The remarkable discreteness of being

BAHRAM HOUCHMANDZADEH
Univ. Grenoble 1 / CNRS, LIPhy UMR 5588, Grenoble, F-38041, France

(Email, bahram.houchmandzadeh@ujf-grenoble.fr)

Life is a discrete, stochastic phenomenon: for a biological organism, the time of the two most important events of its life (reproduction and death) is random and these events change the number of individuals of the species by single units. These facts can have surprising, counterintuitive consequences. I review here three examples where these facts play, or could play, important roles: the spatial distribution of species, the structuring of biodiversity and the (Darwinian) evolution of altruistic behaviour.

[Houchmandzadeh B 2014 The remarkable discreteness of being. *J. Biosci.* **39** 249–258] **DOI** 10.1007/s12038-013-9350-7

1. Introduction

Many quantities in the physical world are continuous and measured by real numbers: position, speed, concentration, weight, etc. In many areas of science, however, it was realized that complex patterns can be explained by supposing the existence of discrete underlying levels that can be described using integers. In chemistry, the various laws of composition of elements such as 'definite proportions' and 'multiple proportions' known around 1800 AD led Dalton to formulate the atomistic theory and give a simple, elegant explanation to all these laws. Around 1900 AD, Planck, Einstein, Bohr and others realized that the most daunting problems of the (then) modern physics such as the radiation spectrum of stars, the universal temperature dependence of the specific heat of solids, the speed of electrons ejected by solids under radiation (photoelectric effect), etc., could be solved elegantly by supposing that the energy (or action) is quantified and varies only in integer units. In biology, the theory of Darwinian evolution was inconsistent with the then obvious blending theory of inheritance. The work of Mendel, and its rediscovery by de Vries, Correns and Tschermak around 1900 AD (a curious coincidence), restored the mathematical consistency of evolution by introducing the concept of genes as the quantum of inheritance information (see the first chapter of Fisher's book (Fisher 1999)).

These are but a few examples where complex patterns could be simply explained by supposing an underlying discrete level. The discreteness hypothesis, and specially its consequences, was in each of these cases unintuitive. Living organisms, on the other hand, do not need the discreteness *hypothesis*, as the most obvious fact about them is that death and birth events change their number by integers only. This obvious fact can have unintuitive consequences (Durrett and Levin 1994), and we will review three such cases: (i) spatial clustering of organisms, which is observed for nearly all living organisms; (ii) the observed biodiversity and many of its general laws such as the species–area relationship; (iii) the emergence of cooperative behaviour during Darwinian evolution. In all these cases, the reason for the importance of discreteness in living organisms is related to the unpredictability (stochasticity) of many phenomenon of life. For example, the moment of death or duplication, the direction of movement for food search, etc., are more or less random for an individual. One of the most fundamental results of stochastic phenomena is that the unpredictability of a system as a whole diminishes as the number of its constituent grows. When individuals can be counted by integers, the number of the constituents is finite and stochasticity can be an important player, inducing counterintuitive effects. The prime examples of this phenomenon are chemical reactions. In a typical test tube reaction, there are ~10^{23} molecules, a number so large that the unpredictability of molecules colliding to form bounds is erased and chemical reactions are perfectly predictable. In a bacteria, however, some molecules are present at such a low level as to be countable and chemical reactions associated to these molecules can become extremely noisy. This phenomenon, first noticed by Novick and Weiner for the transcription of Lac operon (Novick and Weiner 1957), has given rise to the field of epigenetics and non-genetic individuality (Houchmandzadeh and Mihalcescu 2011).

Keywords. Altruism; ecological drift; genetic drift; neutral biodiversity; spatial clustering in ecology

Published online: 15 March 2014

An important remark is in order. In none of the cases reviewed here is it claimed that a simple discrete theory will explain all the phenomena. I only observe that discreteness implies some surprising patterns which always exist. When complex patterns are observed in nature, the contribution of discreteness should be removed and only the remaining part, if any, needs a special theory.

2. Spatial clustering

Since the 1970s and the gathering of large amount of data on spatial distribution of various species, ranging from plants to insects to mammals, it has become obvious that nearly all species tend to have a clustered distribution and to aggregate into some areas (Taylor *et al.* 1978). The study of these spatial distributions has now become an independent field and is called metapopulation biology or ecology (for a review, see Hanski and Gaggiotti 2004; Hanski and Gilpin 1997). If the diffusion of organisms (animals move and plants disperse their seed) were random, one would expect that the distribution of species would soon (in few generations) become homogeneous (everything is everywhere) (O'Malley 2007). This is analogous, for example, to the dilution of a drop of ink in water. Common sense, therefore, requires that if we observe aggregation of individuals in one place, we should look for deterministic causes. There is no shortage of deterministic causes: (i) species are adapted to some environment, nature is heterogeneous, and therefore each species tends to concentrate in places (sub-habitats) to which it is best adapted; (ii) many species are social and their social interaction could be a cause of aggregation. These are two of the most studied explanations of clustering of organisms.

Plain common sense, however, is wrong in this case: the sole fact of discreteness of life is enough to cause clustering and no amount of random movement can counteract this agglomeration. This is what we show below. However, before going further, we should define precisely what we mean by clustered distribution and how we can measure it. The most practical way of measuring patchiness is to divide space into squares (quadrats), count the number n_i of individuals in each square, compute the mean $\langle n \rangle$ and the variance V of these numbers, and calculate the variance to mean ratio (VMR) $V/\langle n \rangle$ (figure 1).

For homogeneous random distributions, the VMR is equal to unity (figure 1A); for clustered distributions, VMR>1. VMR<1 denotes regularities (such as in human plantations) and is seldom encountered in nature. The VMR is a robust measure of patchiness; it is a function of the grid size, and this size dependence can be used as a test of the underlying theory causing the clustering. This measure can also be related to other popular metrics such as box counting and fractal dimension measurements.

