

The paradox of the clumps mathematically explained

Hugo Fort · Marten Scheffer · Egbert H. van Nes

Received: 3 November 2008 / Accepted: 23 January 2009 / Published online: 27 February 2009
© Springer Science + Business Media B.V. 2009

Abstract The lumpy distribution of species along a continuous one-dimensional niche axis recently found by Scheffer and van Nes (Scheffer and van Ness 2006) is explained mathematically. We show that it emerges simply from the eigenvalue and eigenvectors of the community matrix. Both the transient patterns—lumps and gaps between them—as well as the asymptotic equilibrium are explained. If the species are evenly distributed along the niche axis, the emergence of these patterns can be demonstrated analytically. The more general case, of randomly distributed species, shows only slight deviations and is illustrated by numerical simulation. This is a robust result whenever the finiteness of the niche is taken into account: it can be extended to different analytic dependence of the interaction coefficients with the distance on the niche axis (i.e., different kernel interactions), different boundary conditions, etc. We also found that there is a critical value both for the width of the species distribution σ and the number of species n below which the clusterization disappears.

Keywords Self-organized similarity · Niche construction · Biodiversity · Competition

H. Fort (✉)
Facultad de Ciencias, Instituto de Física,
Universidad de la República,
Igua 4225,
11.400 Montevideo, Uruguay
e-mail: hugo@fisica.edu.uy

M. Scheffer · E. H. van Nes
Aquatic Ecology and Water Quality Management Group,
Wageningen Agricultural University,
P.O. Box 47, 6700 AA Wageningen, The Netherlands

M. Scheffer
e-mail: Marten.Scheffer@wur.nl

E. H. van Nes
e-mail: Egbert.vanNes@wur.nl

Introduction

In a recent work, Scheffer and van Nes (2006) found that community interactions alone can create lumpy distributions of species along a niche axis. This is an important finding as it implied that there are to different ways for species to coexist: either by being sufficiently different or by being sufficiently similar. This is in agreement with the conjecture of Bonsall et al. (2004) that coexisting species may be more similar than one would expect by chance alone and that their distribution falls into a number of distinct clusters. More recently, Roelke and Eldridge (2008) found similar patterns in different resource competition model. Furthermore, they suggest that this mechanism is not very fragile. This counterintuitive result that reconciles niche theory (MacArthur & Levins 1967) with neutrality (Allen et al. 2006; Holt et al. 2006) has received quite some attention (Herault et al. 2007; Hubbel et al. 2007) but remains hard to understand exactly.

Along a niche axis, self-organized clusters of look-a-likes emerge spontaneously from coevolution of competitors described by n -species Lotka–Volterra competition model (LVCM):

$$\frac{dx_i}{dt} = r_i x_i \left(1 - \sum_{j=1}^n \alpha_{ij} x_j \right) \quad i = 1, \dots, n \quad (1)$$

where x_i is the density of species i , normalized by its carrying capacity K_i ($x_i = N_i/K_i$); r_i is its maximum per capita growth rate; and the coefficients α_{ij} represent the competitive interaction of species j over species i . Or in matrix notation:

$$\frac{d\mathbf{x}}{dt} = \mathbf{r}\mathbf{x}(1 - \mathbf{A}\mathbf{x}) \quad (1')$$

This LVCM is parameterized in such a way that it mimics competition between species along a niche gradient

(see “Methods”) and so the elements α_{ij} of the \mathbf{A} matrix are given by

$$\alpha_{ij} = e^{-\frac{(\mu_i - \mu_j)^2}{4\sigma^2}} \tag{2}$$

where μ_i is the position on the niche axis of species i and σ stands for the width of the niche.

The model produces rather long transients with clumps of similar species that can become stable if there is density dependent predation (Scheffer and van Ness 2006).

Here, we explain this so-called paradox of the clumps (Nee and Colegrave 2006) from a mathematical point of view, showing that it emerges simply from the associated linear system of equations.

