Math 442

Exercise set 5 — solutions

1. Directly from the definition of the norm, for any $x, y \in E$,

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle\,,$$
$$\|x-y\|^2 = \langle x,x\rangle - \langle x,y\rangle - \langle y,x\rangle + \langle y,y\rangle\,, \qquad \text{and so}$$
$$(*) \qquad \|x+y\|^2 - \|x-y\|^2 = 2(\langle x,y\rangle + \langle y,x\rangle) = 4\,\Re(\langle x,y\rangle)\,. \quad \text{Hence,}$$
$$\begin{aligned}\|x+y\|^2 - \|x-y\|^2 &+ i(\|x+iy\|^2 - \|x-iy\|^2)\\ &= 2(\langle x,y\rangle + \langle y,x\rangle) + 2i(\langle x,iy\rangle + \langle iy,x\rangle)\\ &= 2(\langle x,y\rangle + \langle y,x\rangle) + 2i(-i\langle x,y\rangle + i\langle y,x\rangle) = 4\langle x,y\rangle\,.\end{aligned}$$

2. Granted that the Apollonian identity holds, define tentatively for $x, y \in E$

$$\langle x,y\rangle := \tfrac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2\big) \tag{1}$$

(which must be true, by ($*$), if the norm does come from a real inner product). We must prove that $\langle\, ,\rangle$, as so defined, is an inner product.

Firstly, $\langle x,y\rangle = \langle y,x\rangle$, which was 8.11*(a)*. Secondly, $\langle x,x\rangle = \|x\|^2$, so that 8.11*(c), (d)* are automatic. Neither of these properties requires the Apollonian identity. The difficulty is with linearity in the first argument, 8.11*(b)*.

Let $x, y, z \in E$. Then, using the Apollonian identity,

$$\begin{aligned}\|x+y+z\|^2 &+ \|z\|^2 + \|x-z\|^2 + \|y-z\|^2\\ &= \tfrac{1}{2}\big(\|x+y+2z\|^2 + \|x+y\|^2\big) + \tfrac{1}{2}\big(\|x+y-2z\|^2 + \|x-y\|^2\big)\\ &= \tfrac{1}{2}\big(\|x+y+2z\|^2 + \|x+y-2z\|^2\big) + \tfrac{1}{2}\big(\|x+y\|^2 + \|x-y\|^2\big)\\ &= \|x+y\|^2 + 4\|z\|^2 + \|x\|^2 + \|y\|^2\,.\end{aligned} \tag{2}$$

Since the expression (2) is unchanged if z is substituted by $-z$,

$$\begin{aligned}\|x+y+z\|^2 &+ \|z\|^2 + \|x-z\|^2 + \|y-z\|^2\\ &= \|x+y-z\|^2 + \|-z\|^2 + \|x+z\|^2 + \|y+z\|^2\,,\end{aligned}$$

from which, after cancellation and rearrangement,

$$\|x+y+z\|^2 - \|x+y-z\|^2 = \|x+z\|^2 - \|x-z\|^2 + \|y+z\|^2 - \|y-z\|^2\,,$$

that is, by (1),

$$\langle x+y, z\rangle = \langle x,z\rangle + \langle y,z\rangle\,. \tag{3}$$

To complete the proof of linearity in the first argument, we must show $\langle \lambda x, y\rangle = \lambda\langle x,y\rangle$ for any $\lambda \in \mathbb{R}$ and $x, y \in E$. If $\lambda = -1$, this is true:

$$\langle -x, y\rangle = \tfrac{1}{4}\big(\|-x+y\|^2 - \|-x-y\|^2\big) = \tfrac{1}{4}\big(-\|x+y\|^2 + \|x-y\|^2\big) = -\langle x, y\rangle\,.$$

Consequently, we need only consider *nonnegative* λ. By induction from (3),

$$\langle mx, y\rangle = m\langle x, y\rangle \qquad \text{for } m \in \mathbb{N}\,,\ x, y \in E\,.$$

Substituting x by $\frac{1}{m}x$, we deduce that $\langle x, y\rangle = m\langle \frac{1}{m}x, y\rangle$, $\langle \frac{1}{m}x, y\rangle = \frac{1}{m}\langle x, y\rangle$. Putting these facts together, $\langle \frac{m}{n}x, y\rangle = \frac{m}{n}\langle x, y\rangle$ for any positive rational m/n. However, (1) implies that $\langle\, ,\rangle$ is continuous (as a function of two variables, but we only need continuity in the first variable); thus the equality $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$ which has been established for positive rational λ must also be valid for all nonnegative real λ. This completes the proof.

