Emergent Behavior in Flocks

Felipe Cucker *
Department of Mathematics
City University of Hong Kong
83 Tat Chee Avenue, Kowloon
HONG KONG
e-mail: macucker@math.cityu.edu.hk

Steve Smale †
Toyota Technological Institute at Chicago
1427 East 60th Street, Chicago, IL 60637
U.S.A.
e-mail: smale@math.berkeley.edu

July 13, 2005

PRELIMINARY VERSION.

1 Introduction

As a motivating example we consider a population, say of birds or fish, whose members are moving in $\mathbb{R}^3$. It has been observed that under some initial conditions, for example on their positions and velocities, the state of the flock converges to one in which all birds fly with the same velocity. A goal of this paper is to provide some justification of this observation. To do so, we will postulate a model for the evolution of the flock and exhibit conditions on the initial state under which a convergence as above is established. In case these conditions are not satisfied, dispersion of the flock may occur.

There has been a large amount of literature on flocking, herding and schooling. Much of it is descriptive, most of the remaining proposes models, which are then studied via computer simulations, e.g., [3, 7]. A starting point for this paper is the model proposed in the latter of these references which, for convenience, we will call

*Partially funded by a grant from the Research Grants Council of the Hong Kong SAR (project number CityU 1085/02P). Thanks are due to the Toyota Technological Institute at Chicago for a very pleasant stay there in May 2005.

†Partially supported by an NSF grant. This author also wants to thank City University of Hong Kong for inviting me for a month during June 2005.

Vicsek's model. Its analytic behavior was subsequently studied in [4] and this paper, brought to our attention by Ali Jadbabaie, has been helpful for us.

Vicsek's model is motivated by the idea that bird i adjusts its velocity towards the average of its neighbors' velocities. With (our first modification) $\mathbb{R}^2$ replaced by Euclidean space $\mathbb{E}^3$ and the heading θ replaced by the velocity v. The model is

$$\begin{aligned} x_i(t+1) &= x_i(t) + v_i(t) \\ v_i(t+1) &= \frac{1}{n_i(t)} \sum_{j \in \mathcal{N}_i(t)} v_j(t). \end{aligned} \tag{1}$$

where $x_i, v_i \in \mathbb{E}^3$ for $i = 1, \ldots, k$ and time $t = 0, 1, 2 \ldots$. Here $\mathcal{N}_i(t) = \{j \leq k \mid \|x_i(t) - x_j(t)\| \leq r\}$ and $n_i(t) = \#\mathcal{N}_i(t)$ for some $r > 0$.

Let A_x be the $k \times k$ binary matrix given by

$$a_{ij} = \begin{cases} 1 & \text{if } \|x_i(t) - x_j(t)\| \leq r \\ 0 & \text{otherwise.} \end{cases}$$

Denoting by A_i the ith row of A_x and by $\mathbf{e}_i$ the vector $(0, \ldots, 0, 1, 0, \ldots, 0)$ with the 1 in the ith place, we have

$$v_i(t+1) - v_i(t) = \left(\frac{1}{n_i(t)} A_i - \mathbf{e}_i \right) v(t) = \frac{1}{n_i(t)} [A_i - n_i \mathbf{e}_i] v(t).$$

The last two expressions should be understood as linear combinations of elements in $\mathbb{E}^3$. Extending this notation to matrix form,

$$v(t+1) - v(t) = -D_x^{-1} L_x v(t) \tag{2}$$

where D_x is the $k \times k$ diagonal matrix whose ith diagonal entry is $\sum_{j \leq k} a_{ij}$ and L_x is the matrix whose ith row is $A_i - n_i \mathbf{e}_i$, that is,

$$L_x = D_x - A_x. \tag{3}$$

We found it convenient to modify (2) by scaling L_x in a slightly different way, namely,

$$\begin{aligned} x(t+1) &= x(t) + \Delta t v(t) \\ v(t+1) &= \left(\mathrm{Id} - \frac{L_x}{\|L_x\|} \right) v(t). \end{aligned} \tag{4}$$

Here $\|L_x\|$ is the operator norm of L_x (with respect to the norm in $(\mathbb{E}^3)^k$ induced by the norm of $\mathbb{E}^3$).

Our third modification proceeds as follows. It is reasonable to assume that birds influence each other as a function of their distance. We give form to this assumption via a non-increasing function $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ such that the *adjacency matrix* A_x has entries

$$a_{ij} = \eta(\|x_i - x_j\|^2). \tag{5}$$

In this paper we will take, for some fixed $K, \sigma > 0$ and $\beta \geq 0$,

$$\eta(y) = \frac{K}{(\sigma^2 + y)^\beta}. \tag{6}$$

Vicsek's adjacency matrix is also of this form where, for some $r > 0$,

$$\eta(y) = \begin{cases} 1 & \text{if } y \leq r^2 \\ 0 & \text{otherwise.} \end{cases}$$

Note that, in contrast with the abrupt behavior of this last function, the function in (6) decreases continuously with y and the rate of decay is given by $\beta > 0$.

We also consider evolution for continuous time. The corresponding model can be given by the system of differential equations

$$\begin{aligned} x' &= v \\ v' &= -L_x v. \end{aligned} \tag{7}$$

Our first two main results give conditions to ensure that the birds' velocities converge to a common one and the distance between birds remain bounded for both continuous and discrete time. They can be stated as follows (more precise statements are in Theorems 2 and 3 below).

