To do overlap screening, we need to estimate the value of certain overlap integrals. Consider two Gaussians a distance d apart, with type l, we consider the product

$$n_1 n_2 x^{l_1} \exp\left(-\alpha_1 x^2\right)\cdot\left(x-d\right)^{l_2} \exp\left(-\alpha_2\left(x-d\right)^2\right) \tag{1.1}$$

where the normalization factors are:

$$n=\sqrt{\left(\frac{2\alpha}{\pi}\right)^{\frac{1}{2}}\frac{\left(4\alpha\right)^{\ell}}{(2\ell-1)!!}} \tag{1.2}$$

According to the Gaussian product theorem, we can rewrite this as:

$$\sqrt{\frac{2}{\pi}\sqrt{\alpha_1\alpha_2}\left(\frac{\left(4\alpha_1\right)^{l_1}\left(4\alpha_2\right)^{l_2}}{(2l_1-1)!!(2l_2-1)!!}\right)}x^{l_1}\left(x-d\right)^{l_2}\exp\left[-\left(\alpha_1+\alpha_2\right)\left(x-\frac{\alpha_2 d}{\alpha_1+\alpha_2}\right)^2\right]\exp\left(-\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2}d^2\right) \tag{1.3}$$

The first Gaussian factor is narrowly peaked near $x=\frac{\alpha_2}{\alpha_1+\alpha_2}d$. It is convenient then to choose

$$\sqrt{\frac{2}{\pi}\sqrt{\alpha_1\alpha_2}\left(\frac{\left(4\alpha_1\right)^{l_1}\left(4\alpha_2\right)^{l_2}}{(2l_1-1)!!(2l_2-1)!!}\right)}\left(\frac{\alpha_2 d}{\alpha_1+\alpha_2}\right)^{l_1}\left(\frac{\alpha_2 d}{\alpha_1+\alpha_2}-d\right)^{l_2}\left[\int_{-\infty}^{\infty}\exp\left(-\left(\alpha_1+\alpha_2\right)x^2\right)dx\right]\exp\left(-\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2}d^2\right)$$

as an estimate for the overlap. (We assume that the x axis is aligned with the segment that connects the Gaussian centers. This is the “most evil case.”)

Simplifying we have:

$$\begin{aligned}
&\sqrt{\frac{2}{\pi}\sqrt{\alpha_1\alpha_2}\left(\frac{\left(4\alpha_1\right)^{l_1}\left(4\alpha_2\right)^{l_2}}{(2l_1-1)!!(2l_2-1)!!}\right)}\left(\frac{d}{\alpha_1+\alpha_2}\right)^{l_1+l_2}\left(\alpha_2\right)^{l_1}\left(\alpha_1\right)^{l_2}\left[\int_{-\infty}^{\infty}\exp\left(-\left(\alpha_1+\alpha_2\right)x^2\right)dx\right]\exp\left(-\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2}d^2\right)\\
&=\sqrt{\frac{2}{\pi}\sqrt{\alpha_1\alpha_2}\left(\frac{\left(4\alpha_1\right)^{l_1}\left(4\alpha_2\right)^{l_2}}{(2l_1-1)!!(2l_2-1)!!}\right)}\frac{\left(d\alpha_2\right)^{l_1}\left(d\alpha_1\right)^{l_2}\pi^{\frac{1}{2}}}{\left(\alpha_1+\alpha_2\right)^{l_1+l_2+\frac{1}{2}}}\exp\left(-\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2}d^2\right)\\
&=2^{l_1+l_2}\sqrt{2}\sqrt{\frac{\alpha_1^{l_1+\frac{1}{2}}\alpha_2^{l_2+\frac{1}{2}}}{(2l_1-1)!!(2l_2-1)!!}}\frac{\alpha_1^{l_1}\alpha_2^{l_2}}{\left(\alpha_1+\alpha_2\right)^{l_1+l_2+\frac{1}{2}}}d^{l_1+l_2}\exp\left(-\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2}d^2\right)\\
&=\frac{2^{l_1+l_2+\frac{1}{2}}}{\sqrt{(2l_1-1)!!(2l_2-1)!!}}\frac{\alpha_1^{\frac{3}{2}l_1+\frac{1}{4}}\alpha_2^{\frac{3}{2}l_1+\frac{1}{4}}}{\left(\alpha_1+\alpha_2\right)^{l_1+l_2+\frac{1}{2}}}d^{l_1+l_2}\exp\left(-\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2}d^2\right)\\
&=\frac{1}{\sqrt{(2l_1-1)!!(2l_2-1)!!}}\frac{\alpha_1^{\frac{3}{2}l_1+\frac{1}{4}}\alpha_2^{\frac{3}{2}l_1+\frac{1}{4}}}{\left(\frac{\alpha_1+\alpha_2}{2}\right)^{l_1+l_2+\frac{1}{2}}}d^{l_1+l_2}\exp\left(-\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2}d^2\right)
\end{aligned} \tag{1.4}$$

If I rearrange this, I have:

$$\frac{\left(d\sqrt{\alpha_1}\right)^{l_1}\left(d\sqrt{\alpha_2}\right)^{l_2}}{\sqrt{(2l_1-1)!!(2l_2-1)!!}}\frac{\alpha_1^{l_1+\frac{1}{4}}\alpha_2^{l_2+\frac{1}{4}}}{\left(\frac{\alpha_1+\alpha_2}{2}\right)^{l_1+l_2+\frac{1}{2}}}\exp\left(-\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2}d^2\right) \tag{1.5}$$

The first part of this is relatively "constant" because $d\sqrt{\alpha_1}$ is the "natural length scale" for that Gaussian. I.e., the Gaussian will be diffuse (large *d*, small $\sqrt{\alpha_1}$) or sharp (small *d*, large $\sqrt{\alpha_1}$) but these two things go together. The second term is less than 1, but maybe not *much* less than one in some cases, due to the geometric-arithmetic mean inequality. So the exponential factor is the dominant term.

At the *simplest* level then, one screens based on the exponential factor alone. That is, one can ignore all terms for which $\exp\left(-\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2}d^2\right)$ is sufficiently small. This can be treated with a distance cutoff. One first computes a cutoff for atom-type and shell. So,

```
for every atom-type A
    for every atom-type B
        for every shell s on atom-type A
            find the smallest exponent in the shell, α_min(α_s,A)
        for every shell t on atom-type B
           find the smallest exponent in the shell, α_min(t,B)
```

$$d_{As;Bt} > \sqrt{-\frac{\alpha_{\min(\alpha_{s,A})}+\alpha_{\min(\alpha_{t,B})}}{\alpha_{\min(\alpha_{s,A})}\alpha_{\min(\alpha_{t,B})}}\ln(\varepsilon)}$$ `!shell-and-atom-dependent-cutoff`

One also computes a cutoff for every atom type,

$$d_{AB} = \max_{s,t} d_{As,Bt} \tag{1.6}$$

Now, when one is looping over atoms in overlap/kinetic-energy/moment/electron-nuclear attraction integrals, one can ignore atoms that are further apart than d_{AB} and ignore shells whose centers are further apart than $d_{As,Bt}$. A typical value for ε is about 10^{-20}.

Paul W. Ayers
March 3, 2019