Notes on

Mathematical Physics

The Sturm-Liouville problem

Mathematical review and preliminary definitions

Field (Körper, cuerpo): a set $\mathbb{K}$ is a field when an addition $\mathbb{K} \times \mathbb{K} \to \mathbb{K}, (x, y) \to x + y$ and multiplication $\mathbb{K} \times \mathbb{K} \to \mathbb{K}, (x, y) \to xy$ can be defined in such a way that both operations are associative, commutative and there are neuter and inverse elements. (Examples: $\mathbb{R}, \mathbb{Q}, \mathbb{C}$). Don't confuse this definition of a "field" with a scalar or vector field. In this document, $\mathbb{N}_0$ is the set $\mathbb{N} \cup \{0\}$.

Function: a function f from a set A to a set B is a correspondence such that for each element $x \in A$ there is exactly one corresponding element $y \in B$, $y = f(x)$. For example, A and B can be $\mathbb{R}$, $\mathbb{R}^n$or $\mathbb{C}$. Example of a function definition: $f: \mathbb{R} \to \mathbb{R}, f(x) = x^2$.

Comment: We have to be careful in defining the domain of a function; also with programming (we must tell the compiler whether the arguments are real or complex, and the function return is real or complex). Example of the function $f: \mathbb{R} \to \mathbb{C}, f(x) = i + x$ in Fortran:

```
function f(x)
      real x
      complex f
      f = (0,1) + x
end function
```

Leibniz rule for differentiation: generalization of the product rule for the nth derivative. Consider two functions $f(x), g(x): \mathbb{R} \to \mathbb{R}$, and $n, k \in \mathbb{N}$. Then, $\frac{d^n}{dx^n}[f(x) \cdot g(x)] = \sum_{k=0}^{n} \binom{n}{k} \frac{d^{n-k}}{dx^{n-k}} f(x) \cdot \frac{d^k g}{dx^k}$. One can visualize this as $\frac{d^n(fg)}{dx^n} = \left(D_f + D_g\right)^n (fg)$, where D_f, D_g are derivatives operating only on f and g, respectively. Then, the derivative operators form a binomial raised to the nth power.

Example: $(fg)'' = f''g + 2f'g' + fg''$.

Vector space: a set V is a vector space over a field $\mathbb{K}$ if there is an addition $V \times V \to V: (x, y) \to x + y$ and a scalar multiplication $\mathbb{K} \times V: (\lambda, x) \to \lambda x$. The scalar multiplication should satisfy: a) $\lambda(x + y) = \lambda x + \lambda y$; b) $(\lambda + \mu)x = \lambda x + \mu x$; c) $(\lambda\mu)x = \lambda(\mu x)$; d) $1x = x$, for all $\lambda, \mu \in \mathbb{K}$ and $x, y \in V$.

Comment: the definition of a vector space does not define what a vector is. Several mathematical objects over a given field can be behave as a vector (not only an "arrow" linking two points in physical space is a vector). Polynomials and functions can also form vector spaces (see below), as long as they satisfy the properties of a vector space.

Polynomial: a real polynomial is a function $f: \mathbb{R} \to \mathbb{R}$ of the form $f(x) = a_0 + a_1 x + \cdots + a_n x^n = \sum_{k=0}^{n} a_k x^k$. $a_k \in \mathbb{R}$ are the coefficients of the polynomial.

Linear independence: The vectors $\psi_1, \psi_2, \ldots$ are linearly independent when the equation $\sum c_n \psi_n = 0$ only has the trivial solution $c_1 = c_2 = \cdots = 0$. The maximum number of linearly independent vectors is called the *dimension* of the vector space.

Basis and coordinates: there are linearly independent vectors $\{e_1, e_2, \dots\} \subset V$ for which each vector $v \in V$ can be written as a linear combination of basis vectors: $v = \sum_i x_i e_i$. There is always a bijective transformation between two given bases. The x_i are the coordinates of the vector, usually written as a column vector (or transposed row vector to save space). Example: polynomials $f(X)$ can form a vector space; for the polynomial $f(X) = 1 + 2X + 3X^2$, the coordinates in the basis $\{1, X, X^2\}$ are $(1,2,3)^T$, but in the basis $\{1, 1+X, 1+X+X^2\}$, they are $(-1,-1,3)^T$.

Comment: continuing with the previous example, we see that for the set (sequence) of linearly-independent polynomials $\{1, X, X^2, \dots, X^n, \dots\}$ (until infinity), we obtain a vector space (of functions) of infinite dimensions.