Theorem 1 *Let $(x(t), v(t))$ be a solution of (4) with initial conditions $x(0) = x_0$ and $v(0) = v_0$. If $\beta < 1/2$ then, when $t \to \infty$ the velocities $v_i(t)$ tend to a common limit $\widehat{v} \in \mathbb{E}^3$ and the vectors $x_i - x_j$ tend to a limit vector $\widehat{x_{ij}}$, for all $i, j \leq k$. The same happens if $\beta \geq 1/2$ provided the initial values x_0 and v_0 satisfy a given, explicit, relation.*

The same holds for a solution of (7).

2 Some preliminaries

Given a nonnegative, symmetric, $k \times k$ matrix A the *Laplacian* L of A is defined to be

$$L = D - A$$

where $D = \mathsf{diag}(d_1, \ldots, d_k)$ and $d_\ell = \sum_{j=1}^k a_{\ell j}$. Some features of L are immediate. It is symmetric and it does not depend on the diagonal entries of A. The Laplacian as just defined has its origins in graph theory where the matrix A is the adjacency matrix of a graph G and many of the properties of G can be read out from L (see [6]).

The matrix L_x in (4) and (7) is thus the Laplacian of A_x. It acts on $(\mathbb{E}^3)^k$ and satisfies the following:

(a) For all $v \in \mathbb{E}^3$, $L_x(v, \ldots, v) = 0$.

(b) If $\lambda_1, \ldots, \lambda_k$ are the eigenvalues of L_x then

$$0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_k = \|L_x\|.$$

(c) For all $v \in (\mathbb{E}^3)^k$,

$$\langle L_x v, v \rangle = \sum_{i,j=1}^{k} a_{ij} \|v_i - v_j\|^2.$$

Note that (b) implies L_x is positive semidefinite.

The quantity $E_x(v) = \sum_{i,j=1}^{k} a_{ij} \|v_i - v_j\|^2$ is the *energy* of the flock (at a position $x \in (\mathbb{E}^3)^k$ and a velocity $v \in (\mathbb{E}^3)^k$). Note that $E_x(v) = 0$ when all birds are flying with the same velocity. That is, they fly with the same heading and at the same speed.

The matrix $\mathrm{Id} - \frac{1}{\|L_x\|} L_x$ in (4) acts on $(\mathbb{E}^3)^k$. The fixed points for this action are easily characterized.

Proposition 1 *Let $v \in (\mathbb{E}^3)^k$. The following are equivalent:*

(1) *v is a fixed point (i.e., $\left(\mathrm{Id} - \frac{1}{\|L_x\|} L_x\right) v = v$).*

(2) $L_x(v) = 0$.

(3) $E_x(v) = 0$.

PROOF. The equivalence between (1) and (2) is obvious. The implication (2) $\Longrightarrow$ (3) is trivial. Finally, note that (3) implies that $v_i = v_j$ for all $i \neq j$ and this, together with (a) above, implies (2). $\square$

The second eigenvalue λ_2 of L_x is called the *Fiedler number* of A_x. We denote the Fiedler number of A_x by ϕ_x.

Remark 1 One difference between Vicsek's model and (4) lies in the way in which L_x is scaled. In (2) the scaling used is $D_x^{-1} L_x$. The product matrix $H = D_x^{-1} L_x$ satisfies that, for all $i \leq k$, $\|H_i\|_\infty = 1$. Also, the product matrix $S = D_x^{-1} A_x$ is stochastic (i.e., nonnegative and such that $\|S_i\|_1 = 1$). We note, however, that it destroys the symmetry of both A_x and L_x.

The scaling in (4) considers instead $\frac{L_x}{\|L_x\|}$. This does not lead to stochasticity but preserves symmetry.

We end these preliminaries introducing some concepts which will be useful in this paper.

Let Δ be the diagonal of $(\mathbb{E}^3)^k$, i.e.,

$$\Delta = \{(v, v, \ldots, v) \mid v \in \mathbb{E}^3\}$$

and $\Delta^\perp$ be the orthogonal complement of Δ in $(\mathbb{E}^3)^k$. Then, every point $x \in (\mathbb{E}^3)^k$ decomposes in a unique way as $x = x_\Delta + x_\perp$ with $x_\Delta \in \Delta$ and $x_\perp \in \Delta^\perp$. Note that if $x(t+1) = x(t) + \Delta t v(t)$ then $x(t+1)_\perp = x(t)_\perp + \Delta t v(t)_\perp$. Similarly, if $v(t+1) = -\left(\mathrm{Id} - \frac{L_x}{\|L_x\|}\right) v(t)$ then

$$v(t+1)_\perp = -\left(\mathrm{Id} - \frac{L_x}{\|L_x\|}\right) v(t)_\perp$$

since $L_x(\Delta) = 0$ and $L_x(\Delta^\perp) \subseteq \Delta^\perp$. Finally, note that for all $x \in (\mathbb{E}^3)^k$ the matrices A_x and $A_{x_\perp}$ are equal. It follows that the projections over $\Delta^\perp$ of the solutions of (4) are the solutions of the restriction of (4) to $\Delta^\perp$. A similar remark holds for (7).

These projections over $\Delta^\perp$ are of the essence since we are interested on the differences $x_i - x_j$ and $v_i - v_j$, for $i \neq j$, rather than on the x_i or v_i themselves.

