18.369 Problem Set 1 Solutions

Problem 1: Adjoints and operators (5+10+5 points)

(a) If † is conjugate-transpose of a matrix or vector, we are just using the usual linear-algebra rule that $(AB)^\dagger = B^\dagger A^\dagger$, hence $\langle h, Oh'\rangle = h^\dagger(Oh') = (O^\dagger h)^\dagger h' = \langle O^\dagger h, h'\rangle$ for the Euclidean inner product.

More explicitly, if h is a column-vector and we let $h^\dagger$ be its conjugate transpose, then $h^\dagger$ is a row vector and $h^\dagger h' = \sum_m h_m^* h'_m = \langle h, h'\rangle$ by the usual row-times-column multiplicaton rule. 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, as claimed in class, the abstract $\langle h, Oh'\rangle = \langle O^\dagger h, h'\rangle$ definition of $O^\dagger$ implies the usual conjugate transpose definition of $O^\dagger$ for matrices.

(b) If $\hat{O}$ is unitary and we send $u \to \hat{O}u$ and $v \to \hat{O}v$, then $\langle u, v\rangle \to \langle u, \hat{O}^\dagger \hat{O} v\rangle = \langle u, v\rangle$, and thus inner products are preserved. Consider now two eigensolutions $\hat{O}u_1 = \lambda_1 u_1$ and $\hat{O}u_2 = \lambda_2 u_2$. Then $\langle u_1, \hat{O}^\dagger \hat{O} u_2\rangle = \langle u_1, u_2\rangle$ by the unitarity of $\hat{O}$ and $\langle u_1, \hat{O}^\dagger \hat{O} u_2\rangle = \langle \hat{O}u_1, \hat{O}u_2\rangle = \lambda_1^* \lambda_2 \langle u_1, u_2\rangle$ by the eigenvector property (where we let $\hat{O}^\dagger$ act to the left, and conjugate the eigenvalue when we factor it out, as in class). Combining these two expressions, we have $(\lambda_1^*\lambda_2 - 1)\langle u_1, u_2\rangle = 0$. There are three cases, just like for Hermitian operators. If $u_1 = u_2$, 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 $\langle u_1, u_2\rangle = 0$ and the eigenvectors are orthogonal. If $\lambda_1 = \lambda_2$ but have linearly independent $u_1 \neq u_2$ (degenerate eigenvectors, i.e. geometric multiplicity > 1), then we can form orthogonal linear combinations (e.g. via Gram–Schmidt).

(c) Take two vectors u and v, and consider their inner product. Then $\langle u, \hat{O}^{-1}\hat{O}v\rangle = \langle u, v\rangle$. By definition of the adjoint, however, if we move first $\hat{O}^{-1}$ and then $\hat{O}$ to act to the left, then we get $\langle u, v\rangle = \langle \hat{O}^\dagger(\hat{O}^{-1})^\dagger u, v\rangle$. For this to be true for all u and v, we must have $\hat{O}^\dagger(\hat{O}^{-1})^\dagger = \mathbf{1}$ and thus $(\hat{O}^{-1})^\dagger = (\hat{O}^\dagger)^{-1}$. Q.E.D.

Problem 2: Maxwell eigenproblems (5+5+5+5+5 points)

(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 under the usual inner product $\langle \mathbf{E}_1, \mathbf{E}_2\rangle = \int \mathbf{E}_1^* \cdot \mathbf{E}_2$—integrating by parts as in class, assuming boundary conditions such that the boundary terms vanish, 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. [Alternatively, we could simply quote the Hermitian results from class once we prove part (c).] Consider two eigensolutions u_1 and u_2 (where $\hat{A}u = \lambda\hat{B}u$, and $u \neq 0$), and take $\langle u_2, \hat{A}u_1\rangle$. Since $\hat{A}$ is Hermitian, we can operate it to the left or to the right in the inner product, and get $\lambda_2^* \langle u_2, \hat{B}u_1\rangle = \lambda_1 \langle u_2, \hat{B}u_1\rangle$, or $(\lambda_2^* - \lambda_1)\langle u_2, \hat{B}u_1\rangle = 0$. There are three cases. First, if $u_1 = u_2$ then we must have $\lambda_1 = \lambda_1^*$ (real eigenvalues), since $\langle u_1, \hat{B}u_1\rangle > 0$ by definition if $\hat{B}$ is positive definite. Second, if $\lambda_1 \neq \lambda_2$ then we must have $\langle u_2, \hat{B}u_1\rangle = 0$, which is our modified orthogonality condition. Finally, if $\lambda_1 = \lambda_2$ but $u_1 \neq u_2$, then we can form a linear combination that is orthogonal (since any linear combination still is an eigenvector); e.g.