3. If E is a *complex* normed space in which the Apollonian identity holds, then the previous exercise constructs a *real* inner product (that is to say, it is real-valued, symmetric and positive definite, and linear in the first argument with respect to *real* scalars), which I shall call $\langle\, ,\rangle_{\mathbb{R}}$. Notice that, for any $x, y \in E$,

$$\begin{aligned} 4\langle x, iy\rangle_{\mathbb{R}} &= \|x+iy\|^2 - \|x-iy\|^2 = \|i(-ix+y)\|^2 - \|(-i)(y+ix)\|^2 \\ &= \|y-ix\|^2 - \|y+ix\|^2 = -4\langle y, ix\rangle_{\mathbb{R}}\,. \end{aligned} \tag{4}$$

Therefore, for any $x \in E$, $\langle x, ix\rangle_{\mathbb{R}} = -4\langle x, ix\rangle_{\mathbb{R}}$, so that $\langle x, ix\rangle = 0$. (5)

Using the hint provided by the polarization identity, define for $x, y \in E$

$$\langle x, y\rangle_{\mathbb{C}} := \langle x, y\rangle_{\mathbb{R}} + i\langle x, iy\rangle_{\mathbb{R}}\,. \tag{6}$$

Because of (5), $\langle x, x\rangle_{\mathbb{C}} = \langle x, x\rangle_{\mathbb{R}}$. So $\langle\, ,\rangle_{\mathbb{C}}$ is positive definite, 8.11*(c)* and *(d)*, as in the previous exercise. Also

$$\langle y, x\rangle_{\mathbb{C}} = \langle y, x\rangle_{\mathbb{R}} + i\langle y, ix\rangle_{\mathbb{R}} = \langle x, y\rangle_{\mathbb{R}} - i\langle x, iy\rangle_{\mathbb{R}}\,,$$

by symmetry of $\langle\, ,\rangle_{\mathbb{R}}$ and (4). So $\langle\, ,\rangle_{\mathbb{C}}$ is Hermitian, 8.11*(a)*. But also

$$\begin{aligned} \langle x+y, z\rangle_{\mathbb{C}} &= \langle x+y, z\rangle_{\mathbb{R}} + i\langle x+y, iz\rangle_{\mathbb{R}} \\ &= \langle x, z\rangle_{\mathbb{R}} + \langle y, z\rangle_{\mathbb{R}} + i\langle x, iz\rangle_{\mathbb{R}} + i\langle y, iz\rangle_{\mathbb{R}} = \langle x, z\rangle_{\mathbb{C}} + \langle y, z\rangle_{\mathbb{C}}\,, \end{aligned}$$

so it only remains to prove that, for any *complex* scalar λ and any $x, y \in E$, $\langle \lambda x, y\rangle_{\mathbb{C}} = \lambda\langle x, y\rangle_{\mathbb{C}}$. Now, from (6) and (4) and symmetry,

$$\begin{aligned} \langle ix, y\rangle_{\mathbb{C}} &= \langle ix, y\rangle_{\mathbb{R}} + i\langle ix, iy\rangle_{\mathbb{R}} = -\langle x, iy\rangle_{\mathbb{R}} - i\langle i^2x, y\rangle_{\mathbb{R}} \\ &= i(\langle x, y\rangle_{\mathbb{R}}) + i\langle x, iy\rangle_{\mathbb{R}}) = i\langle x, y\rangle_{\mathbb{C}}\,, \end{aligned}$$

whilst, for $\lambda \in \mathbb{R}$, $\langle \lambda x, y\rangle_{\mathbb{C}} = \lambda\langle x, y\rangle_{\mathbb{C}}$ directly from (6). Linearity for all complex scalars follows in the obvious way.

4. Consider first the real case. Then, if S is skew-adjoint,

$$\langle Sx, x\rangle = \langle x, Sx\rangle = -\langle x, Sx\rangle$$

and, therefore, must be 0. On the other hand, if $\langle Sx, x\rangle = 0$ for all x, then for all $x, y \in E$

$$0 = \langle S(x+y), x+y \rangle = \langle Sx, x \rangle + \langle Sx, y \rangle + \langle Sy, x \rangle + \langle Sy, y \rangle$$
$$= \langle Sx, y \rangle + \langle Sy, x \rangle \,, \quad \text{since } \langle \, , \rangle \text{ is symmetric;}$$

this establishes that S is skew-adjoint.

