Peeter Joot
peeterjoot@protonmail.com

Jackson's electrostatic self energy analysis

Motivation   I was reading my Jackson [1], which characteristically had the statement "the [...] integral *can easily be shown* to have the value 4π", in a discussion of electrostatic energy and self energy. After a few attempts and a couple of pages of calculations, I figured out how this can be easily shown.

Context   Let me walk through the context that leads to the "easy" integral, and then the evaluation of that integral. Unlike my older copy of Jackson, I'll do this in SI units.

The starting point is a statement that the work done (potential energy) of one charge q_i in a set of n charges, where that charge is brought to its position $\mathbf{x}_i$ from infinity, is

$$W_i = q_i \Phi(\mathbf{x}_i), \tag{1.1}$$

where the potential energy due to the rest of the charge configuration is

$$\Phi(\mathbf{x}_i) = \frac{1}{4\pi\epsilon} \sum_{i \neq j} \frac{q_j}{|\mathbf{x}_i - \mathbf{x}_j|}. \tag{1.2}$$

This means that the total potential energy, making sure not to double count, to move all the charges in from infinity is

$$W = \frac{1}{4\pi\epsilon} \sum_{1 \leq i < j \leq n} \frac{q_i q_j}{|\mathbf{x}_i - \mathbf{x}_j|}. \tag{1.3}$$

This sum over all unique pairs is somewhat unwieldy, so it can be adjusted by explicitly double counting with a corresponding divide by two

$$W = \frac{1}{2}\frac{1}{4\pi\epsilon} \sum_{1 \leq i \neq j \leq n} \frac{q_i q_j}{|\mathbf{x}_i - \mathbf{x}_j|}. \tag{1.4}$$

The point that causes the trouble later is the continuum equivalent to this relationship, which is

$$W = \frac{1}{8\pi\epsilon} \int \frac{\rho(\mathbf{x})\rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3\mathbf{x} d^3\mathbf{x}', \tag{1.5}$$

or

$$W = \frac{1}{2} \int \rho(\mathbf{x})\Phi(\mathbf{x}) d^3\mathbf{x}. \tag{1.6}$$

1

There's a subtlety here that is often passed over. When the charge densities represent point charges $\rho(\mathbf{x}) = q\delta^3(\mathbf{x} - \mathbf{x}')$ are located at, notice that this integral equivalent is evaluated over all space, including the spaces that the charges that the charges are located at.

Ignoring that subtlety, this potential energy can be expressed in terms of the electric field, and then integrated by parts

$$
\begin{aligned}
W &= \frac{1}{2} \int (\boldsymbol{\nabla} \cdot (\epsilon \mathbf{E})) \Phi(\mathbf{x}) d^3\mathbf{x} \\
&= \frac{\epsilon}{2} \int (\boldsymbol{\nabla} \cdot (\mathbf{E}\Phi) - (\boldsymbol{\nabla}\Phi) \cdot \mathbf{E}) \, d^3\mathbf{x} \\
&= \frac{\epsilon}{2} \oint dA \hat{\mathbf{n}} \cdot (\mathbf{E}\Phi) + \frac{\epsilon}{2} \int \mathbf{E} \cdot \mathbf{E} d^3\mathbf{x}.
\end{aligned}
\tag{1.7}
$$

The presumption is that $\mathbf{E}\Phi$ falls off as the bounds of the integration volume tends to infinity. That leaves us with an energy density proportional to the square of the field

$$
w = \frac{\epsilon}{2}\mathbf{E}^2.
\tag{1.8}
$$

Inconsistency   It's here that Jackson points out the inconsistency between eq. (1.8) and the original discrete analogue eq. (1.4) that this was based on. The energy density is positive definite, whereas the discrete potential energy can be negative if there is a difference in the sign of the charges.

