640:192:03 | **How big is a ball?** | **9/23/2005**

How big is a ball?

I want to compute the volume of the unit ball in $\mathbb{R}^n$, n-dimensional space. The unit ball will be the collection of points whose distance to the origin is less than or equal to 1. I should define the terms used in this paragraph before starting the computation.

The computation because it is a significant application of a reduction formula obtained using integration by parts. It also provides some interesting and almost surely unsuspected asymptotic information about higher dimensional geometry. A few other interesting intellectual morsels are gained, including some hints about the factorial function.

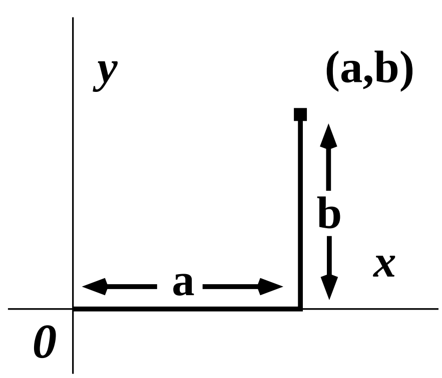
Learning more strengthens intuition!

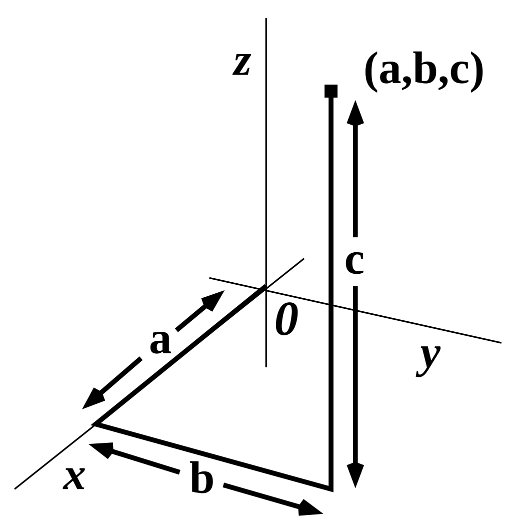
• *Some formulas to start with* •

Definition $\mathbb{R}^n$ is the collection of $n-$tuples of real numbers. A typical point in $\mathbb{R}^n$, P, has coordinates $(x_1, x_2, \ldots, x_n)$.

Math is made up of definitions and deductions from the definitions, all structured to make sense of examples given by applications and by past work. It makes a great deal of sense to investigate the appropriateness of any definition, especially when seeking to link it to ideas already known!

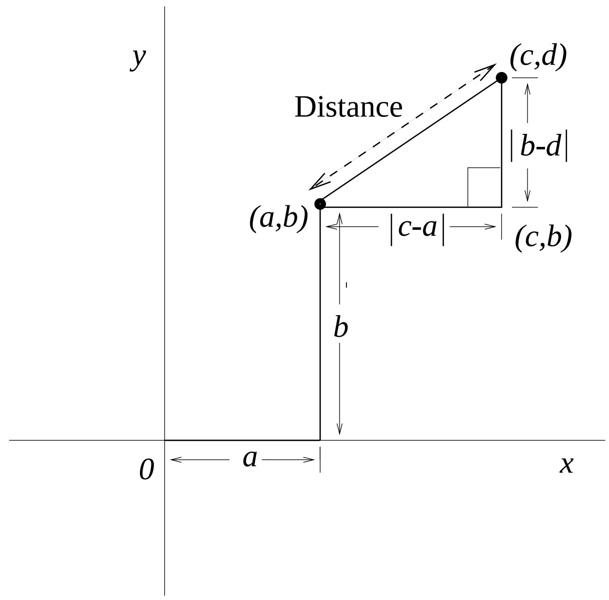
Discussion of the definition $\mathbb{R}^2$ is the Euclidean plane, as described to us by Descartes and similar thinkers: every point "is" a pair of numbers (a, b). The point can be located by moving a directed distance a along a first axis (the $x-$axis) and then moving perpendicular to that a directed distance b, as indicated below. $\mathbb{R}^3$ is three-dimensional space, and a point is a triple of numbers, (a, b, c). Such a point can be located using three mutually perpendicular axes as indicated in the picture below.