We start by considering an equilibrium of the system specified by a set of equilibrium densities x_i^* for each species i . To check the stability of this equilibrium, the standard procedure is to study the fate of small disturbances y_i from the equilibrium values x_i^* . It can be shown that the temporal variation of the perturbation \mathbf{y} is approximately proportional to the matrix \mathbf{A} : $d\mathbf{y}(t)/dt \propto -\mathbf{A}\mathbf{y}(t)$ (see “Methods”). This is a linear model than can be solved easily. For large times only the maximal eigenvalue λ_m of

the matrix $-\mathbf{A}$, and its associated eigenvector \mathbf{v}^m determines the rate and direction of change of the disturbance \mathbf{y} , respectively. As we will show, λ_m is positive and the initially small perturbation, instead of being smoothed out, yields for large times (see “Methods”).

$$\mathbf{x} \approx e^{-\mathbf{A}\mathbf{v}^m t} = e^{\lambda_m \mathbf{v}^m t} \text{ (for } t \text{ large)} \tag{3}$$

Analytic expressions for the eigenvalues and eigenvectors of $-\mathbf{A}$ are not known for the general case of random distributions of species on the niche axis. However, for the simpler case when the n species are evenly spaced along the niche axis, that is $\mu_j = (j-1)/n$, the eigenvalues λ_k and the components v_j^k of the eigenvectors \mathbf{v}^k ($k=1, \dots, n$) are given, for n even and PBC, by (Berlin & Kac 1952):

$$\lambda_k = - \left(c_1 + 2 \sum_{j=2}^{n/2} c_j(n, \sigma) \cos [2\pi(k-1)(j-1)/n] + c_{n/2+1}(n, \sigma) \cos [\pi(k-1)] \right), \tag{4}$$

with $c_j(n, \sigma) = e^{-\frac{(j-1)^2}{2\sigma n}}$ for $1 \leq j \leq n/2 + 1$,

$$v_j^k = \begin{cases} \left(\frac{2}{n}\right)^{\frac{1}{2}} \sin \frac{2\pi(k-1)(j-1)}{n} = \left(\frac{2}{n}\right)^{\frac{1}{2}} \sin [2\pi(k-1)\mu_j] & \text{if } k \leq \frac{n}{2} + 1 \\ \left(\frac{2}{n}\right)^{\frac{1}{2}} \cos \frac{2\pi(n-k+1)(j-1)}{n} = \left(\frac{2}{n}\right)^{\frac{1}{2}} \cos [2\pi(n-k+1)\mu_j] & \text{if } k > \frac{n}{2} + 1 \end{cases} \quad j = 1, \dots, n \tag{5}$$

From Eqs. 4 and 5, one can see that, except for $k=1$ and $k=n/2+1$, the eigenvalues come in pairs: $\lambda_{n-k+2} = \lambda_k$, one corresponding to the sine and the other to the cosine eigenvector. Equation 4 can be used to find numerically the index $k=m$ with the maximal eigenvalue $\lambda_m(n, \sigma)$ (as we have just seen, $k=n-m+2$ has the same value). The specific value of this index m is important since, as turns out from Eq. 5, the corresponding dominant eigenvector \mathbf{v}^m has $m-1$ peaks and $m-1$ valleys.

The maximal eigenvalue λ_m was determined from expression Eq. 4 for a grid of values of n and σ : $2 \leq N \leq 200$, and $0.05 \leq \sigma \leq 0.5$. The surface depicted in Fig. 1 corresponds to $\lambda_m(n, \sigma)$, showing that λ_m is positive except for small values of n ($n < 8$) or when σ is below a critical value σ_c which is approximately 0.075. This means that for higher values of n and s than those threshold values the equilibrium in which all species have the same biomass is not stable, implying that a pattern can be formed.

