



© 1997–2009, Millennium Mathematics Project, University of Cambridge.

Permission is granted to print and copy this page on paper for non-commercial use. For other uses, including electronic redistribution, please contact us.

---

May 2001

Features

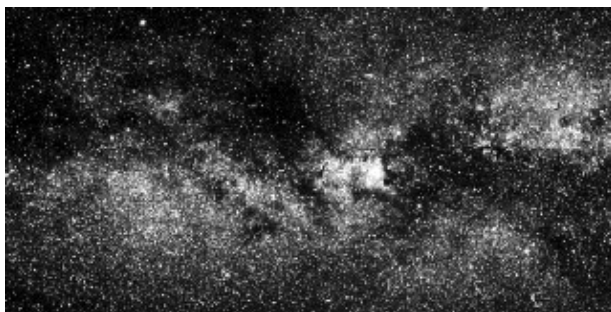


## How big is the Milky Way?

by Toby O'Neil



### Introduction



A photograph of part of the Milky Way, courtesy of NASA

All objects attract each other and when we look up at a (clear) night sky we see many massive stars all exerting forces on each other. A question which has been vexing astronomers for a long time is whether these forces of attraction between stars and galaxies will eventually result in the universe collapsing back into a single point, or whether it will expand forever with the distances between stars and galaxies growing ever larger. The answer to this question turns out to depend on how much matter the universe contains: the more matter, the more attraction and the more likely the "collapse" theory.

So how can we find how much matter there is in the universe? In this article, I'm going to describe how the mathematical theory of dimension gives us one way of approaching this question, and helps us to estimate

# How big is the Milky Way?

how much visible matter there is in the universe.

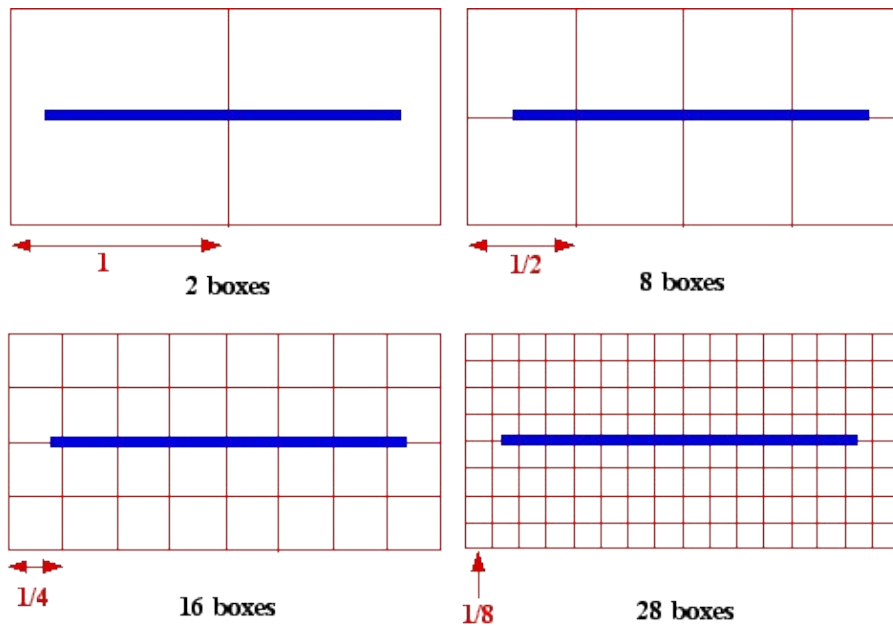
## What is dimension?

We all think of a line segment as being one-dimensional, a square two-dimensional and a cube three-dimensional, but what does this really mean? Intuitively, the dimension of an object should measure how well it fills space, how "wiggly" it is: the problem is to turn this intuitive idea into a mathematical definition.

In order to reduce problems with visualising what is happening, we consider objects sitting in the plane. We begin by looking at a line and a square more closely.

### The line

Consider a line segment sitting in the plane – we expect this to be one-dimensional. If we split the plane up into squares of the same side lengths  $r$  and count how many hit the line segment, what do we find? In the figure below we have divided the plane up into boxes of side lengths, 1, 1/2, 1/4 and 1/8, and counted how many of the boxes hit (or intersect) a particular line segment.