$$u_2 \to u_2 - u_1 \frac{\langle u_2, \hat{B}u_1\rangle}{\langle u_1, \hat{B}u_1\rangle},$$

where we have again relied on the fact that $\hat{B}$ is positive definite (so that we can divide by $\langle u_1, \hat{B}u_1\rangle$). This is certainly true for $\hat{B} = \varepsilon$, since $\langle E, \hat{B}E\rangle = \int \varepsilon|\mathbf{E}|^2 > 0$ for all $\mathbf{E} \neq 0$ (almost everywhere) as long as we have a real $\varepsilon > 0$ as we required in class.

(c) First, let us verify that $\langle \mathbf{E}, \mathbf{E}'\rangle_B = \langle \mathbf{E}, \hat{B}\mathbf{E}'\rangle$ is indeed an inner product. Because $\hat{B}$ is self-adjoint, we have $\langle \mathbf{E}', \mathbf{E}\rangle_B = \langle \mathbf{E}', \hat{B}\mathbf{E}\rangle = \langle \hat{B}\mathbf{E}', \mathbf{E}\rangle = \langle \mathbf{E}, \hat{B}\mathbf{E}'\rangle^* = \langle \mathbf{E}, \mathbf{E}'\rangle_B^*$. Bilinearity follows from bilinearity of $\langle\cdot,\cdot\rangle$ and linearity of $\hat{B}$. Positivity $\langle \mathbf{E}, \mathbf{E}\rangle_B = \langle \mathbf{E}, \hat{B}\mathbf{E}\rangle > 0$ except for $\mathbf{E} = 0$ (almost everywhere) follows from positive-definiteness of $\hat{B}$. All good!

Now, Hermiticity of $\hat{B}^{-1}\hat{A}$ follows almost trivially from Hermiticity of $\hat{A}$ and $\hat{B}$: $\langle \mathbf{E}, \hat{B}^{-1}\hat{A}\mathbf{E}'\rangle_B = \langle \mathbf{E}, \cancel{\hat{B}\hat{B}^{-1}}\hat{A}\mathbf{E}'\rangle = \langle \hat{A}\mathbf{E}, \mathbf{E}'\rangle = \langle \hat{A}\mathbf{E}, \hat{B}^{-1}\hat{B}\mathbf{E}'\rangle = \langle \hat{B}^{-1}\hat{A}\mathbf{E}, \hat{B}\mathbf{E}'\rangle = \langle \hat{B}^{-1}\hat{A}\mathbf{E}, \mathbf{E}'\rangle_B$, where we have used the fact, from problem 1, that Hermiticity of $\hat{B}$ implies Hermiticity of $\hat{B}^{-1}$. Q.E.D.

(d) 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). 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 eigenfunction H. Then $\langle H, \hat{\Theta}H\rangle = \frac{\omega^2}{c^2}\langle H, \mu H\rangle$ 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).

Alternatively, as in part (c), we can write them as ordinary Hermitian eigenproblems with a modified inner product, e.g. $\frac{1}{\varepsilon}\nabla \times \frac{1}{\mu}\nabla \times \mathbf{E} = \frac{\omega^2}{c^2}\mathbf{E}$, where $\frac{1}{\varepsilon}\nabla \times \frac{1}{\mu}\nabla\times$ is Hermitian and positive semidefinite under the $\langle \mathbf{E}, \mathbf{E}'\rangle_B = \int \mathbf{E}^* \cdot \varepsilon\mathbf{E}\prime$ inner product as above. The results then follow.

(e) 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 $\mathbf{x}$ (that is, equal to its conjugate transpose, from problem 1). (Technically, these requirements hold "almost everywhere" rather than at every point, but as usual I will gloss over this distinction as well as other technical restrictions on the functions to exclude crazy functions.) 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 (almost) every $\mathbf{x}$. (The analysis for the $\mathbf{E}$ eigenproblem is identical.)

Optional: Technically, there are a few 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 so-called "left-handed 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 $\mu^\dagger = -\mu$). More generally, for any complex number z, if we replace ε and μ by $z\varepsilon$ and μ/z, then ω is unchanged (e.g. making $z = i$ gives anti-Hermitian matrices).

