International Journal of Astrobiology 11 (2): 103–108 (2012)
doi:10.1017/S1473550411000413

A stochastic process approach of the drake equation parameters

Nicolas Glade[1], Pascal Ballet[2] and Olivier Bastien[3]

[1]*Joseph Fourier University, AGeing, Imagery and Modeling (AGIM) Laboratory, CNRS FRE3405, Faculty of Medicine of Grenoble, 38700 La Tronche, France*
[2]*European University of Brittany (UEB) – University of Brest, Complex Systems and Computer Science Laboratory (LISyC) – EA3883, 20 Avenue LeGorgeu, 29238 Brest Cedex, France*
[3]*Laboratoire de Physiologie Cellulaire Végétale. UMR 5168 CNRS-CEA-INRA-Université Joseph Fourier, CEA Grenoble, 17 rue des Martyrs, 38054, Grenoble Cedex 09, France*
e-mail: olivier.bastien@cea.fr

Abstract: The number N of detectable (i.e. communicating) extraterrestrial civilizations in the Milky Way galaxy is usually calculated by using the Drake equation. This equation was established in 1961 by Frank Drake and was the first step to quantifying the Search for ExtraTerrestrial Intelligence (SETI) field. Practically, this equation is rather a simple algebraic expression and its simplistic nature leaves it open to frequent re-expression. An additional problem of the Drake equation is the time-independence of its terms, which for example excludes the effects of the physico-chemical history of the galaxy. Recently, it has been demonstrated that the main shortcoming of the Drake equation is its lack of temporal structure, i.e., it fails to take into account various evolutionary processes. In particular, the Drake equation does not provides any error estimation about the measured quantity. Here, we propose a first treatment of these evolutionary aspects by constructing a simple stochastic process that will be able to provide both a temporal structure to the Drake equation (i.e. introduce time in the Drake formula in order to obtain something like $N(t)$) and a first standard error measure.

Received 13 November 2011, accepted 6 December 2011, first published online 9 January 2012

Key words: astrobiology, drake formula, Poisson processes, SETI.

Introduction

The number of detectable (i.e. communicating) extraterrestrial civilizations in the Milky Way galaxy is usually calculated by using the Drake equation (Burchell 2006). This equation was established in 1961 by Frank Drake and was the first step to quantifying the Search for ExtraTerrestrial Intelligence (SETI) field (Drake 1965). This formula is broadly used in the fields of exobiology and the SETI. Practically, this equation is rather a simple algebraic expression and its simplistic nature leaves it open to frequent re-expression (Walters *et al.* 1980; Shermer 2002; Burchell 2006; Forgan 2009). While keeping in mind that other equivalent forms exist, we investigate the following form:

$$N^* = R^* f_p n_e f_l f_i f_c L. \quad (1)$$

In this expression, the symbols have the following meanings: N is the number of Galactic civilizations that can communicate with Earth; R^* is the average rate of star formation per year in our galaxy; f_p is the fraction of stars that host planetary systems; n_e is the number of planets in each system which are potentially habitable; f_l is the fraction of habitable planets where life originates and becomes complex; f_i is the fraction of life-bearing planets that bear intelligence; f_c is the fraction of intelligence bearing planets where technology can develop; and L is the mean lifetime of a technological civilization within the detection window.

An additional problem of the Drake equation is the time-independence of its terms (Cirkovic 2004a, b), which for example excludes the effects of the physico-chemical history of the galaxy (Forgan 2009). Indeed, Cirkovic (2004a, b) shows that the main shortcoming of the Drake equation is its lack of temporal structure, i.e., it fails to take into account various evolutionary processes that form a prerequisite for anything quantified by f parameters and n_e. This Drake equation's drawback was mentioned earlier by Franck Drake but the discussion of systematic biases following such simplification was avoided (Drake & Sobel 1991).

In particular, not only some difficulties arise from changing one or more parameters values in equation (1) with time, but also the Drake equation does not provide any error estimation about the measured quantity. To be short, a estimation of $N=5$ with a standard deviation (SD) $SD(N) << 1$ is radically different from an estimation of $N=10$ with a standard error of $SD(N)=10$. Recently, Maccone (2010) derived the first statistical Drake equation by associating each parameter with a random variable and then, given some assumptions, applying the Theorem Central Limit. However, this important new result does not take into account the temporal aspect of the processes of civilizations appearance. Here, we propose a first treatment for these evolutionary aspects by constructing a simple stochastic process that will be able to provide both a temporal structure to the Drake equation (i.e. introduce time in the Drake formula in order to obtain something like $N(t) = (R^* f_p n_e f_l f_i f_c L)(t))$ and a first standard error on $N(t)$.

