AstroTech MOOC supplementary material

Big Numbers

Megaparsecs, Petabytes, and all that Jazz

Very large numbers are bandied around in astronomy, and you will have seen people using expressions like "10 to the power 15", and "Petabytes" and so on. Also, many of you may be thinking "I thought a Megabyte was 1024 Kilobytes, not 1000 Kilobytes". Lets step through these issues and try to clear them up.

Millions and billions etc

One thing to clear up first... a million is a thousand thousand; a billion is a thousand million; a trillion is a thousand billion. It used to be that in British English, as opposed to US English, a "billion" meant a million million, but this has fallen out of use. There are words in English for larger numbers, but we recommend that when you get to bigger numbers it is better to use one of the methods below - either scientific notation, or standard prefixes.

Scientific Notation: powers of ten

Writing two thousand as 2,000 is ok, but writing seven trillion as 7,000,000,000,000 gets a bit laborious, and prone to error, especially as you get to even bigger numbers - you have to count the zeros to check what you've got. The standard scientific method is to write 7×10^{12} . You can interpret this as "seven followed by twelve zeros", and speak it as "seven times ten to the twelve".

The beauty of this method is that you can use it for numbers as big as you like - e.g. 3×10^{37} is three with thirty seven zeros. It doesn't need a special name. If you want to say it out loud you can just call it "three times ten to the thirty seven". You can also think of this method as "shifting the decimal point", so that 3.176×10^{9} is the same as $3,176,000,000.0$.

This is the "powers of ten" method. In maths, raising something to the power N means multiplying it by itself N times, so that $24 = 2 \times 2 \times 2 \times 2 = 2^4$. In the same vein $10^3 = 10 \times 10 \times 10 = 1000$ and so $7 \times 10^3 = 7000$.

Prefixes: Mega, Giga etc

The "power" notation is best if you are doing calculations, but for scientific discussion its often better to use standard prefixes. These go up in jumps of 1000. For example, suppose we are measuring the power of something, in Watts, normally written with just a W. Then the standard notations are as set out in Table 1.

As you will doubtless know, there is a corresponding set of standard prefixes for small numbers - milli, micro, nano, pico, femto. Astronomers tend to use this method even with non-standard units, in order to make discussion easier. For example, they often measure distances in units of "parsecs", abbreviated pc. (One pc is the distance at which a star has a parallax of one arc second... yes, its weird, but useful). The nearest stars are a few pc away. But once we started measuring distances to galaxies, they turned out to be typically millions of pc away, so astronomers just started quoting things in Mega-parsecs - e.g. "I estimate that galaxy SplargXYZ is 237Mpc away".

Table 1

1 Watt	= not a lot	
1 kiloWatt	= 1kW	= 10^3W
1 MegaWatt	= 1MW	= 10^6W
1 GigaWatt	= 1GW	= 10^9W
1 TeraWatt	= 1TW	= 10^{12}W
1 PetaWatt	= 1PW	= 10^{15}W
1 ExaWatt	= 1EW	= 10^{18}W

Powers of two: Bits and Bytes

Computer scientists tend to think in powers of two rather than powers of ten. This is because computers store numbers by lots of separate electrical locations which are either on or off - just two choices. Then just as 305 really means "5×1 plus 0×10^1 plus 3×10^2", the binary number 101 means "1×1 plus 0×2^1 plus 1×2^2 " and so on. (In other words, the number 5 in decimal speak.) Each digit in the binary number is a "bit".

Obviously you need quite a long sequence of ones and zeros to store big numbers, but thats ok - we humans don't have to write them down - its all dealt with inside the computer. Now... computers tend to store these long sequences in chunks. Historically, this started by using chunks of eight bits at a time, which became known as a "byte". Next, 1024 of these bytes became known in the computing world as a "kilobyte". On the one hand, 1024 is exactly 2^{10} , but its also roughly a thousand...

Good practice is to use "b" to indicate bits and "B" to indicate bytes, but beware - people are often sloppy about this! Sometimes you might see an internet connection advertised as say 12 MBps and you think it means twelve megabytes per second... but actually they meant twelve megabits per second .. eight times smaller!

But what do we mean by a Megabyte anyway? Lets look at that next.

Decimal vs binary prefixes

So as we have seen above, scientists always use "kilo" to mean a thousand of something, "mega" to mean a million of something, and so on, whereas the computer industry started using "kilo" to mean 1024 of something, and "mega" to mean 1024×1024 of something. Even more confusingly, some manufacturers would mix the two, using "megabyte" to mean a thousand kilobytes, but with those kilobytes still being 1024 bytes. This mess was officially cleared up in 1998 by various standards bodies, but it seems to be taking a long time to sink in! Here is the official position...

A byte is still eight bits. That won't change.

A kilobyte (kB) is 10^3 bytes, a megabyte (MB)is 10^6 bytes, and so on, following the standard scientific usage.

A kibibyte (KiB) is 1024 bytes, a mebibyte (MiB) is 1024^2 bytes, a gibibyte (GiB) is 1024^3 bytes, and so on.

So you can use either, but make sure you say which you mean! You can read more on this nice Wikipedia page