IS THE “STAR TREK” CONCEPT OF ALIENS REALLY NAIVE? – A COMMENT ON DURIC AND FIELD

M. M. Ćirković

Astronomical Observatory, Volgina 7, 11160 Belgrade 74, Serbia and Montenegro

(Received: December 18, 2003; Accepted: April 13, 2004)

SUMMARY: Several assumptions used in the recent review paper on the search for extraterrestrial intelligence (SETI) by Duric and Field (2003) are critically discussed. New astrobiological results—in particular those on the age distribution of terrestrial planets—enable relaxing of some of these assumptions, which gives further justification to optimism regarding SETI projects.

Key words. Astrobiology – Extraterrestrial intelligence – Radio lines: general – Galaxy: stellar content

In recent fascinating Invited Review paper in this journal, Duric and Field (2003) have surveyed the field of Search for ExtraTerrestrial Intelligence (henceforth SETI) and suggested several novel points and strategies. Although the basic tone of the review is mildly optimistic toward SETI, we wish to argue here that this optimism in fact receives a significant boost when several unwarranted assumptions are identified and rejected. In particular, we wish to argue that there is a significant chance that an appreciable number of extraterrestrial societies is more similar to ours than it is a priori expected. This shift occurs on purely physical level, without taking explicitly into account any aspect of the communication itself.

(I) The first point is largely of formal nature. In calculations performed in Sec. 5.1, Duric and Field use the volume of the Galactic disk, $V_G \approx 4 \times 10^{11} \text{ pc}^3$ in order to obtain the density of extraterrestrial civilizations. However, the relevant density is the one associated with the Galactic Habitable Zone (henceforth GHZ; Gonzalez, Brownlee and Ward 2001), an annular ring comprising the potential sites for genesis of complex metazoans in the Milky Way. Unfortunately, the exact boundaries of GHZ are not yet determined, but the following rough estimate suggests that the difference is relevant for the present purposes.

The outer GHZ boundary is set by the metallicity gradient in the Milky Way disk. According to most observations, this gradient is about $\nabla Z = 0.09 \text{ dex kpc}^{-1}$ (e.g. Tadross 2003). In order to determine the outer boundary, we need an assumption on the mass scaling of terrestrial planets, as well as the minimal mass of the terrestrial planet viable from the astrobiological point of view. The former is suggested by Gonzalez et al. (2001) to be the scaling $(M) \propto 10^{1.5Z}$ (we assume that [Fe/H] represents overall metallicity sufficiently well). The latter is a topic of considerable debate in astrobiological circles, but we can take as a prototype often expressed opinion that an object of the size of Mars $(M \approx 0.1 \oplus)$ is the minimum, for several reason (development of plate tectonics, retention of $O_2$ and $O_3$, etc.). In principle, we shall have

$$R_{\text{out}} = R_s - \frac{1}{\nabla Z} \log \left( 1 + \frac{2}{3} \log \frac{M_{\text{min}}}{M_{\oplus}} \right), \quad (1)$$

where $R_s = 8.5 \text{ kpc}$ is the Solar galactocentric distance. In this approximation, we obtain $R_{\text{out}} = 13.5 \text{ kpc}$, which is significantly less than the Holmberg radius of the Milky Way.
As far as the inner boundary of GHZ is concerned, there are two major factors influencing it: dynamical stability and radiation field (due to high rate of supernovae and $\gamma$-ray bursts). Let us briefly and very roughly consider only the first one (for calculations of the latter see below). The rate of stellar encounters is $\Gamma = \langle n_\sigma \rangle \approx \langle n \rangle / \sigma(h)$, where $\sigma$ is the cross-section of stars to be exponential, with the disk lengthscale of $R_d = 3$ kpc (e.g. Drimmel and Spergel 2001).

Let us define the critical value of $\Gamma$ as $\Gamma_{\text{crit}} = 1$ (age of the Earth)$^{-1}$. According to most of recent calculations and simulations, an encounter disruptive enough to destroy terrestrial planets either directly or indirectly (through perturbing orbits of outer planets sufficiently to create high eccentricities driving the habitable planets into their sun) occurs with an average cross-section of $\langle \sigma \rangle = \pi (30 \text{ AU})^2$. At the Solar circle, the average relative stellar velocity is $\langle v \rangle \approx 40$ km s$^{-1}$; however, in the interior of the Galactic bulge, where the inner boundary is located, we should use larger value, say $\langle v \rangle \approx 122$ km s$^{-1}$ suggested by Freeman et al (1988). In this approximation, we have

$$R_{\text{inn}} = R_s - R_d \ln \frac{\Gamma_{\text{crit}}}{\langle \sigma \rangle (v)(n_0)}.$$  \hspace{1cm} (2)

Using the assumptions above we obtain that for purely dynamical constraints we have $R_{\text{inn}} = 1.3$ kpc.