We denote $\Gamma = \frac{1}{2}\sum_{i\neq j} \|x_i - x_j\|^2$ and $\Lambda = \frac{1}{2}\sum_{i\neq j} \|v_i - v_j\|^2$. To better deal with these functions consider $Q : (\mathbb{E}^3)^k \times (\mathbb{E}^3)^k \to \mathbb{R}$ defined by

$$Q(u,v) = \frac{1}{2}\sum_{i,j=1}^{k} \langle u_i - u_j, v_i - v_j\rangle.$$

Then Q is bilinear, symmetric, and, when restricted to $\Delta^\perp \times \Delta^\perp$, positive definite. It follows that it defines an inner product $\langle\ ,\ \rangle_Q$ on $(\mathbb{E}^3)^k/\Delta \simeq \Delta^\perp$. Now note that $\Lambda = \|v\|_Q^2$ and $\Gamma = \|x\|_Q^2$ and that $\Gamma(x) = \Gamma(x_\perp)$ and $\Lambda(v) = \Lambda(v_\perp)$.

Let $\boldsymbol{\nu}, \overline{\boldsymbol{\nu}} > 0$ be such that, restricted to $\Delta^\perp$,

$$\boldsymbol{\nu}\|\ \|^2 \leq \|\ \|_Q^2 \leq \overline{\boldsymbol{\nu}}\|\ \|^2.$$

Note that $\boldsymbol{\nu}, \overline{\boldsymbol{\nu}}$ depend only on k. We now show bounds for them in terms of k.

Lemma 1 *For all $k \geq 2$, $\boldsymbol{\nu}(k) \geq \frac{1}{3k}$ and $\overline{\boldsymbol{\nu}}(k) \leq 2k(k-1)$.*

Proof. By definition, $\overline{\boldsymbol{\nu}} \leq \max_{\|x\|=1} \|x\|_Q^2$. Since $\|x\| = 1$, $\|x_i\| \leq 1$ for $i = 1, \ldots, k$ and therefore, $\|x_i - x_j\|^2 \leq 4$ for all $i \neq j$. This implies

$$\|x\|_Q^2 \leq \frac{1}{2}k(k-1)4 = 2k(k-1).$$

Also by definition, $\frac{1}{\boldsymbol{\nu}} \leq \max_{\|x\|_Q=1} \|x\|^2$. Let $x \in \Delta^\perp$ such that $\|x\|_Q = 1$. We claim that, for all $i \leq k$ and $\ell \leq 3$, $|x_{i\ell}| < 1$. Assume the contrary. Then there exists i_0 and ℓ such that $|x_{i_0\ell}| \geq 1$. Without loss of generality, $x_{i_0\ell} \geq 1$. Since $\sum x_i = 0$, there exists i_1 such that $x_{i_1\ell} < 0$. But then

$$\|x\|_Q^2 = \frac{1}{2}\sum_{i\neq j} \|x_i - x_j\|^2 \geq \|x_{i_0} - x_{i_1}\|^2 \geq (x_{i_0\ell} - x_{i_1\ell})^2 > 1$$

contradicting $\|x\|_Q^2 = 1$. So the claim is proved. Finally

$$\|x\|^2 = \sum_{i=1}^{k} \sum_{\ell=1}^{3} x_{i\ell}^2 \leq 3k$$

which shows $\frac{1}{\boldsymbol{\nu}} \leq 3k$. □

Remark 2 The condition "the velocities $v_i(t)$ tend to a common limit $\widehat{v} \in \mathbb{E}^3$" in Theorem 1 is equivalent to the condition "$v_\perp(t) \to 0$." Also, the condition "the vectors $x_i - x_j$ tend to a limit vector $\widehat{x_{ij}}$, for all $i, j \leq k$" is equivalent to "$x_\perp(t)$ tend to a limit vector $\widehat{x}$ in $\Delta^\perp$." This suggests that we are actually interested on the solutions of the systems induced by (4) and (7), respectively, on the space $\Delta^\perp \times \Delta^\perp$. Since, as we mentioned, these induced systems have the same form as (4) and (7), we will keep referring to them but we will consider them on $\Delta^\perp \times \Delta^\perp$. Actually, we will consider positions in

$$X := (\mathbb{E}^3)^k / \Delta \simeq \Delta^\perp$$

and velocities in

$$V := (\mathbb{E}^3)^k / \Delta \simeq \Delta^\perp.$$

3 Convergence in continuous time

Proposition 2 *Let A be a symmetric matrix, $L = D - A$ its Laplacian, ϕ its Fiedler number, and $\mu = \min_{i \neq j} a_{ij}$. Then $\phi \geq \boldsymbol{\nu}\mu$. In particular, if $a_{ij} = \eta(\|x_i - x_j\|^2)$ then*

$$\phi \geq \boldsymbol{\nu}\eta(\Gamma_x).$$

PROOF. For all $v \in V$

$$\|Lv\|\|v\| \geq \langle Lv, v \rangle = \sum_{i,j=1}^{k} a_{ij}\|v_i - v_j\|^2 \geq \mu\|v\|_Q^2 \geq \boldsymbol{\nu}\mu\|v\|^2.$$

It follows that $\|Lv\| \geq \boldsymbol{\nu}\mu\|v\|$ and thus the statement. □

In the following we fix a solution (x, v) of (7). At a time $t \in \mathbb{R}_+$, $x(t)$ and $v(t)$ are elements in X and V, respectively. In particular, $x(t)$ determines an adjacency matrix $A_{x(t)}$. For notational simplicity we will denote this matrix by A_t and its Laplacian and Fiedler number by L_t and ϕ_t, respectively. Similarly, we will write $\Lambda(t)$ and $\Gamma(t)$ for the values of Λ and Γ, respectively, at $(v(t), x(t))$. Finally, we will write Γ_0 for $\Gamma(0)$ and similarly for Λ_0.

Denote $\Phi_t = \min_{\tau \in [0,t]} \phi_\tau$.

