Reading

KLM stands for Kaye, Laflamme, & Mosca. Here we use $\langle \cdot, \cdot \rangle$ to denote the inner product; it is the same as $\langle \cdot | \cdot \rangle$.

- KLM, Chapter 1
- KLM, Chapter 2, Sections 2.1–2.2
- Course notes, first three lectures
- Referring to the Gram-Schmidt procedure described in class, here is a proof that $\langle v_k, u_k \rangle > 0$, for all $1 \leq k \leq n$: We start with the definition of $v_k$:

$$v_k = \frac{u_k - \sum_{j<k} \langle v_j, u_k \rangle v_j}{N},$$

where $N := \left| u_k - \sum_{j<k} \langle v_j, u_k \rangle v_j \right|$ is the norm of the numerator. Note that $N > 0$, because $u_k$ is outside the span of $\{v_j\}_{j<k}$, making the numerator a nonzero vector. Solving for $u_k$, we get

$$u_k = N v_k + \sum_{j<k} \langle v_j, u_k \rangle v_j,$$

and thus

$$\langle v_k, u_k \rangle = N \langle v_k, v_k \rangle + \sum_{j<k} \langle v_j, u_k \rangle \langle v_k, v_j \rangle = N > 0.$$

(Why? By linearity and orthonormality of the $v_j$.)

Exercises

- KLM, Exercise 1.5.1
- Course notes, Exercises 2.1, 2.4, 2.6, 2.7
Show that any orthogonal set of nonzero vectors is linearly independent. [Hint: Let $v$ be any linear combination of such vectors, and consider $\langle v, v \rangle$. You’ll need the fact that $\langle \cdot, \cdot \rangle$ is positive definite.]

Let $B = \langle b_1, \ldots, b_n \rangle$ be an orthonormal basis for a Hilbert space $\mathcal{H}$, and let $x$ be a vector in $\mathcal{H}$. Show that, for $1 \leq i \leq n$, $\langle b_i, x \rangle$ is the coefficient of $b_i$ in the unique expansion of $x$ with respect to $B$. Use this to show that if $\langle b_i, u \rangle = \langle b_i, v \rangle$ for all $i$, then $u = v$.

Extra credit: KLM Exercise 2.3.1, Course notes, Exercises 2.5, 3.1

Note: The notations used in the book and in my course notes are pretty similar, but beware of the following two discrepancies:

1. I denote the adjoint of an operator $A$ by $A^*$, which is the general convention in pure math. The KLM book uses $A^\dagger$ to denote the adjoint of $A$, which is used more in the physics community. (Discrepancies like this are inevitable in such a multidisciplinary subject.)

2. I don’t introduce Dirac notation until later, whereas KLM uses it early on.