Here Jackson uses a two particle charge distribution to help resolve this conundrum. For a superposition $\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2$, we have

$$
\mathbf{E} = \frac{1}{4\pi\epsilon} \frac{q_1(\mathbf{x} - \mathbf{x}_1)}{|\mathbf{x} - \mathbf{x}_1|^3} + \frac{1}{4\pi\epsilon} \frac{q_2(\mathbf{x} - \mathbf{x}_2)}{|\mathbf{x} - \mathbf{x}_2|^3},
\tag{1.9}
$$

so the energy density is

$$
w = \frac{1}{32\pi^2\epsilon} \frac{q_1^2}{|\mathbf{x} - \mathbf{x}_1|^4} + \frac{1}{32\pi^2\epsilon} \frac{q_2^2}{|\mathbf{x} - \mathbf{x}_2|^4} + 2\frac{q_1 q_2}{32\pi^2\epsilon} \frac{(\mathbf{x} - \mathbf{x}_1)}{|\mathbf{x} - \mathbf{x}_1|^3} \cdot \frac{(\mathbf{x} - \mathbf{x}_2)}{|\mathbf{x} - \mathbf{x}_2|^3}.
\tag{1.10}
$$

The discrete potential had only an interaction energy, whereas the potential from this squared field has an interaction energy plus two self energy terms. Those two strictly positive self energy terms are what forces this field energy positive, independent of the sign of the interaction energy density. Jackson makes a change of variables of the form

$$
\begin{aligned}
\boldsymbol{\rho} &= (\mathbf{x} - \mathbf{x}_1)/R \\
R &= |\mathbf{x}_1 - \mathbf{x}_2| \\
\hat{\mathbf{n}} &= (\mathbf{x}_1 - \mathbf{x}_2)/R,
\end{aligned}
\tag{1.11}
$$

for which we find

$$
\mathbf{x} = \mathbf{x}_1 + R\boldsymbol{\rho},
\tag{1.12}
$$

so

$$
\mathbf{x} - \mathbf{x}_2 = \mathbf{x}_1 - \mathbf{x}_2 + R\boldsymbol{\rho}R(\hat{\mathbf{n}} + \boldsymbol{\rho}),
\tag{1.13}
$$

2

and

$$d^3\mathbf{x} = R^3 d^3\boldsymbol{\rho},$$

(1.14)

so the total interaction energy is

$$
\begin{aligned}
W_{\text{int}} &= \frac{q_1 q_2}{16\pi^2\epsilon} \int d^3\mathbf{x} \frac{(\mathbf{x} - \mathbf{x}_1)}{|\mathbf{x} - \mathbf{x}_1|^3} \cdot \frac{(\mathbf{x} - \mathbf{x}_2)}{|\mathbf{x} - \mathbf{x}_2|^3} \\
&= \frac{q_1 q_2}{16\pi^2\epsilon} \int R^3 d^3\boldsymbol{\rho} \frac{R\boldsymbol{\rho}}{R^3|\boldsymbol{\rho}|^3} \cdot \frac{R(\hat{\mathbf{n}} + \boldsymbol{\rho})}{R^3|\hat{\mathbf{n}} + \boldsymbol{\rho}|^3} \\
&= \frac{q_1 q_2}{16\pi^2\epsilon R} \int d^3\boldsymbol{\rho} \frac{\boldsymbol{\rho}}{|\boldsymbol{\rho}|^3} \cdot \frac{(\hat{\mathbf{n}} + \boldsymbol{\rho})}{|\hat{\mathbf{n}} + \boldsymbol{\rho}|^3}.
\end{aligned}
$$

(1.15)

Evaluating this integral is what Jackson calls easy. The technique required is to express the integrand in terms of gradients in the $\boldsymbol{\rho}$ coordinate system

$$
\begin{aligned}
\int d^3\boldsymbol{\rho} \frac{\boldsymbol{\rho}}{|\boldsymbol{\rho}|^3} \cdot \frac{(\hat{\mathbf{n}} + \boldsymbol{\rho})}{|\hat{\mathbf{n}} + \boldsymbol{\rho}|^3} &= \int d^3\boldsymbol{\rho} \left(-\boldsymbol{\nabla}_{\boldsymbol{\rho}} \frac{1}{|\boldsymbol{\rho}|} \right) \cdot \left(-\boldsymbol{\nabla}_{\boldsymbol{\rho}} \frac{1}{|\hat{\mathbf{n}} + \boldsymbol{\rho}|} \right) \\
&= \int d^3\boldsymbol{\rho} \left(\boldsymbol{\nabla}_{\boldsymbol{\rho}} \frac{1}{|\boldsymbol{\rho}|} \right) \cdot \left(\boldsymbol{\nabla}_{\boldsymbol{\rho}} \frac{1}{|\hat{\mathbf{n}} + \boldsymbol{\rho}|} \right).
\end{aligned}
$$

