Mathematical Methods

Problem Sheet 1: "Normed and inner product vector spaces"

Andre Lukas, MT 2022

1) (Examples of function vector spaces)
Consider the set $\mathcal{F}(S,V)$ of all functions $S \to V$ from a set S to a vector space V.

(a) Write down the definitions of vector addition and scalar multiplication for $\mathcal{F}(S,V)$ and argue that this makes $\mathcal{F}(S,V)$ a vector space.

(b) Sub-vector spaces of $\mathcal{F}(S,V)$ can be obtained by restricting to functions with certain properties which are preserved under vector addition and scalar multiplication. Based on this observation find three vector spaces of functions.

(c) Discuss possible norms and scalar product which can be introduced on the vector spaces found in part (b).

2) (Polarisation identities)
Let V be a (complex) inner product vector space with inner product $\langle \cdot, \cdot \rangle$ and associated norm $\| \cdot \|$ defined in the usual way as $\| \mathbf{v} \| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$.

(a) Show that

$$\langle \mathbf{v}, \mathbf{w} \rangle = \frac{1}{4}\left(\| \mathbf{v} + \mathbf{w} \|^2 - \| \mathbf{v} - \mathbf{w} \|^2 - i \| \mathbf{v} + i\mathbf{w} \|^2 + i \| \mathbf{v} - i\mathbf{w} \|^2 \right) , \tag{1}$$

for all $\mathbf{v}, \mathbf{w} \in V$. (This identity, also called polarisation identity, implies that the scalar product is completely determined by its associated norm.)

(b) Find the polarisation identity if V is a real inner product space.

(c) Let $T : V \to V$ be a linear operator on the complex inner product space V. Show that the following generalisation of the polarisation identity holds:

$$\begin{aligned} \langle \mathbf{v}, T\mathbf{w} \rangle &= \frac{1}{4} \left(\langle \mathbf{v} + \mathbf{w}, T(\mathbf{v} + \mathbf{w}) \rangle - \langle \mathbf{v} - \mathbf{w}, T(\mathbf{v} - \mathbf{w}) \rangle \right. \\ &\qquad \left. - i \langle \mathbf{v} + i\mathbf{w}, T(\mathbf{v} + i\mathbf{w}) \rangle + i \langle \mathbf{v} - i\mathbf{w}, T(\mathbf{v} - i\mathbf{w}) \rangle \right) . \end{aligned} \tag{2}$$

(d) Suppose $T : V \to V$ is a linear operator on a complex inner product space V with $\langle \mathbf{v}, T\mathbf{v} \rangle = 0$ for all $\mathbf{v} \in V$. Use the result from part c) to show that $T = 0$.

(e) Show that the linear operator $T : V \to V$ is hermitian if and only if $\langle \mathbf{v}, T\mathbf{v} \rangle \in \mathbb{R}$ for all $\mathbf{v} \in V$. (Hint: Apply the statement of part (d) to $T - T^\dagger$.)

3) (The normed vector space ℓ^p and the parallelogram identity)

(a) Let V be an inner product vector space with inner product $\langle\cdot,\cdot\rangle$ and associated norm $\|\cdot\|$. Show that the associated norm satisfies the parallelogram identity

$$\| \mathbf{v}+\mathbf{w} \|^2 + \| \mathbf{v}-\mathbf{w} \|^2 = 2\left(\| \mathbf{v} \|^2 + \| \mathbf{w} \|^2\right) , \tag{3}$$

for all $\mathbf{v},\mathbf{w} \in V$.

(b) Consider the normed vector space ℓ^p of sequences $(x_i)_{i=1}^{\infty}$, where $x_i \in \mathbb{C}$, with $\left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p}$ finite. Show that, for $1 < p < \infty$, the expression

$$\| (x_i) \| := \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p} \tag{4}$$

defines a norm on ℓ^p.

(c) For the normed space ℓ^p consider the vectors $\mathbf{v} = (1,0,0,\ldots)$ and $\mathbf{w} = (0,1,0,0,\ldots)$ together with the parallelogram identity from part (a) to show that the above norm on ℓ^p is not associated to an inner product for $p \neq 2$.

4) (Recap of Gram-Schmidt procedure)
Consider the Hilbert space $L^2_w(\mathbb{R})$ with weight function $w(x) = e^{-x^2}$ and scalar product

$$\langle f,g\rangle := \int_{\mathbb{R}} dx\, w(x) f(x) g(x) . \tag{5}$$

(a) Apply the Gram-Schmidt procedure to the sub-space $V \subset L^2_w(\mathbb{R})$ with basis $1, x, x^2$ to obtain ortho-normal polynomials p_i of degree i, where $i = 0,1,2$.

