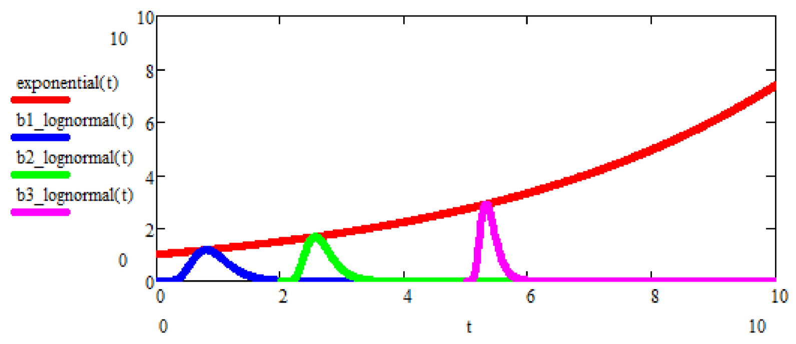
MOLECULAR CLOCK on Earth and Exoplanets *as ENTROPY* changing linearly in time

Claudio Maccone, INAF (Istituto Nazionale di Astrofisica - Italy) and International Academy of Astronautics (IAA). E-mail: clmaccon@libero.it

ABSTRACT. In a series of recent papers (refs. [1] through [9]) and in a book (ref. [6]), this author gave a mathematical model that may be applied to the analysis of Mass Extinctions as a stochastic process of the time. His work is based on lognormal probability distributions, b-lognormals (i.e. lognormals starting at a time b>0 higher than zero) and the stochastic process called Geometric Brownian Motion (GBM). The key feature of GBM is that its mean value increases exponentially in time. Thus, GBM may be applied to represent the number of living species on Earth at a certain time during the last 3.5 billion years, with a total increase from 1 (i.e. the first living being appeared on Earth 3.5 billion years ago (i.e. RNA)) to 50 million (the estimated current number of living species on Earth). In this mathematical scenario, Mass Extinctions are just times in the past where the GBM value did go down a lot from its own mean exponential curve and, a certain time after the impact that caused the Mass Extinction, did go up again to regain its previous high values or even higher values still. In addition, this author proved the so-called "Peak-Locus Theorem" stating that exponential GBM mean value (solid red curve on the right) also is the geometric locus of the peaks of the family of b-lognormals (three of them in color on the right) each starting at a different time b, having a different peak at p, and then "dying off at infinity on the right", but in such a way that the AREA under each b-lognormal always equals 1 (normalization condition of all b-lognormals).



THE DRAKE EQUATION (1961). The foundational equation of SETI (the Search for Extraterrestrial Intelligence) is the Drake equation. In fact, the problem of finding how close the nearest ET Civilization may possibly be "approximately solved" by reducing it to the computation of *N*, the total number of Extraterrestrial Civilizations now existing in this Galaxy, and then diving by the Galaxy volume. In this short introductory section the famous Drake equation is summarized. *N* was written by Frank Drake (1961) as the product of seven factors, each a kind of filter, every one of which must be sizable for there to be a large number of civilizations:

Ns, the number of stars in the Milky Way Galaxy; *fp*, the fraction of stars that have planetary systems; *ne*, the number of planets in a given system that are ecologically suitable for life; *fl*, the fraction of otherwise suitable planets on which life actually arises (by Darwinian Evolution); *fi*, the fraction of inhabited planets on which an intelligent form of life evolves (Human History); *fc*, the fraction of planets inhabited by intelligent beings on which a communicative technical civilization develops (as we have it today); and *fL*, the fraction of planetary lifetime graced by a technical civilization (a totally unknown factor). Written out, the equation reads $N = Ns \cdot fp \cdot ne \cdot fl \cdot fi \cdot fc \cdot fL$ (1). All of the *f*'s are fractions, having values between 0 and 1; they will pare down the large value of *Ns*. To derive *N* we must estimate each of these quantities. We know a fair amount about the early factors in the equation, the number of stars and planetary systems. We know very little about the later factors, concerning the evolution of life, the evolution of intelligence or the lifetime of technical societies. In these cases our estimates will be little better than guesses.