Norm: a norm over a real or complex vector field V is a map $\|\cdot\| : V \to \mathbb{R}, v \to \|v\|$ such that its a) positive definite, b) homogeneous, $\|\lambda x\| = |\lambda| \|x\|$; c) satisfies the *triangle inequality*. There are many possible (equivalent) norms that can be defined in a vector space. Example: the Euclidean norm of a *n*-dimensional vector space $\|x\| = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}$ (L2 norm).

Inner product of two functions: $\langle g, f \rangle = \int_a^b \bar{g}(x) f(x) w(x) dx$, given functions $f, g : [a, b] \in \mathbb{R} \to \mathbb{C}$ (i.e, the functions are defined within an interval), and $w(x)$ is a *weigh function*. The bar means complex conjugation. Similar to other vector fields (like forces or velocities), two functions can be orthogonal when their scalar product is zero. The *normalization* of a function is computed as $N^2 = \langle f, f \rangle = \int_a^b |f(x)|^2 w(x) dx$.

Gram-Schmidt orthogonalization: just as there is a Gram-Schmidt orthogonalization method for vectors in their matrix representation, there is an equivalent for functions. A set of functions $\{f_n\}$are orthonormal to each other when $\langle f_n, f_m \rangle = \delta_{nm}$. The Gram-Schmidt orthogonalization method creates an orthonormal set of functions from a nonorthogonal set. We won't study it here in detail.

Example: given the sequence of functions $\{x^n\}, n \in \mathbb{N}_0$, and an inner product with interval $[-1,1]$ and $w(x) = 1$, one can start with $x^0 = 1$, which normalized makes $f_0 = 1/\sqrt{2}$. Then, one takes as an Ansatz the next order polynomial, $f_1 = x + (\text{const}) f_0 = x + (\text{const}) \frac{1}{\sqrt{2}}$ orthonormalizing with f_0 one arrives at $f_1 = \sqrt{3/2}\, x$. Repeating the process one more time, one has $f_2 = x^2 + (\text{const})_1 f_1 + (\text{const})_0 f_0$, and orthonomalizing against f_1, f_0 one obtains $f_2 = \sqrt{5/2} \cdot (1/2) \cdot (3x^2 - 1)$, etc. Those are the *orthonormalized Legendre polynomials*.

Hilbert space: a (complete metric) vector space with a scalar product (warning: informal definition!). The idea of completeness requires defining a *Cauchy sequence* of vectors $\{|a_i\rangle\}_{i=1}^{\infty}$, and it tests whether the limit of the sequence as $i \to \infty$ is also within the vector space. In a *complete Hilbert space*, a vector basis can be defined.

Lebesgue space: the space of square-integrable functions (see "Inner product" definition) over the interval $[a, b]$ is denoted as $\mathscr{L}^2_w(a, b)$, where 2 means the power of f in the integral and w is the weight function. It can be shown that:

- $\mathscr{L}^2_w(a, b)$ is a complete space ($\Longrightarrow$ a basis can be defined)
- The sequence of monomials $\{x^k\}, k \in \mathbb{N}_0$ forms a basis of $\mathscr{L}^2_w(a, b)$ (Stone-Weiertrass approximation theorem).

Orthogonal polynomials: given that any function can be written as $f(x) = \sum_{k=0}^{\infty} a_k x^k$, we can also use a linear combination of those vectors as a basis to also expand $f(x)$. The linear combination of those vectors is then a set of certain polynomials $C_n(x)$, orthogonal to one another, and that span $\mathscr{L}_w^2(a,b)$. If n is the degree of the polynomial $C_n(x)$, for another polynomial $p_{\leq n-1}(x)$, $\langle C_n(x), p_{\leq n-1}(x)\rangle = 0$. Warning: remember that this time, our definition of the inner product includes a weigh function.

Differential equations and eigenvalue problems: given a second-order differential operator $\hat{\mathscr{L}}$, a function $\psi(x) : \mathbb{R} \to \mathbb{C}$ and $\lambda \in \mathbb{C}$, the differential equation

$$\hat{\mathscr{L}}\psi(x) = \lambda\psi(x)$$

subject to *boundary* conditions, defines an *eigenvalue problem* (a problem such that a scalar λ has the same effect on ψ than an operator $\hat{\mathscr{L}}$). Compared to linear algebra of matrices, the differential operator $\hat{\mathscr{L}}$ is equivalent to a matrix ad the functions $\psi(x)$ are equivalent to eigenvectors (here called *eigenfunctions*). The functions $\psi(x)$ subject to the boundary conditions form a Hilbert space. The operator $\hat{\mathscr{L}}$ has a general form

$$\hat{\mathscr{L}} = p_0(x)\frac{d^2}{dx^2} + p_1(x)\frac{d}{dx} + p_2(x).$$

where we limit ourselves to the case $p_j(x) : \mathbb{R} \to \mathbb{R}, j \in \{0,1,2\}$.

