18.369 Problem Set 1 Solutions

Problem 1: Adjoints and operators

(a) If O is a matrix then $Oh' = \sum_n O_{mn} h'_n$ by the usual matrix-vector product. Then the dot product of h with Oh' is given by $\sum_m h^*_m (\sum_n O_{mn} h'_n) = \sum_n (\sum_m O^*_{mn} h_m)^* h'_n$, which is the same thing as the dot product of $O^\dagger h$ with h' where $O^\dagger$ is the conjugate transpose of O. Thus, interpreting $\dagger$ as the conjugate transpose in this finite-dimensional case is consistent with the abstract definitions given in class.

(b) We simply rely on two required properties of inner products: conjugacy $\langle \phi, \psi \rangle = \langle \psi, \phi \rangle^*$, and linearity $\langle \phi, \alpha\psi \rangle = \alpha \langle \phi, \psi \rangle$. Then, $\langle \phi, \hat{O}\psi \rangle = \langle \phi, o\psi \rangle = o\langle \phi, \psi \rangle = o\langle \psi, \phi \rangle^* = \langle \psi, o^*\phi \rangle^* = \langle o^*\phi, \psi \rangle$. Thus, inspecting the definition of $\hat{O}^\dagger$, we see that $\hat{O}^\dagger = o^*$.

(c) If $\hat{O}$ is unitary and we send $\psi \to \hat{O}\psi$ and $\phi \to \hat{O}\phi$, then $\langle \phi, \psi \rangle \to \langle \hat{O}\phi, \hat{O}\psi \rangle = \langle \hat{O}^\dagger \hat{O}\phi, \psi \rangle = \langle \phi, \psi \rangle$, and thus inner products are preserved. Consider now two eigenstates $\hat{O}\psi_1 = \lambda_1$ and $\hat{O}\psi_2 = \lambda_2$. Then $\langle \psi_1, \psi_2 \rangle = \langle \hat{O}\psi_1, \hat{O}\psi_2 \rangle = \lambda_1^* \lambda_2^* \langle \psi_1, \psi_2 \rangle$, and thus $(\lambda_1^* \lambda_2 - 1)\langle \psi_1, \psi_2 \rangle = 0$. There are three cases, just like for Hermitian operators. If $\psi_1 = \psi_2$, then we must have $\lambda_1^* \lambda_1 = 1 = |\lambda_1|^2$, and thus the eigenvalues have unit magnitude. This also implies that $\lambda_1^* = 1/\lambda_1$. If $\lambda_1 \neq \lambda_2$, then $(\lambda_1^* \lambda_2 - 1) = (\lambda_2/\lambda_1 - 1) \neq 0$, and therefore $\langle \psi_1, \psi_2 \rangle = 0$ and the eigenstates are orthogonal. If $\lambda_1 = \lambda_2$ but ψ_1 and ψ_2 are linearly independent (degenerate eigenstates), then we can form orthogonal linear combinations by Gram-Schmidt.

(d) Take two states ϕ and ψ, and consider the inner product. Then $\langle \phi, \psi \rangle = \langle \phi, \hat{O}^{-1}\hat{O}\psi \rangle = \langle (\hat{O}^{-1})^\dagger \phi, \hat{O}\psi \rangle = \langle \hat{O}^\dagger (\hat{O}^{-1})^\dagger \phi, \psi \rangle$. Since this must be true for all ψ and ϕ, then $\hat{O}^\dagger (\hat{O}^{-1})^\dagger$ must be the identity operation, and thus $(\hat{O}^{-1})^\dagger = (\hat{O}^\dagger)^{-1}$. Q.E.D.

Problem 2: Completeness

(a) Prove that the eigenvectors H_n of a *finite-dimensional* Hermitian operator $\hat{O}$ (a $d \times d$ matrix) are *complete*:

  (i) Given a Hermitian matrix O, the eigenvalues λ are the roots of the characteristic polynomial $\det |O - \lambda \mathbf{1}|$. This polynomial is of degree $d > 0$, which means that it has at least one root λ_1, and correspondingly at this point of zero determininant there is a solution H_1 of $OH_1 = \lambda_1 H_1$. (λ_1 must, of course, be real because O is Hermitian.)