Let us come back to the problem of clustering. Consider a collection of simple and similar organisms all moving randomly with diffusion coefficient D, reproducing at rate α and dying at rate μ. A naive model of the distribution of these 'Brownian bugs' would use a diffusion equation for their concentration $c(x,t)$ of the form of

$$\partial_t c = D\Delta c + (\alpha - \mu)c \tag{1}$$

The diffusion equation used by Fick to model concentration propagation is widely used in ecology (Okubo and Levin 2002). It is a consequence of 'particles flow from high to low concentrations'. The left-hand term denotes changes (per unit of time) in the concentration. On the right-hand side, the first term (Δc) is the balance of particle movements at one position due to flow from neighbouring positions; the second term states that the number of birth and death at one position is proportional to the number of individual present at this position.

Consider, for example, the particular case where birth and death rates are equal. Then we will have a plain diffusion equation and any spatial heterogeneity will be smoothed out after some time, as in the example of the drop of ink in water. Young *et al.* (2001) used such a simple model of Brownian bugs to study the phenomenon of plankton blooms, but instead of resorting to equation (1), they numerically simulated these bugs and found the exact contrary of the expected phenomena: the distribution, which was homogeneous at the initial time, would get more and more patchy as time passed. Something is grossly wrong with the use of the continuous approach of equation (1): such equations are written for averaged quantities and assume that fluctuations (deviations from the mean) are small compared to the mean. However, the noise of reproduction and death due to the discreteness of living organism violates these assumptions; as we will see below, fluctuations become much larger than the averages, and hence the error in using continuous differential equations in modelling ecological systems.

To understand reproduction/death-induced fluctuations, we need to include the second important aspect of life: for an individual, the moment of its death or reproduction is a random variable. For a collection of n individuals, we can speak only about the probability of a death/birth occurring during a given time interval. If the probabilities induce small fluctuations, the stochastic process can be approximated by a differential equation (mean field approach; see appendix); if not, one has to resort to the Master equation approach (and/or its numerical resolution) in order to estimate various statistical quantities such as the mean, the variance and the correlations (see appendix). Let us forget about the spatial aspect of the problem at hand for the moment. Consider the space divided into non-communicating habitats and place exactly n_0 individuals in each cell at time $t_0 = 0$ (figure 2). These individuals are capable of only reproducing/dying. Let $n(t)$ be the number of individuals in a cell at time t. In the simplest possible model when birth at rate α and death at rate μ are constant and independent of density, age structure,

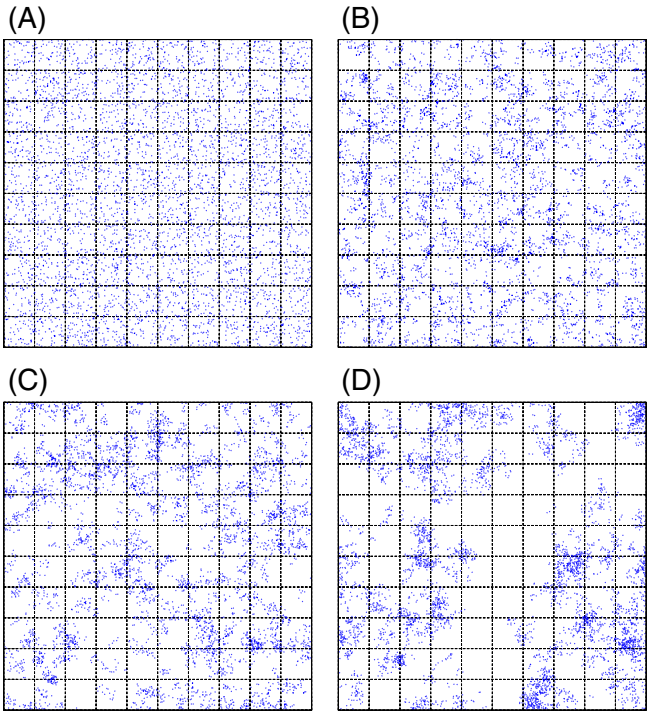

Figure 1. Spatial distribution of 5000 individuals (dots) with increasing patchiness. The VMR computed over the displayed grid is (**A**) 1 (homogeneous); (**B**) 5.1; (**C**) 13.8 and (**D**) 49.1.

previous births, etc., the probability density for one birth/death to occur is proportional to the number of of individuals n and independent of time. It reads

$$W^+(n) = \alpha n \; ; \; W^-(n) = \mu n \qquad (2)$$

From the above transition rates, the probability $P(n,t)$ of observing n individuals at time t is deduced through the book-keeping (Master) equation (see appendix). We can very easily numerically simulate the above stochastic process and observe that as time increases, many cells will become empty and a few will harbour a very large number of individuals. The average number of individuals per cell will remain constant, but the variance and hence the VMR will increase linearly as a function of time. This is a system which, after a few generations, will display a huge amount of clustering. In fact, in this simple stochastic process, the time evolution of the mean $\langle n \rangle$ and the variance $V = \langle n^2 \rangle - \langle n \rangle^2$ of the number of individuals in cells can be deduced exactly from the Master equation. For the case where $\alpha = \mu$ and n_0 individuals per cell at $t = 0$:

$$\langle n(t) \rangle = n_0 \; ; \; V(t) = 2n_0 t$$

Therefore, even though the average number of individuals per cell remain constant, the variance increases as a function of time, a behaviour shown in figure 2 (Houchmandzadeh 2002).

What is special about this reproductive noise is the fact that it cannot be cancelled out by diffusion. Removing the barriers between the cells and letting the individuals diffuse from high-density to low-density sites only slows down the clustering phenomenon, but does not inhibit it (Houchmandzadeh 2002) and the VMR still increases with time for one- and two-dimensional ecosystems. In the case where $\alpha = \mu$:

$$\begin{aligned} \text{VMR} \sim & \sqrt{t} \quad \text{for} \; d = 1 \\ & \sim \log t \quad \text{for} \; d = 2 \end{aligned}$$

The two-dimensional case corresponds exactly to what Young *et al.* (2001) observed in their numerical simulations. Note that the prediction of a pure diffusion equation is that the VMR will stay at the value 1 at all times.