A stochastic process approach of the Drake equation

Grouping the Drake parameters

When looking at the Drake equation given by equation (1), it is obvious that a kind of Bayesian structure underlying its construction (Shklovsky & Sagan 1966). While the Bayesian structure of the SETI equation has been extensively described by Wilson (1984), the Drake equation has not been analysed in this way. To begin with a heuristics approach, let us consider the three terms of the product $f_l f_i f_c$. For instance, f_c is the estimate (because it is a frequency) of the probability that a technology arise on a planet, knowing the fact that intelligence has appeared. Without worrying about formalism, it is something like $f_c = P(\text{technology}|\text{intelligence})$, where $P(A|B)$ is the conditional probability measure of the event A given B (Capiński & Kopp 2002). In a similar way, f_i is the estimate of the probability that intelligence arise on a planet, knowing the fact that life has appeared ($f_i = P(\text{intelligence}|\text{life})$) and f_l is the estimate of the probability that life arise on a planet, knowing the fact that we considered only potentially habitable planet ($f_l = P(\text{life}|\text{potentially habitable planet})$). More rigorously, if we consider the three set of events (i) E_T (planet bears technology), (ii) E_I (planet bears intelligence) and (iii) E_L (planet bears life), all are subsets of the sample space of all potentially habitable planet, and then it is straightforward that we have $E_T \subset E_I \subset E_L$. As a consequence of this underlying conditional structure, the product of these three previous terms is simply:

$$\begin{aligned} f_c f_i f_l &= P(\text{technology}|\text{intelligence}) P(\text{intelligence}|\text{life}) \\ &\quad \times P(\text{life}|\text{potentially habitable planet}) \\ &= P(\text{technology}|\text{potentially habitable planet}). \end{aligned}$$

This expression is an estimate of the probability that a technological civilization develops on a potentially habitable planet. All these preliminary remarks and heuristics approaches suggest that equation (1) parameters can be grouped together into two new parameters for which the meaning is straightforward:

1. $A = R^* f_p$ is the number of new planetary systems produced in the galaxy per year.
2. $B = n_e f_l f_i f_c$ is the number of advanced intelligent civilizations (AIC) that are able to communicate (and for which we can detect their communication) per planetary systems. AIC can be interpreted as instantiations of the sixth Dick's mega trajectory (Dick 2003), in the same way Cirkovic's ATC (i.e. advanced technological civilizations), are instantiations of the seventh one (Cirkovic & Bradbury 2006).

In other words, AB is the number of new AIC produced per year. L is the average AIC lifetime. Historically, the Drake equation was rather written $N^* = R f_s L$, where R is the average rate of life-supportable star production, f_s is the number of civilizations per suitable star and L is still the average lifetime of an AIC. Equation (1) was established by Shklovsky & Sagan (1966) by expanding f_s.

AIC appearance occurs in a space that is by definition the galactic habitable zone (GHZ). The concept of GHZ was introduced a few years ago as an extension of the much older concept of Circumstellar Habitable Zone (Lineweaver *et al.* 2004). This location is usually considered to be an annulus, with an inner radius of 7 kpc and an outer radius of 9 kpc (1 kpc = 1000 pc ≈ 3000 light years) (Lineweaver *et al.* 2004; Forgan 2011). However, other authors are pointing out that the physical processes underlying the former concept are hard to identify and that the entire Milky Way disk may well be suitable for complex life (Prantzos 2008; Gowanlock *et al.* 2011). In this paper, we will consider the entire Milky Way disk to be suitable for the complex life, i.e. to be the GHZ. This will also allow us to use estimation of R^* for the entire galaxy (Diehl *et al.* 2006).

This above new parameter grouping, the need for a temporal structure to the Drake equation and the reasonable assumption that AIC appearance should be roughly random in time and in space (this assumptions are discussed in the following part), strongly suggest that an AIC appearance mathematical model could be made by using a stochastic process like a Poisson process $\{N(t){:}t \geqslant 0\}$ with rate parameter $\lambda = AB$.