  (ii) If H has $\langle H_1, H \rangle = 0$, then $\hat{O}H$ is also orthogonal to H_1: $\langle H_1, \hat{O}H \rangle = \langle \hat{O}H_1, H \rangle = \lambda_1 \langle H_1, H \rangle$, since $\hat{O} = \hat{O}^\dagger$ for a Hermitian $\hat{O}$. Thus, $\hat{O}$ transforms states orthogonal to H_1 into states orthogonal to H_1. The set of states orthogonal to H_1 has dimension $d - 1$, and we have shown that $\hat{O}$ is a linear operator within this space—thus, $\hat{O}$ within this space can be represented by a Hermitian matrix $O^{(1)}$ of size $(d-1) \times (d-1)$.

In particular, let S be a similarity transform that rotates $(1, 0, \cdots, 0)$[1] into H_1 and $(0, x_2, x_3, \cdots, x_d)$ into V_1 (i.e. the columns of S are H_1 followed by a basis for V_1—without loss of generality, we can use an orthonormal basis for V_1). Then, $\tilde{O} = S^{-1}OS$ has the same eigenvalues as O and[2] has the block-diagonal form:

$$\tilde{O} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & O^{(1)} \end{pmatrix}$$

where $O^{(1)}$ is $(d-1) \times (d-1)$ (the zeros in the first row come from our lemma, that vectors in V_1 are transformed by O into vectors in V_1; the zeros in the first column come from the fact that H_1 must be transformed into itself). If $\mathbf{v}$ is an eigenvector of $O^{(1)}$, then $S \cdot (0, \mathbf{v})$ is clearly an eigenvector of O orthogonal to H_1. We also need to show that $O^{(1)}$ is Hermitian—this follows automatically if $\tilde{O}$ is Hermitian, which in turn follows if S is unitary (so that $\tilde{O} = S^\dagger OS = \tilde{O}^\dagger$). S is automatically unitary, however, as long as we chose an orthonormal basis for V_1, since $S^\dagger S$ is then simply the dot products of the columns of S, which gives the identity matrix $\mathbf{1}$ and thus $S^\dagger = S^{-1}$.

(iii) We proceed inductively: if we assume that all $(d-1) \times (d-1)$ Hermitian operators have a complete orthonormal set of eigenvectors $\mathbf{v}$, then from above we can construct a complete orthonormal basis of eigenvectors for a $d \times d$ Hermitian matrix O: $S \cdot (0, \mathbf{v})$ are eigenvectors of O and must form a complete basis for V_1, which combined with H_1 forms a complete basis for the whole space (which can be chosen orthonormal as we showed in class). The base case of the induction is $d = 1$, where O must be simply a real number, and the corresponding eigenvector (1) is clearly complete.

(b) Consider the 2×2 matrix $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$, with $a \neq 0$. The characteristic polynomial is $\lambda^2 - 2\lambda + 1$, which obviously has a double root of $\lambda = 1$. The corresponding eigenvectors $\begin{pmatrix} x \\ y \end{pmatrix}$ must satisfy $x + ay = \lambda x = x$, $y = \lambda y = y$, and thus are satisfied for any $\begin{pmatrix} x \\ 0 \end{pmatrix}$. But these are *all* the eigenvectors of the matrix, and they obviously do not span the whole xy plane (they only span the x axis). Q.E.D.

(c) If a particular Hermitian operator did *not* have a complete basis of eigenfunctions, that means that there are some solutions that cannot be represented as a linear combination of eigenfunctions. However, when we discretize the operator into a finite Hermitian problem on a computer, the resulting finite matrix

[1] By $(a, b, c, \cdots)$ here, we mean a *column* vector.

[2] A similarity transform $v' = Sv$ preserves the eigenvalues, since if $Ov = \lambda v$ then $(S\tilde{O}S^{-1})v' = \lambda SS^{-1}v'$ by inspection.

is always diagonalizable from above. We would potentially, therefore, have a problem simulating such a problem accurately on a computer—the computer's eigenfunctions are always a complete basis for the discretized problem, so it is unclear how they could converge to the eigenfunctions of the continuous problem. Conversely, if we have a problem that can be simulated accurately on a computer, one might suspect that it is not one of those pathological operators lacking a complete basis of eigenfunctions. This is not a proof, of course, but the fact is that most interesting eigenproblems that arise in physics *can* be simulated well on a computer and *do* have a complete basis of eigenfunctions (the *spectral theorem* applies to them), and the point is that these two properties seem to be related.