So, all living organisms will naturally form spatial clusters, at least in the simplest, neutral models. Can this clustering be observed experimentally? The answer is yes. Spread some microorganism capable of movement, reproduction and death on a Petri dish, measure the position of *each* one at each time step and compare it to the predicted auto-correlation function or VMR (figure 3). In such a controlled experiment, all the parameters (α, μ, D) are measured and there is no room for free fitting parameters. The only difficult step is to measure the position of all microorganisms, which can be achieved by an automatized microscope and image analysis. The experiment that was indeed performed (Houchmandzadeh 2008, 2009) showed perfect agreement of the spatial autocorrelation function with the theoretical computations.

Real ecosystems are density dependent and the density of individuals cannot grow very large. We can incorporate this density dependence into the model in its most stringent case: the ecosystem is composed of many species, the total density is fixed, and when one individual dies, it is replaced by the progeny of a neighbouring one, whatever its species. In the framework of neutral ecosystems, i.e. all individuals similar in their birth, death and dispersal properties, it can be shown that the major features exhibited above do not change and

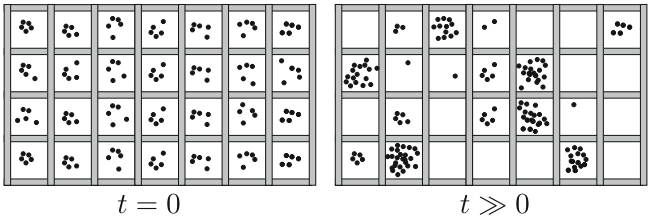

Figure 2. Reproduction/death noise: placing n_0 individuals in each cell at time $t = 0$ and letting them die and reproduce at the same rate. As time passes, some cells will contain a large number of individuals, while others become empty.

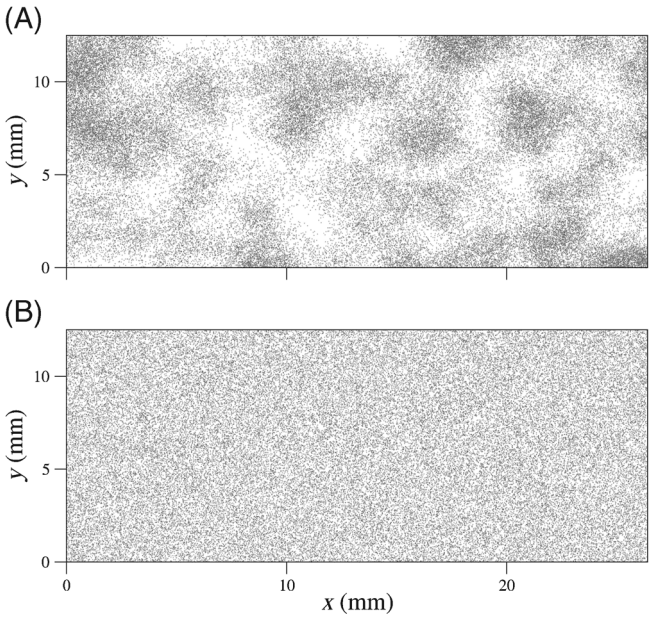

Figure 3. Neutral clustering of microorganisms: (**A**) ~70000 *Dictyostelium discoïdum* (size ~10 μm) after approximately 7 generations. Each dot represents a single individual and its position was measured from an actual experiment (Houchmandzadeh 2008); the VMR ≈ 40 for squares of size 0.7 mm; (**B**) the same number of individuals spread homogeneously.

individuals still form increasingly large clusters, uniform in their composition of species (Houchmandzadeh and Vallade 2003). This clustering can also be (painfully) measured in real ecosystems (for example, in rain forests) (Condit *et al.* 2002) and shown to be compatible with the theoretical computations, although in real ecosystems, many parameters cannot be measured.

Spatial clustering of organisms is one of the most fundamental problems in ecological studies. The message of this section is the following: observing a patchy distribution should not be considered *per se* as surprising and one should not rush to find deterministic causes for it. The very nature and discreteness of life naturally leads to clustering. Of course, all clustering are not caused by discrete effects. Before looking for other causes, however, one must remove the effect of neutral causes and use deterministic causes only for the remaining (if any) patchiness.

3. Neutral biodiversity

Observation of the stunning biodiversity in various ecosystems is one of the factors that led Darwin and Wallace to formulate the theory of evolution. The finches of Galapagos are the standard example cited in any textbook of the field (Lomolino 2006). Even at a single trophic level, i.e. considering species which use the same resources, the biodiversity is always large. In spite of many competing theories, the question of the causes of biodiversity is still unanswered today. The adaptationist program criticized by Gould and Lewontin (1979) is still predominant: each species is adapted to its local environment and biodiversity is just a reflection of the heterogeneity of nature. Neutral biodiversity can exist only because of geographical barriers between close ecotypes. The possibility of having speciation at the same trophic level at the same geographical location has been ruled out by Ernst Mayr in his famous book (Mayr 1942), with far-reaching consequences on evolutionary thinking.

Ecologists, however, began to gather large amount of data on biodiversity and observed general patterns everywhere. One of the most striking observed 'law' is the species–area relationship, which states that the number of species S in an area exhibits a power law dependence on the size A of the area considered: $S = kA^z$ with z in the [0.2,0.3] range for most habitats (Williamson *et al.* 2008). An alternative and more precise measure of biodiversity for a fixed area is the abundance curve: collecting species in a given area and measuring the abundance of each species leads to the abundance curve $\phi(n)$, which is the histogram of the number of species having abundance n. Abundance curves taken from very different habitats began to show very similar patterns (for a review, see Hubbel 2001). The third observation came from measurements of biodiversity in islands close to a continent. It was observed that the number of species in islands decreased as a function of its distance from the continent and increased with the size of the island.

To explain the third observation, MacArthur and Wilson (1963, 1967) took a bold approach. They supposed that (i) all species at the same trophic level are *equivalent*; (ii) species migrate from continent to islands, with the rate of migration a decreasing function of the distance; (iii) due to random sampling from one generation to the other, species become extinct in islands, with the extinction rate a decreasing function of the size of the island. The number of species present on the island is then a dynamic equilibrium between migration and extinction.