Proposition 3 *For all $t \geq 0$*

$$\Lambda(t) \leq \Lambda_0 e^{-2t\Phi_t}.$$

PROOF. Let $\tau \in [0, t]$. Then

$$\begin{aligned} \Lambda'(\tau) &= \frac{d}{d\tau}\langle v(\tau), v(\tau)\rangle_Q \\ &= 2\langle v'(\tau), v(\tau)\rangle_Q \\ &= -2\langle L_\tau v(\tau), v(\tau)\rangle_Q \\ &\leq -2\phi_{x(\tau)}\Lambda(\tau). \end{aligned}$$

Here we have used that L_τ is symmetric positive definite on V. Using this inequality,

$$\ln(\Lambda(\tau))\Big|_0^t = \int_0^t \frac{\Lambda'(\tau)}{\Lambda(\tau)}d\tau \leq \int_0^t -2\phi_\tau d\tau = -2t\Phi_t$$

i.e.,

$$\ln(\Lambda(t)) - \ln(\Lambda_0) \leq -2t\Phi_t$$

from which the statement follows. □

Proposition 4 *For $T > 0$*

$$\Gamma(T) \leq 2\left(\Gamma_0 + \frac{\Lambda_0}{\Phi_T^2}\right).$$

PROOF. We have $|\Gamma'(t)| = |2\langle v(t), x(t)\rangle_Q| \leq 2\|v(t)\|_Q\|x(t)\|_Q$. But $\|x(t)\|_Q = \Gamma(t)^{1/2}$ and $\|v(t)\|_Q^2 = \Lambda(t) \leq \Lambda_0 e^{-2t\Phi_t}$, by Proposition 3. Therefore,

$$\Gamma'(t) \leq |\Gamma'(t)| \leq 2\left(\Lambda_0 e^{-2t\Phi_t}\right)^{1/2}\Gamma(t)^{1/2} \tag{8}$$

and, using that $t \mapsto \Phi_t$ is non-increasing,

$$\begin{aligned} \int_0^T \frac{\Gamma'(t)}{\Gamma(t)^{1/2}}dt &\leq 2\int_0^T \left(\Lambda_0 e^{-2t\Phi_t}\right)^{1/2} dt \\ &\leq 2\int_0^T \Lambda_0^{1/2} e^{-t\Phi_T} dt \\ &= 2\Lambda_0^{1/2}\left(-\frac{1}{\Phi_T}\right) e^{-t\Psi(T)}\Big|_0^T \leq \frac{2\Lambda_0^{1/2}}{\Phi_T} \end{aligned}$$

which implies

$$\Gamma(t)^{1/2}\Big|_0^T = \frac{1}{2}\int_0^T \frac{\Gamma'(t)}{\Gamma(t)}dt \leq \frac{\Lambda_0^{1/2}}{\Phi_T}$$

from which it follows that

$$\Gamma(T) \leq \left(\Gamma_0^{1/2} + \frac{\Lambda_0^{1/2}}{\Phi_T}\right)^2.$$

The statement now follows from the elementary inequality $(\alpha+\beta)^2 \leq 2(\alpha^2+\beta^2)$. □

A proof of the following lemma is in [1, Lemma 7].

Lemma 2 *Let $c_1, c_2 > 0$ and $s > q > 0$. Then the equation*

$$F(z) = z^s - c_1 z^q - c_2 = 0$$

has a unique positive zero z_*. In addition

$$z_* \leq \max\left\{(2c_1)^{\frac{1}{s-q}}, (2c_2)^{\frac{1}{s}}\right\}$$

and $F(z) \leq 0$ for $0 \leq z \leq z^*$. □

Theorem 2 *Assume that there are constants $K, \sigma > 0$ and $\beta \geq 0$ and*

$$a_{ij} = \frac{K}{(\sigma^2 + \|x_i - x_j\|^2)^\beta}.$$

Assume also that one of the three following hypothesis hold:

(i) $\beta < 1/2$,

(ii) $\beta = 1/2$ *and* $\Lambda_0 < \frac{(\boldsymbol{\nu}K)^2}{2}$,

(iii) $\beta > 1/2$ *and*

$$\left[\left(\frac{1}{2\beta}\right)^{\frac{1}{2\beta-1}} - \left(\frac{1}{2\beta}\right)^{\frac{2\beta}{2\beta-1}}\right]\left(\frac{(\boldsymbol{\nu}K)^2}{2\Lambda_0}\right)^{\frac{1}{2\beta-1}} > 2\Gamma_0 + \sigma^2.$$

Then there exists a constant B_0 (independent of t, made explicit in the proof of each of the three cases) such that $\Gamma(t) \leq B_0$ for all $t \in \mathbb{R}_+$. In addition, $\Lambda(t) \to 0$ when $t \to \infty$. Finally, there exists $\widehat{x} \in X$ such that $x(t) \to \widehat{x}$ when $t \to \infty$.