In the complex case, $\langle Sx, x \rangle = -\langle x, Sx \rangle = -\overline{\langle Sx, x \rangle}$, so that $\langle Sx, x \rangle$ is purely imaginary. Conversely, if $\langle Sx, x \rangle$ is purely imaginary for every $x \in E$, then as before

$$\langle S(x+y), x+y \rangle = \langle Sx, x \rangle + \langle Sx, y \rangle + \langle Sy, x \rangle + \langle Sy, y \rangle$$

is purely imaginary, but as $\langle Sx, x \rangle$ and $\langle Sy, y \rangle$ are too, it follows that $\langle Sx, y \rangle + \langle Sy, x \rangle$ must be purely imaginary for all x, y. So

$$0 = \Re\langle Sx, y \rangle + \Re\langle Sy, x \rangle = \Re\langle Sx, y \rangle + \Re\overline{\langle Sy, x \rangle} = \Re\langle Sx, y \rangle + \Re\langle x, Sy \rangle \,.$$

That is, $\Re\langle Sx, y \rangle = -\Re\langle x, Sy \rangle$ for all x, y. However,

$$\Im\langle Sx, y \rangle = -\Re(i\langle Sx, y \rangle) = -\Re\langle S(ix), y \rangle = \Re\langle ix, Sy \rangle \quad \text{by above}$$
$$= \Re(i\langle x, Sy \rangle) = -\Im\langle x, Sy \rangle \,.$$

Thus, the equality $\langle Sx, y \rangle = -\langle x, Sy \rangle$ holds (as it holds for real and imaginary parts separately).

5. For $f, g \in E$,

$$\langle Df, g \rangle = \int_0^{2\pi} f'(t)\overline{g(t)}\,dt = \left[f(t)\overline{g(t)}\right]_0^{2\pi} - \int_0^{2\pi} f(t)\overline{g}'(t)\,dt \,.$$

Since f and g are periodic with period 2π, the first expression vanishes, and, of course, the derivative of the conjugate is the conjugate of the derivative, so that

$$\langle Df, g \rangle = -\int_0^{2\pi} f(t)\overline{g'(t)}\,dt = -\langle f, Dg \rangle \,.$$

For f to belong to the kernel of D, $f' = 0$, which means that f is constant. So $\ker D$ is one-dimensional, and consists only of the constant functions.

For g to be in the image of D, it must be the derivative of a C^∞ function with period 2π: $g = f'$. Thus $\int_0^{2\pi} g(t)\,dt = \int_0^{2\pi} f'(t)\,dt = f(2\pi) - f(0) = 0$. It is clear that this condition is sufficient as well as necessary, since it ensures that any indefinite integral of g is C^∞ and periodic with period 2π. Thus, $D(E)$ consists exactly of the C^∞ functions that are periodic with period 2π and have zero integral over a period. (Notice that it is in fact of codimension 1 — indeed, $\ker D$ and $D(E)$ are complementary subspaces of E.)

[This very elementary example has remarkable generalizations.]

6. I claim that the graph of D is closed in $E \times E$. Suppose, in fact, that $((f_n, Df_n))$ is a sequence in $G(D)$ which converges in $E \times E$ to (f, g). Then

$$f_n \to f \quad \text{in } E, \qquad Df_n \to g \quad \text{in } E.$$

Take any $h \in E$, and then from ex. 5

$$\langle Df_n, h\rangle \to \langle g, h\rangle\,, \quad \langle Df_n, h\rangle = -\langle f_n, Dh\rangle \to -\langle f, Dh\rangle = \langle Df, h\rangle\,,$$

so that $\langle g, h\rangle = \langle Df, h\rangle$, or $\langle g - Df, h\rangle = 0$. As this is true for any $h \in E$, it follows that $g = Df$ (we could take $h := g - Df$, for instance). But this tells us that $(f, g) \in G(D)$. Therefore, the limit in $E \times E$ of a sequence in $G(D)$ that converges in $E \times E$ itself belongs to $G(D)$, or $G(D)$ is closed in $E \times E$.

If E is a Hilbert space, the closed graph theorem will apply, and we may conclude that D is continuous as a linear mapping from E to E. However, it manifestly is *not* continuous. For instance, consider the function $s_n := \sin(nt)$, with $n \in \mathbb{N}$. We know

$$\|s_n\|^2 = \int_0^{2\pi} \sin^2(nt)\,dt = \pi\,, \quad \|Ds_n\|^2 = \int_0^{2\pi} n^2\cos^2(nt)\,dt = n^2\pi\,.$$

So D is not a bounded linear map. (Equivalently, $s_n/n \to 0$, but $D(s_n/n) \not\to 0$.) This shows that E cannot be a Hilbert space.