MacArthur and Wilson's article, considered as a cornerstone of biogeography, was a radical departure from Mayr and the adaptationist program, and proved extremely successful. The next radical step then was taken by Hubbel (2001), who applied the same idea to the whole continent: all species at a given trophic level are equivalent; new species appear by mutation and become extinct by genetic drift. The biodiversity curve is then a function of a single number that takes into account the mutation rate and the size of the community. Hubbell's book founded what is called the neutral theory of biodiversity and provoked an incredibly wide and heated debate in the ecological community, which is still ongoing.

I review here some of the mathematical consequences of the neutral theory to which we contributed. In retrospect, it seems

strange that the idea of neutrality, considered very early by population geneticists such as Malécot (1948) and Kimura (1983), took so much time to permeate the ecological/evolutionary thinking; I believe that this is partly due to the influence of Mayr's book (Butlin *et al.* 2008) and the prevalence of the competitive exclusion principle (Hardin 1960). The main idea, however, is very simple: neutral macroecology is similar to neutral population genetics (Hu *et al.* 2006). The latter deals with alleles of a gene, their frequency and its change because of genetic drift and (neutral) mutations, where the former deals with the equivalent concepts of species, their abundance and its change because of ecological drift and neutral speciation (presumably because of the accumulation of many mutations at the individual levels) and so on. New species emerge with a rate ν. It takes some times τ_a for a new species to become abundant by pure genetic drift. If the arrival time of new species τ_e is much shorter than τ_a, many equivalent species will coexist at the same geographical location and their abundance will be a dynamic interplay between emergence of new species and extinction of existing one.

Consider a community consisting of N individuals and S species, with species i having n_i individuals (figure 4A).

All individuals, regardless of their species, are equivalent in their reproductive/death rate. When an individual dies, it is immediately replaced by the progeny of another one. Because of mutations, the progeny can differ from its parent with probability ν, thus forming a new species appearing with abundance 1. After its appearance, the species abundance is a stochastic function of time; if an individual is the sole representative of a species and dies, then this species disappears. As in the previous section, the probability $P(n,t|1,t_0)$ for species i to have n individuals at time t, knowing the species appeared at time t_0, obeys a Master equation where the transition rates are (Vallade and Houchmandzadeh 2003) (see appendix):

$$W^+(n) = \mu(N-n)n(1-\nu)/N \tag{3}$$

$$W^-(n) = \mu n(N-n+\nu(n-1))/N \tag{4}$$

The increase rate $W^+(n)$ is the probability density of death of an individual that does not belong to the considered species $\mu(N-n)$ multiplied by the probability of birth of an individual that belongs to the considered species n/N, times the probability of no mutation $(1-\nu)$. The decrease rate is similar, but takes also into account the probability of an individual dying and being replaced by the progeny of a member of its own species with a mutation.

Let us set the origin of time at $t_0 = 0$. The master equation gives the fate of one particular species. $\langle\phi(n)\rangle$, the average number of species having population size n at time t is the sum of all those who have been generated at an earlier time τ and have reached abundance n at time t:

$$\begin{aligned}\langle\phi(n)\rangle &= \int_0^t f(\tau)P(n,t|1,\tau)d\tau \\ &= \nu\int_0^t P(n,t-\tau|1,0)d\tau \\ &= \nu\int_0^t P(n,\tau|1,0)d\tau\end{aligned}$$

where $f(\tau)$ is the probability per unit of time of generating a mutant and is equal to ν (time is measured in units of generations $1/\mu$). Defining the mutation pressure as $\theta = N\nu$, the quantity ϕ can be obtained at the limit of large times and shows that an equilibrium is reached. For large communities, using proportions $\omega = n/N$ and abundances $g(\omega)=N\langle\phi(n)\rangle$, the result takes a simple form (Vallade and Houchmandzadeh 2003):

$$g(\omega) = \theta\omega^{-1}(1-\omega)^{\theta-1} \tag{5}$$

The above computations ignore spatial distances: an individual can be replaced only by the progeny of its neighbour rather than by everyone in the community. A self-consistent model of geographical dispersal is incredibly difficult. We can, however, go one step further and apply the above model to the case of island biogeography, where a small island of size M is close to a continent of size N ($M<<N$). The population of the island is affected by migration from the continent, but given the large size of the continent, the reverse is not true. We can also neglect mutation inside the island as the mutation pressure is small. So the transition rates in the island are similar to equations 3 and 4 except that a local individual can be replaced by a migrant from the continent with probability m, where the abundances are

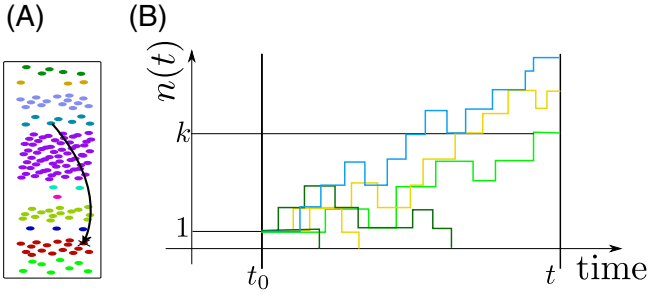

Figure 4. (**A**) The Moran model of a neutral community composed of various species (distinguished here by their colours), where an individual is replaced upon its death by the progeny of another regardless of its species. (**B**) Each new species appears with abundance 1 by mutation at some time t_0; stochastic dynamics of the number of individuals $n(t)$ of few species appeared at time t_0.

given by expression (5). Defining the migration pressure as $\xi = Mm$, in the limit of large sizes of both the island and the continent, we can compute the relative abundance $g_I(\omega)$ inside the island as (Vallade and Houchmandzadeh 2003)

$$g_I(\omega) = \xi\theta \int_0^1 (1-\omega)^{\xi u-1}\omega^{\xi(1-u)-1}u^{\theta}du \qquad (6)$$

This expression may seem cumbersome, but it can be easily plotted and depends on only two parameters: θ, which itself can be seen as a function of biodiversity on the continent, and ξ, which is a simple decreasing function of the distance between the continent and the island. This expression was also obtained by Volkov *et al*. (2003) and in a slightly modified form by Etienne (2005). Expression (6) is the mathematical expression of the original MacArthur and Wilson model and can be put to experimental verification.