Poisson process

A Poisson process is a continuous-time stochastic process in which events occur continuously and independently of each others. Examples that are well-modelled as Poisson processes include the radioactive decay of atoms (Foata & Fuchs 2002), Turing machine rules mutations (Glade *et al.* 2009), the arrival of customers in a queue n telephone calls arriving at a switchboard and proteins evolution (Bastien 2008; Ortet & Bastien 2010). The Poisson process is a collection $\{N(t){:}t \geqslant 0\}$ of random variables, where $N(t)$ is the number of events, often called "top", that have occurred up to time t (starting from time 0). The number of events between time a and time b is given as $N(b) - N(a)$ and has a Poisson distribution. Each realization of the process $\{N(t){:}t \geqslant 0\}$ is a non-negative integer-valued step function that is non-decreasing in time. In our case, each "top" could be an AIC appearance. Hence, $N(t)$ would be the number of AIC that has appeared up to time t.

Definition of a homogeneous Poisson process

The homogeneous Poisson process is one of the most well-known Lévy processes (Itô 2004). A continuous-time counting process $\{N(t){:}t \geqslant 0\}$ will be called a Poisson process if it possesses the following properties:

1. $N(0) = 0$.
2. Independent increments (the numbers of occurrences counted in disjoint intervals are independent from each other).
3. Stationary increments (the probability distribution of the number of occurrences counted in any time interval only depends on the length of the interval).
4. No counted occurrences are simultaneous. More precisely, the process is locally continuous in probability, i.e., for all $t \geqslant 0$, $\lim_{h \to 0}\{P(N(t+h) - N(t)\} = 0$.

In our model, condition 1 means that we must begin the AIC count at a time when no previous AIC exists. Condition 2 means that each AIC evolves independently from each others. Condition 3 means that the number of AIC in a time interval does not depend on the date at which we sample this interval, that is to say invariance of physical law and global homogeneity of the space–time in the considered galaxy region, i.e. the GHZ (Gonzales *et al.* 2001; Cirkovic 2004a, b; Lineweaver *et al.* 2004; Vukotic & Cirkovic 2007; Prantzos 2008; Gowanlock *et al.* 2011). If these conditions are satisfied, then we can deduce the following results.

1. Consequences of this definition include: The probability distribution of $N(t)$ is a Poisson distribution. That is to say $P(N(t+\tau)-N(t)=k)=(\lambda\tau)^k e^{-\lambda\tau}/k!$ where $N(t+\tau)-N(t)$ is the number of events between the time interval $[t,\ t+\tau]$ and λ is the stochastic process rate parameter also called density or intensity. The product $\lambda\tau$ is called the parameter of the Poisson distribution.
2. The probability distribution of the waiting time until the next event occurs is an exponential distribution.
3. The occurrences are distributed uniformly on any interval of time. (Note that $N(t)$, the total number of occurrences, has a Poisson distribution over $[0, t]$, whereas the location of an individual occurrence on t in $[a, b]$ is uniform.)

Homogeneous and non-homogeneous Poisson process

As recalled above, a homogeneous Poisson process is characterized by its rate parameter λ, which is the expected number of events (also called arrivals) that occur per unit time. Nevertheless, in general, the rate parameter may change over time; such a process is called a non-homogeneous Poisson process or inhomogeneous Poisson process. In this case, the generalized rate function is given as $\lambda(t)$, where $\lambda(t)$ is a real continuous function of time (and hence, defined on the positive part of the real axis). In this case, the above definition of a Poisson process remains unchanged except for the third condition (stationary increments). Then, the three Poisson process conditions are:

1. $N(0)=0$.
2. Independent increments (the numbers of occurrences counted in disjoint intervals are independent from each other).
3. Let $\rho(t)=\int_0^t \lambda(u)\mathrm{d}u$; then for all pair (s,t) with $0 \leqslant s < t < +\infty$, the number $N(t)-N(s)$ of events occurring in $[s,\ t]$ is a Poisson random variable with parameter $\mu(]s,\ t])=\rho(t)-\rho(s)=\int_s^t \lambda(u)\mathrm{d}u$.