Thus, the real density of ETI sites is increased by a factor $V_G / V_{\text{GHZ}} \approx 4$. However, since (although it is not entirely clear from the derivation that this is the case) $N_c$ is perhaps derived from an assumption about the total stellar census of the Milky Way, the reduction of the area of interest to GHZ will have smaller effect. It will not entirely cancel the change in volume, since the distribution of stars is inhomogeneous, notably decreasing exponentially with the galactocentric radius. Thus, stellar density inside GHZ is—for the same normalization constants and the same value of $R_s$—larger than the average Galactic stellar density by the factor of $\langle n \rangle_{\text{GHZ}} / \langle n \rangle_{\text{tot}} \approx 1.3$.

Calculations of Lineweaver, Fenner and Gibson (2004) — which improve on the study of Lineweaver (2001) — indicate that an inner bound due to probability of supernovae exploding too close and too often to life-bearing planetary systems is much more conservative. Their study indicates that the probability of a planet being undisturbed for sufficient amount of time to evolve complex life is rather small within an inner boundary of about 7 kpc or so. If that is the case, the conclusions of this section are strongly reinforced, since the Galactic disk is still farther from being uniform in the astrobiological sense. (Although, of course, the exact value depends on still poorly understood effects of close supernovae on planetary environments.) The same general conclusion follows from the estimate of Peña-Cabrera and Durand-Manterola (2004), who find $R_{\text{inn}} \gtrsim 4$ kpc, on the basis of the assumption that metal-rich planets in the inner Galaxy would grow too fast to be habitable: this assumption, however, is highly questionable (see, e.g. Léger et al. 2004).

(II) The second issue deals with the nature of spatial size of a "typical" Galactic civilization vs. the duration of the "window of opportunity". The underlying assumption of the Duric and Field analysis is that no matter how small or big the window is, an ETI civilization stays local, i.e. is to be found only in the vicinity of its domicile star.\footnote{With this assumption applied consequently, the (in)famous Fermi’s paradox (Brin 1983, Webb 2002) is automatically solved, so there must be something wrong with it, or we would not bother with the puzzle for 50+ years!} However, this is a highly doubtful assumption, probably only for very small values of the window of opportunity $\tau_c$. For larger values, we expect other planetary systems to be colonized and to behave as the secondary sources of detectable artificial emissions. In other words, propagation of ETI signals would be guided by a rough analogue of the famous Huygens’ principle of wave optics: a point struck by a wavefront, becomes a secondary source of waves of the same wavelength.

The cumulative probability of colonizing other planetary systems certainly increases with time. For windows as large as $10^8$ yrs, it may well be close to unity. In any case, we should expect that the graph of the number of SETI targets per ETI civilization vs. $\tau_c$ becomes (after lingering at unity for some time) an ever-ascending line.

(III) Finally, the most important problem with most of the theoretical SETI studies so far (including the one of Duric and Field) is the unsupported underlying gradualism (also known as uniformitarianism of rate; e.g. Gould 1987), which is rarely even stated as an assumption, but which critically impacts conclusions of any such study. It’s most clearly seen from the statements in Sec. 5.1.: “Under the optimistic scenario, there may be as many as $10^6$ technologically advanced civilizations in the Galaxy. However, these societies are at various stages of development. The probability that two extraterrestrial societies are at the same stage of evolution, to say within a million years, is very small.”

This sounds perfectly reasonable at first glance, but contains a hidden key assumption; namely, it would have been true only if civilizations’ ages were picked at random from a uniform (or at least broad enough) distribution. This gradualist assumption is assumed in most of astrobiological and SETI treatises, usually without a slightest attempt to justify it. In a field as young and immature as SETI studies certainly are, it is necessary to identify and justify each assumption. But one can go even further; is fact, we wish to suggest here that the gradualist assumption is indeed unjustified, and that there is a whole class of viable astrobiological models ("phase transition" ones, after Annis 1999) that offer clear physical reason for rejection of this assumption.
Gradualism, parenthetically, has not shone as a brilliant guiding principle in astrophysics and cosmology. It is well-known, for instance, how the strictly gradualist (and from many points of view methodologically superior) steady-state theory of the universe of Bondi and Gold (1948), as well as Hoyle, has after the "great controversy" of 1950s and early 1960s succumbed to the rival evolutionary models, now known as the standard ("Big Bang") cosmology (Kragh 1996). Balashov (1994) has especially stressed this aspect of the controversy by showing how deeply justified was the introduction | by the Big Bang cosmologists | of events and epochs never seen or experienced. Similar arguments are applicable in the nascent discipline of astrobiology, which might be considered to be in an analogous state today as cosmology was half a century ago (Čirković 2004a).

