Boundaries of the habitable zone: unifying dynamics, astrophysics, and astrobiology

Milan M. Ćirković

Astronomical Observatory Belgrade,
Volgina 7, 11160 Belgrade-74, Serbia and Montenegro
email: mcirkovic@aob.aob.bg.ac.yu

Department of Physics, University of Novi Sad,
Trg Dositeja Obradovića 4, 21000 Novi Sad,
Serbia and Montenegro

Abstract. We are witnessing tremendous progress in the nascent multidisciplinary field of astrobiology, encompassing the origin and evolution of life in the cosmic context. One of the key concepts recently introduced in this field is the Galactic Habitable Zone (GHZ): an interval of galactocentric distances convenient for formation of stars possessing habitable planets. The boundaries of the GHZ are still poorly understood, however. Here we present a comparative analysis of various proposals for the mechanisms determining the GHZ boundaries, as well as different numerical values obtained. When joined with the models of Galactic stellar distribution, this gives us a better handle on the number of potential life-bearing sites.

Keywords. Astrobiology, Galaxy: structure, Galaxy: kinematics and dynamics, extraterrestrial intelligence

1. Introduction

We are lucky enough to live in an epoch of great progress in the nascent discipline of astrobiology, which deals with the three canonical questions: How does life begin and develop? Does life exist elsewhere in the universe? What is the future of life on Earth and in space? A host of important discoveries have been made during the last decade or so, the most important certainly being a large number of extrasolar planets, but also the existence of many extremophile organisms possibly comprising "deep hot biosphere" of the late Thomas Gold; the discovery of subsurface water on Mars and the huge ocean on Europa, and possibly also Ganymede and Callisto; the unequivocal discovery of amino-acids and other complex organic compounds in meteorites; modelling organic chemistry in Titan's atmosphere; the quantitative treatment of the Galactic habitable zone; the development of a new generation of panspermia theories, spurred by experimental verification that even terrestrial microorganisms easily survive conditions of an asteroidal or a cometary impact; progress in methodology of SETI studies, etc. (for recent beautiful reviews see Des Marais and Walter 1999; Darling 2001; Grinspoon 2003).

One of the most important and fruitful concepts recently introduced in astrobiology is Galactic habitable zone (henceforth GHZ). Proposed by Gonzalez, Brownlee and Ward (2001), it represents an annular ring comprising the potential sites for genesis of complex metazoans in the Milky Way. It is easy to see intuitively why is that so: locations at either large or very small galactocentric distances will be inhospitable to life as we know it. At very large galactocentric distance, for example, the metallicity of too low, and terrestrial planets will be very difficult to form (not to mention suppression of the complicated

chemical evolution necessary for supporting biochemistry). Close to the Galactic Center, the number of close stellar encounters, as well as threats of supernovae or even nuclear outbursts are likely to decrease habitability.

But the exact boundaries of GHZ remain somewhat mysterious: different authors suggested not only different particular values, but also different criteria for the zone boundary. It is the purpose of this contribution to try to sort out the confusion and evaluate different estimates. We shall also briefly discuss the issue of relevance of the GHZ concept for the SETI projects.

2. Outer boundary

Outer boundary is, rather incontrovertibly, set by the metallicity gradient of the Milky Way disk (Hou, Prantzos, and Boissier 2000; Tadross 2003). The original study of Gonzalez *et al.* (2001) was not very quantitative on this issue, but it is quite easy to see how a reliable estimate can be achieved using the appropriate scaling relationship.

According to most observations, $\nabla Z = 0.09$ 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 (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 ($M_{\min} \approx 0.1 \, M_{\oplus}$) is the minimum, for several reasons (development of plate tectonics, retention of O_2 and O_3, etc.). In principle, we shall have

$$R_{\text{out}} = R_{\odot} - \frac{1}{\nabla Z} \log \left(1 + \frac{2}{3} \log \frac{M_{\min}}{M_{\oplus}} \right), \qquad (2.1)$$

where $R_{\odot} = 8.5$ kpc is the Solar galactocentric distance. In this approximation, we obtain $R_{\text{out}} = 13.5$ kpc, which is significantly less than the Holmberg radius of the Milky Way.

Of course, since the overall metallicity grows in time, the net effect on the outer boundary is to move slowly outward, which is fairly uncontroversial (compare §3 below): $\frac{dR_{\text{out}}}{dt} > 0$. Thus, regions of the Galaxy devoid of habitable planets will slowly become habitable over time, and this process can, in principle, continue until the asymptotic maximal metallicity is reached in the physical eschatological context (Adams and Laughlin 1997). However, due to the exponential stellar distribution, the net gain in number of habitable sites due to the expansion of the GHZ outer boundary will become less and less important as time passes.