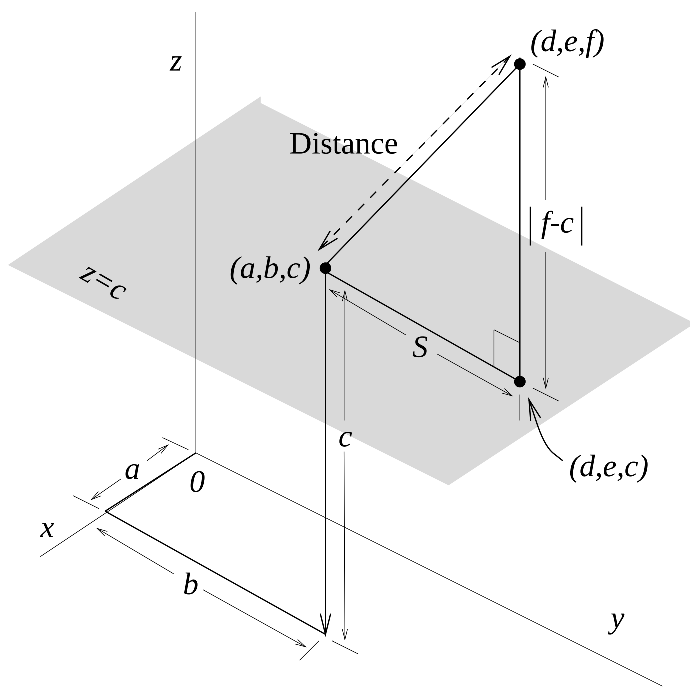


Definition If $P = (x_1, x_2, \ldots, x_n)$ and $Q = (y_1, y_2, \ldots, y_n)$ are two points in $\mathbb{R}^n$, then the *distance from P to Q* is $\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}$.

Discussion of the definition Of course if P and Q are points in $\mathbb{R}^2$, say $P = (a, b)$ and $Q = (c, d)$ respectively, then the distance we've known for years between P and Q is $\sqrt{(a-c)^2 + (b-d)^2}$. This "is" the distance because the Theorem of Pythagoras (see the picture below) combined with the observation that $\sqrt{(|c-a|)^2 + (|b-d|)^2} = \sqrt{(c-a)^2 + (b-d)^2}$ convince us that it is. It is Euclidean distance.



If P and Q are points of $\mathbb{R}^3$, with $P = (a, b, c)$ and $Q = (d, e, f)$, then a slightly more elaborate argument combined with a more complex picture is necessary. Here's the picture:



"Drop" a perpendicular from Q to the plane described by $z = c$. The perpendicular intersects the plane in the point whose coordinates at (d, e, c). The distance S in the diagram is a two-dimensional distance between points (a, b, c) and (d, e, c) given by $S = \sqrt{(e-b)^2 + (d-a)^2}$. The Theorem of Pythagoras again applies to give us the distance between P and Q as

$$\sqrt{S^2 + (|f-c|)^2} = \sqrt{(e-b)^2 + (d-a)^2 + (f-c)^2}$$

Notice first that we begin to run out of letters for two points after a few dimensions, so we should call the coordinates $x_1, x_2, \ldots$, etc. Second, pictures become difficult to draw (see Edward Abbot's *Flatland* for an implicit discussion about how to draw pictures in higher dimensions).

But these two geometrically reasonable demonstrations are supposed to convince you that the distance formula above is "correct": it is intuitive and obvious. It is only one of several candidates for distance in higher dimensions, but it is the one most often considered. The computation we are about to do, however, will show that some of the geometric intuition we have needs to be extended suitably. The intuition isn't wrong, but it needs more examples!