(1.16)

I found it somewhat non-trivial to find the exact form of the chain rule that is required to simplify this integral, but after some trial and error, figured it out by working backwards from

$$\boldsymbol{\nabla}_{\boldsymbol{\rho}}^2 \frac{1}{|\boldsymbol{\rho}||\hat{\mathbf{n}} + \boldsymbol{\rho}|} = \boldsymbol{\nabla}_{\boldsymbol{\rho}} \cdot \left(\frac{1}{|\boldsymbol{\rho}|} \boldsymbol{\nabla}_{\boldsymbol{\rho}} \frac{1}{|\hat{\mathbf{n}} + \boldsymbol{\rho}|} \right) + \boldsymbol{\nabla}_{\boldsymbol{\rho}} \cdot \left(\frac{1}{|\hat{\mathbf{n}} + \boldsymbol{\rho}|} \boldsymbol{\nabla}_{\boldsymbol{\rho}} \frac{1}{|\boldsymbol{\rho}|} \right).$$

(1.17)

In integral form this is

$$
\begin{aligned}
\oint dA' \hat{\mathbf{n}}' \cdot \boldsymbol{\nabla}_{\boldsymbol{\rho}} \frac{1}{|\boldsymbol{\rho}||\hat{\mathbf{n}} + \boldsymbol{\rho}|} &= \int d^3\boldsymbol{\rho}' \boldsymbol{\nabla}_{\boldsymbol{\rho}'} \cdot \left(\frac{1}{|\boldsymbol{\rho}' - \hat{\mathbf{n}}|} \boldsymbol{\nabla}_{\boldsymbol{\rho}'} \frac{1}{|\boldsymbol{\rho}'|} \right) + \int d^3\boldsymbol{\rho} \boldsymbol{\nabla}_{\boldsymbol{\rho}} \cdot \left(\frac{1}{|\hat{\mathbf{n}} + \boldsymbol{\rho}|} \boldsymbol{\nabla}_{\boldsymbol{\rho}} \frac{1}{|\boldsymbol{\rho}|} \right) \\
&= \int d^3\boldsymbol{\rho}' \left(\boldsymbol{\nabla}_{\boldsymbol{\rho}'} \frac{1}{|\boldsymbol{\rho}' - \hat{\mathbf{n}}|} \cdot \boldsymbol{\nabla}_{\boldsymbol{\rho}'} \frac{1}{|\boldsymbol{\rho}'|} \right) + \int d^3\boldsymbol{\rho}' \frac{1}{|\boldsymbol{\rho}' - \hat{\mathbf{n}}|} \boldsymbol{\nabla}_{\boldsymbol{\rho}'}^2 \frac{1}{|\boldsymbol{\rho}'|} \\
&\quad + \int d^3\boldsymbol{\rho} \left(\boldsymbol{\nabla}_{\boldsymbol{\rho}} \frac{1}{|\hat{\mathbf{n}} + \boldsymbol{\rho}|} \right) \cdot \boldsymbol{\nabla}_{\boldsymbol{\rho}} \frac{1}{|\boldsymbol{\rho}|} + \int d^3\boldsymbol{\rho} \frac{1}{|\hat{\mathbf{n}} + \boldsymbol{\rho}|} \boldsymbol{\nabla}_{\boldsymbol{\rho}}^2 \frac{1}{|\boldsymbol{\rho}|} \\
&= 2\int d^3\boldsymbol{\rho} \left(\boldsymbol{\nabla}_{\boldsymbol{\rho}} \frac{1}{|\hat{\mathbf{n}} + \boldsymbol{\rho}|} \right) \cdot \boldsymbol{\nabla}_{\boldsymbol{\rho}} \frac{1}{|\boldsymbol{\rho}|} \\
&\quad - 4\pi \int d^3\boldsymbol{\rho}' \frac{1}{|\boldsymbol{\rho}' - \hat{\mathbf{n}}|} \delta^3(\boldsymbol{\rho}') - 4\pi \int d^3\boldsymbol{\rho} \frac{1}{|\boldsymbol{\rho} + \hat{\mathbf{n}}|} \delta^3(\boldsymbol{\rho}) \\
&= 2\int d^3\boldsymbol{\rho} \left(\boldsymbol{\nabla}_{\boldsymbol{\rho}} \frac{1}{|\hat{\mathbf{n}} + \boldsymbol{\rho}|} \right) \cdot \boldsymbol{\nabla}_{\boldsymbol{\rho}} \frac{1}{|\boldsymbol{\rho}|} - 8\pi.
\end{aligned}
$$