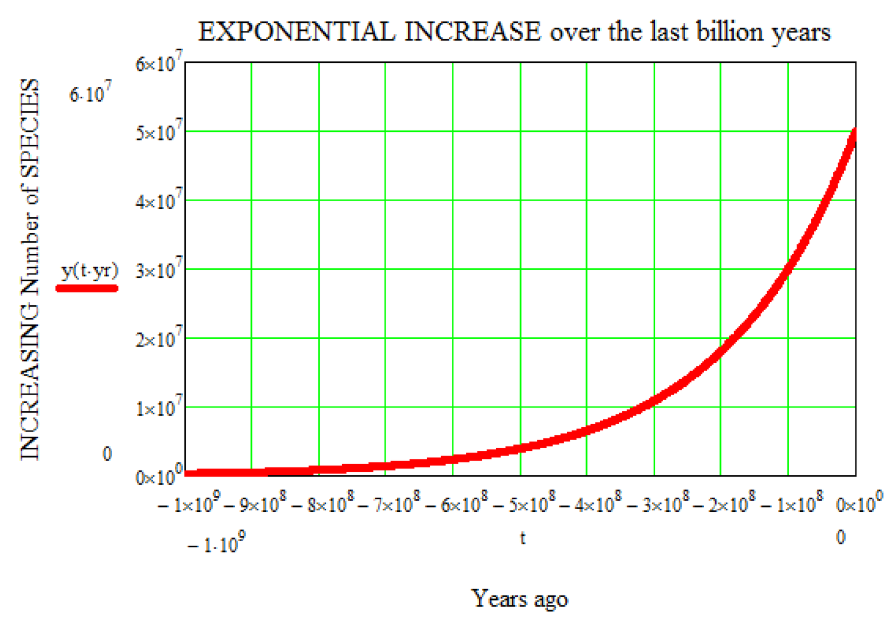
THE STATISTICAL DRAKE EQUATION (2008). Consider the first independent variable in the Drake equation (1), i.e., *Ns*, the number of stars in the Milky Way Galaxy. Astronomers tell us that *approximately* there should be about 350 billions stars in the Galaxy. Of course, nobody has counted (or even seen in the photographic plates) *all* the stars in the Galaxy! There are too many practical difficulties preventing us from doing so: just to name one, the dust clouds that don't allow us to see even the Galactic Bulge (i.e. the central region of the Galaxy) in the visible light (although we may "see it" at radio frequencies like the famous neutral hydrogen line at 1420 MHz). So, it doesn't make much sense to say that $Ns = 350 \times 10^9$, or, say (even worse) that the number of stars in the Galaxy is (say) 354,233,321,123, or similar fanciful exact integer numbers. That is just silly and non-scientific. Much more scientific, on the contrary, is to say that the number of stars in the Galaxy is 350 billion plus or minus, say, 50 billions (or whatever values the astronomers may regard as more appropriate, since this is just an example to let the reader understand the difficulty).

Thus, it makes sense to REPLACE each of the seven independent variables in the Drake equation (1) by a MEAN VALUE (350 billions, in the above example) PLUS OR MINUS A CERTAIN STANDARD DEVIATION (50 billions, in the above example). By doing so, we have made a great step ahead: we have abandoned the too-simplistic equation (1) and replaced it by something more sophisticated and scientifically more serious: the STATISTICAL Drake equation. In other words, we have transformed the simplistic classical Drake equation (1) into an advanced statistical tool to investigate of a host of facts hardly known to us in detail. In other words still: We replace each independent variable in (1) by a RANDOM VARIABLE, labelled (from Drake); We assume that the MEAN VALUE of each is the same numerical value previously attributed to the corresponding independent variable in (1); But now we also ADD A STANDARD DEVIATION on each side of the mean value, provided by the knowledge gathered by scientists in each discipline encompassed by each . Having so done, the next question is: How can we find out the PROBABILITY DISTRIBUTION for each ? For instance, shall that be a Gaussian, or what? This is a difficult question, for nobody knows, for instance, the probability distribution of the number of stars in the Galaxy, not to mention the probability distribution of the other six variables in the Drake equation (1). There is a brilliant way to get around this difficulty, though, as we see in the next section.