Example: the spatial part of a standing wave must satisfy the differential equation $\frac{d^2\psi}{dx^2} + k^2\psi = 0$ subject to the boundary conditions $\psi(0) = \psi(l) = 0$. Defining $\hat{\mathscr{L}} = \frac{d^2}{dx^2}$, one has $\hat{\mathscr{L}}\psi = -k^2\psi$. The eigenvalues are then $\lambda = -k^2$. The solution to this equation are the eigenfunctions $\psi_n(x) = A\sin(n\pi x/l)$, with $k^2 = n^2\pi^2/l^2$, $n \in \mathbb{N}$. We can show that those functions are orthogonal, i.e., $\langle \psi_n, \psi_m\rangle = 0 \iff n \neq m$ within the interval $[0,l]$.

Hermitian adjoint operator: consider the inner product of two functions, $\langle \hat{\mathscr{D}}f, g\rangle$, where $\hat{\mathscr{D}}$ is a differential operator acting on g. An adjoint operator $\hat{\mathscr{D}}^\dagger$ is defined such that $\langle \hat{\mathscr{D}}f, g\rangle = \langle f, \hat{\mathscr{D}}^\dagger g\rangle$ + extra terms due to the boundaries. If those extra terms are zero, then the operator is called *Hermitian adjoint*. (Remember that the inner product has the complex conjugate defined, as well as an interval and a weight!)

Self-adjoint operator: an operator for which $\hat{\mathscr{D}}^\dagger = \hat{\mathscr{D}}$. A Hermitian self-adjoint operator is simply called Hermitian. It can be shown that $\hat{\mathscr{L}}$ is self-adjoint if $p_0'(x) = p_1(x)$. This means that one can write $\hat{\mathscr{L}} = \frac{d}{dx}\left[p_0(x)\frac{d}{dx}\right] + p_2(x)$.

Ortogonality of the eigenfunctions: consider a Hermitian operator $\hat{\mathscr{L}}$ that satisfies $\hat{\mathscr{L}}u = \lambda_u u$, $\hat{\mathscr{L}}v = \lambda_v v$, where $u, v : \mathbb{R} \to$; $v(x), v(x)$ and no weight. Then, $\int_a^b \bar{v}\hat{\mathscr{L}}u\,dx = \int_a^b [\bar{v}(p_0u')' + \bar{v}p_2u]dx$ = (int. by parts) $= [\bar{v}p_0u']_a^b + \int_a^b [-\bar{v}'p_0u' + \bar{v}p_2u]dx$ = (another int. by parts) = $\int_a^b \bar{v}\hat{\mathscr{L}}u = [\bar{v}p_0u' - \bar{v}'p_0u]_a^b + \int_a^b [(p_0\bar{v}')'u + \bar{v}p_2u]dx = [\bar{v}p_0u' - \bar{v}'p_0u]_a^b + \int_a^b \overline{(\hat{\mathscr{L}}v)}u\,dx$. The boundary terms must vanish if $\hat{\mathscr{L}}$ is Hermitian, by definition. We see that if u, v are eigenfunctions, we can write $(\lambda_u - \lambda_v)\int_a^b \bar{v}u\,dx = [p_0(\bar{v}u' - \bar{v}'u)]_a^b$. For $u \neq v$ and in general $\lambda_u \neq \lambda_v$, and if the boundary terms vanish, then $\langle v, u\rangle = 0$ (i.e, the eigenfunctions must be orthogonal).

Comment: More generally, if the operator $\hat{\mathscr{L}}_{\text{nsa}}$ is not self-adjoint (nsa), then there can be a function $w(x)$ (the weight function) such that, when $w(x)\hat{\mathscr{L}}_{\text{nsa}}\psi(x) = w(x)\lambda\psi(x)$, it makes the operator self-adjoint (sa) (i.e., $\hat{\mathscr{L}}_{\text{sa}} = w\hat{\mathscr{L}}_{\text{nsa}}$). One can show that such a function is $w(x) = p_0^{-1}(x)\exp\left[\int \frac{p_1(x)}{p_0(x)}dx\right]$, and that the orthogonality condition requires the presence of the weight function in the inner product.