Improving the above model by fully taking into account the spatial dimension seems mathematically intractable. We have been able to slightly improve the continent-island model by treating both communities on an equal footing (Vallade and Houchmandzadeh 2006) but going further seems beyond the reach of the mathematical tools we used. Nevertheless, the neutral theory of biodiversity is a falsifiable theory of biodiversity. It has been put to intense test and has been proved successful at interpreting quantitatively available data in island biogeography (Rosindell *et al.* 2011). As in the previous section, the merit of this model is to provide a first approximation for biodiversity which will always be present, even though many data will necessitate the addition of more ingredients, such as density dependence of replacement rates (McGill *et al.* 2006; Jabot and Chave 2011), to explain deviation from this theory.

4. Emergence of altruistic behaviour in Darwinian evolution

Altruistic behaviour is widespread among living organisms. 'Altruism' is an emotionally charged term that many scientists avoid in favour of more neutral terms such as cooperative behaviour. We stick to this word here and define altruistic behaviour as the production of some 'common good' that benefits all individuals of the same species in the community, at a cost to the producer. From the evolutionary point of view, fitness is the mean number of surviving and reproducing descendant of individuals of a given genotype. The cost and benefit here are all measured in units of fitness.

Light production in *Vibrio fischeri* (Foster *et al.* 2004; Visick and Ruby 2006), siderophore production in *Pseudomonas aeruginosa* (West and Buckling 2003), invertase enzyme production in *Saccharomyces cerevisiae* (Gore *et al.* 2009) and stalk formation by *Dictyostelium discoideum* (Kessin 2001; Foster *et al.* 2004) are but a few examples, taken from the microbial world, where individuals in a community help others at their own cost by devoting part of their resources to this help. From the evolutionary point of view, altruists have a lower fitness than other individuals in the community who do not help, but are recipients of the benefits produced by altruists. Throughout this article, we call these latter individuals 'selfish'.

How can altruistic behaviour emerge by natural selection if individuals having a phenotype which displays this trait, and the phenotype is strongly correlated with a particular genotype, have a lower fitness than the selfish ones? This is among the hottest debates of evolutionary biology, and has been ongoing from the inception of the discipline (Dugatkin 2006). In the deterministic view of evolution, genotypes with higher fitness increase their frequency in the population; therefore, if altruism is selected it means that its associated genotype has some hidden benefits that compensate its apparent lower fitness. The only task is to discover the hidden advantage.

The first class of model for the hidden advantage was proposed by Hamilton (1964a, b) and is known as kin selection: the common good is not provided to everybody but only to individuals related by common descent (kins). The original Hamilton model is based on 'frequency-dependent fitness' and was formulated for sexually reproducing organisms. This model and other mathematically equivalent models such as direct or indirect reciprocity (Nowak 2006b) can be easily understood as follow for asexual organisms. Consider the simplest case where an allele of gene confers a fitness r. The deterministic Fisher equation for the change in the frequency (relative abundance) p of this allele is (Ewens 2004)

$$dp/dt = (r-1)p(1-p) \qquad (7)$$

This equation is derived from the original Fisher–Wright model of population genetics: in a community of size N (where N is supposed to be large), each individual produces progeny proportional to its fitness and N individuals are selected randomly among all the progeny to constitute the next generation.

Advantageous mutants have fitness $r>1$ and therefore increase their frequency, where deleterious mutants have fitness $r<1$ and decrease their frequency. This equation supposes that fitness does not depend on the frequency of the allele and is a constant.

However, if the excess fitness $r-1$ is a function of the gene frequency $f(p)$ and the function changes its sign for some intermediate frequency p^*, then the gene will increase its frequency if $p>p^*$ (figure 5). This is precisely the point made by Hamilton: if help is provided and received only among altruists, then at high frequency, the benefits that each altruists

receive from other 'kins' or other altruists can outweigh the cost of the common good production to one individual. Nowadays these models are mostly studied through the game theory language and the frequency dependent fitness *f(p)* is obtained through a payoff matrix; the frequency-dependent equation for the rate of change is called replicator dynamics (Nowak 2006a).

The second class of models, called group selection, supposes that individuals are divided into groups. Not only do individuals compete inside each group in order to increase their frequency, but groups compete among each other at a higher level of selection (Traulsen *et al.* 2005). The idea of group selection goes back to the inception of evolutionary biology and was promoted by the founding fathers of modern evolutionary synthesis, then was discredited by GC Williams (1966), then restored by Lewontin (1970) and Price (1970) and regained respectability again in the 1990s.

These two class of models are nowadays the main explanations for the emergence of altruisms in Darwinian evolution (Nowak 2006b), even though a 'religious' war can erupt between them from time to time (see for example (Nowak *et al.* 2010) and some among many replies to it (Boomsma *et al.* 2011; Strassmann *et al.* 2011)).

As in the two previous sections, I want to review an alternative theory we developed (Houchmandzadeh and Vallade 2012), based on the discreteness and the genetic drift it causes. As in the previous sections, I do not claim that this theory explains all the observed behaviour and replaces the other two. But as in the previous section, I show that a very simple explanation exists which does not rely on some hidden benefits. This simple model applies to cases where the production of common good or the cooperative behaviour increases the carrying capacity of the habitat.