3. Inner boundary

The inner boundary of GHZ is subject to controversy, since at least three possible mechanisms for its determination have been proposed: 1. dynamical instability of planetary orbits under perturbations from stellar encounters, 2. high radiation environment due to frequent SNe and gamma-ray bursts, and 3. cosmogonical problems linked with high metallicity and UV-induced evaporation of protoplanetary material. The first two are obviously contingent mechanisms (i.e. those whose operation affect habitability in a random manner), while the third one can be considered deterministic. The first two has been proposed in a qualitative manner by Gonzalez *et al.* (2001), and subsequently quantified by Lineweaver *et al.* (2004). The third has been proposed by Pena-Cabrera and Durand-Manterola (2004); they have considered only suppression of terrestrial planets

due to very high Z. Photoevaporation of protoplanetary disks have been considered in a different context by Adams *et al.* (2004); its application to GHZ remains to be done.

Clearly, the true GHZ inner boundary will be the value driven by whichever mechanism is dominant at particular epoch: $R_{\text{inn}} = \max[R_{\text{rad}}, R_{\text{dyn}}, R_{\text{cosm}}]$. For example, the dynamical radius can be estimated in a following crude manner. The rate of stellar encounters is $\Gamma = \langle nv\sigma \rangle \approx \langle n \rangle \langle v \rangle \langle \sigma \rangle$, where n, v, and σ are stellar density, velocity and encounter cross-section respectively. We take the distribution of stars to be exponential, with the disk lengthscale of $R_d = 3$ kpc (Drimmer and Spergel 2001). Let us define the critical value of Γ 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 $\langle \sigma \rangle = 100$ AU^2 (Laughlin and Adams 2000). At the Solar circle, the average relative stellar velocity is $\langle v_0 \rangle \simeq 40$ km s^{-1}; however, in the interior of the Galactic bulge, where the inner boundary is located, we should use larger value, say $\langle v \rangle \simeq 122$ km s^{-1} suggested by Freeman *et al.* (1988). In this approximation, we obtain

$$R_{\text{inn}} = R_{\odot} - R_d \ln \frac{\Gamma_{\text{crit}}}{\langle n_0 \rangle \langle v \rangle \langle \sigma \rangle}. \tag{3.1}$$

Inserting numerical values, we obtain negative values for R_{inn}, which indicates that the dynamical inner limit is, in fact, irrelevant, if not non-existent. Even using the most radical assumption that $\langle \sigma \rangle = \pi(30 \text{ AU})^2$ we obtain that for purely dynamical constraints $R_{\text{inn}} = 1.3$ kpc. Calculations of Lineweaver, Fenner, and Gibson (2004)—which greatly improve on the previous 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 (see Fig. 2). Their study indicate 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 3 kpc (today). 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 Pena-Cabrera and Durand-Manterola (2004), who find $R_{\text{inn}} = 4$ kpc, on the basis of the assumption that metal-rich planets in the inner Galaxy would grow too fast to be inhabitable; this assumption, however, is highly questionable since another kind of habitable planets could plausibly form under high-Z conditions (see, e.g. Léger *et al.* 2004). Additional effect not taken into account by Pena-Cabrera and Durand-Manterola (2004) is that in the currently preferred theories of planet formation, protoplanet growth rate scales as the stellar mass. This can lead to an astrobiologically interesting effect, namely that in metal-rich systems, the preferred planet-forming systems (for terrestrial planets) become those with lower mass. Since it has been argued that planets around K and M dwarfs can indeed be habitable (e.g. Heath *et al.* 1999), this indicates the breakdown of this criterion for the inner GHZ boundary.