Sturm-Liouville equation: The second-order linear ordinary differential equation $\hat{\mathscr{L}}[y(x)] = -\lambda w(x)y(x)$ with the Hermitian operator $\hat{\mathscr{L}} := \frac{d}{dx}\left[p(x)\frac{d}{dx}\right] - s(x)$, subject to suitable boundary conditions, is called a *Sturm-Liouville problem*.

Comment: the boundary conditions must be such that $[p(\bar{v}u' - \bar{v}'u)]_a^b$ for two eigenfunctions. There are several possibilities. For real eigenfunctions, here are some examples:

- *Dirichlet boundary conditions*: $u(a) = u(b) = v(a) = v(b) = 0$
- *Neumann boundary conditions:* $u'(a) = u'(b) = v'(a) = v'(b) = 0$
- *Periodic boundary conditions:* $p(a) = p(b)$; $v(a) = v(b)$; $v'(a) = v'(b)$
- Other combinations of the eigenfunctions or their derivatives such that the term in the parenthesis is zero at the boundaries
- Making $p(x)$ to be zero at the boundaries.

Series expansion of a function: one can develop a given function $f(x)$ into a series (i.e., set the function as the limit value to which the series must converge) of a set of (orthogonal) eigenfunctions $\{g_k(x)\}$ ($k \in \mathbb{Z}$) that form the basis of a vector space:

$$f(x) = \sum_k c_k g_k(x).$$

The limit of the series can be, e.g., from $k = -\infty$ to ∞, or from $k = 0$ to ∞. If the eigenfunctions are not orthonormal, one can find the normalization constant N_k as $N_k^2 = \langle g_k, g_k \rangle$. The coefficients $c_k \in \mathbb{C}$ can be found using the inner product (multiplying from the right by $g_\ell(x)$ with a different index $\ell \in \mathbb{Z}$):

$$\langle f(x), g_\ell(x)\rangle = \sum_k c_k \langle g_k(x), g_\ell(x)\rangle = \sum_k c_k N_k^2 \delta_{\ell k} \quad (\delta_{\ell k} \text{ is the Kronecker delta}).$$

$$\Longrightarrow \quad c_\ell = \langle f(x), g_\ell(x)\rangle / N_\ell^2.$$

Classical orthogonal polynomials: there are some orthogonal polynomials $C_n(x)$ that satisfy the Sturm-Liouville equation and are called *classical orthogonal polynomials*. Remember: there are functions that satisfy the Sturm-Liouville equation and that are not polynomials.

Generalized Rodrigues's formula: $C_n(x) = \frac{K_n}{w}\frac{d^n}{dx^n}[wp_0^n]$. Where: $p_0(x) : \mathbb{R} \to \mathbb{R}$ is a polynomial of degree ≤ 2 with only real roots; $w(x) : \mathbb{R} \to \mathbb{R}$ strictly positive function ("weight function"), integrable within $[a, b]$ and such that $w(a)s(a) = 0 = w(b)s(b)$; $C_1(x)$ is a first-degree polynomial in x. The generalized Rodrigues's formula generates classical orthogonal polynomials in the interval $[a, b]$. The constant K_n depends on the standard normalization of the polynomials (chosen so for historical reasons).

Comment: It can be proven that the Rodrigues's formula generates polynomials that satisfy the differential equation $(wp_0C_n')' = w\lambda_n C_n$, i.e., they satisfy a Sturm-Liouville problem.

Generating function: it is also possible to generate all orthogonal polynomials that satisfy a given Sturm-Liouville problem by repeated differentiations of a *generating function* $g(x, t)$ that can be expanded as $g(x, t) = \sum_{n=0}^{\infty} a_n t^n C_n(x)$ with some constants a_n. Note that the ***nth*** derivative w.r.t. t brings out the polynomials and changes of variables to the index in the series create terms like C_{n-1} or C_{n+1}. This means that derivatives of the generating function can create recurrence relations.