The *origin*, 0, in $\mathbb{R}^n$ has coordinates $(0, 0, \ldots, 0)$.

Definition The *unit ball in* $\mathbb{R}^n$ is the collection of points whose distance to the origin is less than or equal to 1. So $P = (x_1, x_2, \ldots, x_n)$ is in the ball if $x_1^2 + x_2^2 + x_3^2 + \ldots + x_n^2 \leq 1$.

The ball of radius R in n dimensions is defined analogously. $V_{n,R}$ will denote the n-dimensional volume of the ball of radius R. If c_n is the volume of the unit ball, then $V_{n,R}$ will be $c_n R^n$ because multiplication by R changes n-dimensional volume by a factor of R^n (which won't be hard to see from what follows). I want to identify the sequence of constants whose n^{th} term is c_n. Here are some volumes from a textbook:

n	$V_{n,R}$	c_n	Approx. value
1	$2R$	2	2.000
2	πR^2	π	3.142
3	$\frac{4}{3}\pi R^3$	$\frac{4}{3}\pi$	4.189
4	$\frac{1}{2}\pi^2 R^4$	$\frac{1}{2}\pi^2$	4.935
5	$\frac{8}{15}\pi^2 R^5$	$\frac{8}{15}\pi^2$	5.264
6	$\frac{1}{6}\pi^3 R^6$	$\frac{1}{6}\pi^3$	5.168

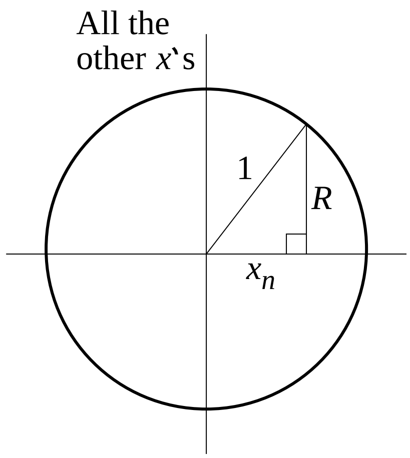
The entries for $n = 2$ and $n = 3$ are well-known. Some explanation might be needed for the others. The one-dimensional "ball" is the interval $[-R, R]$ whose length (or 1-dimensional "volume") is $2R$. The entry for dimension 4 was gotten earlier in class, and the entries for dimensions 5 and 6 were copied from a text. They will be verified below. The pattern of the entries is almost surely *not* clear. There are too many coincidences for small n and we'll need to be more systematic to discover what is happening.

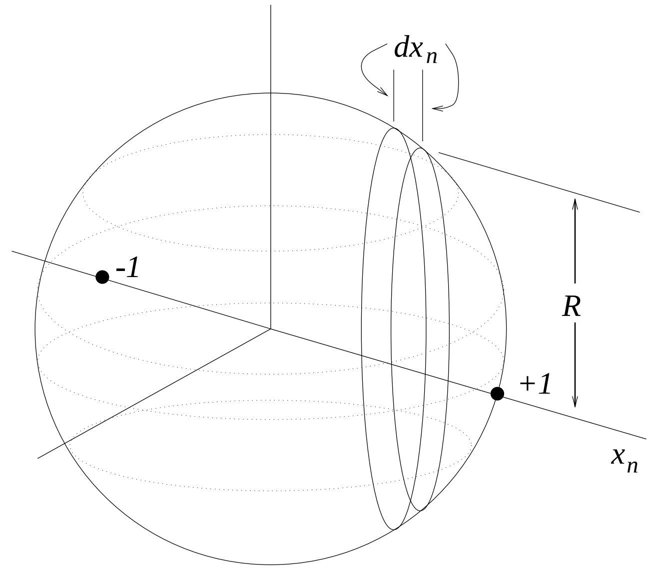
• *Volumes by slicing* •

Consider the inequality for the unit ball:

$$x_1^2 + x_2^2 + x_3^2 + \ldots + x_n^2 \leq 1$$

Since the left-hand side is a sum of squares, we see that x_n^2 must always be less than or equal to 1. That means x_n itself is in the interval $[-1, 1]$. What do $(n-1)$-dimensional cross-sections obtained by holding x_n fixed look like? Since $x_1^2+x_2^2+x_3^2+\ldots+x_{n-1}^2 \leq 1-x_n^2$ in such a slice, we see that the slice is an $(n-1)$-dimensional ball centered at the origin, with radius $R = \sqrt{1 - x_n^2}$. Here's a "picture" of what the situation *might* look like:





The $(n-1)$-dimensional volume of the cross-section is $c_{n-1}R^{n-1}$ so the n-dimensional slice has approximate n-dimensional volume $c_n R^{n-1} dx_n$ and the total volume (inserting our formula for R) is

Equation **ONE** : $$c_n = \int_{x_n=-1}^{x_n=+1} c_{n-1} R^{n-1} \; dx_n = \int_{x_n=-1}^{x_n=+1} c_{n-1} \left(\sqrt{1-x_n^2}\right)^{n-1} \; dx_n$$

• ***Computing the integral*** •

If we make the change of variables (substitution) $x_n = \sin\theta$ then

$$\begin{cases} dx_n \text{ becomes } \cos\theta \; d\theta \\ \sqrt{1-x_n^2} \text{ becomes } \cos\theta \\ x_n = 1 \text{ becomes } \theta = \dfrac{\pi}{2} \\ x_n = -1 \text{ becomes } \theta = -\dfrac{\pi}{2} \end{cases}$$

and equation **ONE** changes to the following equation:

Equation **TWO** : $$c_n = \left(\int_{-\pi/2}^{\pi/2} (\cos\theta)^n \; d\theta \right) c_{n-1}$$

which we'll now proceed to use. Note that **TWO** is valid when n is **an integer greater than 1** and that c_1 is $2R$. So when $n = 2$ we get

$$c_2 = \left(\int_{-\pi/2}^{\pi/2} (\cos\theta)^2 \; d\theta \right) c_1$$

so that, since $c_1 = 2$ and since $\int_{-\pi/2}^{\pi/2} (\cos\theta)^2 \, d\theta = \frac{1}{2}\pi$ (using the double angle trick, as we have done several times in class), we can confirm that $c_2 = (\frac{1}{2})\pi \cdot 2 = \pi$. The next dimension is also fairly easy. We know that $\int_{-\pi/2}^{\pi/2} (\cos\theta)^3 \, d\theta = \frac{4}{3}$ (a direct computation this integral with the substitution $u = \sin\theta$ and a trig identity) and we proved that $c_2 = \pi$. Therefore, **TWO** applies and $c_3 = \frac{4}{3} \cdot \pi = \frac{4}{3}\pi$, as expected.

Let's define I_n as we did previously by the equation

$$I_n = \int_{-\pi/2}^{\pi/2} (\cos\theta)^n \, d\theta$$

since the integral will be important and referred to many times. Then **TWO** above translates to

Equation **THREE** : $\qquad c_n = I_n \cdot c_{n-1}$

which is valid for $n > 1$.

But let's try something a bit more challenging. We can try to compute c_5. By repeatedly using **THREE** we get the following sequence of equalities:

$$c_5 = I_5 \cdot c_4 = I_5 \cdot I_4 \cdot c_3 = I_5 \cdot I_4 \cdot I_3 \cdot c_2 = I_5 \cdot I_4 \cdot I_3 \cdot I_2 \cdot c_1$$

and now we stop. We know that $c_1 = 2$, and "all we need to do" to "compute" c_5 is to learn the value of the various integrals.

