeness. One basically has to show closure of one sort or another, and closure is not immediately obvious. It has long since been proven, however, and is shown in serious books like Hilbert and Courant if you want/need to look over the proof some day.

From this day forth, then, I will assume that you just know that the solutions to nearly every 2OLHODE that we treat in this course (and the rest of your courses form an orthogonal (appropriately normalized, in practice) **basis**, out of which it is perfectly clear that general solutions can be built via superposition.

It is time now for a short interlude, first on tensors (to get the trivia out of the way), then on curvilinear coordinates (to get you to where you can appreciate separation of variables in the Big Three coordinate systems) and we'll hop on back to ODE's, this time in the context of PDE's and "real physics".

Lecture 1

General, Nth Order Linear Homogeneous ODE's

$$a_N(x) \frac{d^N y}{dx^N} + a_{N-1}(x) \frac{d^{N-1} y}{dx^{N-1}} + \dots + a_0(x) y = 0 \quad (101)$$

If we divide out the $a_N(x) \rightarrow A_k(x) = \frac{a_k(x)}{a_N(x)}$, this becomes:

$$\frac{d^N y}{dx^N} + A_{N-1}(x) \frac{d^{N-1} y}{dx^{N-1}} + \dots + A_0(x) y = 0 \quad (102)$$

(making the leading order linearity obvious). Let x be complex - most general abelian algebra. Then a point x_0 is:

- **ordinary** if $A_k(x)$'s are all analytic at $x = x_0$.
- **singular** if not:
 1. **regular singular** if $(x - x_0)^{N-k} A_k(x)$ are all analytic, so that $A_{N-m}(x)$ has no more than an m th order pole.
 2. **irregular point** or **essential singularity** (otherwise).

Fuch's Theorem

- At an ordinary point x_0 there are N linearly independent solutions.
- The Taylor series for these, $\sum_{n=0} \infty a_n (x - x_0)^n$ converges at least up to the nearest singular point.
- At a **regular singular point**, there is at least one solution of the form:

$$(x - x_0)^k \sum_{n=0} \infty a_n (x - x_0)^n \quad (103)$$

where k can be any complex number. k is called the **indicial exponent**. This expansion, too, converges at least to the nearest singular point.

Example: Bessel (Differential) Equation

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + (x^2 - m^2)y = 0 \quad (104)$$

or

$$\frac{d^2 y}{dx^2} + \frac{1}{x} \frac{dy}{dx} + \left(1 - \frac{m^2}{x^2}\right)y = 0 \quad (105)$$

so (obviously) $x = 0$ is a *regular singular point*.

Solutions: Frobenius Method

...follows directly from Fuch's Theorem:

1. If you want to expand around $x_0 \neq 0$, change variables so that you are expanding around 0: $x' = x - x_0$ or (if $x_0 = \infty$) $x' = 1/x$. This set of coordinates defines the **radius of convergence** of the solution derived.
2. Decide which case/kind of point x_0 is:
 - Ordinary: Use $y(x) = \sum_n a_n x^n$.
 - Regular singular: Use $y(x) = x^k \sum_n a_n x^n$ with $a_0 \neq 0$.
 - Irregular singular, essential: Use Famous Mathematician Method (i.e. - look it up if you can find it) or Give Up.
3. Substitute into ODE, also expanding any $A_k(x)$ coefficients (e.g. - $\sin x \rightarrow x - \frac{x^3}{3!} + \dots$. Match up coefficients of each power of x .

- Take especial care with the first N coefficients (usually 2, for our purposes).
- After that, write a recurrence relation and hence the whole solution. Make pithy observations, if any exist to be made.
- For the regular singular point case, the lowest remaining power of x determines k via the **indicial equation**.