Problem 3: Maxwell eigenproblems

(a) To eliminate $\mathbf{H}$, we start with Faraday's law $\nabla \times \mathbf{E} = i\frac{\omega}{c}\mathbf{H}$ and take the curl of both sides. We obtain:

$$\nabla \times \nabla \times \mathbf{E} = \frac{\omega^2}{c^2}\varepsilon\mathbf{E}.$$

If we divide both sides by ε, we get the form of a linear eigenproblem but the operator $\frac{1}{\varepsilon}\nabla\times\nabla\times$ is not Hermitian—integrating by parts, we find that its adjoint is $\nabla \times \nabla \times \frac{1}{\varepsilon}\cdot$, which is not the same operator unless the $\frac{1}{\varepsilon}$ commutes with the curls, which only happens if ε is a constant. However, if we leave it in the form above we have a generalized Hermitian problem with $\hat{A} = \nabla \times \nabla\times$ and $\hat{B} = \varepsilon$. $\hat{A}$ is Hermitian for the same reason that $\hat{\Theta}$ was (it *is* $\hat{\Theta}$ for $\varepsilon = 1$), and $\hat{B}$ is Hermitian as long as ε is real (so that $\mathbf{H}_1^* \cdot \varepsilon\mathbf{H}_2 = (\varepsilon\mathbf{H}_1)^* \cdot \mathbf{H}_2$).

(b) The proof follows the same lines as in class. Take two eigenstates ψ_1 and ψ_2 (where $\hat{A}\psi_{1,2} = \lambda_{1,2}\hat{B}\psi_{1,2}$), and consider $\langle\psi_2, \hat{A}\psi_1\rangle$. Since $\hat{A}$ is Hermitian, we can operate it to the left or to the right in the inner product, and get $\lambda_1\langle\psi_2, \hat{B}\psi_1\rangle = \lambda_2^*\langle\hat{B}\psi_2, \psi_1\rangle = \lambda_2^*\langle\psi_2, \hat{B}\psi_1\rangle$, or $(\lambda_2^* - \lambda_1)\langle\psi_2, \hat{B}\psi_1\rangle = 0$. There are three cases. First, if $\psi_1 = \psi_2$, then we must have $\lambda_1 = \lambda_1^*$ (real eigenvalues), since $\langle\psi_1, \hat{B}\psi_1\rangle > 0$ by definition if $\hat{B}$ is positive definite. Second, if $\lambda_1 \neq \lambda_2$ then we must have $\langle\psi_2, \hat{B}\psi_1\rangle = 0$, which is our modified orthogonality condition. Finally, if $\lambda_1 = \lambda_2$ but ψ_1 and ψ_2 are linearly independent, then we can form a linear combination that is orthogonal (since any linear combination still is an eigenvector); e.g. by Gram-Schmidt

$$\psi_2 \to \psi_2 - \psi_1\frac{\langle\psi_2, \hat{B}\psi_1\rangle}{\langle\psi_1, \hat{B}\psi_1\rangle},$$

where we have again relied on the fact that $\hat{B}$ is positive definite (so that we can divide by $\langle\psi_1, \hat{B}\psi_1\rangle$). This is certainly true for $\hat{B} = \varepsilon$, since $\langle\mathbf{E}, \varepsilon\mathbf{E}\rangle = \int \varepsilon|\mathbf{E}|^2 > 0$ for all $\mathbf{E} \neq 0$ as long as $\varepsilon > 0$ as we assumed in class.

(c) If $\mu \neq 1$ then we have $\mathbf{B} = \mu\mathbf{H} \neq \mathbf{H}$, and when we eliminate $\mathbf{E}$ or $\mathbf{H}$ from Maxwell's equations we get:

$$\nabla \times \frac{1}{\varepsilon}\nabla \times \mathbf{H} = \frac{\omega^2}{c^2}\mu\mathbf{H}$$

$$\nabla \times \frac{1}{\mu} \nabla \times \mathbf{E} = \frac{\omega^2}{c^2} \varepsilon \mathbf{E}$$