[It is possible, and not very hard, to give explicit examples of Cauchy sequences in E that have no limit in E, but they are not so easy to justify.]

7. *(a)* If $a, b \in H$, and $b \neq 0$, then

$$\begin{aligned} 0 \le \left\| a - \frac{\langle a,b\rangle b}{\|b\|^2} \right\|^2 &= \left\langle a - \frac{\langle a,b\rangle}{\langle b,b\rangle} b, a - \frac{\langle a,b\rangle}{\langle b,b\rangle} b \right\rangle \\ &= \langle a,a\rangle - \frac{\langle a,b\rangle\langle b,a\rangle}{\langle b,b\rangle} - \frac{\overline{\langle a,b\rangle}\langle a,b\rangle}{\overline{\langle b,b\rangle}} + \frac{\langle a,b\rangle\overline{\langle a,b\rangle}}{\langle b,b\rangle\overline{\langle b,b\rangle}}\langle b,b\rangle \\ &= \langle a,a\rangle - \frac{|\langle a,b\rangle|^2}{\langle b,b\rangle} \qquad \text{as } \langle b,b\rangle \text{ is real and } \langle a,b\rangle\langle b,a\rangle = |\langle a,b\rangle|^2\,. \end{aligned}$$

It follows that $|\langle a,b\rangle|^2 \le \langle a,a\rangle\langle b,b\rangle = \|a\|^2\|b\|^2$, which is in effect the Cauchy-Schwartz inequality. But the proof shows that this inequality will be an equality if and only if

$$\left\| a - \frac{\langle a,b\rangle b}{\|b\|^2} \right\| = 0\,, \quad \text{i.e. } a = \frac{\langle a,b\rangle}{\|b\|^2} b\,,$$

which says that a is *some* scalar multiple of b — and, conversely, if $a = \mu b$ for some scalar μ, then $\langle a,b\rangle/\langle b,b\rangle = \mu$, so that the inequality will be an equality if a is *any* scalar multiple of b. It is also an equality if $b = 0$. These two cases may be condensed into the condition that a and b are linearly independent.

(b) The inequality $\Re\langle a,b\rangle \le \|a\|\|b\|$ results from the chain of inequalities

$$\Re\langle a,b\rangle \le |\langle a,b\rangle| \le \|a\|\|b\|\,.$$

For it to become an equality, both these inequalities must be equalities. This is true if $b = 0$, when b is indeed a real nonnegative multiple of a; so we may as well assume $b \neq 0$. The second inequality, as we have seen, is an equality if and only if a is a scalar multiple of b, $a = \lambda b$. Then

$$\Re\langle a,b\rangle = (\Re\lambda)\langle b,b\rangle\,, \quad |\langle a,b\rangle| = |\lambda|\langle b,b\rangle\,,$$

and (since $\langle b,b\rangle > 0$) they are equal only if $\Re\lambda = |\lambda\,|$; which, in turn, will be so if and only if λ is real nonnegative.

8. Suppose that $x, y \in K$ and $0 \le t \le 1$. Then

$$\begin{aligned}\|tx + (1-t)y\| &\le \|tx\| + \|(1-t)y\| = t\|x\| + (1-t)\|y\| \\ &\le t.1 + (1-t).1 = 1\,,\end{aligned}$$

so that $tx + (1-t)y \in K$ too. This proves that K is convex.

Certainly $0 \in K$ and 0 is not an extreme point (take any $a \in K$ with $a \ne 0$, and then $-a \in K$ and $0 = \frac{1}{2}a + \frac{1}{2}(-a)$, so that 0 is an interior point of the segment from $-a$ to a.

Next, suppose $b \in K$ and $0 < \|b\| < 1$; then $\left\|\dfrac{b}{\|b\|}\right\| = 1$, so that $\dfrac{b}{\|b\|} \in K$. But $b = \|b\|\dfrac{b}{\|b\|} + (1 - \|b\|)0$; b is an interior point of the segment from 0 to $b/\|b\|$.

