1. Any spin orbital can be written as the product of a spatial orbital and a spin function. We call these the spatial and spin components of the orbital, respectively. We index the spin orbitals by lowercase letters and the spatial component of that orbital by the uppercase letter.

To show the equality of the two sums, it suffices to show a component-wise equality. In particular, we show equalities between (i, j, a, b) and (i, j, b, a) terms and the (I, J, A, B) and (I, J, B, A). Note that the $\epsilon$ denominator between the two is trivial because $\epsilon$ only depends on the combination of spatial components and whether a spatial component is occupied or unoccupied. Spin and whether A is before B are irrelevant to $\epsilon$.

Our lefthand expressions are of the form $\Sigma\frac{1}{4}|\langle ij||ab\rangle|^2$, so consider the possible values of $\langle ij||ab\rangle$. For this to not be 0, the spins on each "side" should biject. Otherwise, an alpha pairs with a beta, and the integral goes to 0. Suppose that both sides have two of one spin and none of the other (all alphas or all betas). Integrating out spin yields $\langle IJ|AB\rangle - \langle IJ|BA\rangle$. Suppose both sides have one of each spin. If they line up initially, this yields $\langle IJ|AB\rangle$, but if they do not align, this yields $-\langle IJ|BA\rangle$. Here, only one term survives, but each of these possibilities can be attained twice.

Noticing that $|\langle ij||ab\rangle|^2 = \langle ij||ab\rangle\langle ab||ij\rangle$ due to the symmetry rules for two-electron operators, we can take advantage of our previous findings to rewrite $|\langle ij||ab\rangle|^2$, but in cases.

In the first case, we assume that our spinorbitals are distinct, so all terms exist:

$\frac{1}{2}[(\langle IJ|AB\rangle-\langle IJ|BA\rangle)(\langle AB|IJ\rangle-\langle AB|JI\rangle)+\langle IJ|AB\rangle\langle AB|IJ\rangle+\langle IJ|BA\rangle\langle AB|JI\rangle] =$
$\frac{1}{2}[\langle IJ|AB\rangle\langle AB|IJ\rangle-\langle IJ|AB\rangle\langle AB|JI\rangle-\langle IJ|BA\rangle\langle AB|IJ\rangle+\langle IJ|BA\rangle\langle AB|JI\rangle+$
$\langle IJ|AB\rangle\langle AB|IJ\rangle + \langle IJ|BA\rangle\langle AB|JI\rangle] =$
$\frac{1}{2}[2\langle IJ|AB\rangle\langle AB|IJ\rangle-\langle IJ|AB\rangle\langle AB|JI\rangle-\langle IJ|BA\rangle\langle AB|IJ\rangle+2\langle IJ|BA\rangle\langle AB|JI\rangle]$

We can add our (i, j, b, a) to our (i, j, a, b) so:
$\frac{1}{2}[2\langle IJ|AB\rangle\langle AB|IJ\rangle-\langle IJ|AB\rangle\langle AB|JI\rangle-\langle IJ|BA\rangle\langle AB|IJ\rangle+2\langle IJ|BA\rangle\langle AB|JI\rangle]+$
$\frac{1}{2}[2\langle IJ|BA\rangle\langle BA|IJ\rangle-\langle IJ|BA\rangle\langle BA|JI\rangle-\langle IJ|AB\rangle\langle BA|IJ\rangle+2\langle IJ|AB\rangle\langle BA|JI\rangle] =$
$\frac{1}{2}[2\langle IJ|AB\rangle\langle AB|IJ\rangle-\langle IJ|AB\rangle\langle AB|JI\rangle-\langle IJ|BA\rangle\langle AB|IJ\rangle+2\langle IJ|BA\rangle\langle AB|JI\rangle]+$
$\frac{1}{2}[2\langle IJ|BA\rangle\langle AB|JI\rangle-\langle IJ|BA\rangle\langle AB|IJ\rangle-\langle IJ|AB\rangle\langle AB|JI\rangle+2\langle IJ|AB\rangle\langle AB|IJ\rangle] =$
$2\langle IJ|BA\rangle\langle AB|JI\rangle-\langle IJ|BA\rangle\langle AB|IJ\rangle-\langle IJ|AB\rangle\langle AB|JI\rangle+2\langle IJ|AB\rangle\langle AB|IJ\rangle =$
$\langle IJ|BA\rangle(2\langle AB|JI\rangle - \langle AB|IJ\rangle) + \langle IJ|AB\rangle(2\langle AB|IJ\rangle - \langle AB|JI\rangle) =$
$\langle IJ|AB\rangle(2\langle AB|IJ\rangle - \langle AB|JI\rangle) + \langle IJ|BA\rangle(2\langle IJ|BA\rangle - \langle AB|IJ\rangle)$