with the constraints $\nabla \cdot \varepsilon \mathbf{E} = 0$ and $\nabla \cdot \mu \mathbf{H} = 0$. These are both generalized Hermitian eigenproblems (since μ and $\nabla \times \frac{1}{\mu} \nabla \times$ are both Hermitian operators for the same reason ε and $\nabla \times \frac{1}{\varepsilon} \nabla \times$ were); we can't make them ordinary Hermitian eigenproblems for the same reason as in the $\mathbf{E}$ eigenproblem above, except in the trivial case of μ or ε constant. Thus, the eigenvalues are real and the eigenstates are orthogonal through μ and ε, respectively, as proved above. To prove that ω is real, we consider an eigenstate $\mathbf{H}$. Then $\langle \mathbf{H}, \hat{\Theta} \mathbf{H} \rangle = \frac{\omega^2}{c^2} \langle \mathbf{H}, \mu \mathbf{H} \rangle$ and we must have $\omega^2 \geq 0$ since $\hat{\Theta}$ is positive semi-definite (from class) and μ is positive definite (for the same reason ε was, above). The $\mathbf{E}$ eigenproblem has real ω for the same reason (except that μ and ε are swapped).

(d) Consider the $\mathbf{H}$ eigenproblem. (To even get this linear eigenproblem, we must immediately require ε to be an invertible matrix, and of course require ε and μ to be independent of ω or the field strength.) For the right-hand operator μ to be Hermitian, we require $\int \mathbf{H}_1^* \cdot \mu \mathbf{H}_2 = \int (\mu \mathbf{H}_1)^* \cdot \mathbf{H}_2$ for all $\mathbf{H}_1$and $\mathbf{H}_2$, which implies that $\mathbf{H}_1^* \cdot \mu \mathbf{H}_2 = (\mu \mathbf{H}_1)^* \cdot \mathbf{H}_2$. Thus, we require the 3×3 $\mu(\mathbf{x})$ matrix to be itself Hermitian at every $\mathbf{x}$ (i.e., equal to its conjugate transpose from problem 1). Similarly, for $\hat{\Theta}$ to be Hermitian we require $\int \mathbf{F}_1^* \cdot \varepsilon^{-1} \mathbf{F}_2 = \int (\varepsilon^{-1} \mathbf{F}_1)^* \cdot \mathbf{F}_2$ where $\mathbf{F} = \nabla \times \mathbf{H}$, so that we can move the ε^{-1} over to the left side of the inner product, and thus $\varepsilon^{-1}(\mathbf{x})$ must be Hermitian at every $\mathbf{x}$. From problem 1, this implies that $\varepsilon(\mathbf{x})$ is also Hermitian. Finally, to get real eigenvalues we saw from above that we must have μ positive definite ($\int \mathbf{H}^* \cdot \mu \mathbf{H} > 0$ for $\mathbf{H} \neq 0$); since this must be true for all $\mathbf{H}$ then $\mu(\mathbf{x})$ at each point must be a positive-definite 3×3 matrix (positive eigenvalues). Similarly, $\hat{\Theta}$ must be positive semi-definite, which implies that $\varepsilon^{-1}(\mathbf{x})$ is positive semi-definite (non-negative eigenvalues), but since it has to be invertible we must have $\varepsilon(\mathbf{x})$ positive definite (zero eigenvalues would make it singular). To sum up, we must have $\varepsilon(\mathbf{x})$ and $\mu(\mathbf{x})$ being positive-definite Hermitian matrices at every $\mathbf{x}$. (The proof for the $\mathbf{E}$ eigenproblem is identical.)

Actually, there are a couple other possibilities. In part (b), we showed that if $\hat{B}$ is positive-definite it leads to real eigenvalues etc. The same properties, however, hold if $\hat{B}$ is *negative*-definite, and if *both* $\hat{A}$ and $\hat{B}$ are negative-definite we still get real, *positive* eigenvalues. Thus, another possibility is for ε and μ to be Hermitian *negative*-definite matrices. (For a scalar $\varepsilon < 0$ and $\mu < 0$, this leads to materials with a *negative* real index of refraction $n = -\sqrt{\varepsilon \mu}$!) Furthermore, ε and μ could both be *anti*-Hermitian instead of Hermitian (i.e., $\varepsilon^\dagger = -\varepsilon$ and $\mu^\dagger = -\mu$), although I'm not aware of any physical examples of this. More generally, for any complex number z, if we replace ε and μ by $z\varepsilon$ and μ/z, respectively, then ω is unchanged (e.g. making $z = i$ gives anti-Hermitian matrices).