Some (definitely not all!) **orthogonal polynomials and eigenfunctions of the Sturm-Liouville problem**

Polynomial/ Function	**Interval**	$w(x)$	$p(x)$ **(nb 1)**	$s(x)$	λ	**Standard normalization** N^2	**Generating function (nb 2)**	K_n **for Rodrigues's formula**
Harmonic $\{e^{inx}\}$	$[-\pi, \pi]$	1	1	0	n^2	π	-	not a polynomial
Legendre $P_l(x)$	$[-1,1]$	1	$1-x^2$	0	$l(l+1)$	$\frac{2}{2l+1}$	$(1-2xt+t^2)^{-1/2}$	$\frac{1}{2^l l!}\cdot(-1)^l$ **(nb 3)**
Associated Legendre $P_l^m(x)$	$[-1,1]$	1	$1-x^2$	$\frac{-m^2}{1-x^2}$	$l(l+1)$	$\frac{2}{2l+1}\cdot\frac{(l+m)!}{(l-m)!}$	Modifications needed. Main definition: $P_l^m(x)=(-1)^m\frac{d^m}{dx^m}P_l(x)$	
Bessel of the first kind $J_n(x)$	$[0,a]$	x	x	$-x$	n^2/x	$\int_0^a\left[J_\nu\left(k_{\nu\mu}\frac{x}{a}\right)\right]^2 x\,dx = \frac{a^2}{2}[J_{\nu+1}(k_{\nu\mu})]^2$ ($k_{\nu\mu}$: μth zero of J_ν) **(nb 4)**	$e^{(x/2)(t-1/t)}$	not a polynomial
Laguerre $L_n(x)$	$[0,\infty]$	e^{-x}	xe^{-x}	0	n	1	$\frac{e^{-xt/(1-t)}}{1-t}$, $a_n=1/n!$	$\frac{1}{n!}$
Associated Laguerre $L_n^k(x)$ **(nb 5)**	$[0,\infty]$	$x^k e^{-x}$	$x^{k+1}e^{-x}$	0	$n-k$	$\frac{(n+k)!}{n!}$	$\frac{e^{-xt/(1-t)}}{(1-t)^{k+1}}$, $a_n=1/n!$	$\frac{1}{n!}$
Hermite $H_n(x)$ **(nb 6)**	$[-\infty,\infty]$	e^{-x^2}	e^{-x^2}	0	$2n$	$2^n\pi^{1/2}n!$	e^{-t^2+2tx}, $a_n=1/n!$	$(-1)^n$
Chebyshev polynomials $T_n(x)$	$[-1,1]$	$\frac{1}{\sqrt{1-x^2}}$	$(1-x^2)^{\frac{1}{2}}$	0	$-n^2$	$\pi/2$, if $n\neq 0$; π, if $n=0$	$\frac{1-tx}{1-2tx+t^2}$	$\frac{(-2)^n n!}{(2n)!}$

Checked against: Weber & Arfken (2003) *Essential Mathematical Methods for Physicists*, Academic Press. Chapters 9-13.
The Chebyshev polynomials are taken from Korn & Korn (1968) *Mathematical handbook...* , table 21.7-1.
However: please be careful and always check.

(nb 1) Warning, when computing the Rodrigues's formula: $p_0(x)=p(x)/w(x)$.
(nb 2) Here, $a_n=1$ in the respective series expansion, unless otherwise specified. Warning: generating functions are not unique.
(nb 3) In most books, they take in Rodrigues's formula $p_0(x)\to -p_0(x)$ and therefore they show no factor $(-1)^l$ in K_l. However, they only do it in the Rodrigues's formula and not in the differential equation, so this is why we prefer to modify the K_l and introduce the minus sign there instead.
(nb 4) Because of recurrence formulae, one can express the normalization in other ways, for example, in terms of $J'_\nu(k_{\mu\nu})$.
(nb5) Similarly to the case of the associated Legendre polynomials, the Associated Laguerre polynomials are defined in terms of a derivative of the Laguerre polynomials, namely, $L_n^k(x)=(-1)^k\frac{d^k}{dx^k}L_{k+n}(x)$.
(nb 6) There is an alternative way of defining the Hermite polynomials and its differential equation. Here we take the "Physicist Hermite polynomials".

References

- Karpfinger, C. (2022) *Höhere Mathematik in Rezepte* (4th edition). Springer Spektrum.
- Hassani, S. (1999) *Mathematical Physics.* Springer.
- Korn, G.A.; Korn, T.M. (1968) *Mathematical handbook for scientists and engineers.* Dover publications.
- Boas, M. (1983) *Mathematical Methods in the Physical Sciences.* Wiley.
- Weber & Arfken (2003) *Essential Mathematical Methods for Physicists.* Academic Press.

Last update: 9 April 2025