The dominant eigenvector \mathbf{v}^m can also be used to predict the species distribution in time, using Eq. 3. In Fig. 2, we can see that the lumps and gaps coincide, respectively, with the $m-1$ peaks and valleys of \mathbf{v}^m . In Fig. 3, we compare this

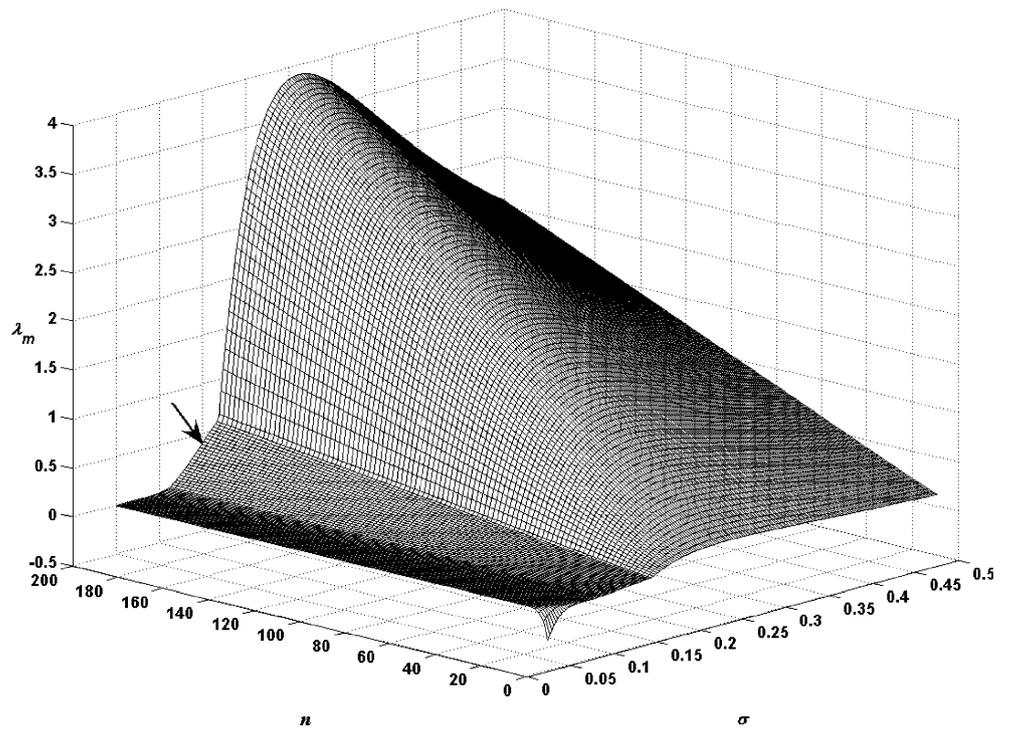
linearized model with simulations. For instance, if $n=200$ and $\sigma=0.15$, we get $m=5$ (and $m=200-5+2=197$), and then $\lambda_m=0.3938$ and \mathbf{v}^m is either $\sin(8\pi\mu_j)/10$ or $\cos(8\pi\mu_j)/10$. Figure 3A is a plot of a typical population distribution, produced by simulation, for $t=1,000$ generations and the expected biomass based on the dominant eigenvector $\mathbf{v}^m = \cos(8\pi\mu_j)/10$ substituted in the linearized model (Eq. 3). Notice that the agreement is quite good and that the quality of the agreement improves with time (Fig. 3B), until it becomes very good when the lumps are thinned to single lines.

So far, the results are for the particular case of the n species evenly spaced along the niche axis. What happens in the general case in which species are randomly distributed? In that case, we do not have an analytical proof. However, it turns out that simulations show that the results are very similar. This is evident from comparison of Fig. 3C and D with Fig. 3A and B, respectively.

Notice that for a given n (even) the maximum possible number of peaks is $n/2$ (one half of the components of the eigenvector pointing up and the other half down).