This last condition implies the fourth of the definition of the homogeneous processes, that is to say: the process is locally continuous in probability. Moreover, it can be demonstrate (Foata & Fuchs 2002) that this condition is similar to the following:

Condition 3bis. For $h \to 0$, we have $P(N(t+h)-N(t)=1)=\lambda(t)+o(h)$ and $P(N(t+h)-N(t)\geqslant 2)=o(h)$.

Usually, $m(t)$ is called the renewable function. A remarkable result is that all non-homogeneous Poisson process can be transformed into a homogeneous Poisson process by a time transformation (Foata & Fuchs 2002). Of course, a homogeneous Poisson process may be viewed as a special case when $\lambda(t)=\lambda$, a constant rate.

AIC birth and death process

Here, we construct a simple stochastic process that will represent the stochastic appearance of AIC in the Milky Way as a function of time. More exactly, we will construct a model that will give us the number $C(t)$ of existing AIC for a given time t. A first reasonable hypothesis for this model is that there exist a time t_0 for which no AIC is present in the galaxy, i.e. $C(t_0)=0$. So, without loss of generality, we can consider a stochastic process for which the first condition, $C(0)=0$, for the process to be a Poisson process is true. A second hypothesis for this model is that AIC appearances (i.e. births) are independent from each other (communication between them does not influence their birth or their lifetime). As a consequence, a possible limitation of the present model could come from the fact that there is a legitimate case to be made that AIC numbers may violate this condition of the Poisson distribution: the longevity, may be significantly affected by the discovery of a long-lived intelligent community, possibly leading to clustering in time. A third hypothesis is that the number of AIC in a time interval does not depend (at least locally) on the date at which we sample this interval.

With these three hypotheses, we can consider the stationary increments Poisson stochastic process $\{N(t){:}t \geqslant 0\}$ with density $\lambda > 0$ (Itô 2004) where each “top” of the Poisson process is an AIC birth event. As the rate parameter λ represents the expected number of events that occur per unit time, it is clear that it is equal to the number of new AIC produced per year, that is to say the product AB of the upper new Drake equation parameter grouping. We also suppose that each AIC lifetime is a random variable X, i.e. all AIC during time are independent, identically distributed and are independent of the process $\{N(t){:}t \geqslant 0\}$. Let S_k be the birth date of the kth AIC and X_k its lifetime. Then, its death date is S_k+X_k. Let $S_0{:}=X_0{:}=0$ and $C(t)$ be the number of AIC present at the time t. The question is to evaluate the probability law of $C(t)$, with the hypothesis $N(0)=0$ (and hence $C(0)=0$). The following theorem for the current AIC number can then be formulated (for the proof, see the Appendix):

Theorem 1. *Theorem for the current AIC number: with the previous hypotheses and notations, the number $C(t)$ of AIC present at the time t is a Poisson distributed random variable with parameter $m(t)=\lambda\int_0^t r(u)\,du$, where λ is the rate parameter of the AIC appearance Poisson stochastic process and $r(u)$ is the survival function of AIC lifetime random variable X.*

To resume, the Poisson stochastic process $\{N(t){:}t \geqslant 0\}$, which represents the appearance of new AIC in the Milky Way is combined with a random variable X, which represents the lifetime of these new civilizations. The main result is that the number $C(t)$ of AIC present at time t is a Poisson distributed random variable with parameter $m(t)$. This general result can be link to the classical Drake equation by the following. As t tends to infinity and using a classical theorem

Table 1. *Drake equation parameter estimations*

Parameters	Significations	1961's estimation (Drake and Sobel 1991)	Recent estimations
N	The number of Galactic civilizations that can communicate with Earth		
R^*	The average rate of star formation per year in our galaxy	10 per year	Seven per year (Diehl *et al.* 2006)
f_p	The fraction of stars that host planetary systems	0.5	0.5 (Maccone 2010)
n_e	The number of planets in each system that are potentially habitable	2	1 (Maccone 2010)
f_l	The fraction of habitable planets where life originates and becomes complex	1	0.5 (Maccone 2010)
f_i	The fraction of life-bearing planets that bear intelligence	0.01	0.2 (Maccone 2010)
f_c	The fraction of intelligence bearing planets where technology can develop	0.01	0.2 (Maccone 2010)
L	The mean lifetime of a technological civilization within the detection window	10 000	See text