Problem 3: Dispersion (5+5+10+5 points)

(a) If $\ddot{\mathbf{P}}$ is proportional to acceleration, then it is also proportional to the net force by Newton's law, and hence the specified forces yield:

$$\ddot{\mathbf{P}} = -\alpha_0\mathbf{P} - \gamma_0\dot{\mathbf{P}} + \sigma_0\mathbf{E}$$

for some unknown proportionality constants $\alpha_0, \gamma_0, \sigma_0$. In the absence of friction ($\gamma_0 = 0$) and the driving force ($\mathbf{E} = 0$ or $\sigma_0 = 0$), this equation $\ddot{\mathbf{P}} = -\alpha_0\mathbf{P}$ is that of a simple harmonic oscillator, which has solutions $e^{-i\omega_0 t}$ for $\omega_0 = \pm\sqrt{\alpha_0}$. Hence, in terms of the oscillation frequency ω_0, the equation becomes

$$\boxed{\ddot{\mathbf{P}} = -\omega_0^2\mathbf{P} - \gamma_0\dot{\mathbf{P}} + \sigma_0\mathbf{E}}.$$

(Note that $\omega_0, \gamma_0, \sigma_0$ may depend on $\mathbf{x}$! That is, you might have different materials at different points in space.) Both γ_0 and σ_0 must be ≥ 0 (for the system to be passive: absorbing rather than supplying energy).

(b) For time-harmonic solutions $\mathbf{E}(\mathbf{x})e^{-i\omega t}$ and $\mathbf{P}(\mathbf{x})e^{-i\omega t}$, the equation becomes $-\omega^2\mathbf{P} = -\omega_0^2\mathbf{P} + i\omega\gamma_0\mathbf{P} + \sigma_0\mathbf{E}$, and hence

$$\mathbf{P} = \underbrace{\frac{\sigma_0}{(\omega_0^2 - \omega^2) - i\omega\gamma_0}}_{\chi_e(\omega,\mathbf{x})}\mathbf{E},$$

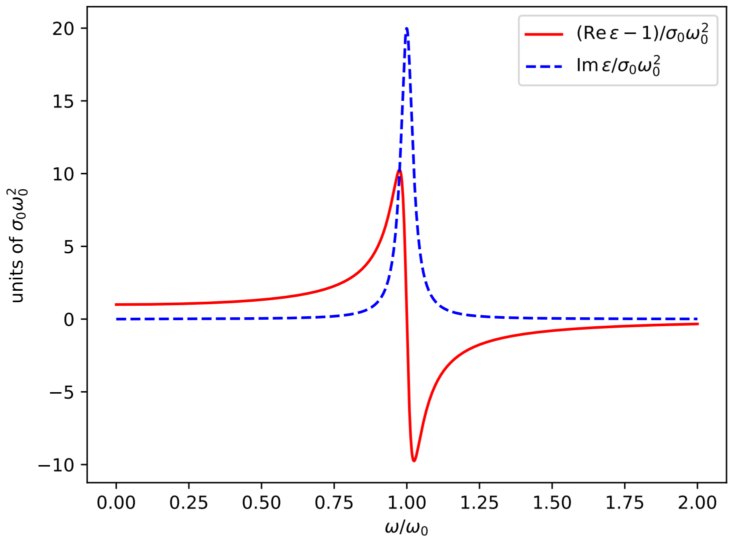

Figure 1: Plot of $\mathrm{Re}\,\varepsilon - 1$ and $\mathrm{Im}\,\varepsilon$, both in "nondimensionalized" units of $\sigma_0\omega_0^2$ for problem 3(c), with a relatively small loss $\gamma_0 = 0.05\omega_0$, as a function of ω/ω_0.

where we have identified the proportionality factor as $\chi_e(\omega)$ (the electric susceptibility) from class. Hence (from class),

$$\boxed{\varepsilon(\omega, \mathbf{x}) = 1 + \chi_e(\omega, \mathbf{x}) = 1 + \frac{\sigma_0}{(\omega_0^2 - \omega^2) - i\omega\gamma_0}}.$$

where the $\mathbf{x}$ dependence (if any) comes from $\omega_0(\mathbf{x}), \gamma_0(\mathbf{x}), \sigma_0(\mathbf{x})$. For example, if we have a single kind of polarizable material in some regions of space and vacuum ($\varepsilon = 1$) in other regions, then ω_0 and γ_0 will be constants while σ_0 will be 0 in vacuum and some positive constant in the material.