• ***A numerical interlude*** ... •

We can and will compute the integrals "exactly", but we can already get some numerical information which has rather disconcerting implications for the volumes of unit balls. Either hand computation or a short session with a computing device produces the following table of three decimal place numerical approximations to I_n for "low" n's (repeat: this is a table of values of integrals!):

n	I_n
1	2.000
2	1.571
3	1.333
4	1.178
5	1.067
6	.982

What can we learn? The numbers displayed are decreasing as n increases. We already knew that the whole sequence of numbers $\{I_n\}$ is a decreasing sequence because most values of cosine in $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ are between 0 and 1. Therefore the integral will be less than the area of the box whose base is π long and whose height is 1. But, in fact, the higher the power, the smaller the value of the function except at those places where cosine is either 1 ($x = 0$) or 0 ($x = \frac{\pi}{2}$ or $x = -\frac{\pi}{2}$). Higher powers make smaller integrals.

These observations mean that for $n \geq 6$, the I_n's all will be less than .982, which means that for large n, many numbers less than .982 will be multiplied to get c_n. But powers of .982 decrease rapidly. Therefore, the sequence of c_n's must go to 0 very quickly. We can already see that the n-dimensional unit balls have *very* small volume when n is large! The strange unit balls are those in the first 5 dimensions, since they are the only ones whose volumes grow with n. After $n = 5$, the volume of the balls *shrinks* quite fast. Let's recall what we learned about the I_n's a few days ago.

$$I_n = \frac{n-1}{n}\, I_{n-2}$$

Explicit values of I_n for $5 \leq n \leq 8$ were given in a previous lecture.

• ***The volume of unit balls*** •

We start with the formula we developed for dimension 5:

$$c_5 = I_5 \cdot I_4 \cdot I_3 \cdot I_2 \cdot c_1$$

Now let's replace the I_n's by their values:

$$c_5 = \overbrace{\left(\frac{4\cdot 2}{5\cdot 3}\right)2}^{I_5} \cdot \overbrace{\left(\frac{3\cdot 1}{4\cdot 2}\right)\pi}^{I_4} \cdot \overbrace{\left(\frac{2}{3}\right)\cdot 2}^{I_3} \cdot \overbrace{\left(\frac{1}{2}\right)\pi}^{I_2} \cdot \overbrace{(2)}^{c_1} = \frac{8}{15}\pi^2$$

Let's check dimension 6:

$$c_6 = I_6 \cdot I_5 \cdot I_4 \cdot I_3 \cdot I_2 \cdot c_1$$

And replace each of the I_n's as before:

$$c_6 = \overbrace{\left(\frac{5\cdot 3\cdot 1}{6\cdot 4\cdot 2}\right)\pi}^{I_6} \cdot \overbrace{\left(\frac{4\cdot 2}{5\cdot 3}\right)2}^{I_5} \cdot \overbrace{\left(\frac{3\cdot 1}{4\cdot 2}\right)\pi}^{I_4} \cdot \overbrace{\left(\frac{2}{3}\right)\cdot 2}^{I_3} \cdot \overbrace{\left(\frac{1}{2}\right)\pi}^{I_2} \cdot \overbrace{(2)}^{c_1} = \frac{1}{6}\pi^3$$

So we agree with the textbook answers for dimensions 5 and 6. In addition, if we examine the structure of the answers, we can see patterns which will clearly persist in general. Here the phrase "will clearly persist in general" is a substitution for "can be proved by mathematical induction." Mathematical induction is a simple proof technique used to prove statements involving positive integers.

Let's look at even n first to discover the general formula. In the answer for c_6 above, all of the fractions cancel except for

$$\frac{1}{6\cdot 4\cdot 2} = \frac{1}{(3\cdot 2\cdot 1)\cdot(2\cdot 2\cdot 2)} = \frac{1}{3!\cdot 2^3}$$

The π's and 2's alternate: there are three of each. Note that $6 = 2 \cdot 3$. This suggests the following formula, which is correct:

$$\text{If } n \text{ is even, then } c_n = \frac{\pi^{\frac{n}{2}}}{\left(\frac{n}{2}\right)!}$$

Odd dimensions seem more complicated. Let's look at c_5 again, and group terms.