SOLVING THE STATISTICAL DRAKE EQUATION BY VIRTUE OF THE CENTRAL LIMIT THEOREM (CLT) OF STATISTICS. The solution to the problem of finding the analytical expression for the probability density function of *N* in the statistical Drake equation was found by this author only in September 2007, and was presented publicly for the first time at a SETI Meeting run by Paul Davies at the "Beyond Center" of the University of Arizona at Phoenix on February 8th, 2008. The key steps are the following: Take the natural logs of both sides of the statistical Drake equation (1). This changes the product into a sum. The mean values and standard deviations of the logs of the random variables may all be expressed analytically in terms of the mean values and standard deviations of the . Recall the Central Limit Theorem (CLT) of statistics, stating that (loosely speaking) if you have a SUM of independent random variables, each of which is ARBITRARILY DISTRIBUTED (hence, also including uniformly distributed), then, when the number of terms in the sum increases indefinitely (i.e. for a sum of random variables infinitely long)… the SUM RANDOM VARIABLE APPROACHES A GAUSSIAN. Thus, the natural log of *N* tends to a Gaussian**. Thus, *N* approaches the LOGNORMAL DISTRIBUTION.** The mean value and standard deviations of this lognormal distribution of *N* may all be expressed analytically in terms of the mean values and standard deviations of the logs of the already found previously, as shown in Table 1. ***This result is fundamental.*** For all the relevant mathematical proofs, more mathematical details and a few numerical examples of how the Statistical Drake Equation works, see Maccone 2010 #1.

DARWINIAN EVOLUTION AS THE EXPONENTIAL INCREASE OF THE NUMBER OF LIVING SPECIES. Consider now Darwinian Evolution. In order to cast it into a mathematically fruitful form, we may regard it as the exponentially increasing number of living species on Earth starting 3.5 billion years ago. In other words, 3.5 billion years ago there was on Earth only one living species (questionably RNA ?) whereas now there may be (say) 50 million living species or more Note that the actual number of species currently living on earth does not really matter as a number for us: we just want to stress the ***exponential*** character of the growth of species. Thus, we shall assume that the number of living species on Earth increases in time as $E(t) = A\, e^{Bt}$ where *A* and *B* are two positive constants. A few steps show then that the numeric values of these two constants are, respectively :

$$A = 50 \text{ million species} = 5 \cdot 10^7 \text{ species} \qquad B = -\frac{\ln\left(5 \cdot 10^7\right)}{-3.5 \cdot 10^9 \text{ year}} = \frac{1.605 \cdot 10^{-16}}{\text{sec}}.$$



***b*-LOGNORMALS AND LOGNORMALS** Just for reference, we write here the equation of the lognormal probability density function starting at any positive time *b*>0, that we call b-lognormal in Maccone 2011 #2. The ordinary lognormal simply is the special case *b*=0 of the *b*-lognormal. The letter *b* stands for "birth instant".

$$\begin{cases} \text{b\_lognormal}(t,\mu,\sigma,b) = \dfrac{1}{\sqrt{2\pi}\,\sigma\,(t-b)}\, e^{-\frac{(\ln(t-b)-\mu)^2}{2\sigma^2}} \\ \textit{holding for } t > b \textit{ and up to } t = \infty. \end{cases}$$