of probability about survival function (which stated that $\int_0^{+\infty} r(u)\mathrm{d}u = E[X]$; Skorokhod & Prokhorov 2004), we observe that $\lim_{t\to+\infty} m(t) = \lambda\int_0^{+\infty} r(u)\mathrm{d}u = \lambda E[X]$. As a consequence, $C(t)$ is going towards a Poisson random variable V with parameter $\lambda E[X]$. This result is exactly the Drake equation with $E[V] = \lim_{t\to+\infty} m(t)$, $E[X]=L$ and $\lambda=AB$. Interestingly, if one assumes that the lifetime of any galactic civilization is finite, that is to say has an upper bound. Then, their exists a number t_M for which $r(t_M)=0$ and so the limit value of $m(t)$ will be reached at finite time, that is to say we will have $\int_0^{t_M} r(u)\mathrm{d}u = E[X]$.

Discussion

Mean and variance of the number of AIC

As stated above, as t becomes larger than the AIC maximum lifetime, the previous approximation becomes exact. So we have $E[V(t)]=\mathrm{VAR}[V(t)]=\lambda E[X]$ (the mean and the variance of a Poisson distributed random variable are equal; Skorokhod & Prokhorov 2004) and the three terms of the equality are time independent. As a consequence, we can study the coefficient of variation ε of the $C(t)$ (also named fluctuation around the mean) which is the ratio between the SD and the mean of the stochastic process. Here, we have

$$\varepsilon = \frac{1}{\sqrt{\lambda \mathrm{E[X]}}} = \frac{1}{\sqrt{R^* f_p n_e f_l f_i f_c E[X]}}.$$

With the 1961's Drake parameters estimation (see Table 1), $\lambda=0.001$ and hence, depending on the AIC lifetime, we can have

1. with $E[X]=200$, we obtain $E[V(t)]=0.2$ and $\varepsilon=2.23$,
2. with $E[X]=10\ 000$, we obtain $E[V(t)]=10$ and $\varepsilon=0.31$.

An error of magnitude one order in any parameter can lead to a estimation of $E[V(t)]$ and ε equal to 1. For example, a lower bound can be estimated for $E[X]$, while considering the span time between now and the invention of the parabolic telescope, i.e. radioastronomy (Reber & Conklin 1938). This gives $E[X]=73$, and so $E[V(t)]=0.072$ and $\varepsilon=3.9$.

The dramatic effect of parameter evolution estimations on a possible value of $N(t)$ can be seen in the following. Indeed, with more recent estimations (Diehl *et al.* 2006; Maccone 2010), we obtain $\lambda=0.07$ and hence, we can have

1. with $E[X]=200$, we obtain $E[V(t)]=14$ and $\varepsilon=0.27$,
2. with $E[X]=10\ 000$, we obtain $E[V(t)]=700$ and $\varepsilon=0.04$,
3. with $E[X]=73$, we obtain $E[V(t)]=5$ and $\varepsilon=0.45$.

All these considerations give a large probability for detection of another AIC, depending on the reliability of the new estimations. Starting from $V(t_0)=0$, the Poisson stochastic process theory allow us to estimate the average mean time for the occurrence of a new AIC appearance, which is given by the inverse of λ (Foata & Fuchs 2002). With the 1961's Drake parameters estimation (respectively, recent estimation), this average time is equal to 1000 years (respectively, 14 years). Apply to the SETI research programme, if we suppose that we are alone in our universe, the average mean time for the occurrence of another AIC should be roughly 14 years with recent parameter estimations (respectively, 1000 years for the 1961's Drake estimation), with a standard error of the same order error.

Conclusion

The proposed model allows a first analytic estimation of the SD of the number of Galactic civilization estimate. In addition, it provides a temporal structure of the Drake equation which can help study the influence of several effects on the number of Galactic civilization estimate. An important case is the notion of global regulation mechanism (i.e. a dynamical process preventing uniform emergence and development of life all over the Galaxy; Annis 1999; Vukotic & Cirkovic 2008). Vukotic & Cirkovic (2007) investigated the effects of a particular global regulation mechanism, the Galactic gamma-ray bursts (GRBs) (colossal explosions caused either by terminal collapse of supermassive objects or mergers of binary neutron stars), on the temporal distribution of hypothetical inhabited planets, using simple Monte Carlo numerical experiments. Here, GRB is clearly just one of the possible physical processes for resetting astrobiological clocks. They obtain that the times required for biological evolution on habitable planets of the Milky Way are highly correlated. More precisely, using simulations cosmological observations (Bromm & Loeb 2002), they demonstrated that the correlation (and so the covariance $\mathrm{cov}(t_b, t^*)$) between the biological timescale t_b and the astrophysical timescale t^* is non-zero. Using the distribution of GRB over the time, an analytic approach would be to compute the random time T since the last GRB event. Assuming that T is independent from both $N(t){:}t \geqslant 0$, the Poisson stochastic process that represents the appearance of new AIC, and X which represents the lifetime of these new civilizations, Theorem 1 can be replaced by a more general form:

Theorem 2. *Theorem for the current AIC number under global regulation mechanisms: with the previous hypotheses and notations, the number $C(T)$ of AIC present at the time T is a Poisson distributed random variable with parameter $m(t) = \lambda \int_0^t r(u)\, du$, where λ is the rate parameter of the AIC appearance Poisson stochastic process, $r(u)$ is the survival function of AIC lifetime random variable X and T is the random variable 'time elapsed since the last major global regulation event'.*

This result follows from the above fact that $E[C(t)] = m(t) = \lambda \int_0^t r(u)\mathrm{d}u$, which can be rewritten as $E[C(T)|T=t]=m(t)$, where $E[.|.]$ is the conditional expectation operator. This last formula is called the conditional expectation of C given $T=t$. Since we do not know the true value of t since the last GRB event, we must consider the new random variable $m \circ T = m(T) = E[C|T]$. Future works will explore this result. Especially, it would be suitable to obtain a formula $E[C(T)] = \lambda E[\int_0^T r(u)\mathrm{d}u]$ where the expectation is computed using the global regulation event (in particular GRB events) distribution over the time. This analytic approach could provide an accurate analysis of classical Monte-Carlo simulations (Vukotic & Cirkovic 2008; Forgan 2009, 2011; Hair 2011). For instance, Hair (2011) and Forgan (2011) had proposed in their two models that the distribution of the civilization arrival times is Gaussian-distributed. In principle, this is equivalent to allowing the AIC appearance rate parameter λ, to varying in time. More precisely, the rate parameter $\lambda(t)$ corresponding to the Forgan model (Forgan 2011) should be an increasing function of $[0, \mu]$, where μ is the mean of the Hair (2011) and Forgan (2011) arrival time Gaussian distribution (μ has the same order of magnitude than the Hubble time, $t_H = 13\,700$ Myr), and a decreasing function of $[\mu, +\infty]$. Nevertheless, this work is mainly the first approach to model AIC appearance, but future studies would have to address the fact that AIC appearance in habitable planets should be correlated with the Galaxy's star formation history (Heavens *et al.* 2004; Juneau *et al.* 2005; Vukotic & Cirkovic 2007) and the location of the GHZ (Gonzalez *et al.* 2001; Lineweaver *et al.* 2004; Prantzos 2008; Gowanlock *et al.* 2011). For instance, Planet formation and star formation could be included in the first AIC appearance model by extracting the original Drake parameters R^*, f_p and n_e from λ and let them to vary in time.

Appendix: Demonstration of the theorem

The proof of the theorem can be divided into three parts (Foata & Fuchs 2002).

Lemma 1. *Let $r(u) := P\{X \geqslant u\}$ be the survival function of X and $\bar{r}(t) := \frac{1}{t}\int_0^t r(u)\mathrm{d}u$. If U is a uniform random variable on $[0, t]$, independent of X, the survival function of $U+X$ is given by $P\{U + X \geqslant t\} = \bar{r}(t)$.*

Proof. Following the fact that the density of U is $\frac{1}{t}I_{[0,t[}$, where $I_{[0,t[}(x)$ is 0 outside the interval [0,t[, we have

$$
\begin{aligned}
P\{U+X \geqslant t\} &= \frac{1}{t}\int_0^t P\{U+X \geqslant t | U = s\}\,\mathrm{d}s \\
&= \frac{1}{t}\int_0^t P\{X \geqslant t-s | U = s\}\,\mathrm{d}s \\
&= \frac{1}{t}\int_0^t P\{X \geqslant t-s\}\,\mathrm{d}s,
\end{aligned}
$$

$$
P\{U+X \geqslant t\} = \frac{1}{t}\int_0^t r(t-s)\mathrm{d}s = \frac{1}{t}\int_0^t r(u)\,\mathrm{d}u,
$$

which are equal to $\bar{r}(t)$ by definition.