The deterministic Fisher equation (7) is not satisfactory at small excess relative fitness: a fitter genotype appearing at one copy number can disappear just by chance and not have the possibility of increasing its frequency at all. Evolutionary dynamics is a stochastic process due to competition between deterministic selection pressure and the inevitable role of chance factors in influencing who reproduces, how many children they have, and so on. In order to capture the main characteristics of this competition, Fisher and Wright introduced a very simple model, which was later slightly modified by Moran (1962) to make it mathematically more tractable (Houchmandzadeh and Vallade 2010). The model consists of a community of fixed size N, composed of wild-type individuals with fitness 1 and mutants with fitness r. Individuals are chosen at random (with rate μ for the wild type and $r\mu$ for mutant) to reproduce. When an individual reproduces, another is chosen at random to die, thus keeping the population at constant size. The probabilities per unit of time for the mutants to increase (W^+) or decrease (W^-) their number by one individual is (see appendix for a general explanation of stochastic modelling)

$$W^+(n) = r\mu(N-n)n/N \tag{8}$$

$$W^-(n) = \mu(N-n)n/N \tag{9}$$

and their mean field approximation leads to the deterministic Fisher equation (7). However, this is a probabilistic process: the number of mutants can fall to zero (extinction) or N (fixation) with finite probability, and if it does so, the system remains in this state. One of the most fundamental concepts of evolutionary dynamics is precisely the *fixation probability*, i.e. the probability that a mutant spreads and takes over the whole community (Patwa and Wahl 2008). In the framework of the Moran model the fixation probability is (Moran 1962; Houchmandzadeh and Vallade 2010)

$$\pi_f = \frac{1-r^{-N_0}}{1-r^{-N}} \tag{10}$$

where N_0 is the original number of mutants. For small selection pressure $Ns \ll 1$, where $s = r - 1$, the fixation probability π_f of a mutant appearing at one copy can be approximated by

$$\pi_f \approx \frac{1}{N} + \frac{s}{2} \tag{11}$$

The fixation probability is composed of two terms: even in the absence of selection, the population will become homogenic via a process known as genetic drift; in the neutral case, all individuals at generation zero have an equal probability $1/N$ of taking over the whole community (being fixed). When a beneficial mutant is present, the fixation probability of its carrier is increased by the relative excess fitness. Note that genetic drift is at the heart of the neutral theory of biodiversity discussed in the previous section.

The Fisher–Wright–Moran model is the most fundamental model of population genetics, displaying the importance

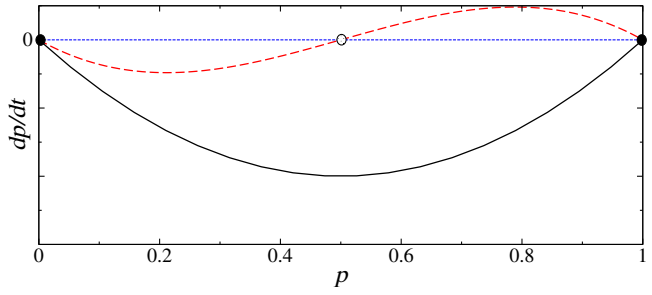

Figure 5. The phase space presentation of the Fisher equation (7). Solid curve: constant fitness $r - 1 < 0$; of the two fixed point, only $p = 0$ is stable. Dashed curve: frequency dependent fitness where $r - 1 = f(p) < 0$ for $p < p^*$. Both fixed points $p = 0$ and $p = 1$ are stable.

of genetic drift. We can complement it to take into account the effect of altruistic individuals, without adding any hidden benefits. The most notable effect of 'common good' production is the increase in the carrying capacity of the habitat, which benefits everybody regardless of its genotype (altruistic or selfish).

Let us suppose that the carrying capacity is N_i when only selfish individuals are present and N_f when only altruistic individuals are present ($N_i < N_f$) and has an intermediate value when the community is a mixture of both genotypes, with the carrying capacity an increasing function of the number of altruistic individuals n (figure 6). Let us suppose now that selfish individuals have fitness 1 and altruistic one have fitness $r < 1$ and let us set s=1−r as the cost of altruism. In a deterministic model, altruists will *always* lose to selfish ones (Houchmandzadeh and Vallade 2012). When taking into account the stochastic nature of this process, the answer can be different. As I stressed above, the quantity of interest in the stochastic process is the fixation probability. Let us compare the fixation probability π^A of one altruistic mutant introduced into a community of selfish individuals to the fixation probability π^S of one selfish mutant introduced into a community of altruistic individuals (figure 6). A back-of-the-envelop computation, according to equation 11, gives

$$\pi^A = \frac{1}{N_i} - \frac{s}{2} \quad ; \quad \pi^S = \frac{1}{N_f} + \frac{s}{2}$$

We see that even though selfish individuals have higher fitness, we can nave $\pi^A > \pi^S$ if

$$s < \frac{1}{N_i} - \frac{1}{N_f}$$

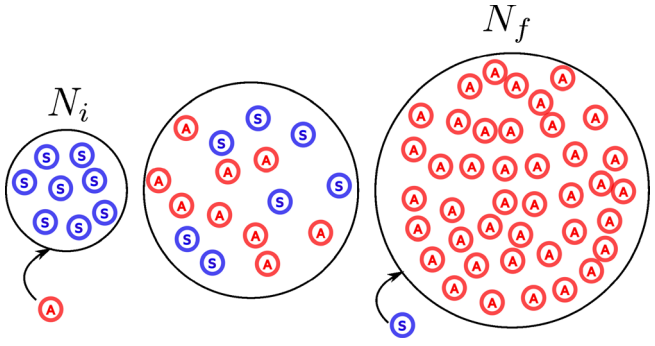

Figure 6. The neutral effect of common good production: the carrying capacity N of the habitat depends on the number of altruists present, ranging from a minimum N_i when only selfish individuals are present to a maximum N_f. In this case, the fixation probability of one A introduced into a community of S can be higher than the fixation probability of one S introduced into a community of A.

Alternatively, by setting $\Delta N = N_f - N_i$ and $\overline{N} = \sqrt{N_i N_f}$, the above criteria can be written in terms of selection pressure

$$\overline{N} s < \Delta N / \overline{N} \qquad (12)$$

which means that if the selection pressure against the altruists is smaller than the relative change in the carrying capacity, then altruists *win*, even though they have a smaller fitness.

The above discussion is an oversimplification in order to give the general idea of the genetic drift favouring altruists. At each time step, the community is formed of a number of n of altruists and m of selfish individuals, and the total size of the community $N = n + m$ itself is a function of altruist numbers. The fixation probability thus has to be computed by taking into account the full stochastic behaviour of both populations. The fixation probabilities can be computed from the backward Kolmogorov (BK) equation and its diffusion approximation (Ewens 2004). This procedure was used by Kimura (1962) to deduce the fixation probability of the Fisher–Wright classical model of population genetics. In this model the population size N is kept constant and the size of one population can be deduced from the size of the other: $m = N - n$. The Kimura equation is thus concerned with only the frequency of one allele. A similar approach can be used to compute the fixation probability of the model presented above, although now the BK equation has to track the frequency of both populations at the same time, as the total population is not constant anymore. The resulting equation, although much more complicated, leads to the result (equation 12) given above (Houchmandzadeh and Vallade 2012).