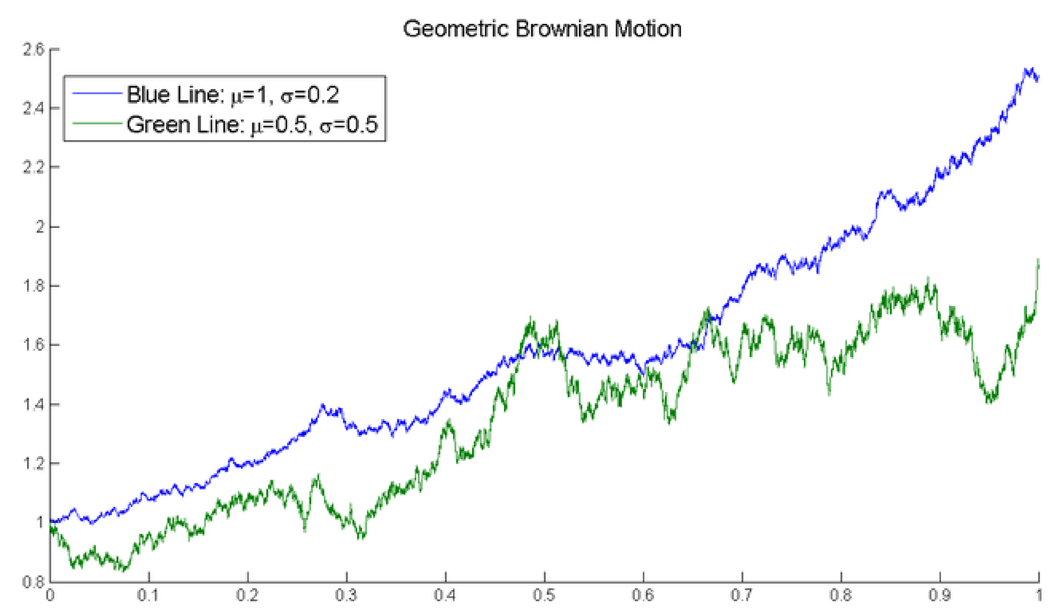
INTRODUCING THE "DARWIN" (d) UNIT, MEASURING THE AMOUNT OF EVOLUTION THAT A GIVEN SPECIES REACHED. In all sciences "to measure is to understand". In physics and chemistry this is done by virtue of units such as the meter, second, kilogram, coulomb, etcetera. So, it appears useful to introduce a new unit measuring the degree of evolution that a certain species has reached at a certain time *t* of Darwinian Evolution, and the obvious name for such a new unit is the "Darwin", denoted by a lower case "d". For instance, if we adopt the exponential evolution curve described in the previous section, we might say that the dominant species on Earth right now (Humans) have reached an evolution level of 50 million darwins. How many darwins may have an alien civilization already reached? Certainly more than 50 millions, i.e. more than 50 Md, but we will not check out until SETI succeeds for the first time. We are not going to discuss further this notion of measuring the "amount of evolution" since we are aware that endless discussions might come out of it. But it is clear to us that such a new measuring unit (and ways to measure it for different species) will sooner or later have to be introduced to make Evolution a fully quantitative science.

GEOMETRIC BROWNIAN MOTION (GBM) AS THE KEY TO STOCHASTIC EVOLUTION WITH AN INCREASING EXPONENTIAL AVERAGE.

We now make a major step ahead: the number *N* in the Statistical Drake Equation, yielding the number of extraterrestrial civilizations now existing and communicating in the Galaxy, is replaced in this section by a ***stochastic process*** *N*(*t*), jumping up and down in time like the number *e* raised to a Brownian motion, but actually in such a way that ***its mean value keeps increasing exponentially in time as*** $\langle N(t) \rangle = N_0\, e^{\mu t}$. This evolution in time of is just what we expect to happen in the Galaxy, where the overall number of ET civilizations does probably ***increase*** (rather than decrease) in time because of the obvious technological evolution of each civilization. But this scenario is a stochastic one, rather than a deterministic one, and certainly does not exclude temporary setbacks, like the end of civilizations due to causes as diverse as: asteroid and comet impacts, rogue planets or stars, arriving from somewhere and disrupting the gravitational stability of the planetary system, supernova explosions that would "fry" entire nearby ET civilizations, ET nuclear wars, and possibly more causes of civilization end that we don't know about yet. The GMB pdf is the lognormal:

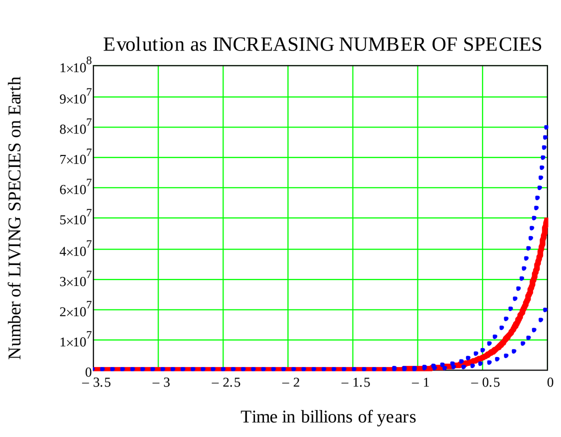
$$N(t)\_\text{pdf}\left(n, N_0, \mu_{GBM}, \sigma_{GBM}, t\right) = \frac{1}{\sqrt{2\pi}\,\sigma_{GBM}\,\sqrt{t}\, n}\, e^{-\frac{\left[\ln(n) - \left(\ln N_0 + \mu_{GBM}\, t - \frac{\sigma_{GBM}^2\, t}{2}\right)\right]^2}{2\sigma_{GBM}^2\, t}} \quad \text{for} \quad 0 \le n \le \infty.$$

GBMs are of paramount importance in the mathematics of finance (Black-Sholes models). We have thus proven that the GBM used in the mathematics of finance is the same thing as the exponentially increasing process yielding the number of communicating ET civilizations in the Galaxy!

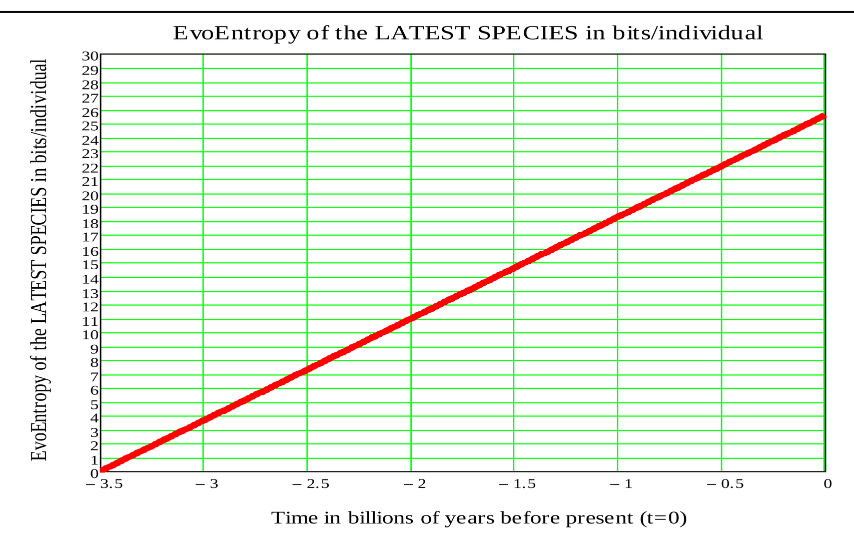


DARWINIAN EVOLUTION AS A GBM IN THE NUMBER OF LIVING SPECIES ON EARTH OVER THE LAST 3.5 BILLION YEARS. Having understood what GBMs are, it is now possible to re-define Darwinian Evolution, as it unfolded on Earth over the last 3.5, as just one single realization of a GBM in the number of living species on Earth over the last 3.5 billion years. All equations are the same, as for the process *N*(*t*), of course: only numbers change.

THE ENTROPY OF EACH b-LOGNORMAL (IN THE PEAK-LOCUS THEOREM) INCREASES LINEARLY IN TIME, JUST AS THE MOLECULAR CLOCK DOES. In Shannon's Information Theory, the ENTROPY of any probability distribution is a measure of how much "poorly peaked" that distribution is. We call "EVO-ENTROPY" the negative of the Shannon Entropy with the further requirement that it started at zero just 3.5 billion years ago. Then, it can be proved mathematically (see refs. [6], [7], [8] and [9]) that THE EVO-ENTROPY of the family of b-lognormals described by the Peak-Locus Theorem INCREASES LINEARLY IN TIME, as shown in the Figure appearing here on the right.