$$c_5 = \left(\frac{4 \cdot 2}{5 \cdot 3}\right) 2 \cdot \left(\frac{3 \cdot 1}{4 \cdot 2}\right) \pi \cdot \left(\frac{2}{3}\right) \cdot 2 \cdot \left(\frac{1}{2}\right) \pi \cdot (2) = \left(\frac{2 \cdot 2 \cdot 2}{5 \cdot 3 \cdot 1}\right) \pi^2 = \frac{1}{\left(\frac{5}{2}\right) \cdot \left(\frac{3}{2}\right) \cdot \left(\frac{1}{2}\right)} \cdot \pi^{\frac{4}{2}}$$

My aim here is a simple formula for c_5. What's the bottom of the fraction look like? $\left(\frac{5}{2}\right) \cdot \left(\frac{3}{2}\right) \cdot \left(\frac{1}{2}\right)$ seems almost to be the beginning of a factorial – a collection of numbers decreasing by 1 and being multiplied. But we would have to consider factorials of *half* integers. There is a generally accepted interpolation to non-integer values of the factorial function. If you're willing to willing to accept (temporarily) the fact that

$$\left(\tfrac{1}{2}\right)! = \left(\tfrac{1}{2}\right) \cdot \sqrt{\pi}$$

that is, the factorial of $\frac{1}{2}$ is or should be $\left(\frac{1}{2}\right) \cdot \sqrt{\pi}$, then "clearly" $\left(\frac{5}{2}\right)! = \left(\frac{5}{2}\right) \cdot \left(\frac{3}{2}\right)! = \left(\frac{5}{2}\right) \cdot \left(\frac{3}{2}\right) \cdot \left(\frac{1}{2}\right)! = \left(\frac{5}{2}\right) \cdot \left(\frac{3}{2}\right) \cdot \left(\frac{1}{2}\right) \cdot \sqrt{\pi}$. Multiply the formula for c_5 by $\sqrt{\pi}$ on both the top and bottom:

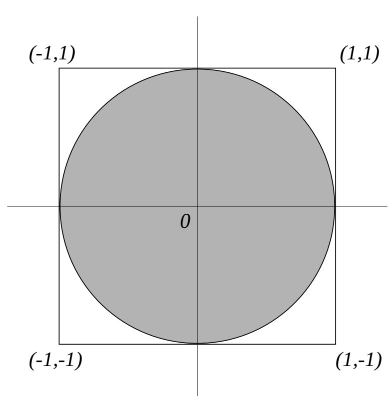
$$c_5 = \frac{1}{\left(\frac{5}{2}\right) \cdot \left(\frac{3}{2}\right) \cdot \left(\frac{1}{2}\right)} \cdot \pi^{\frac{4}{2}} \cdot \frac{\sqrt{\pi}}{\sqrt{\pi}} = \frac{\pi^{\frac{5}{2}}}{\left(\frac{5}{2}\right)!}$$

I must show you how to find factorials of non-integer positive real numbers, and then will verify the result quoted above at that time. If you're willing to accept this future result on factorial interpolation, we get a neat formula, valid in all dimensions:

$$\text{If } n \text{ is } \textbf{any} \text{ positive integer, then } c_n = \frac{\pi^{\frac{n}{2}}}{\left(\frac{n}{2}\right)!}$$

• *A ball in a box* •

The box just containing the unit circle in $\mathbb{R}^2$ is the square with corners $(1,1)$, $(-1,1)$, $(1,-1)$, and $(-1,-1)$.