As far as the evolutionary changes in the inner boundary with time go, the confusion is compounded by different trends exhibited by candidates for the dominant mechanism. Since the overall metallicity increases with time, the maximal mass for a terrestrial plan*et al*so increases, thus yielding $\frac{dR_{\text{cosm}}}{dt} > 0$. However, the decrease in SNe frequency observable in external galaxies (and following predicted star formation history), will lead to the $\frac{dR_{\text{rad}}}{dt} < 0$ behavior, while on the timescales smaller than physical eschatological ones, we expect to encounter $\frac{dR_{\text{dyn}}}{dt} \approx 0$. Thus, the direction of the temporal evolution
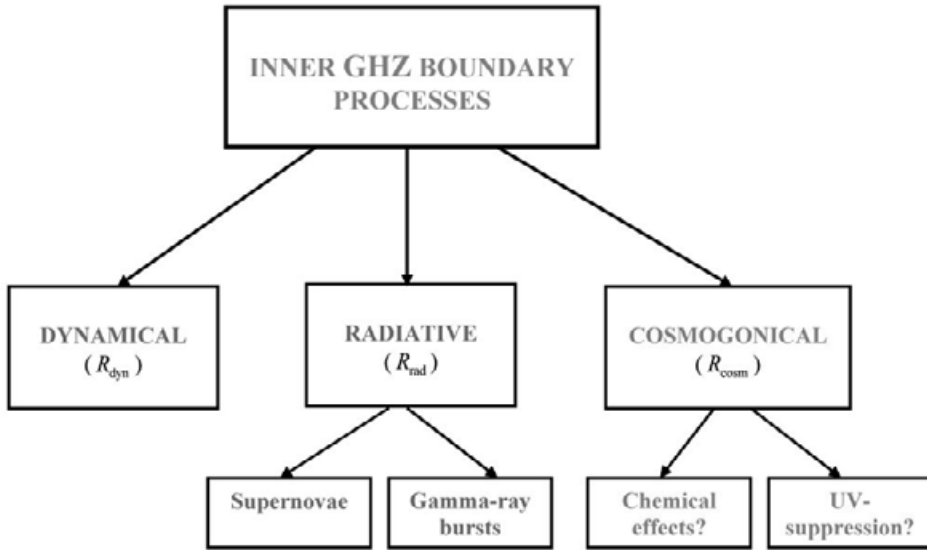
Figure 1. A scheme of various physical mechanisms determining the inner boundary of GHZ. Contingent mechanisms are listed in blue and deterministic in red color. Note that both cosmogonical mechanisms are still quite controversial, at least when applied to GHZ stratification.

of the inner GHZ boundary is not clear at present. In future research, it will be important to explicate the assumption of "biological invariance": that intrinsic probability of emergence/complexification of life at life-friendly sites does not vary with time.

4. An interlude: GHZ and SETI studies

There are at least two conclusions following from the studies of GHZ so far which are relevant to SETI studies. First and foremost, the 4-D structure of GHZ is such that it comprises a large number of Sun-like stars and habitable planets much older than the Solar System. According to Lineweaver (2001), Earth-like planets around other stars in the Galactic habitable zone are, on average, 1.8 ± 0.9 Gyr older than our planet, and mostly located at somewhat smaller galactocentric distance than the Sun. Applying the Copernican assumption naively, we would expect that correspondingly complex lifeforms on those others to be on the average 1.8 Gyr older.† This makes Fermi's paradox (for a review, see Webb 2002) more serious than hitherto thought. Obviously, this spells failure for any meaningful SETI-type communication; instead, the attention should be directed toward the youngest members of the GHZ-family, notably the outer rim of it.

The second issue of some practical importance in SETI targeting is the size of the volume to be search in radio eavesdropping and similar projects (e.g. Tarter 2001). Suppose that somehow (using a crystal ball or Drake equation) we ascertain the number of

† Strictly speaking, we need to take into account the width of the age distribution as well; it is conceivable, for instance, that most habitats in the Galaxy last shorter than the required biological timescale, Earth being exceptionally long-lived compared to the average (non-Copernican!). This leads us to the so-called Carter argument, very interesting topic in astrobiology and SETI studies, which is beyond the scope of the present paper.
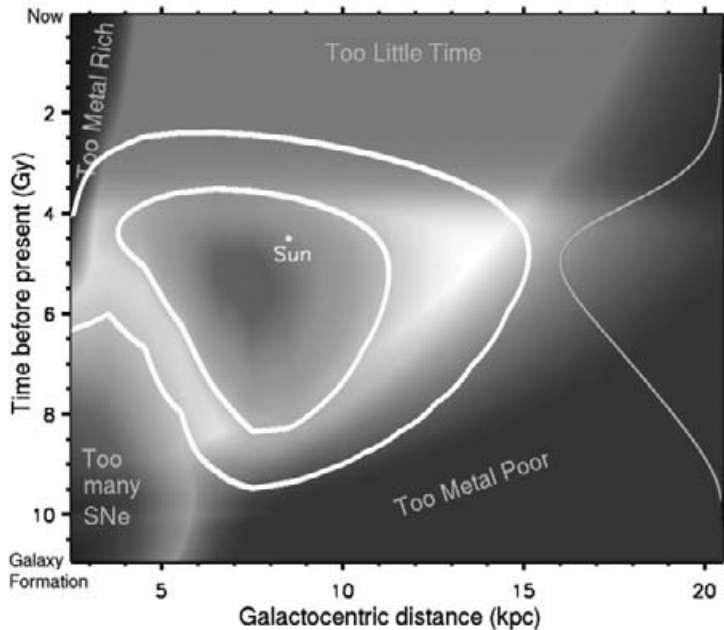
Figure 2. GHZ according to Lineweaver *et al.* (2004) in the disk of the Milky Way based on the star formation rate, metallicity, sufficient time for evolution, and freedom from life-extinguishing supernova explosions. The white contours encompass 68% (inner) and 95% (outer) of the origins of stars with the highest potential to be harboring complex life today (reproduced with kind permission of Prof. Charles Lineweaver).