(1.18)

This used the Laplacian representation of the delta function $\delta^3(\mathbf{x}) = -(1/4\pi)\boldsymbol{\nabla}^2(1/|\mathbf{x}|)$. Back-substitution gives

3

$$\int d^3\rho \, \frac{\rho}{|\rho|^3} \cdot \frac{(\hat{\mathbf{n}} + \rho)}{|\hat{\mathbf{n}} + \rho|^3} = 4\pi + \oint dA' \hat{\mathbf{n}}' \cdot \nabla_\rho \frac{1}{|\rho||\hat{\mathbf{n}} + \rho|}. \tag{1.19}$$

We can argue that this last integral tends to zero, since

$$\begin{aligned}
\oint dA' \hat{\mathbf{n}}' \cdot \nabla_\rho \frac{1}{|\rho||\hat{\mathbf{n}} + \rho|} &= \oint dA' \hat{\mathbf{n}}' \cdot \left(\left(\nabla_\rho \frac{1}{|\rho|} \right) \frac{1}{|\hat{\mathbf{n}} + \rho|} + \frac{1}{|\rho|} \left(\nabla_\rho \frac{1}{|\hat{\mathbf{n}} + \rho|} \right) \right) \\
&= -\oint dA' \hat{\mathbf{n}}' \cdot \left(\frac{\rho}{\frac{1}{|\rho|}^3} \frac{1}{|\hat{\mathbf{n}} + \rho|} + \frac{1}{|\rho|} \frac{(\rho + \hat{\mathbf{n}})}{|\hat{\mathbf{n}} + \rho|^3} \right) \\
&= -\oint dA' \frac{1}{|\rho||\rho + \hat{\mathbf{n}}|} \left(\frac{\hat{\mathbf{n}}' \cdot \rho}{|\rho|^2} + \frac{\hat{\mathbf{n}}' \cdot (\rho + \hat{\mathbf{n}})}{|\rho + \hat{\mathbf{n}}|^2} \right).
\end{aligned} \tag{1.20}$$

The integrand in this surface integral is of $O(1/\rho^3)$ so tends to zero on an infinite surface in the ρ coordinate system. This completes the "easy" integral, leaving

$$\int d^3\rho \, \frac{\rho}{|\rho|^3} \cdot \frac{(\hat{\mathbf{n}} + \rho)}{|\hat{\mathbf{n}} + \rho|^3} = 4\pi. \tag{1.21}$$

The total field energy can now be expressed as a sum of the self energies and the interaction energy

$$W = \frac{1}{32\pi^2\epsilon} \int d^3\mathbf{x} \frac{q_1^2}{|\mathbf{x} - \mathbf{x}_1|^4} + \frac{1}{32\pi^2\epsilon} \int d^3\mathbf{x} \frac{q_2^2}{|\mathbf{x} - \mathbf{x}_2|^4} + \frac{1}{4\pi\epsilon} \frac{q_1 q_2}{|\mathbf{x}_1 - \mathbf{x}_2|}. \tag{1.22}$$

The interaction energy is exactly the potential energies for the two particles, the this total energy in the field is biased in the positive direction by the pair of self energies. It is interesting that the energy obtained from integrating the field energy density contains such self energy terms, but I don't know exactly what to make of them at this point in time.

Bibliography

[1] JD Jackson. *Classical Electrodynamics*. John Wiley and Sons, 2nd edition, 1975. 1