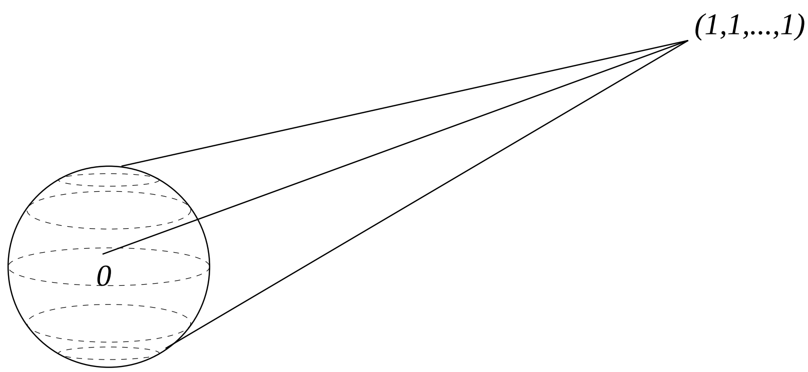


A deceptive two-dimensional picture

The proportion of the box occupied by the ball is $\frac{\pi}{4}$, approximately 78.5%. In $\mathbb{R}^3$, the box just containing the ball has 8 corners, all those points whose coordinates are either $+1$ or -1. The volume of that box is $2^3 = 8$, and the proportion of the box occupied by the unit ball is $\left(\frac{4}{3}\pi\right)/8$, about 52.4%. But remember that low dimensions are very exceptional in what we're doing here. Indeed, they are almost weird!

In dimension n, the box just containing the unit ball has 2^n corners (all the points whose n coordinates are either $+1$ or -1), and every side has length 2. So the n-dimensional volume of this box is 2^n. The volume of the ball is given by the formula above. The top of the ball formula is $\pi^{\frac{n}{2}}$ which is $\left(\sqrt{\pi}\right)^n$, approximately $(1.772)^n$. Therefore the top grows more slowly than 2^n all by itself. But the factorial "downstairs" in the ball formula makes the proportion of the box occupied by the ball get small *very* quickly.

A example computed with the help of `Maple` may illustrate this more concretely. The volume of the 56-dimensional ball is reported to be approximately $2.729 \cdot 10^{-16}$. The 56-dimensional box just enclosing this ball has 56-dimensional volume 2^{56}, which is approximately $7.206 \cdot 10^{16}$. The quotient of the volume of the ball divided by the volume of the box is approximately $3.788 \cdot 10^{-33}$, a *really* tiny number! What's happening "geometrically" and how can we hope to understand it? Here are some suggestions. A typical "corner" of the box is a point with coordinates $(\pm 1, \pm 1, \ldots, \pm 1)$. The distance from the origin to this corner is $\sqrt{(\pm 1)^2 + (\pm 1)^2 + \ldots + (\pm 1)^2} = \sqrt{56} > 7$ (since the coordinates of a point are a 56-tuple). The corners are very far from the origin, more than 7 times the radius of the unit ball, while the ball stays very close since points in the ball have distance ≤ 1 to the origin. Not only are the corners of the box far away, but there are many of them (2^{56}, in fact). So our low dimensional intuition has not seen enough examples. More directly: dimensions 2 and 3 are exceptional. Below is a poetic picture, an "impression", of part of what the ball in the box *might* look like if we had better perception. I tried to show only one corner and the ball. Of course, in the picture the ball is much too big and the corner is much too close!



A vision of higher dimensional truth (?)

As I mentioned in class, I first saw these computations significantly used in an electrical engineering text. Let me give a a brief and much simplified version of the discussion I read. The question of how to analyze "random" strings of signals was being explored. Each signal had "strength" between -1 and $+1$. One possible measure of the total strength of an n long string of signals was the square root of the sum of the squares of the strengths – our distance to the origin. In this context, a signal string of length 56 doesn't seem excessive at all to me! The analysis I described above was done in just a few lines of the text (many implied uses of “clearly” were needed!) and the notion of total signal strength gotten from Euclidean distance (the square root of a sum of squares) was then abandoned. An alternative measure of strength derived by considering boxes instead of balls was proposed and used: measure the strength of the signal by taking the maximum of the absolute values of the individual signals. In our language, the recommendation was to discard $\sqrt{\sum_{j=1}^{n}(x_j)^2}$ as a measure of the strength of the entire signal and to use $\max\{|x_j|, 1 \leq j \leq n\}$ instead.