Proof. By Proposition 2, for all $x \in X$,

$$\phi_x \geq \frac{\boldsymbol{\nu}K}{(\sigma^2 + \max_{i\neq j}\|x_i - x_j\|^2)^\beta} \geq \frac{\boldsymbol{\nu}K}{(\sigma^2 + \Gamma_x)^\beta}.$$

Let $t^* \in [0,t]$ be the point maximizing Γ in $[0,t]$. Then

$$\Phi_t = \min_{\tau\in[0,t]} \phi_\tau \geq \min_{\tau\in[0,t]} \frac{\boldsymbol{\nu} K}{(\sigma^2+\Gamma(\tau))^\beta} \geq \frac{\boldsymbol{\nu} K}{(\sigma^2+\Gamma(t^*))^\beta}.$$

By Proposition 4

$$\Gamma(t) \leq 2\Gamma_0 + 2\Lambda_0 \frac{(\sigma^2+\Gamma(t^*))^{2\beta}}{(\boldsymbol{\nu} K)^2}. \tag{9}$$

Since t^* maximizes Γ in $[0,t]$ it also does so in $[0,t^*]$. Thus, for $t=t^*$, (9) takes the form

$$(\sigma^2+\Gamma(t^*)) - 2\Lambda_0 \frac{(\sigma^2+\Gamma(t^*))^{2\beta}}{(\boldsymbol{\nu} K)^2} - (2\Gamma_0+\sigma^2) \leq 0. \tag{10}$$

Let $z=\Gamma(t^*)+\sigma^2$,

$$\boldsymbol{a} = \frac{2\Lambda_0}{(\boldsymbol{\nu} K)^2}, \quad \text{and} \quad \boldsymbol{b} = 2\Gamma_0+\sigma^2.$$

Then (10) can be rewritten as $F(z)\leq 0$ with

$$F(z) = z - \boldsymbol{a} z^{2\beta} - \boldsymbol{b}.$$

(i) Assume $\beta < 1/2$. By Lemma 2, $F(z) \leq 0$ implies that $z = (\sigma^2+\Gamma(t^*)) \leq U_0$ with

$$U_0 = \max\left\{ \left(\frac{4\Lambda_0}{(\boldsymbol{\nu} K)^2}\right)^{\frac{1}{1-2\beta}}, 2(2\Gamma_0+\sigma^2) \right\}.$$

That is $\Gamma(t^*) \leq B_0 := U_0 - \sigma^2$. Since B_0 is independent of t, we deduce that, for all $t \in \mathbb{R}_+$, $\Gamma(t) \leq B_0$. But this implies that $\phi_t \geq \frac{\boldsymbol{\nu} K}{(\sigma^2+B_0)^\beta}$ for all $t\in\mathbb{R}_+$ and therefore, the same bound holds for Φ_t. By Proposition 3

$$\Lambda(t) \leq \Lambda_0 e^{-2\frac{\boldsymbol{\nu} K}{(\sigma^2+B_0)^\beta} t} \tag{11}$$

which shows that $\Lambda(t)\to 0$ when $t\to\infty$. Finally, for all $T>t$,

$$\begin{aligned}
\|x(T)-x(t)\| &= \left\|\int_t^T v\right\| \leq \int_t^T \|v\| \leq \int_t^T \frac{1}{\boldsymbol{\nu}}\Lambda^{1/2} \\
&\leq \int_t^T \frac{1}{\boldsymbol{\nu}}\Lambda_0^{1/2} e^{-\frac{\boldsymbol{\nu} K}{(\sigma^2+B_0)^\beta}s} ds = \frac{1}{\boldsymbol{\nu}}\Lambda_0^{1/2}\left(-\frac{(\sigma^2+B_0)^\beta}{\boldsymbol{\nu} K} e^{-\frac{\boldsymbol{\nu} K}{(\sigma^2+B_0)^\beta}s}\right)\Bigg|_t^T \\
&= \frac{\Lambda_0^{1/2}(\sigma^2+B_0)^\beta}{\boldsymbol{\nu}^2 K}\left(e^{-\frac{\boldsymbol{\nu} K}{(\sigma^2+B_0)^\beta}t} - e^{-\frac{\boldsymbol{\nu} K}{(\sigma^2+B_0)^\beta}T}\right) \\
&\leq \frac{\Lambda_0^{1/2}(\sigma^2+B_0)^\beta}{\boldsymbol{\nu}^2 K} e^{-\frac{\boldsymbol{\nu} K}{(\sigma^2+B_0)^\beta}t}.
\end{aligned}$$

Since the last tend to zero with t and is independent of T we deduce that there exists $\widehat{x} \in X$ such that, $x \to \widehat{x}$.

(ii) Assume now $\beta = 1/2$. Then (10) takes the form

$$(\sigma^2 + \Gamma(t^*))\left(1 - \frac{2\Lambda_0}{(\boldsymbol{\nu} K)^2}\right) - (2\Gamma_0 + \sigma^2) \le 0$$

which implies that

$$\Gamma(t^*) \le B_0 := \frac{2\Gamma_0 + \sigma^2}{1 - \frac{2\Lambda_0}{(\boldsymbol{\nu} K)^2}} - \sigma^2.$$

Note that $B_0 > 0$ since $\Lambda_0 < \frac{(\boldsymbol{\nu} K)^2}{2}$. We now proceed as in case (i).

(iii) Assume finally $\beta > 1/2$ and let $\alpha = 2\beta$ so that $F(z) = z - \boldsymbol{a} z^\alpha - \boldsymbol{b}$. The derivative $F'(z) = 1 - \alpha \boldsymbol{a} z^{\alpha-1}$ has a unique zero at $z_* = \left(\frac{1}{\alpha \boldsymbol{a}}\right)^{\frac{1}{\alpha-1}}$ and

$$\begin{aligned}
F(z_*) &= \left(\frac{1}{\alpha \boldsymbol{a}}\right)^{\frac{1}{\alpha-1}} - \boldsymbol{a}\left(\frac{1}{\alpha \boldsymbol{a}}\right)^{\frac{\alpha}{\alpha-1}} - \boldsymbol{b} \\
&= \left(\frac{1}{\alpha}\right)^{\frac{1}{\alpha-1}}\left(\frac{1}{\boldsymbol{a}}\right)^{\frac{1}{\alpha-1}} - \left(\frac{1}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}\left(\frac{1}{\boldsymbol{a}}\right)^{\frac{1}{\alpha-1}} - \boldsymbol{b} \\
&= \left(\frac{1}{\boldsymbol{a}}\right)^{\frac{1}{\alpha-1}}\left[\left(\frac{1}{\alpha}\right)^{\frac{1}{\alpha-1}} - \left(\frac{1}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}\right] - \boldsymbol{b} \\
&\ge 0
\end{aligned}$$

the last by our hypothesis. Since $F(0) = -\boldsymbol{b} < 0$ and $F(z) \to -\infty$ when $z \to \infty$ we deduce that the shape of F is as follows:

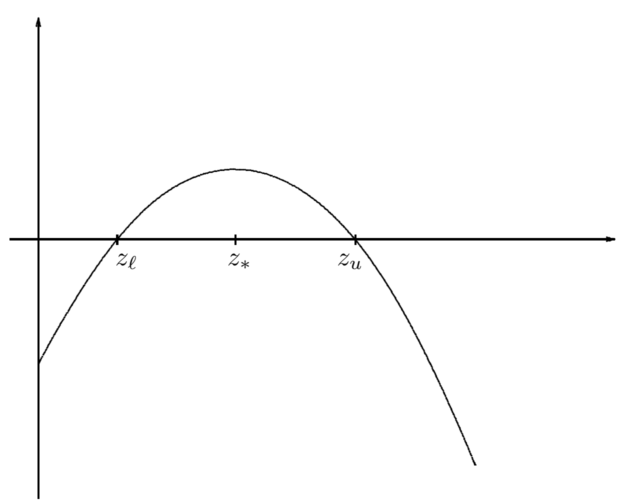

Figure 1

Even though t^* is not continuous as a function of t, the mapping $t \mapsto \Gamma(t^*) + \sigma^2$ is continuous and therefore, so is the mapping $t \mapsto F(\Gamma(t^*) + \sigma^2)$. This fact, together with (10), shows that, for all $t \geq 0$, $F(\Gamma(t^*) + \sigma^2) \leq 0$. In addition, when $t = 0$ we have $t^* = 0$ as well and

$$\begin{aligned} \Gamma_0 + \sigma^2 &\leq 2\Gamma_0 + \sigma^2 = \boldsymbol{b} \\ &< \left(\frac{1}{\boldsymbol{a}}\right)^{\frac{1}{\alpha-1}} \left[\left(\frac{1}{\alpha}\right)^{\frac{1}{\alpha-1}} - \left(\frac{1}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}\right] \\ &< \left(\frac{1}{\boldsymbol{a}}\right)^{\frac{1}{\alpha-1}} \left(\frac{1}{\alpha}\right)^{\frac{1}{\alpha-1}} \\ &= z_*. \end{aligned}$$

This implies that $\Gamma_0 + \sigma^2 < z_\ell$ (the latter being the smallest zero of F on $\mathbb{R}_+$, see the figure above) and the continuity of the map $t \mapsto \Gamma(t^*) + \sigma^2$ implies that, for all $t \geq 0$,

$$\Gamma(t^*) + \sigma^2 \leq z_\ell \leq z_*.$$

Therefore

$$\Gamma(t^*) \leq B_0 := \left(\frac{1}{\alpha \boldsymbol{a}}\right)^{\frac{1}{\alpha-1}} - \sigma^2 = \left(\frac{(\boldsymbol{\nu} K)^2}{2\alpha\Lambda_0}\right)^{\frac{1}{\alpha-1}} - \sigma^2.$$

We now proceed as in case (i). □

Remark 3 **(i)** In Theorem 2, the condition that $a_{ij} = \frac{K}{(\sigma^2 + \|x_i - x_j\|^2)^\beta}$ may be relaxed to $a_{ij} \geq \frac{K}{(\sigma^2 + \|x_i - x_j\|^2)^\beta}$.

(ii) The bound $\beta < 1/2$ for unconditional convergence in Theorem 2 is essentially sharp. We will indicate this in Remark 4 below by studying the special case of a flock with two birds flying on a line.

4 A flock of two birds

We give here a more detailed analysis of the case of two birds flying on a line (i.e., we take $\mathbb{R}$ instead of $\mathbb{E}^3$ for both positions and velocities).

We define $\boldsymbol{x} = x_1 - x_2$ and $\boldsymbol{v} = v_1 - v_2$ and assume that the state $(\boldsymbol{x}, \boldsymbol{v})$ of the pair satisfies the system of ODE's

$$\begin{aligned} \boldsymbol{x}' &= \boldsymbol{v} \\ \boldsymbol{v}' &= -\frac{\boldsymbol{v}}{(1+\boldsymbol{x})^\alpha}. \end{aligned} \tag{12}$$

This is not exactly (7) but it is easier to dealt with and, we will see below, it is close to this system.

The arguments used in the preceding section show that when $\alpha < 1$, for all initial $\boldsymbol{x}_0$ and $\boldsymbol{v}_0$, we have that $\boldsymbol{x}$ is bounded and $\boldsymbol{v} \to 0$ when $t \to \infty$. The next proposition gives conditions on $\boldsymbol{x}_0$ and $\boldsymbol{v}_0$ for such a convergence to hold when $\alpha > 1$.