Lemma 2. *The generating function $h(u) := E[u^{C(t)}]$ of $C(t)$ is given by $e^{-\lambda t \bar{r}(t)(1-u)}$.*

Proof. For $k \geqslant 0$, let $Y_k := I_{\{S_k + X_k \geqslant t\}}$. Obviously we have $C(t) = \sum_{k=0}^{N(t)} Y_k$, and so

$$
\begin{aligned}
h(u) &= E\left[u^{C(t)}\right] = \sum_{n\geqslant 0} E\left[u^{C(t)}|N(t)=n\right]P\{N(t)=n\} \\
&= \sum_{n\geqslant 0} E\left[u^{Y_0+Y_1+\ldots+Y_n}|N(t)=n\right]P\{N(t)=n\},
\end{aligned}
$$

$$
h(u) = \sum_{n\geqslant 0} E\left[\prod_{k=0}^{n} u^{Y_k}|N(t)=n\right]P\{N(t)=n\}.
$$

Conditionally to the event $\{N(t)=n\}$ $(n \geqslant 1)$, the system $(S_1, S_2, \ldots, S_n)$ has the same distribution than the system $(U_1, U_2, \ldots, U_n)$ of independent and uniformly distributed on $[0, t]$ random variables.

For $1 \leqslant k \leqslant n$, let $Z_k := I_{\{U_k + X_k \geqslant t\}}$. We have

$$
h(u) = \sum_{n\geqslant 0} E\left[\prod_{k=0}^{n} u^{Z_k}\right]P\{N(t)=n\} = \sum_{n\geqslant 0}\left(E\left[u^{Z_1}\right]\right)^n P\{N(t)=n\}.
$$

But Z_1 is a Bernoulli random variable with parameter $P\{U_1 + X_1 \geqslant t\} = \bar{r}(t)$. Hence, $E\left[u^{Z_1}\right]$ is the generating function of the Z_1 random variable and is given by $E\left[u^{Z_1}\right] = \sum_{k=0}^{\infty} P(z=k) z_1^k = 1 - \bar{r}(t) + u.\bar{r}(t)$ (Korolouk 1978). Finally, we obtain

$$
h(u) = \sum_{n\geqslant 0} \frac{e^{-\lambda t}(\lambda t)^n}{n!}\left(1 - \bar{r}(t) + u.\bar{r}(t)\right)^n = e^{-\lambda t \bar{r}(t)(1-u)}.
$$

Proof of the theorem. We can write $h(u) = e^{-m(t)(1-u)}$ which is the generating function of a composed Poisson distribution (Korolouk 1978) and so, is a Poisson distributed random variable with parameter $m(t) = \lambda \int_0^t r(u)\mathrm{d}u$.

Acknowledgements

O.B. was supported by the French Agence Nationale de la Recherche, as part of the ReGal project.

Author disclosure statement

No competing financial interests exist.

References

Annis, J. (1999). An astrophysical explanation for the great silence. *J. Br. Interplanet. Soc*. **52**, 19–22.

Bastien, O. (2008). A simple derivation of the distribution of pairwise local protein sequence alignment scores. *Evol. Bioinf*. **4**, 41–45.

Bromm, V. & Loeb, A. (2002). The expected redshift distribution of gamma-ray bursts. *Astrophys. J*. **575**, 111.

Burchell, M.J. (2006). W(h)ither the drake equation? *Int. J. Astrobiol*. **5**(3), 243–250.

Capiński, M. & Kopp, E. (2002). *Measure, Integral and Probability*. Springer-Verlag, Berlin.

Cirkovic, M.M. (2004a). Earths: rare in time, not space? *J. Br. Interplanet. Soc*. **57**, 53.

Cirkovic, M.M. (2004b). The temporal aspect of the drake equation and SETI. *Astrobiology* **4**(2), 225–231.

Cirkovic, M.M. & Bradbury, R.J. (2006). Galactic gradients, postbiological evolution and the apparent failure of SETI. *New Astron*. **11**, 628–639.

Dick, S.J. (2003). Cultural evolution, the postbiological universe and SETI. *Int. J. Astrobiol*. **2**, 65–74.