The integer m that gives the maximal eigenvalue is only a function of the width σ of the niche, $m=m(\sigma)$, and is

Fig. 1 λ_m , determined from Eq. 3 as a function of n and σ . The arrow denotes the point $n=200$ and $\sigma=0.15$



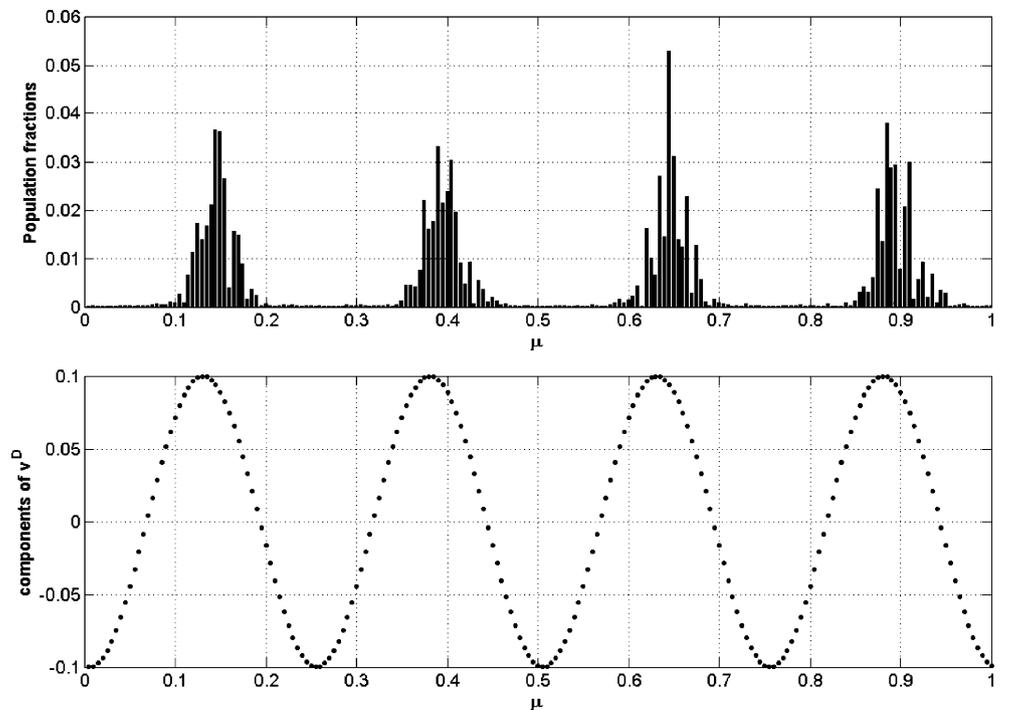
independent from n , provided n is large enough ($n/2 \geq m - 1$). It turns out that m is always an odd number (and then the number of clumps is even). The reason for this can be traced from the cosines appearing in Eq. 4 making contributions to the eigenvalues of opposite signs: positive for odd k and negative for even k . As a consequence the number of peaks, equal to $m - 1$, is

always even. For $\sigma=0.15$, $m - 1 = 4$ for all even n greater or equal than 8. The number of species that survive for asymptotically long times is then given simply by

$$n_\infty(\sigma) = m(\sigma) - 1 \tag{6}$$

In summary, for LVCM with MacArthur–Levins niche overlap, whenever the finiteness of the niche is taken into

Fig. 2 The coincidence of lumps and gaps with, respectively, the peaks and valleys of the dominant eigenvector for $n=200$ and $\sigma=0.15$. Above: results from a simulation after $t=1,000$ generations. Below: the components of the dominant eigenvector v^m