Proposition 5 *Let $\alpha > 1$. Assume that $\boldsymbol{x}_0 > 0$ and $\boldsymbol{v}_0 > 0$ and that*

$$\boldsymbol{x}_0 < \widehat{\boldsymbol{x}_0} := \left(\frac{\alpha - 1}{\boldsymbol{v}_0}\right)^{\frac{1}{\alpha-1}} + 1.$$

Then $\boldsymbol{x}$ is bounded and increasing. In addition, when $t \to \infty$, $\boldsymbol{v}(t) \to 0$ and

$$\boldsymbol{x}(t) \to \left(\frac{\alpha - 1}{\boldsymbol{v}_0 - \frac{\alpha-1}{(1+\boldsymbol{x}_0)^{\alpha-1}}}\right)^{\frac{1}{\alpha-1}} + 1.$$

Proof. It follows from the system (12) that, for all $t \geq 0$,

$$\int_0^t \boldsymbol{v}' = \int_0^t \frac{\boldsymbol{x}'}{(1+\boldsymbol{x})^\alpha}$$

and therefore, integrating both sides between 0 and t, that

$$\boldsymbol{v}(t) - \boldsymbol{v}_0 = \frac{\alpha - 1}{(1 + \boldsymbol{x}(t))^{\alpha-1}} - \frac{\alpha - 1}{(1 + \boldsymbol{x}_0)^{\alpha-1}}$$

or yet, that

$$\boldsymbol{v}(t) = \frac{\alpha - 1}{(1 + \boldsymbol{x}(t))^{\alpha-1}} - \gamma_0 \tag{13}$$

where $\gamma_0 > 0$ since $\boldsymbol{x}_0 < \widehat{\boldsymbol{x}_0}$.

If, for some t_*, $\boldsymbol{v}(t_*) = 0$ then $\boldsymbol{v}'(t_*) = 0$. But then the pair $(\widetilde{x}, \widetilde{v})$ defined by $\widetilde{x}_1(t) = x_1(t_*)$, $\widetilde{x}_2(t) = x_2(t_*)$ and $\widetilde{v}(t) = 0$, for all $t \geq 0$, is a solution of (12) satisfying the conditions $\widetilde{\boldsymbol{x}}(t_*) = \boldsymbol{x}(t_*)$ and $\widetilde{\boldsymbol{v}} = 0$. By the unicity of the solutions of (12) it follows that $\widetilde{\boldsymbol{v}} = \boldsymbol{v}$ and hence that $\widetilde{\boldsymbol{v}}_0 = 0$ in contradiction with our assumptions. We conclude that $\boldsymbol{v}(t) > 0$ for all $t \geq 0$. But then

$$0 < \boldsymbol{v}(t) = \frac{\alpha - 1}{(1 + \boldsymbol{x}(t))^{\alpha-1}} - \gamma_0$$

implies that

$$\boldsymbol{x}(t) < \left(\frac{\alpha - 1}{\gamma_0}\right)^{\frac{1}{\alpha-1}} + 1.$$

Thus, $\boldsymbol{x}$ remains bounded on $\mathbb{R}_+$. Furthermore $\boldsymbol{x}$ is increasing since $\boldsymbol{v} > 0$. This implies that there exists $\boldsymbol{x}_* > 0$ such that $\boldsymbol{x}(t) \to \boldsymbol{x}_*$ and $\boldsymbol{x}'(t) \to 0$ when $t \to \infty$. It follows from $\boldsymbol{x}' = \boldsymbol{v}$ and (13) that $\boldsymbol{x}_*$ is as claimed. □

Remark 4 It follows from the proof of Proposition 5 that, for all $\alpha > 1$, $\boldsymbol{v}$ fails to converge if $\boldsymbol{x} \geq \widehat{\boldsymbol{x}_0}$. Also, since

$$\frac{1}{(1+\boldsymbol{x})^\alpha} \leq \frac{1}{(1+\boldsymbol{x}^2)^\beta} \leq 2\frac{\sqrt{2}}{(1+\boldsymbol{x})^\alpha}$$

the system (7) is tightly bounded in between two versions of (12) differing only by a constant factor. This indicates that convergence may fail as well in (7) for $\alpha > 1$.

5 Convergence in discrete time

We now focus on discrete time. The model is thus (4). A motivation to consider discrete time is that we want to derive (possibly a small variation of) our model from a mechanism based on exchanges of signals. The techniques to do so, learning theory, are better adapted to discrete time. Also, we want our model to include noisy environments and this issue becomes more technically involved in continuous time.

We assume as before that there are constants $K, \sigma > 0$ and $\beta \geq 0$ such that

$$a_{ij} = \frac{K}{(\sigma^2 + \|x_i - x_j\|^2)^\beta}.$$

Note that, by Proposition 2, this implies that $\phi_x > 0$ for all $x \in X$. This, in turn, shows that L_x is a self-adjoint, positive definite linear map, whose smallest eigenvalue is ϕ_x. We denote by $\kappa(x)$ its condition number, i.e.,

$$\kappa(x) = \frac{\|L_x\|}{\phi_x}.$$

Lemma 3 *For all $x \in X$,*

$$\kappa(x) \leq \frac{k(\sigma^2 + \Gamma_x)^\beta}{\boldsymbol{\nu}\sigma^{2\beta}}.$$

Proof. By Proposition 2

$$\phi_x \geq \boldsymbol{\nu} \min_{i,j} a_{ij} \geq \boldsymbol{\nu}\frac{K}{(\sigma^2 + \Gamma_x)^\beta}.$$

Also, $\|L_x\| \leq \frac{Kk}{\sigma^{2\beta}}$ since all of its entries are bounded by $\frac{K}{\sigma^{2\beta}}$. □

In the following we fix a solution (x, v) of (4). At a time $t \in \mathbb{N}$, $x(t)$ and $v(t)$ are elements in X and V, respectively. The meaning of expressions like ϕ_t, L_t, $\kappa(t)$, or $\Gamma(t)$ is as described in Section 3.