Finding the box dimension of a line segment

If we tabulate the results we find the following:

<b>Side length</b>	1	1/2	1/4	1/8	...	$2^{-k}$
<b>Number of boxes</b>	2	8	16	28	...	approx $2^k$

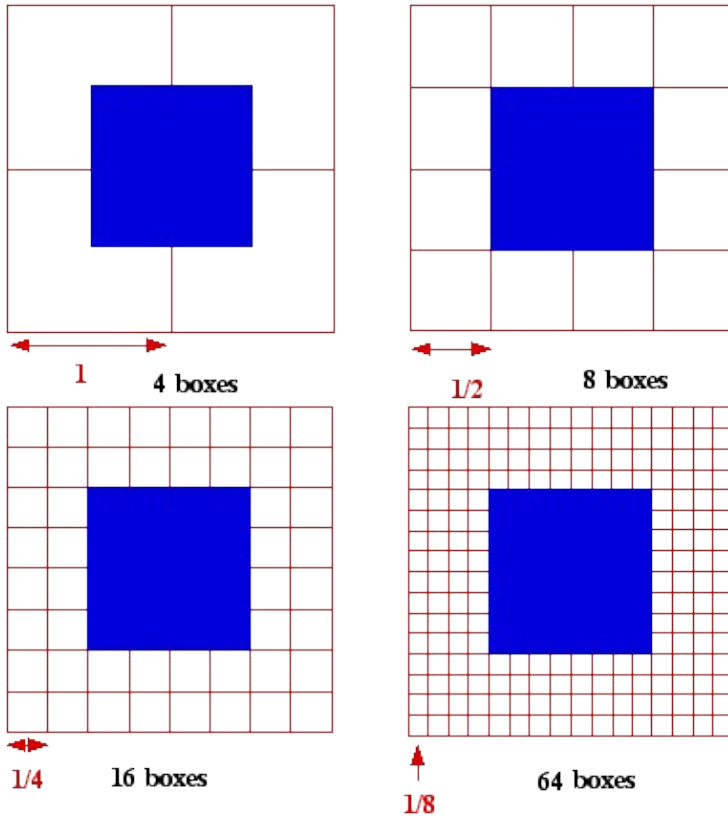
Thus, if  $N_r$  denotes the number of boxes of side-length  $r$  required to cover the line segment, then

$$N_r \approx r^{-1} = (1/r)^1.$$

What is dimension?

## The square

Now consider a square. This is something which we would expect to be two dimensional. Again, if we split the plane up into squares of the same side lengths  $r$  and count how many hit our square, what do we find? In the figure below we have divided the plane up into boxes of side lengths, 1, 1/2, 1/4 and 1/8, and counted how many of the boxes hit (or intersect) a particular line segment.



Finding the box dimension of a square

<b>Side length</b>	1	1/2	1/4	1/8	...	$2^{-k}$
<b>Number of boxes</b>	4	16	64	...	...	approx $2^{2k}$

Thus, if  $N_r$  denotes the number of boxes of side-length  $r$  required to cover the square, then

$$N_r \approx r^{-2} = (1/r)^2.$$

Thus for these examples we find that when  $r$  is very small, then the number of boxes of side length  $r$  required to cover the set is roughly  $(1/r)^2$  to the power of the dimension. This suggests that we make the following definition:

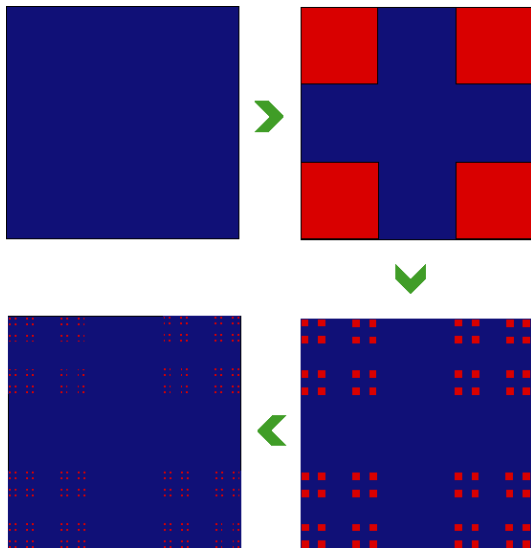
The **(box) dimension** of a set  $E$  is the number  $d$  such that the number of boxes of side length  $r$  required to cover  $E$  is proportional to  $r^{-d}$ .

## A strange set

Now let us consider something a little more complicated, a set which is somewhere between being a line and a square in that its dimension will turn out to be between one and two.