One could think that natural communities are composed of large number of individuals; so, even for small costs s, criterion (12) is violated. This argument, however, is not correct because populations are geographically structured: individuals can be replaced only by their neighbours, and so the effective populations entering into expression (12) are indeed much smaller than the total size of the community. In fact, at small migration rate, the altruistic advantage is amplified: at constant and uniform migration rate, the number of migrants is proportional to the local population size; therefore, places with high carrying capacity (composed of altruists) send out more migrants than places with a lower carrying capacity (composed of selfish). This amplification mechanism can be computed at small migration rates and it can be shown that large, geographically structured population are indeed *immune* to invasion by selfish individuals ($\pi^S = 0$) (Houchmandzadeh and Vallade 2012).

Let us again stress that this simple advantage of altruists is a pure effect of the discreteness of life which cannot exist if living organisms were part of a continuum. I do not claim that kin or group selection do not exist or are irrelevant, but there is an inherent advantage in producing a common good that, when this trait increases, the carrying capacity of the habitat and benefits everybody. It may not overcome the cost

associated with this behaviour in some living ecosystems, and then other more elaborate schemes have to be considered. But, before resorting to these 'hidden advantage' theories, one should subtract the contribution of discreteness and the increase in the carrying capacity.

5. Conclusion

There are many other biological systems where the discreteness of underlying processes has come to the forefront. The most notable example is noise-driven chemical reactions taking place inside living cells giving rise to non-genetic individuality, which has been thoroughly investigated during the last decade (Davidson and Surette 2008). The message I intended to carry through the three examples reviewed in this article is that, as in many other areas of science, the discrete nature of life has important consequences, which have all too often been neglected. The main reason for this neglect may be the counterintuitive nature of these consequences: a drop of ink in water tends to dilute, and it is not evident that by adding neutral reproduction, the ink should reverse its course and *concentrate*. I hope, however, that this very fundamental and important aspect of life will become more a part of the general culture of scientists.

Appendix

Stochastic modelling

A *deterministic* behaviour is perfectly predictable: knowing that at time t a system (for example, a projectile) is in a state x (from example, its position), we know its position $x' + x$ at a time $t + t'$. This knowledge is modelled by a function $x' = f(x, t, t')$ and completely characterizes the temporal dynamics of this particular system. For many laws, it is sufficient to know the function f for a very short (infinitesimal) time increment t', dt. The evolution of the system for the small increment $x' = dx$ is then often obtained as $dx = f(x, t)dt$ and is called a differential equation. It is enough to add these short increments to determine the state of the system for any long time.

A *stochastic* behaviour is only partially predictable: knowing the system is in a state x at time t, we cannot predict its state at a later time $t + t'$; but we can only give a probability $P(x')$ that it will be at $x + x'$ at this time. The function $P(x')$ can be experimentally measured by making a large number of measurement: let N similar systems at time t be in the state x, and we measure their state $x + x'$ at time $t + t'$, then make a normalized histogram of all this measurements, which is a function $P(x'|x,t,t')$. Again, if we know this function for very short time t', we can know the probability function for longer times by summing up the short-term evolution. This is, for example, how meteorology works: knowing the evolution of probabilities on time scales of seconds, we can predict that, for example, there is 70% chance that tomorrow will be rainy and 30% that it will be sunny.

For many stochastic systems (called Markov chains), it is enough to know the temporal evolution of probabilities for infinitesimal time increments $t' = dt$, written as

$$P(x'|x, t, dt) = W(x'|x, t)dt$$

where the function W is called the transition rate. Equations (2), (3) and (4), and (8) and (9) are examples of such transition rates. $W^{\pm}(n)$ are shorthand for $W(\pm 1|n)$, i.e. the probability (per unit of time) that the system (here number of individuals of a given type) will be in state $n + 1$ at time $t + dt$, knowing that it is in state n at time t. The knowledge of these short time transition rates allows for the determination of probabilities at long times, through an evolution equation called the Master equation when the states are discrete.

It not always easy to solve a Master equation. A crude approximation, called mean field, is to write a *deterministic* equation for the evolution of the *mean* increment $\langle dx \rangle$ during a short time dt, in the form of

$$\langle dx \rangle = f(x, t)dt$$

and then use deterministic procedure to predict the *mean* state of the system at long times. The function $f(x,t)$ is readily obtained from the transition rates $W(x'|x,t)$. Mean field approximation can be useful or misleading, depending on the stochastic nature of the system.

References

Boomsma J, Beekman M, Cornwallis C, Griffin A, Holman L, Hughes W, Keller L, Oldroyd B and Ratnieks F 2011 Only full-sibling families evolved eusociality. *Nature* **471** E4–E5

Butlin R, Galindo J and Grahame J 2008 Review. Sympatric, parapatric or allopatric: the most important way to classify speciation? *Philos. Tras. R. Soc. London Ser. Biol. Sci.* **363** 2997–3007

Condit R, Pitman N, Leigh E, Chave J, Terborgh J, Foster R, Núñez P, Aguilar S, Valencia R, Villa G, Muller-Landau H, Losos E, *et al*. 2002 Beta diversity in tropical forest trees. *Science* **288** 1414–1418

Davidson C and Surette M 2008 Individuality in bacteria. *Annu. Rev. Genet.* **42** 253–268

Dugatkin LA 2006 *The altruism eaquation: Seven scientists search for the origins of goodness* (Princeton NJ: Princeton University Press)

Durrett R and S Levin 1994 The importance of being discrete (and spatial). *Theor. Population Biol.* **46** 363–394

Etienne R 2005 A new samping formula for biodiversity. *Ecol. Lett.* **8** 253–260

Ewens W 2004 *Mathematical population genetics* volume I, 2nd edition (Springer-Verlag)