civilizations in the "communication window" in the Galaxy as N_c (Duric and Field 2003). Efficiency of any particular SETI project depends on the density of prospective targets (i.e. inhabited sites). In the first approximation, then, taking into account GHZ acts to increase the real density of possible SETI targets by a factor $V_{MW}/V_{GHZ} \approx 5$. However, since (although it is not entirely clear from the derivation for Duric and Field 2003 study) 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. Roughly speaking, the number of stars within radius R scales as $N \propto R^2 \exp(-R/R_d)$, with maximum at $R \sim 6$ kpc, within GHZ. Thus, the average stellar density inside GHZ is—for the same normalization constants and the same value of the disk radius—larger than the average Galactic stellar density by the factor of $\langle n \rangle_{GHZ}/\langle n \rangle_{MW} \approx 1.4$. In future, we may expect significant improvement in precision of this type of estimate.

5. Prospects

Clearly, much work remains to be done in order to pin-point the extent of GHZ. In particular, research on the influence of close supernovae and local GRBs on Earth-like biospheres or prebiotic environments seems a promising way to go. But the purely dynamical constraints, especially in the long-term view, including the secular evolution of the Galactic gravitational potential, are no less important, particularly at the heuristical level. A unified modelling of dynamical, astrochemical and astrophysical constraints,

both deterministic and contingent ones, is likely to be within reach of state-of-the-art astrobiological research within a decade. Of course, this should not preclude further research toward limitations of the concept of the habitable zone itself. Here, the stellar case again serves as a useful prototype: astrobiologists are investigating mechanisms acting to define the boundaries of the circumstellar habitable zones (like the carbon-silicate cycle, etc.), while keeping an eye open on the possibility of life outside of it (for instance, life in subglacial oceans on Europa, etc.).

Acknowledgements

References

Adams, F. C. and Laughlin, G. 1997, *Rev. Mod. Phys.* 69, 337

Adams, F. C., Hollenbach, D., Laughlin, G., and Gorti, U. 2004, *ApJ* 611, 360

Ćirković, M. M. 2004, *J. Br. Interplan. Soc.* 57, 53

Darling, D. 2001, *Life Everywhere* (Basic Books, New York)

Des Marais, D. J. and Walter, M. R. 1999, *Annu. Rev. Ecol. Syst.* 30, 397

Drimmer, R. and Spergel, D. N. 2001 *ApJ* 556, 181

Duric, N. and Field, L. 2003, *Serb. Astron. J.* 167, 1

Freeman, K. C., de Vaucouleurs, G., Wainscoat, R. J., and de Vaucouleurs, A. 1988, *ApJ* 325, 563

Gonzalez, G., Brownlee, D., and Ward, P. 2001, *Icarus* 152, 185

Grinspoon, D. 2003, *Lonely Planets: The Natural Philosophy of Alien Life* (HarperCollins, New York)

Heath, M. J., Doyle, L. R., Joshi, M. M., and Haberle, R. M. 1999, *Orig. Life Evol. Biosph.* 29, 405

Hou, J. L., Prantzos, N., and Boissier, S. 2000, *Astron. Astrophys.* 362, 921

Irion, R. 2004, *Science* 303, 27

Laughlin, G. and Adams, F. C. 2000, *Icarus* 145, 614

Léger, A., Selsis, F., Sotin, C., Guillot, T., Despois, D., Mawet, D., Ollivier, M., Labéque, A., Valette, C., Brachet, F., Chazelas, B., and Lammer, H. 2004, *Icarus* 169, 499

Lineweaver, C. H. 2001, *Icarus* 151, 307

Lineweaver, C. H., Fenner, Y., and Gibson, B. K. 2004, *Science* 303, 59

Pena-Cabrera, G. V. Y., and Durand-Manterola, H. J. 2004, *Adv. Space Res.* 33, 114

Scalo, J. and Wheeler, J. C. 2002, *ApJ* 566, 723

Tadross, A. L. 2003, *New Ast.* 8, 737

Tarter, J. 2001, *Annu. Rev. Astron. Astrophys.* 39, 511

Ward, P. D. and Brownlee, D. 2000, *Rare Earth: Why Complex Life Is Uncommon in the Universe* (Springer, New York)

Webb, S. 2002, *Where is Everybody? Fifty Solutions to the Fermi's Paradox* (Copernicus, New York)