We next consider the possibility of indistinct spatial orbitals. This eliminates the case of all four orbitals having the same spin, as this would mean our two spinorbitals with the same spatial component have the same spin component as well, which contradicts Pauli exclusion. We assume that A and B are indistinct.

Should A and B be distinct while I and J are not, the same proof works with minimal modifications.

$\frac{1}{2}[\langle IJ|AB\rangle\langle AB|IJ\rangle + \langle IJ|BA\rangle\langle AB|JI\rangle] =$
$\frac{1}{2}[\langle IJ|AA\rangle\langle AA|IJ\rangle + \langle IJ|AA\rangle\langle AA|JI\rangle] =$
$\frac{1}{2}[\langle IJ|AA\rangle\langle AA|IJ\rangle + \langle IJ|AA\rangle\langle AA|IJ\rangle] = \langle IJ|AA\rangle\langle AA|IJ\rangle =$
$\langle IJ|AA\rangle(2\langle AA|IJ\rangle - \langle AA|IJ\rangle) =$
$\langle IJ|AA\rangle(2\langle AA|IJ\rangle - \langle AA|JI\rangle) =$
$\langle IJ|AB\rangle(2\langle AB|IJ\rangle - \langle AB|JI\rangle).$

Given our right hand sides of, excepting the trivial energy denominator, $\langle\langle IJ|AB\rangle(2\langle AB|IJ\rangle - \langle AB|JI\rangle)\rangle$, the proof is complete. (For the case of distinct spins, remember to swap A and B in the second term.)

2. The desired equation comes from substituting in the expansions and remembering that the inner product is linear in the second term but conjugate linear in the first.

Scaling means how the execution time changes in the limit of large-data. Imagine enclosing everything but the first for loop in a function. That would be called $n_o$ times, so that first loop tells us the scaling is $n_o$ times whatever the inner function's scaling is. If we iterate this, we get $n_o^2 n_v^2 n_{b.f.}^4$. Because $n_{b.f.} = n_o + n_v$, and $n_o$ is constant for a given molecule, $n_v$ increases as $n_{b.f.}$ does. This means that our scaling for a given molecule is about $n_{b.f.}^6$.

```
sum = nocc + nvir
g1 = g2 = g3 = g4 zeros((sum, sum, sum, sum))

for mu in ao_basis
    for nu in ao_basis
        for rho in ao_basis
            for sigma in ao_basis
                for I in occupied
                    g1[I, nu, rho, sigma] = g[mu, nu, rho, sigma] * C[mu, I].conjugate()
                for J in occupied
                    # We pull some elements from g1 that weren't defined in the
                    # above loop, but that's fine.
                    g2[mu, J, rho, sigma] = g1[mu, nu, rho, sigma] * C[nu, J].conjugate()
                for A in unoccupied
                    g3[mu, nu, A, sigma] = g2[mu, nu, rho, sigma] * C[rho, A]
                for B in unoccupied
                    g4[mu, nu, rho, B] = g3[mu, nu, rho, sigma] * C[sigma, B]
```

Using the same loop logic as before, this scales to the fifth. This trick works because while the computer "needs" to know mu, nu, rho, and sigma at the same time, it never needs to know all of the I, J, A, B at once. It can work fine knowing just one at a time.