We construct it as follows:

We start with a square of side-length 1, and we split it into 9 smaller subsquares of side-lengths  $1/3$ . We then keep the four corner subsquares and throw the rest away. For each of these four squares we then repeat the process: we divide each up into 9 smaller squares of side-lengths  $1/9$  and keep only the four corner squares of these new squares. We repeat this process forever..., and in the end obtain the object illustrated below. This sort of set is known as a *Cantor Dust* and is named after the German mathematician Georg Cantor, who was one of the pioneers of modern mathematical analysis.



### A Cantor Dust

Let's now try to find the box dimension of this set. If, instead of boxes of side-lengths  $1/2, 1/4, 1/8, \dots$ , we use boxes of side-lengths  $(1/3)^k, k = 1, 2, 3, \dots$ , then we need exactly  $(4^k)$  boxes to cover the set.

Hence

$$N_{(1/3)^k} = 4^k$$

and we can calculate that

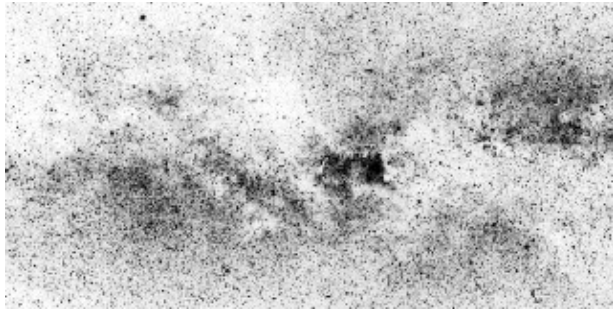
$$\frac{\log(N_{(1/3)^k})}{-\log(1/3^k)} = \frac{\log(4^k)}{k \log(3)} = \frac{k \log(4)}{k \log(3)} = \frac{\log(4)}{\log(3)} \approx 1.26.$$

We deduce that the box dimension of this set is 1.26. This is an example of what is known as a *fractal set* since its dimension is not a whole number. Comparing this object to our picture of the Milky Way, we see that the Milky Way appears to have a far denser distribution of matter, and so we expect that our calculation of the dimension of the Milky Way will give an answer which is larger than 1.26.

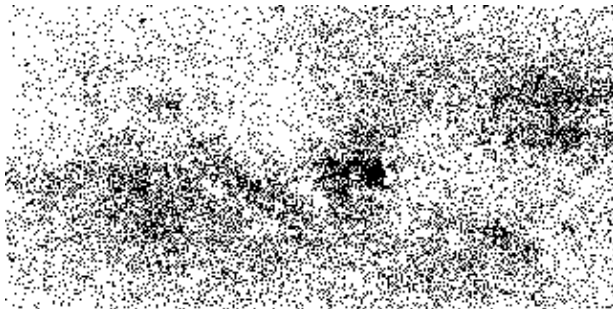
## The box dimension of our photo of the Milky Way

Now we'll look at the photo at the beginning of this article, and use it to try to estimate the box dimension of the Milky Way.

We begin by modifying the photo to create an image which is amenable to analysis: calculations of box dimension require a completely black and white image – no greyscale is allowed. We do this by first inverting the greyscale so that the stars are black and the empty space is white. We then convert the picture to a black and white image, by making anything above a certain level of grey, black, and colouring the remainder, white, as above. There is a certain amount of freedom as to the choices we make here and we shall need to consider how this affects the accuracy of our results later on.



Inverting the greyscale



Changing greyscale to black and white

We now divide the picture up into a grid of squares (of side length  $r$ , say) and count how many contain any black parts of the image. We do this for grids of various sizes and record the results. There are many computer programs which can do this rather tedious work for us and I used the [Fractal Dimension Calculator](#) to find the counts for boxes of various sizes. The table below shows the number of boxes of various sizes required to cover the stars in our photo. The size of a box is given as a fraction of the image's height.