Examples

1. Airy's Equation

$$\frac{d^2 y}{dx^2} - xy = 0 \quad (106)$$

Note that $x = 0$ is ordinary, no singularities but $x \rightarrow \infty$.

$$y(x) = \sum_{n=0}^{\infty} a_n x^n \quad (107)$$

$$\frac{d^1 y}{dx^1}(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} \quad (108)$$

$$\frac{d^2 y}{dx^2}(x) = \sum_{n=2}^{\infty} n(n-1) a_n x^{n-2} \quad (109)$$

or

$$\left(\sum_{n=2}^{\infty} n(n-1) a_n x^{n-2} \right) = \left(\sum_{n=0}^{\infty} a_n x^{n+1} \right) = 0 \quad (110)$$

Examine the coefficients of:

- x^0 : $(2) \times (1) a_2 = 0$ or $a_2 = 0$.
- x^1 : $(3) \times (2) a_3 - a_0 = 0$
- x^{m+1} : $(m+3) \times (m+2) a_{m+3} - a_m = 0$

or

$$\frac{a_{m+3}}{a_m} = \frac{1}{(m+3)(m+2)} \quad (111)$$

We can now reconstruct the entire solution:

(112)

$$y(x) = a_0 \left(1 + \frac{x^3}{6} + \frac{x^6}{360} + \dots \right) + a_1 \left(x + \frac{x^4}{12} + \frac{x^7}{504} + \dots \right) + (a_2 = 0 \dots)$$

Remarks:

- **Two** linearly independent solutions with coefficients a_0 and a_1 .
- Converges everywhere (ratio test for x^3 or Fuch's Theorem).
- $Ai(x)$ and $Bi(x)$ (Airy functions/Airy integrals) correspond to two particular choices for a_0, a_1 . These are not too important in E&M, but are important in quantum theory.

2. Legendre Equation:

$$(1 - x^2) \frac{d^2 y}{dx^2} - 2x \frac{dy}{dx} + \ell(\ell + 1)y = 0 \quad (113)$$

$x = 0$ is ordinary. $x = \pm 1$ are regular singular points. Start with $x = 0$. Using same expansions for $y(x)$ as above:

$$(1 - x^2) \sum_{n=2}^{\infty} n(n-1)a_n x^{n-2} - 2x \sum_{n=1}^{\infty} n a_n x^{n-1} + \ell(\ell+1) \sum_{n=0}^{\infty} a_n x^n = 0 \quad (114)$$

With a little work (and treating the first two carefully) we find that

$$\begin{aligned} \frac{a_{n+2}}{a_n} &= \frac{-\ell^2 - \ell + n^2 + n}{(n+2)(n+1)} \\ &= -\frac{(\ell - n)(\ell + n + 1)}{(n+1)(n+2)} \end{aligned} \quad (115)$$

or

$$\begin{aligned} y(x) = & a_0 \left(1 - \ell(\ell+1) \frac{x^2}{2!} + \ell(\ell+1)(\ell-2)(\ell+3) \frac{x^4}{4!} + \dots \right) \\ & + a_1 \left(x - (\ell-1)(\ell+2) \frac{x^3}{3!} + (\ell-1)(\ell+2)(\ell-3)(\ell+4) \frac{x^5}{5!} + \dots \right) \end{aligned} \quad (116)$$

Remarks:

Iff ℓ is an integer, **one** of the two independent series terminates, giving a finite polynomial solution. We can tabulate these solutions for each ℓ (including some optional normalization so that $P_\ell(x) = 1$):

$$1P_0(x) = 1 \quad (117)$$

$$1P_1(x) = x \quad (118)$$

$$-2P_2(x) = 1 - 3x^2 \quad (119)$$

$$-\frac{2}{3}P_3(x) = x - \frac{5}{3}x^3 \quad (120)$$

$$\dots \quad (121)$$

or in more familiar form we have the **Legendre Polynomials**:

$$P_0(x) = 1 \quad (122)$$

$$P_1(x) = x \quad (123)$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1) \quad (124)$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x) \quad (125)$$

$$\dots \quad (126)$$

-
- [About this document ...](#)
-



Next: [About this document ...](#)

Robert G. Brown 2003-02-09