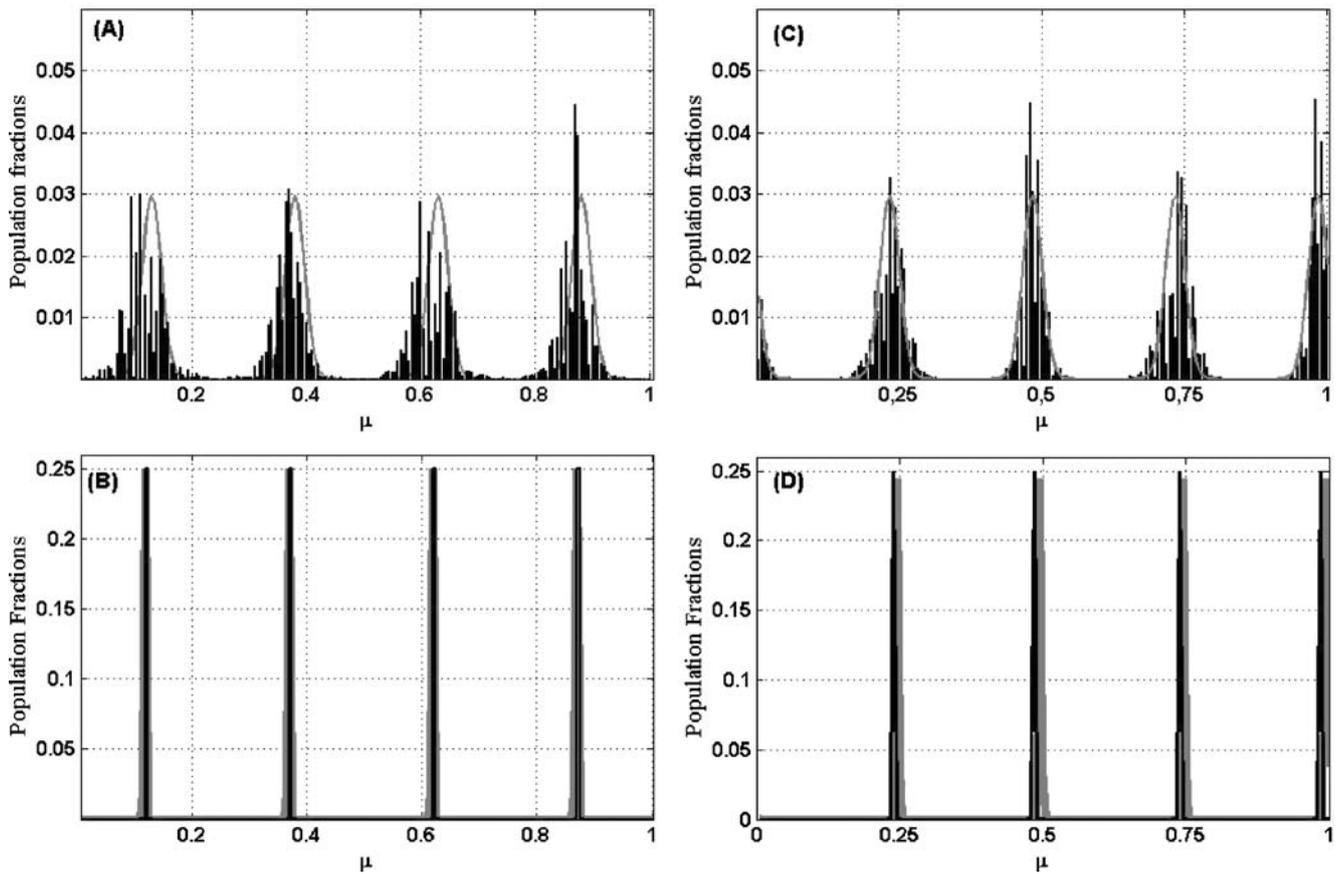


Fig. 3 Distribution of species for $n=200$ and $\sigma=0.15$. In *black*, results from a simulation after t generations and in *gray*, $\exp[\lambda_m v^m t]$. **a, b** Species evenly spaced along the niche axis for $t=1,000$ and $10,000$