(b) Write a general quadratic polynomial $q(x) = \sum_{k=0}^{2} a_k x^k$ as a linear combination of the ortho-normal polynomials in part (a).

(c) Use the result from (b) to work out $\| q \|^2$.

(d) Consider the linear operator $T = \frac{1}{w}\frac{d}{dx} w \frac{d}{dx}$ and show that it is hermitian, relative to the above scalar product. Compute the matrix which describes the restriction of T to V relative to the basis of ortho-normal polynomials from part (a).

5) (Ortho-normal basis)
Let V be a (finite-dimensional) inner product space and $(\boldsymbol{\epsilon}_i)_{i=1}^n$ and $(\boldsymbol{\epsilon}'_i)_{i=1}^n$ be two ortho-normal basis sets.

(a) For any two vectors $\mathbf{v},\mathbf{w} \in V$ show that $\langle \mathbf{v}|\mathbf{w}\rangle = \sum_i \langle \mathbf{v}|\boldsymbol{\epsilon}_i\rangle\langle\boldsymbol{\epsilon}_i|\mathbf{w}\rangle$.

(b) Show that the matrix P with entries $P_{ij} = \langle \boldsymbol{\epsilon}'_i|\boldsymbol{\epsilon}_j\rangle$ is unitary.

(c) For a linear operator $\hat{T} : V \to V$ consider the matrices T and T' which consist of the entries $T_{ij} = \langle \boldsymbol{\epsilon}_i|\hat{T}|\boldsymbol{\epsilon}_j\rangle$ and $T'_{ij} = \langle \boldsymbol{\epsilon}'_i|\hat{T}|\boldsymbol{\epsilon}'_j\rangle$, respectively. Show that $T' = PTP^\dagger$, where P is the matrix from part (b).

6) (Rotations and unitary matrices)
Consider three-dimensional rotation matrices, that is, 3×3 real matrices R which satisfy $R^T R = \mathbb{1}_3$ and $\det(R) = 1$, as well as two-dimensional special unitary matrices, that is, complex 2×2 matrices U which satisfy $U^\dagger U = \mathbb{1}_2$ and $\det(U) = 1$.

(a) Write $R = \mathbb{1}_3 + iT + \cdots$ (where T is purely imaginary) and show that, to linear order in T, R satisfies the defining relation for rotations iff T is anti-symmetric. Such matrices T can be seen as "infinitesimal rotations".

(b) Convince yourself that the space of infinitesimal rotations is three-dimensional and that the matrices $\tilde{T}_i$ with entries $(\tilde{T}_i)_{jk} = -i\epsilon_{ijk}$ form a basis. Show that $[\tilde{T}_i, \tilde{T}_j] = i\epsilon_{ijk}\tilde{T}_k$, where the square bracket denotes the commutator.

(c) Find the set of matrices S such that $U = \mathbb{1}_2 + iS + \cdots$ satisfies the special unitary constraints to linear order and show that these are precisely the 2×2 hermitian matrices with vanishing trace.

(d) Convince yourself that the matrices $\tau_i := \frac{\sigma_i}{2}$, where σ_i are the Pauli matrices, form a basis of the space of 2×2 hermitian traceless matrices and that $[\tau_i, \tau_j] = i\epsilon_{ijk}\tau_k$.

(e) Given the above results, which statements can you make about the linear map r which maps anti-symmetric 3×3 matrices into hermitian, traceless 2×2 matrices and is defined by $r(\tilde{T}_i) := \tau_i$?

7) (Convergence and completeness)
Let V be the inner product space of polynomials in x, where $x \in [0, a]$ and $0 < a < 1$, with the standard inner product defined by an integral.

(a) Show that the sequence $(x^k)_{k=1}^\infty$ is a Cauchy sequence and show that it converges to the zero function.

(b) Consider the sequence $(s_k)_{k=0}^\infty$, where $s_k = \sum_{i=0}^k x^i$. Show that this is a Cauchy sequence.

(c) Argue from the result in (b) that V is not complete and, hence, that it is not a Hilbert space.

Additional computational problems

Computational methods, both numerical and symbolic, are of increasing importance in physics and symbolic computational tools have become significantly more powerful over the past decade or so. This is changing the way physicists work. Much as the introduction of the pocket calculator some 50 years ago has made by-hand numerical calculations unnecessary, modern systems such as Mathematica, can now take over standard symbolic calculations, such as algebraic manipulations or integration. This facilitates powerful checks of by-hand calculations but also allows for calculations which are virtually intractable with a pen-and-paper approach. The following problems present an opportunity to practice some of these methods in the context of topics from the Mathematical Methods course. They are supplementary and voluntary but strongly recommended and hopefully a fun way to engage with symbolic computations early on. The problems are meant for realisation in Mathematica which can be downloaded from the university server. Mathematica is easy to use, has good built-in documentation and many high-level mathematical functions - you can start to experiment immediately.