First, it is worth noticing that the uniform age distribution assumption is indeed untenable at closer look by the very nature of physical environment of ETIs. The age distribution of terrestrial planets in the Galaxy is not uniform. The seminal study of Lineweaver (2001) has demonstrated, for instance, that Earth-like planets around other stars in the Galactic habitable zone are, on the average 1.8 (±0.9) Gyr older than our planet (see also Franck et al. 2001). His calculations are based on chemical enrichment as the basic precondition for the existence of terrestrial planets. But this result—the beginning of possible restrictions of gradualism—reflects only the most fundamental constraint. It is entirely plausible that the history of the Galaxy is divided still finer into several distinct periods with radically different conditions for life. The impact on the parameters of the Drake equation of such a partition is drastic (Čirković 2004b). In that case, only weighted relative durations are relevant, not the overall age. The net effect of such weighting would be to significantly increase the value of \(N_c\) (the number of civilizations capable of communicating with us in Duric and Field 2003).

Exactly such a picture is presented by a class of phase transition models (Clarke 1981, Annis 1999, see also Norris 2000), which assume a global regulation mechanism for preventing the formation of complex life forms and technological societies early in the history of the Galaxy. Such a global mechanism could have the physical form of \(\gamma\)-ray bursts, if it can be shown that they exhibit sufficient lethality to cause mass biological extinctions over a large part of the volume of the Galactic habitable zone (Scalo and Wheeler 2002, see also Thorsett 1995, Melott et al. 2004). If, as maintained in these models, continuous habitability is just a myth, the validity of the Drake equation (and the spirit in which it was constructed and used) is seriously undermined. Carter’s argument (Sec. 2.3.3. of Duric and Field 2003) is also diffused by this strategy: it is obvious that the timescales are not correlated, due to the forcing introduced by the timescale of the global regulation mechanism (for instance, the local frequency of \(\gamma\)-ray bursts).

Thus, what can be jovially called the Star Trek model of Galactic ETIs (independently arising civilizations of roughly the same age) is not at all so preposterous or naive as usually assumed in the academic discussions of SETI. Phase transition models would indeed imply such a nearly constant age of ETI enclaves.

Intuitively, it is clear that in such models it is a very sensible policy for humanity to engage in serious SETI efforts: we expect practically all ETI societies to be roughly of the same age as ours, and to be our competitors for Hart-Tiplerian colonization of the Milky Way. The price to be paid for bringing the arguments of "optimists" and "pessimists" into accord is, obviously, the assumption that we are living in a rather special epoch in Galactic history—i.e. the epoch of "phase transition". That such an assumption is entirely justifiable (by an observation-selection effect) in the astrobiological context will be argued in a subsequent study. Parenthetically, this is entirely in accord with the tenets of the currently much-discussed "rare Earth" hypothesis; see Ward and Brownlee (2000).

We may conclude that the optimism with regard to SETI expressed by Duric and Field (2003) is justified to an even higher degree than the authors concluded.

Acknowledgements — The author thanks Viera Mišović, Maja Bulatović, Alan Robertson, Nikola Milutinović, Vesna Milošević-Zdjelar, Samir Salim, Branimir K. Nikolić, Zoran Knežević, Srdjan Samurović, Mark A. Walker, and Nick Bostrom for help in finding some of the references or enjoyable discussions on the subject of this Comment. This is an opportunity to thank KoBSON Consortium of Serbian libraries, which at last enabled overcoming of the gap in obtaining the scientific literature during the tragic 1990s. The author also acknowledges support of the Ministry of Science and Environment of the Republic of Serbia through the projects no. 1196, "Astrophysical Spectroscopy of Extragalactic Objects" and no. 1468, "Structure and Kinematics of the Milky Way."

REFERENCES


DA LI JE KONEČNOST VANCEMAJUCA IZ "ZVEZDANIH STAZA" ZAISTA NAIVAN? – KOMENTAR NA ČLANAK ĐURIĆA I FILDA

M. M. Đirković

Astronomsko opštna observatorija, Volgina 7, 11160 Beograd 74, Srbija i Crna Gora

UDK 52-37

Komentar

Критични се разматра неколико претпоставки које се користе у недавном прегледном раду Бурича и Филда (Đurić and Field 2003) у овом часопису. Нови астробиолошки резултати – нарочито у вези расподеле старости земљоликих планета у Галаксији – омогућују нам да релаксирамо неке од ових претпоставки, што води повећању оптимизму у погледу SETI пројекта.