generations, respectively. **c, d** Species randomly distributed along the niche axis for $t=1,000$ and $10,000$ (v^m is obtained now numerically from the matrix $-A$)

account (instead of considering a virtually infinite niche, see a discussion on this in “Methods”), the linearized model is enough to get: (1) qualitatively, the transient pattern—lumps and gaps between them—and (2) quantitatively, the asymptotic state. A key element behind this finding is that provided n is large enough and σ above a critical value, the maximum eigenvalue of the matrix $-A$ is always positive (or equivalently, the minimum eigenvalue of the community matrix A is always negative). As a consequence, the dominant eigenvector ends controlling the state of the ecosystem. This result is robust since it does not depend on taking different growth coefficients for the each species or the same r for all the species; it is quite independent on the interaction kernel and on the initial conditions.

Methods

Parameterization of the model

In the model, each species has an arbitrary random position μ_i on the niche axis ξ , between 0 and 1 (Scheffer & van Ness 2006). The widths of the niches is character-

ized by normal distributions on the niche axis with standard deviation σ (MacArthur & Levins 1967): $P(\xi) = (1/\sqrt{2\pi})e^{-(\xi-\mu)^2/(2\sigma^2)}$. The competition intensity between species i and species j is assumed to be related to their overlap and, thus, to the probability P that individuals of the two species are at the same position on the niche axis, which is the product of both probabilities. Then, the α_{ij} can be computed following MacArthur and Levins (1967), and we get the Eq. 2:

$$\alpha_{ij} \equiv \frac{\int_{-\infty}^{+\infty} P_i(\xi)P_j(\xi)d\xi}{\int_{-\infty}^{+\infty} P_i(\xi)^2 d\xi} = e^{-\frac{(\mu_i-\mu_j)^2}{4\sigma^2}}$$

$$i = 1, \dots, n; j = 1, \dots, n,$$

To avoid border effects, periodic boundary conditions (PBC) are imposed to the niche axis, so that each species has equal numbers of competitors on both sides. We implement this by just taking the minimal distance between $|\mu_i-\mu_j|$ and $1-|\mu_i-\mu_j|$ (see below for a discussion on alternative implementations of boundary conditions and their consequences).

Placing a large number of species at random locations on the niche axis, for example $n=200$, with random maximum per capita growth rates r_i between 0 and 1 and $\sigma=0.15$, and starting from a random configuration of fractions x_i , the above model gives rise, after typically some hundreds of generations, to a pattern of four lumps equally spaced by four gaps. In fact, these lumps are transient and each shall ultimately be thinned out to a single species for asymptotic times (after some thousands of generations). Since we checked that a fixed r for all the species does not introduce noticeable changes, we work in what follows with $r_i=1$ independent of the index i . This is easy to understand since what is relevant for the equilibrium values x_i^* are the terms between brackets in the Lotka–Volterra equation (Eq. 1), and this does not change.

For other values of σ , the number of lumps (or gaps) may change: the smaller σ , the greater the number of lumps.

The population densities for long times obtained from a linear approximation

The equilibrium fractions, x_i^* , follow from Eq. 1 equating the right sides to 0:

$$x_i^* \left(1 - \sum_{j=1}^n \alpha_{ij} x_j^* \right) = 0 \quad i = 1, \dots, n \tag{7}$$

Taylor expanding about this equilibrium for each population, i.e.,

$$x_i = x_i^* + y_i(t) \tag{8}$$

where y_i denotes an initially small perturbation to the i th population and discarding all terms which are of second or higher order in the perturbations y , a linear approximation is obtained as:

$$\frac{dy_i(t)}{dt} \cong - \sum_{j=1}^n x_i^* \alpha_{ij} y_j(t) \quad i = 1, \dots, n \tag{9}$$