Diehl, R., Halloin, H., Kretschmer, K., Lichti, G.G., Schönfelder, V., Strong, A.W., von Kienlin, A., Wang, W., Jean, P., Knödlseder, J. *et al.* (2006). Radioactive ^{26}Al from massive stars in the Galaxy. *Nature* **439**, 45–47.

Drake, F. (1965). The radio search for intelligent extraterrestrial life. In *Current Aspects of Exobiology*, ed. Mamikunian, G. & Briggs, M.H., pp. 323–345. Pergamon, New York.

Drake, F. & Sobel, D. (1991). *Is anyone out there?* Simon and Schuster, London.

Foata, D. & Fuchs, A. (2002). *Processus Stochastiques. Processus de Poisson, chaînes de Markov et martingales*. Dunod, Paris.

Forgan, D.H. (2009). A numerical testbed for hypotheses of extraterrestrial life and intelligence. *Int. J. Astrobiol*. **8**(2), 121–131.

Forgan, D.H. (2011). Spatio-temporal constraints on the zoo hypothesis, and the breakdown of total hegemony. *Int. J. Astrobiol*. **10**, 341–347.

Glade, N., Ben Amor, H.M. & Bastien, O. (2009). Trail systems as fault tolerant wires and their use in bio-processors. In *Modeling Complex Biological Systems in the Context of Genomics, Proceedings of the Spring School, Nice 2009*, ed. Amar, P. *et al.*, pp. 85–119.

Gonzalez, G., Brownlee, D. & Ward, P. (2001). The galactic habitable zone: galactic chemical evolution. *Icarus* **152**(1), 185–200.

Gowanlock, M.G., Patton, D.R. & McConnell, S.M. (2011). A model of habitability within the milky way galaxy. *Astrobiology* **11**(9), 855–873.

Hair, T.W. (2011). Temporal dispersion of the emergence of intelligence: an inter-arrival time analysis. *Int. J. Astrobiol*. **10**, 131–135.

Heavens, A., Panter, B., Jimenez, R. & Dunlop, J. (2004). The star-formation history of the Universe from the stellar populations of nearby galaxies. *Nature* **428**, 625–627.

Itô, K. (2004). *Stochastic Processes*. Springer-Verlag, Berlin.

Juneau, S., Glazebrook, K., Crampton, D., McCarthy, P.J., Savaglio, S., Abraham, R., Carlberg, R.G., Chen, H.W., Borgne, D.L., Marzke, R.O *et al.* (2005). Cosmic star formation history and its dependence on galaxy stellar mass. *Astrophys. J. Lett*. **619**(2), L135.

Korolíouk, V. (1978). Aide-Mémoire de théorie des probabilités et de statistique mathématique. Edition de Moscou, Moscou. pp. 51.

Lineweaver, C.H., Fenner, Y. & Gibson, B.K. (2004). The galactic habitable zone and the age distribution of complex life in the Milky way. *Science* **303**(5654), 59.

Maccone, C. (2010). The statistical drake equation. *Acta Astronaut*. **67**(11–12), 1366–1383.

Ortet, P. & Bastien, O. (2010). Where does the alignment score distribution shape come from? *Evol. Bioinf*. **6**, 159–187.

Prantzos, N. (2008). On the "galactic habitable zone". *Space Sci. Rev*. **135**, 313–322.

Reber, G. & Conklin, E.H. (1938). UHF receivers. *Radio* **225**, 112.

Shermer, M. (2002). Why ET Hasn't Called. *Scientific American* **8**, 21.

Shklovsky, I.S. & Sagan, C. (1966). *Intelligent Life in the Universe*. Holden-Day, San Francisco.

Skorokhod, A.V. & Prokhorov, I.U.V. (2004). *Basic Principles and Applications of Probability Theory*. Springer-Verlag, Berlin.

Vukotic, B. & Cirkovic, M.M. (2007). On the timescale forcing in astrobiology. *Serb. Astron. J*. **175**, 45–50.

Vukotic, B. & Cirkovic, M.M. (2008). Neocatastrophism and the Milky way astrobiological landscape. *Serb. Astron. J*. **176**, 71–79.

Walters, C., Hoover, R.A. & Kotra, R.K. (1980). Inter- stellar colonization: A new parameter for the Drake. Equation? *Icarus* **41**, 193–197.

Wilson, T.L. (1984). Bayes' theorem and the real SETI equation. *Q. J. R. Astron. Soc*. **25**, 435–448.