(c) The real and imaginary parts of ε are

$$\mathrm{Re}\,\varepsilon = 1 + \frac{\sigma_0(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + (\omega\gamma_0)^2},$$

$$\mathrm{Im}\,\varepsilon = \frac{\sigma_0(\omega\gamma_0)}{(\omega_0^2 - \omega^2)^2 + (\omega\gamma_0)^2}.$$

Let's plot this in the regime of *small damping*, where $\gamma_0 \ll \omega_0$, e.g. for $\gamma_0 = 0.05\omega_0$ (note that this is dimensionally correct—both have units of frequency). To make things nice and dimensionless, we'll plot $(\mathrm{Re}\,\varepsilon - 1)/\sigma_0\omega_0^2$ and $\mathrm{Im}\,\varepsilon/\sigma_0\omega_0^2$, both versus ω/ω_0.

By inspection of the formula, when γ_0 is small, it's pretty easy[1] to see that $\operatorname{Im}\varepsilon$ peaks in the vicinity of $\omega = \pm\omega_0$, with a peak amplitude proportional to $\sigma_0/(\omega_0\gamma_0)$ and a peak width proportional to γ_0. That is, it peaks near the resonance frequency, with a "linewidth" proportional to the loss.

We can solve for the peak more precisely by setting $\frac{d}{d\omega}\operatorname{Im}\varepsilon = 0$ and solving for ω. After some tedious algebra, keeping terms only to lowest order in γ_0/ω_0, I find that there are peaks at $\omega = \pm\omega_0\left[1 + \frac{\gamma_0^2}{8}\right]$ (plus terms proportional to γ_0^4 or larger).

Similarly, we can solve for the peak width (the "linewidth") more precisely by realizing that near the peak for small γ_0 (narrow peaks), $|\omega|\gamma_0 \approx \omega_0\gamma_0$. With this approximation, we can compute the "width at half maximum" — i.e., over what width 2Γ is $\operatorname{Im}\varepsilon$ at least 1/2 its peak? That is, what Γ solves $\operatorname{Im}\varepsilon(\omega_0 \pm \Gamma) \approx \frac{1}{2}\operatorname{Im}\varepsilon(\omega_0) \approx \sigma_0/2\omega_0\gamma_0$? After a little algebra, we find that this occurs at $\omega_0^2 - (\omega_0 \pm \Gamma)^2 = \pm\omega_0\gamma_0 \approx \mp 2\omega_0\Gamma$, and hence we find that the linewidth 2Γ is actually γ_0 at half-maximum! (Again, we could be even more precise by doing an expansion in powers of γ_0/ω_0.)

From the plot, notice also that $\operatorname{Re}\varepsilon - 1$ *oscillates in sign* around the resonance frequency ω_0. If this oscillation is large enough (if $\sigma_0/\omega_0\gamma_0$ is large enough), then $\operatorname{Re}\varepsilon$ can actually flip sign and become < 0 — it is totally possible to have a negative permittivity near a strong resonance (although this typically comes hand in hand with large absorption)!

(d) From the formula in the previous part, it is obvious that

$$\omega \operatorname{Im}\varepsilon = \frac{\sigma_0(\omega^2\gamma_0)}{(\omega_0^2 - \omega^2)^2 + (\omega\gamma_0)^2} \geq 0$$

for all ω if and only if $\sigma_0\gamma_0 \geq 0$. Since $\gamma_0 \geq 0$ is required for our harmonic-oscillator equations for $\mathbf{P}$ to decay rather than grow (the friction force $-\gamma_0\dot{\mathbf{P}}$ must *oppose* the motion), it follows that we must have $\sigma_0 \geq 0$ as claimed above.

We will see later that $\sigma_0 \geq 0$ corresponds to the electric field $\mathbf{E}$ "doing net work" *on* the polarization $\mathbf{P}$ (rather than vice versa), which is called *passivity* of the medium, and that this ultimately implies $\omega \operatorname{Im}\varepsilon \geq 0$ much more generally.

[1] For small $\gamma_0 \ll \omega_0$, the denominator of $\operatorname{Im}\varepsilon$ becomes very small when $(\omega_0^2 - \omega^2)^2 \lesssim (\omega\gamma_0)^2$, i.e. for $\omega \approx \omega_0$. Furthermore, in the vicinity of $\omega = \omega_0$ (similarly for $-\omega_0$), we can approximate $(\omega\gamma_0)^2 \approx (\omega_0\gamma_0)^2$ and $(\omega_0^2 - \omega^2)^2 = (\omega_0 - \omega)^2(\omega_0 + \omega)^2 \approx 4\omega_0^2(\omega_0 - \omega)^2$, so that the latter is smaller than the former for $|\omega - \omega_0| \lesssim \gamma_0/2$. This corresponds to a peak for $\omega \in [\omega_0 - \gamma_0/2, \omega_0 + \gamma_0/2]$, or a peak width $\sim \gamma_0$.