$$R = \frac{D_M}{2x_0 + 2x_A}$$

where x_0 is emittance contribution and x_A is aberration contribution.

$$x_0 = \epsilon/\theta, x_A = D_M(\kappa\theta)^n$$

where n is lowest uncorrected order. E.g. if corrected to octupole, $n = 4$.

This has an optimum at some θ which we plug into R to get:

$$R \propto \left(\frac{D_M}{\epsilon}\right)^{\frac{n}{n+1}}$$

So it's easy to see that correcting to higher order yields diminishing returns. On the other hand, because of "feed-down" effects, difficulty of tuning rises very steeply with n.

D_M is proportional to the size of the magnet and cost is proportional to the cube of size. Therefore

$$\text{Cost} \propto R^{3+3/n}$$