We have now shown that an extreme point of K, if any exist, must have norm exactly 1. Let $a, b \in K$, $a \ne b$, and $t \in (0,1)$. In that case $\|a\| \le 1$, $\|b\| \le 1$, and

$$1 = \|ta + (1-t)b\| \le t\|a\| + (1-t)\|b\| \le 1$$

is only possible (granted that $t > 0$ and $1 - t > 0$) if $\|a\| = \|b\| = 1$. However, then

$$\begin{aligned}\|ta + (1-t)b\|^2 &= \langle ta + (1-t)b, ta + (1-t)b\rangle \\ &= t^2\langle a, a\rangle + t(1-t)(\langle a, b\rangle + \langle b, a\rangle) + (1-t)^2\langle b, b\rangle \\ &= t^2\langle a, a\rangle + 2t(1-t)\Re\langle a, b\rangle + (1-t)^2\langle b, b\rangle \\ &\le t^2\|a\|^2 + 2t(1-t)\|a\|\|b\| + (1-t)^2\|b\|^2 \\ &= (t\|a\| + (1-t)\|b\|)^2 = 1\,,\end{aligned}$$

where the intermediate inequality derives from the Cauchy-Schwartz inequality; but it must, therefore, be an equality, $\Re\langle a, b\rangle = \|a\|\|b\|$ (recall again that $t(1-t) > 0$). This, in turn, is only possible if a and b are positive scalar multiples of each other (question 7 above). But, if $\lambda > 0$ and $\lambda a = b$ and $\|a\| = \|b\| = 1$, necessarily $\lambda = 1$ and $a = b$.

This proves that, if $\|x\| = 1$, x cannot be an interior point of any line segment in K; any point of the "sphere" $\{x \in H : \|x\| = 1\}$ is an extreme point of K.

9. There are several ways of doing this. One is to imitate, with obvious alterations, the *proof* of the standard result 13.3. But it is more economical to use 13.3 itself, as follows. Let E be the direct sum Hilbert space $H \oplus H$ or $H \times H$, with the inner product

$$\langle (x, y), (a, b)\rangle_{\oplus} := \langle x, a\rangle + \langle y, b\rangle$$

(which is trivially an inner product, and induces the norm $\|(x, y)\|_{\oplus} := \sqrt{\|x\|^2 + \|y\|^2}$). Then $C_1 \times C_2$ is closed and convex in $H \times H$, and $(x, x) \notin C_1 \times C_2$. By 13.3, there is some $(y_1, y_2) \in C_1 \times C_2$ such that the distance from (x, x) to (y_1, y_2) is least possible (for elements of $C_1 \times C_2$). This is precisely what is wanted.

10. *(a)* Suppose $x \notin \Lambda(A)$. By 13.4 and 13.3, there is a unique $z \in \Lambda(A)$ such that $0 \ne y := x - z \perp \Lambda(A)$. Then $y \in \Lambda(A)^\perp \subseteq A^\perp$ and $\langle x, y\rangle = \langle x - z, y\rangle = \langle y, y\rangle > 0$, as $\langle z, y\rangle = 0$. So $x \notin (A^\perp)^\perp$. This proves $(A^\perp)^\perp \subseteq \Lambda(A)$. (7)

If $p \in A^\perp$, then $A \subseteq \{p\}^\perp$, which is a closed linear subspace of H. So $\Lambda(A) \subseteq \{p\}^\perp$. This implies that $p \in \Lambda(A)^\perp$, and (p being arbitrary in $A^\perp$) that $A^\perp \subseteq \Lambda(A)^\perp$. And now

$$\Lambda(A) \subseteq (\Lambda(A)^\perp)^\perp \subseteq (A^\perp)^\perp .$$

This gives the opposite inclusion to (7). [That $Q \subseteq Q^{\perp\perp}$ is always trivially true — why?]

(b) Suppose that the finite-dimensional subspace E of H has an orthonormal basis $\{p_1, p_2, \ldots, p_n\}$, such as may be constructed by the Gram-Schmidt process. For any $y \in H$,

$$y - \sum_{k=1}^n \langle y, p_k\rangle p_k \in E^\perp ,$$

since it is trivially orthogonal to each of the p_k. Suppose that $x \in E^{\perp\perp}$; then, for any y,

$$\left\langle x - \sum_{k=1}^n \langle x, p_k\rangle p_k, y \right\rangle = \left\langle x, y - \sum_{k=1}^n \langle y, p_k\rangle p_k \right\rangle = 0$$

by simple algebraic manipulation. However, as y may be arbitrary, it follows that

$$x = \sum_{k=1}^n \langle x, p_k\rangle p_k \in E .$$

This shows that $E^{\perp\perp} \subseteq E$, and, from *(a)*, that E must be closed.