Fisher R 1999 *The genetical theory of natural selection, a complete variorum edition* (Oxford: Oxford University Press)
Foster K, Shaulsky G, Strassmann J, Queller D and Thompson C 2004 Pleiotropy as a mechanism to stabilize cooperation. *Nature* **431** 693–696
Gore J, Youk H and van Oudenaarden A 2009 Snowdrift game dynamics and facultative cheating in yeast. *Nature* **459** 253–256
Gould S and Lewontin R 1979 The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. R. Soc. Lond. B* **205** 581–598
Hamilton W 1964a J The genetical evolution of social behaviour I. *J. Theor. Biol.* **7** 1–16
Hamilton W 1964b The genetical evolution of social behaviour II . *J. Theor. Biol.* **7** 17–52
Hanski I and Gaggiotti O 2004 *Ecology, genetics and evolution of metapopulations* (Academic Press)
Hanski I and Gilpin M 1997 *Metapopulation Biology* (Academic Press Inc)
Hardin G 1960 The competitive exclusion principle. *Science* **131** 1292–1297
Houchmandzadeh B 2002 Clustering of diffusing organisms. *Phys. Rev. E* **66** 052902
Houchmandzadeh B 2008 Neutral clustering in a simple experimental ecological community. *Phys. Rev. Lett.* **101** 078103
Houchmandzadeh B 2009 Theory of neutral clustering for growing populations. *Phys. Rev. E* **80** 051920
Houchmandzadeh B and Mihalcescu I 2011 Fluctuations importance and control in biological systems. *Europhys. News* **42** 36–39
Houchmandzadeh B and Vallade M 2003 Analytical solution of a neutral model of biodiversity. *Phys. Rev. E* **68** 061912
Houchmandzadeh B and Vallade M 2010 Alternative to the diffusion equation in population genetics. *Phys. Rev. E* **82** 051913
Houchmandzadeh B and Vallade M 2012 Selection for altruism through random drift in variable size populations. *BMC Evol. Biol.* **12** 61
Hu X-S, He F and Hubbell S 2006 Neutral theory in macroecology and population. *Oikos* **113** 548–556
Hubbel S 2001 *The unified neutral theory of biodiversity and biogeography* (Princeton, NJ: Princeton University Press)
Jabot F and Chave J 2011 Analyzing tropical forest tree species abundance distributions using a nonneutral model and through approximate Bayesian inference. *Am. Nat.* **178** E37–E47
Kessin R 2001 *Dictyostelium: evolution, cell biology and the development of multicellularity* (Cambridge: Cambridge University Press)
Kimura M 1962 On the probability of fixation of mutant genes in a population. *Genetics* **47** 713–719
Kimura M 1983 *The neutral theory of molecular evolution* (Cambridge University Press)
Lewontin R 1970 The units of selection. *Annu. Rev. Ecol. System.* **1** 1–14
Lomolino M 2006 *Biogeography* (US: Sinauer Associates Inc)
MacArthur R and Wilson E 1963 An equilibrium theory of insular zoogeography. *Evolution* **17** 373–387
MacArthur RH and Wilson EO 1967 *The theory of island biogeography* (Princeton: Princeton University Press)
Malécot G 1948 *Les mathématiques de l'hérédité* (Paris: Masson)
Mayr E 1942 *Systematics and the origin of species from the viewpoint of a zoologist* (New York: Columbia University Press)
McGill B, Maurer B and Weiser M 2006 Empirical evaluation of neutral theory. *Ecology* **87** 1411–1423
Moran P 1962 *The statistical processes of evolutionary theo*ry (Oxford: Oxford University Press)
Novick A and Weiner M 1957 Enzyme induction as an all-or-none phenomenon. *PNAS* **43** 553–566
Nowak, MA (2006a) *Evolutionary dynamics: Exploring the equations of life* (Cambridge, MA: Belknap Press)
Nowak M 2006b Five rules for the evolution of cooperation. *Science* **314** 1560–1563
Nowak M, Tarnita C and Wilson E 2010 The evolution of eusociality. *Nature* **466** 1057–1062
O'Malley M 2007 Nature reviews. *Microbiology* **5** 647–651
Okubo A and Levin S 2002 *Diffusion and ecological problems: Modern perspectives* 2nd edition (Springer: New York)
Patwa Z and Wahl L 2008 The fixation probability of beneficial mutations. *J. R. Soc. Interface* **5** 1279–1289
Price G 1970 Selection and covariance. *Nature* **227** 520–521
Rosindell J, Hubbell S and Etienne R 2011 The unified neutral theory of biodiversity and biogeography at age ten. *Trends Ecol. Evol.* **26** 340–348
Strassmann J, Page R, Robinson G and Seeley T 2011 Kin selection and eusociality. *Nature* **471** E5–E6
Taylor L, Woiwod I and Perry J 1978 The density dependence of spatial behaviour and the rarity of randomness. *J. Anim. Ecol.* **47** 383–406
Traulsen A, Sengupta A and Nowak M 2005 Stochastic evolutionary dynamics on two levels. *J. Theor. Biol.* **235** 393–401
Vallade M and Houchmandzadeh B 2003 Analytical solution of a neutral model of biodiversity. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **68** 61902
Vallade M and Houchmandzadeh B 2006 Species abundance distribution and population dynamics in a two-community model of neutral ecology. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **74** 51914
Visick K and Ruby E 2006 Vibrio fischeri and its host: it takes two to tango. *Curr. Opin. Microbiol.* **9** 632–638
Volkov I, Banavar J, Hubbell S and Maritan A 2003 Neutral theory and relative species abundance in ecology. *Nature* **424** 1035–1037
West S and Buckling A 2003 Cooperation, virulence and siderophore production in bacterial parasites. *Proc. Biol. Sci.* **270** 37–44
Williams G 1966 *Adaptation and natural selection: A critique of some current evolutionary thought* (Princeton NJ: Princeton University Press)
Williamson M, Gaston K and Lonsdale W 2008 The species–area relationship does not have an asymptote! *J. Biogeogr.* **28** 827–830
Young W, Roberts A and Stuhne G 2001 Reproductive pair correlations and the clustering of organisms. *Nature* **412** 328–331