C1) (Computational realisation of Gram-Schmidt procedure)
Work within the setting of question 4), above, so in the space $L^2_w(\mathbb{R})$ with scalar product (5) and weight function $w(x) = \exp(-x^2)$.

(a) Define, within Mathematica, the scalar product (5).

(b) Write a short piece of code which realises the Gram-Schmidt procedure for a finite list of functions. In particular, this code should be capable of taking the monomials $(1, x, x^2, \ldots, x^n)$ as an input and returning the orthogonal polynomials $(p_0, \ldots, p_n)$.

(c) Use this code to check your by-hand calculation in question 4) and also apply it to larger monomial lists for $n > 2$.

(d) Choose some functions in $f \in L^2_w(\mathbb{R})$ and expand them in the monomial basis $(p_0, \ldots, p_n)$ to obtain an approximation $\tilde{f} = \sum_{k=0}^{n} \langle p_k, f \rangle p_k$. Compare f and $\tilde{f}$ by producing a plot.

C2) (Computing with three-dimensional rotations)

(a) Define three three-dimensional matrices $R_i \in SO(3)$, where $i = 1, 2, 3$, which represent rotations around the coordinates axes with rotation angles $\theta_1, \theta_2, \theta_3$. Obtain the most general rotation matrix $R \in SO(3)$ as the product $R = R_1 R_2 R_3$.

(b) Taylor expand the matrix R to linear order in the angles θ_i and show that the result is of the form $R = \mathbb{1}_3 + T + \mathcal{O}(\theta_i^2)$, where $T = \sum_i \theta_i T_i$. Find the matrices T_i and check that $[T_i, T_j] = \epsilon_{ijk} T_k$. For a vector $\mathbf{v} \in \mathbb{R}^3$, show that infinitesimal rotations $\delta\mathbf{v} := T\mathbf{v}$ can be written as $\delta\mathbf{v} = \mathbf{w} \times \mathbf{v}$, where $\mathbf{w} = (\theta_1, \theta_2, \theta_3)^T$.

(c) A different parametrisation of $SO(3)$ is given by the Euler angles (ϕ, θ, ψ). To write down this parametrisation introduce R_ψ, a rotation about the z-axis by an angle ψ, R_θ, a rotation about the x axis by an angle θ and R_ϕ, a rotation about the z-axis by an angle ϕ. Then work out $R = R_\psi R_\theta R_\phi$. Think of the angles (ϕ, θ, ψ) as time-dependent. For such time-dependent rotations the corresponding angular velocity $\mathbf{w}$ can be obtained from $W_{ij} = \epsilon_{ijk} w_k$, where $W = \dot{R} R^T$ and the dot denotes the time derivative. Use these relations to work out $\mathbf{w}$ in terms of the Euler angles and their derivatives. (All this is very useful to study the motion of tops but quite tedious to work out by hand.)

C3) (Metropolis-Hastings algorithm)
The problem is somewhat outside our main narrative, but it is related to the discussion of measures. Consider the space $L^1_w(\mathbb{R})$ with measure $w(x)dx$. The problem is to generate a sample of n points $p_i \in \mathbb{R}$ which are distributed according to the measure w, so that integrals $\int_{\mathbb{R}} dx\, w(x) f(x)$, for $f \in L^1_w(\mathbb{R})$, can be numerically evaluated by the sum $\frac{1}{n} \sum_{i=1}^{n} f(p_i)$. The

Metropolis-Hastings algorithm is a method to generate precisely such a point sample and it is an example of a Markow-chain Monte-Carlo method. Such methods are applied in many areas of physics.

(a) Read up on the Metropolis-Hastings algorithm - this is not complicated and a short explanation will do.

(b) Write a simple piece of Mathematica code which generates a sample of n points p_i which are distributed according to a given measure $w(x)$. Use a Gaussian proposal distribution.

(c) Apply this code to the probability distribution $w(x) = \frac{1}{5\pi}(1 + 5x^2 + 2x^4)e^{-x^2}$. Verify the result by plotting w together with a histogram of the point sample p_i.

(d) Choose a function $f \in L^1_w(\mathbb{R})$ and evaluate its integral $\int_{\mathbb{R}} dx\, w(x)f(x)$ numerically, by computing $\frac{1}{n}\sum_{i=1}^{n} f(p_i)$. Compare the result with Mathematica's built-in numerical integration.