Size of box, $s$	Number of boxes, $N_s$
1	6
0.5	18
0.4375	28
0.375	40
0.3125	45

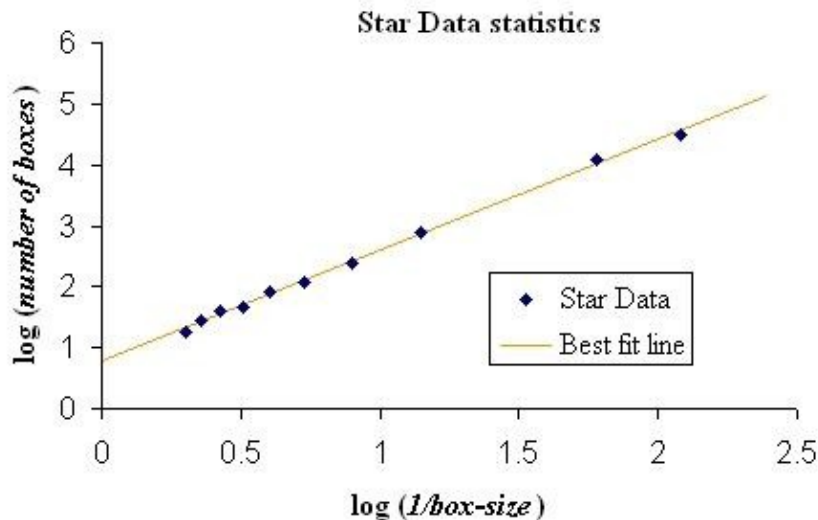
## How big is the Milky Way?

0.25	84
0.1875	120
0.125	252
0.07083	760
0.01667	12066
0.00833	31701
0.00417	47364

If, as we expect, for (small) boxes of side-length  $r$ ,  $N_r \approx cr^{-\dim(\text{Milky Way})}$ , then taking the  $\log$  of both sides, we find that

$$\log(N_r) \approx \dim(\text{Milky Way}) \log(1/r) + \log(c).$$

Hence, if we plot  $\log(N_r)$  against  $\log(1/r)$ , then provided the resulting graph is a straight line, then the box dimension of our picture of the Milky Way will be given by the line's slope. In the figure below, we show the resulting graph.



A plot of  $\log N(r)$  against  $\log(1/r)$ . The line of best fit has gradient approximately 1.82

As you can see, the result is a rather convincing straight line with slope about 1.82 and we conclude that the box dimension of our photo of the Milky Way is about 1.82. There is a problem, however – we haven't examined the relationship between the dimension of the image in the photo and the dimension of the Milky Way itself! This is our final task.

## Projections

In a photograph, the distances between various stars are distorted and it is also very likely that some stars which are in the Milky Way are not visible on the photo. Thus we expect the photo to show us less than is actually there. Hence any estimate we make of the dimension is likely to be too small. The question is whether we can say anything sensible about what the error is likely to be.

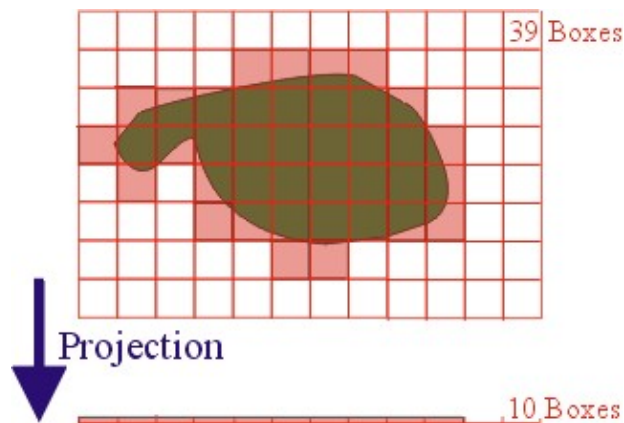
## How big is the Milky Way?



Our photo of the Milky Way is a projection, not an accurate representation

Our photo of the Milky Way can be viewed as a projection of the Milky Way onto a sheet of paper, see the figure above. Thus it would be useful if, given an object  $E$  in space, and a projection  $P(E)$  of it, we could find a relationship between their dimensions.

If we draw a grid of boxes over  $E$  of size  $r$  and then look at the corresponding grid on the projection of  $E$ , then the number of boxes we need to cover the projection is always less than or equal to the number of boxes we needed to cover  $E$ , see the figure below. (We are ignoring the fact that the projected boxes may be distorted – when the argument is done more carefully, it turns out that this doesn't matter.)