Then one has to find the eigenvalues of the Jacobian or, in the ecological language, the community matrix (May 1974) $J_{ij} = -x_i^* \alpha_{ij}$. For simplicity, let us consider the equilibrium in which all populations densities are equal, $x_i^*=x^*$ for all i .¹ (We checked that the same results are obtained when

starting from a completely random assignation of densities.) Next, In this case, the equilibrium condition Eq. 7 reduces to:

$$1 - x^* \sum_{j=1}^n \alpha_{ij} = 0 \quad i = 1, \dots, n \tag{7'}$$

In addition, Eq. 9 becomes

$$d\mathbf{y}(t)/dt = -\mathbf{x}^* \mathbf{A}\mathbf{y}(t) \tag{9'}$$

i.e., J_{ij} becomes proportional to $-\alpha_{ij}$. This implies that the stability of the equilibrium depends simply on the eigenvalues of the matrix $-\mathbf{A}$. Since the coefficients of \mathbf{A} , given by Eq. 2, are symmetric, then all its eigenvalues are real. Then, for asymptotic times, \mathbf{y} becomes proportional to the dominant eigenvector \mathbf{v}^m , the one associated with the maximum eigenvalue of $-\mathbf{A}$ (or, equivalently, the minimum eigenvalue of \mathbf{A}) λ_m

$$\mathbf{y}(t) \propto e^{\lambda_m t} \mathbf{v}^m \text{ (for } t \text{ large)} \tag{10}$$

By Eqs. 8 and 7', we have for the expression between brackets in Eq. 1:

$$\left(1 - \sum_{j=1}^n \alpha_{ij} x_j \right) = - \sum_{j=1}^n \alpha_{ij} y_j \tag{11}$$

Thus, Eq. 1 can be rewritten in matrix form as

$$\frac{d \ln \mathbf{x}}{dt} = -\mathbf{A}\mathbf{y} \tag{12}$$

Notice that from Eq. 10, due to the fact that λ_m is positive, for large times, \mathbf{y} becomes proportional to \mathbf{v}^m (otherwise, if λ_m would be negative, the equilibrium would be stable and the perturbation would be smoothed out for large times). We then have

$$\frac{d \ln \mathbf{x}}{dt} \approx -\mathbf{A}\mathbf{v}^m = \lambda_m \mathbf{v}^m \text{ (for } t \text{ large)} \tag{13}$$

and by integration, we get the approximated solution given by Eq. 3.

Different implementations of the boundary conditions and their effects

Let us analyze the robustness of the lumpy pattern. For the case of the Gaussian kernel (*Normal* probability distributions and MacArthur–Levins niche overlap), it is sensitive on the implementation of boundary conditions: when using the traditional periodic implementation of PBC described above, as we do, the minimal eigenvalue is positive and then the lumpy pattern occurs only for σ above a critical value σ_c . Furthermore, there are lumps for kernels with non-integer exponents $p > 1$. In fact, the formula for the

¹ Of course, in the general case of species randomly distributed, this is not an exact equilibrium. However, for n large enough—say $n = 200$ —the sum of the competition coefficients along any row (or column), by a probabilistic argument, is more or less equal to a constant Σ , so $x_i^* = x^* = 1/\Sigma$ is close to an equilibrium. For the case of evenly distributed species, $x_i^* = x^* = 1/\Sigma$ is an exact equilibrium although in general nonstable.

eigenvectors of Berlin–Kac is valid for any cyclic matrix α of the form

$$\alpha = \begin{bmatrix} c_1 c_2 c_3 \dots \dots c_n \\ c_n c_1 c_2 \dots \dots c_{n-1} \\ \dots \dots \dots \dots \dots \dots \\ c_2 c_3 \dots \dots \dots c_n c_1 \end{bmatrix}. \tag{14}$$

Therefore, independently of the kernel, provided the dominant eigenvalue is negative, there is clusterization of species. This happens for kernels with arbitrary real exponent p greater than 1. So in this case, the pattern is robust against changes in the kernel.