This MATHEMATICAL DISCOVERY has profound implications for life on Exoplanets just as it did on Earth for 3.5 billion years.
In fact, the exponential mean value of the GBM in the number of Species living on Earth over the last 3.5 billion years, shown on the left, reveals that, during the first three billion years the growth of the number of Species was VERY SLOW: basically just MOLECULAR EVOLUTION, and NOT Darwinian Selection. Only when the Cambrian Explosion arrived 542 million years (i.e. half a billion years ago) did the Darwinian Selection proper start to act in the struggle for life among the various competing Species. Darwin himself could hardly imagine that the first three billion years of evolution were NEUTRAL MOLECULAR EVOLUTION in the sense of Moto-o Kimura's Theory (1968): the **neutral theory of molecular evolution** holds that at the **molecular** level most **evolutionary** changes and most of the variation within and between species is not caused by natural selection but by random drift of mutant alleles that are **neutral**. **Now APPLY ALL THAT TO EXTRASOLAR PLANETS: we maintain that 2 or 3 billion years of NEUTRAL MOLECULAR EVOLUTION ARE NECESSARY EVERYWHERE IN THE UNIVERSE ! Thus, SETI may possibly succeed only on exoplanets say 2 or more billion years old !!!**



EvoEntropy (in bits per individual) of the latest species appeared on Earth during the last 3.5 billion years. This shows that a Man (now) is 25.575 bits more evolved than just RNA 3.5 billion years ago.

CONCLUSION. We have provided a new mathematical model capable of accounting for Darwinian Evolution as THE ONE PARTICULAR REALIZATION OF GEOMETRIC BROWNIAN MOTION in the number of living species on Earth that occurred over the last 3.5 billion years of life evolution. In this view, THE SAME MUST HAVE HAPPENED ON ALL OTHER EXOPLANETS, thus restricting the search for Alien Civilizations in SETI and possibly explaining the Fermi paradox.

REFERENCES

[1] Maccone C. (2008), "The Statistical Drake Equation", paper #IAC-08-A4.1.4 presented on October 1st, 2008, at the 59th International Astronautical Congress (IAC) held in Glasgow, Scotland, UK, Sept. 29th thru Oct. 3rd, 2008.
[2] Maccone C. (2010 #1), "The Statistical Drake Equation", Acta Astronautica, Vol. 67 (2010), pages 1366-1383.
[3] Maccone C. (2010 #2), "The Statistical Fermi Paradox", Journal of the British Interplanetary Society, Vol. 63 (2010), pages 222-239.
[4] Maccone C. (2011 #1), "SETI and SEH (Statistical Equation for Habitables)", Acta Astronautica, Vol. 68 (2011), pages 63-75.
[5] Maccone C. (2011 #2), "A Mathematical Model for Evolution and SETI", Origins of Life and Evolution of Biospheres (OLEB), Vol. 41 (2011), pages 609-619.
[6] Maccone C. (2012), "Mathematical SETI", a 724-pages book published by Praxis-Springer in the fall of 2012. ISBN, ISBN-10: 3642274366 | ISBN-13: 978-3642274367 | Edition: 2012.
[7] Maccone C. (2013), "SETI, Evolution and Human History Merged into a Mathematical Model", International Journal of Astrobiology, Vol. 12, issue 3, 2013, pages 218-245.
[8] Maccone C. (2014), "Evolution and Mass Extinctions as Lognormal Stochastic Processes", International Journal of Astrobiology, Vol. 13, issue 4, 2014, pages 290-309.
[9] Maccone C. (2015),), "Evo-SETI Entropy Identifies with Molecular Clock", Acta Astronautica, Vol. 115 (2015), pages 286-290.