Projecting cannot increase the number of boxes needed

Thus

$$N_r(P(E)) \leq N_r(E)$$

and so

$$\log[N_r(P(E))] \leq \log[N_r(E)]$$

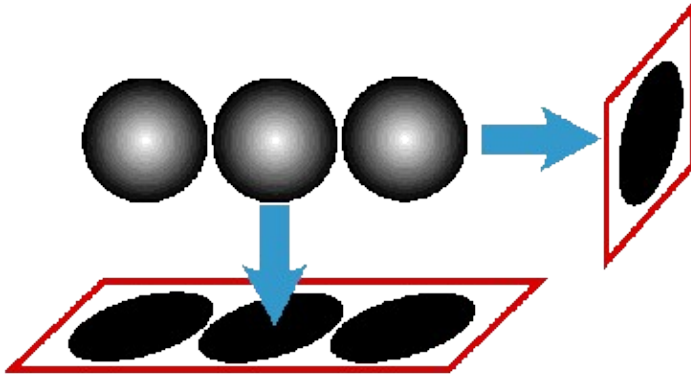
and hence

$$\frac{\log[N_r(P(E))]}{-\log(r)} \leq \frac{\log[N_r(E)]}{-\log(r)}$$

## How big is the Milky Way?

if  $0 < r < 1$ . (We are using the fact that for  $0 < r < 1$ ,  $-\log(r) > 0$ .) Thus  $\dim(P(E)) \leq \dim(E)$ , as we expected.

It is always possible that the photo we have of the Milky Way is taken from a very special direction where there is a lot of overlapping, see the figure below.



In the projection to the right, there is an exceptional amount of overlapping. The projection downwards has less.

However, we would expect that if our photograph of the Milky Way is taken from a typical position, then we would not get an exceptional amount of overlap. Our problem is how to measure how much overlap we typically get. Fortunately for us, a very recent result by John Howroyd (a mathematician at Goldsmiths College in London) tells us what to expect. He showed that for the projection from space onto a *typical* plane of an object,  $E$  with  $\dim(E) > 0$ , then the box dimension of the projection,  $\dim(P(E))$ , satisfies

$$\frac{1}{\dim(P(E))} - \frac{1}{2} \leq \frac{1}{\dim(E)} - \frac{1}{3}.$$

If we rearrange this in order to estimate  $\dim(E)$  in terms of  $\dim(P(E))$ , then we find that

$$\frac{1}{\dim(P(E))} - \frac{1}{6} \leq \frac{1}{\dim(E)}$$

and so, if  $\dim(E) > 0$ , then

$$\dim(E) \leq \frac{1}{[1/\dim(P(E))] - [1/6]},$$

which simplifies to give

$$\dim(E) \leq \frac{\dim(P(E))}{1 - [\dim(P(E))/6]}.$$

Hence, on the assumption that the image in our photo of the Milky Way is essentially just a projection of the Milky Way onto a plane in a typical position, we estimate that

## How big is the Milky Way?

$$\dim(\text{Milky Way}) \leq \frac{1.82}{1 - [1.82/6]} \approx 2.61.$$

That is, the dimension of the Milky Way is no larger than about 2.6. Since we know that its dimension must be at least as large as that estimated from the photo, we conclude that the dimension of the Milky Way lies somewhere between 1.82 and 2.6.

## Conclusions

Our analysis suggests that the dimension of the Milky Way is between 1.82 and 2.6. However there are many potential problems with our approach which may mean these estimates are meaningless! Here are a few of them (you may be able to find others):

1. There was an arbitrary judgement involved in converting our photo into a pure black and white image – we had to decide which shades of grey were to be black and which were to be white;
2. The definition of box dimension involves looking at how many boxes cover a set for arbitrarily small boxes. Since our image of the Milky Way was just a computer graphic, it had a scale below which we would learn nothing useful. This is a problem for calculating box dimension of many things in science.
3. We haven't take any account of the fact that astronomers believe that much of the universe is made up of "dark matter" which is not visible to us on Earth – this could increase the dimension of the universe substantially.

Despite these problems, though, the method does give us some insight into the distribution of matter in the Milky Way. Many of these problems can be overcome and the only one which will always be present, no matter the problem under investigation, is the fact that the calculation of box dimension requires us to look at what happens as the boxes used become arbitrarily small. In practice, this is impossible, and it is a matter of judgement as to whether one has enough data to make a sensible calculation.

For many stars, we also know how far away from us they are, and you may wonder whether it is possible to use this information to get better estimates on the dimension. Of course from our viewpoint on Earth, there will be many stars we can't see, and it is unknown if this extra information does enable better estimates to be made.

## About the author



Toby O'Neil is a Lecturer in Analysis at the Department of Pure Mathematics of the Open University, and when he's not busy counting boxes, he likes to spend his time staring vacantly into space.

---

How big is the Milky Way?



*Plus* is part of the family of activities in the Millennium Mathematics Project, which also includes the NRICH and MOTIVATE sites.