On the other hand, when using an alternative procedure to implement PBC, consisting in taking a periodic array of copies of the same system, the lumpy pattern disappear (Pigolotti et al. 2007, 2008) independently of the value of σ . This “perfectly periodic” boundary conditions mimic an infinite (but periodic) niche. Then, since you can always set $\sigma=1$ by rescaling the positions μ on the niche axis, there is no transition in the behavior for a critical value of σ .

Since it seems a reasonable assumption that things in ecosystems do depend on the value of σ , and we are interested *precisely* in measuring this effect, we consider the ordinary procedure for implementing PBC and then considering a non-infinite niche.

We also analyzed what happens when the finite length L of the niche axis is taken into account explicitly. The competition coefficients α_{ij} for these finite boundary conditions (FBC) are now given by

$$\alpha_{ij} = e^{-\frac{(\mu_i - \mu_j)^2}{4\sigma^2}} \frac{\operatorname{erf}\left(\frac{2L - \mu_i - \mu_j}{2\sigma}\right) + \operatorname{erf}\left(\frac{\mu_i + \mu_j}{2\sigma}\right)}{\operatorname{erf}\left(\frac{L - \mu_i}{\sigma}\right) + \operatorname{erf}\left(\frac{\mu_i}{\sigma}\right)} \tag{15}$$

$$i = 1, \dots, n; j = 1, \dots, n$$

Again, for the FBC interaction coefficients given by Eq. 15, a lumpy pattern emerges although it shows some differences.

For instance, for $s=0.15$, the species organize into three lumps instead of four².

Acknowledgments H.F. thanks Raúl Donangelo for his thorough criticism and constructive suggestions and acknowledges financial support from PEDECIBA, Uruguay. Work was supported, in part, by Project PDT63-013.

References

Allen CR (2006) Discontinuities in ecological data. Proc Natl Acad Sci U S A 103:6083–6084. doi:10.1073/pnas.0601668103

Berlin TH, Kac M (1952) The spherical model of a ferromagnet. Phys Rev 86:821–835. doi:10.1103/PhysRev.86.821

Bonsall MB, Jansen VAA, Hassell MP (2004) Life history trade-offs assemble ecological guilds. Science 306:111–114. doi:10.1126/science.1100680

Herault B (2007) Reconciling niche and neutrality through the Emergent Group approach. Perspect Plant Ecol Evol Syst 9:71–78. doi:10.1016/j.ppees.2007.08.001

Holt RD (2006) Emergent neutrality. Trends Ecol Evol 21:531–533. doi:10.1016/j.tree.2006.08.003

Hubbell SP (2001) The Unified Neutral Theory of Biodiversity and Biogeography.. Princeton University Press, Princeton, p 375

MacArthur RH, Levins R (1967) The limiting similarity, convergence and divergence of coexisting species. Am Nat 101:377–385. doi:10.1086/282505

May RM (1974) Stability and Complexity in Model Ecosystems. Princeton University Press, Princeton

Nee S, Colegrave N (2006) Ecology—Paradox of the clumps. Nature 441:417–418. doi:10.1038/441417a

Pigolotti S, López C, Hernández-García E (2007) Species clustering in competitive Lotka–Volterra models. Phys Rev Lett 98:258101. doi:10.1103/PhysRevLett.98.258101

Pigolotti S, López C, Hernández-García E., Andersen KH (2008). How competition leads to lumpy or uniform species distribution. arXiv:0802.3274v1

Roelke DL, Eldridge PM (2008) Mixing of supersaturated assemblages and the precipitous loss of species. Am Nat 171:162–175. doi:10.1086/524955

Scheffer M, van Nes EH (2006) Self-organized similarity, the evolutionary emergence of groups of similar species. Proc Natl Acad Sci U S A 103:6230–6235. doi:10.1073/pnas.0508024103

² In this case, although the dominant eigenvalue is still negative, due to the factor multiplying the Gaussian, the matrix A is no longer symmetric and so there appear complex eigenvalues.