Proposition 6 *For all $t \in \mathbb{N}$, $\|v(t+1)\| \leq \left(1 - \frac{1}{\kappa(t)}\right)\|v(t)\|$. In particular, $\|v\|$ is decreasing as a function of t.*

PROOF. The linear map $\mathrm{Id} - \frac{1}{\|L_t\|} L_t$ is self-adjoint and its eigenvalues are in the interval $[0, 1)$. In addition, its largest eigenvalue is $1 - \frac{\phi_t}{\|L_t\|}$. Therefore

$$\|v(t+1)\| = \left\| \left(\mathrm{Id} - \frac{1}{\|L_t\|} L_t \right) v(t) \right\| \leq \left(1 - \frac{\phi_t}{\|L_t\|} \right) \|v(t)\|.$$

□

Corollary 1 *For all $t \in \mathbb{N}$, $\|v(t)\| \leq \prod_{i=0}^{t-1} \left(1 - \frac{1}{\kappa(i)} \right) \|v(0)\|$.* □

Theorem 3 *Assume that there are constants $K, \sigma > 0$ and $\beta \geq 0$ such that*

$$a_{ij} = \frac{K}{(\sigma^2 + \|x_i - x_j\|^2)^\beta}.$$

Assume also that one of the three following hypothesis hold:

(i) $\beta < 1/2$,

(ii) $\beta = 1/2$ *and* $\|v(0)\| \leq \frac{\boldsymbol{\nu} \sigma^{2\beta}}{k \overline{\boldsymbol{\nu}}^{1/2} \Delta t}$,

(iii) $\beta > 1/2$ *and*

$$\left(\frac{1}{\boldsymbol{a}} \right)^{\frac{2}{\alpha-1}} \left[\left(\frac{1}{\alpha} \right)^{\frac{2}{\alpha-1}} - \left(\frac{1}{\alpha} \right)^{\frac{\alpha+1}{\alpha-1}} \right] > \overline{\boldsymbol{\nu}} \left(V_0^2 + 2V_0((\alpha \boldsymbol{a})^{-\frac{2}{\alpha-1}} - \sigma^2) \overline{\boldsymbol{\nu}}^{-1/2} \right) + \boldsymbol{b}.$$

Here $\alpha = 2\beta$, $V_0 := \Delta t \|v(0)\|$,

$$\boldsymbol{a} = \frac{k \overline{\boldsymbol{\nu}}^{1/2}}{\boldsymbol{\nu} \sigma^{2\beta}} V_0, \qquad \text{and} \qquad \boldsymbol{b} = \overline{\boldsymbol{\nu}}^{1/2} \|x(0)\| + \sigma.$$

Then there exists a constant B_0 (independent of t, made explicit in the proof of each of the three cases) such that $\|x(t)\| \leq B_0$ for all $t \in \mathbb{N}$. In addition, $\|v(t)\| \to 0$ when $t \to \infty$. Finally, there exists $\widehat{x} \in X$ such that $x(t) \to \widehat{x}$ when $t \to \infty$.

PROOF. For $t \in \mathbb{N}$ let t^* be the point maximizing $\|x\|$ in $\{0, 1, \ldots, t\}$. Then, by Lemma 3, for $i \in \{0, 1, \ldots, t\}$,

$$\kappa(i) \leq \frac{k(\sigma^2 + \Gamma(i))^\beta}{\boldsymbol{\nu} \sigma^{2\beta}} \leq \frac{k(\sigma^2 + \overline{\boldsymbol{\nu}} \|x(i)\|^2)^\beta}{\boldsymbol{\nu} \sigma^{2\beta}} \leq H(t^*) := \frac{k(\sigma^2 + \overline{\boldsymbol{\nu}} \|x(t^*)\|^2)^\beta}{\boldsymbol{\nu} \sigma^{2\beta}}.$$

Using Corollary 1 we obtain

$$\begin{aligned}
\|x(t)\| &\leq \|x(0)\| + \sum_{j=0}^{t-1} \|x(j+1) - x(j)\| \\
&\leq \|x(0)\| + \Delta t \sum_{j=0}^{t-1} \|v(j)\| \\
&\leq \|x(0)\| + \Delta t \left(\|v(0)\| + \sum_{j=1}^{t-1} \|v(j)\| \right) \\
&\leq \|x(0)\| + \Delta t \left(\|v(0)\| + \sum_{j=1}^{t-1} \prod_{i=1}^{j} \left(1 - \frac{1}{\kappa(i)}\right) \|v(0)\| \right) \\
&\leq \|x(0)\| + \Delta t \sum_{j=0}^{t-1} \left(1 - \frac{1}{H(t^*)}\right)^j \|v(0)\| \\
&\leq \|x(0)\| + \Delta t H(t^*) \|v(0)\| \\
&= \|x(0)\| + \Delta t \frac{k(\sigma^2 + \boldsymbol{\overline{\nu}} \|x(t^*)\|^2)^\beta}{\boldsymbol{\nu} \sigma^{2\beta}} \|v(0)\|.
\end{aligned}$$

Since t^* maximizes $\|x\|$ in $\{0, 1, \ldots, t\}$ it also does so in $\{0, 1, \ldots, t^*\}$. For $t = t^*$, the inequality above takes then the following equivalent form

$$\sigma + \boldsymbol{\overline{\nu}}^{1/2} \|x(t^*)\| \leq \left(\boldsymbol{\overline{\nu}}^{1/2} \|x(0)\| + \sigma \right) + \boldsymbol{\overline{\nu}}^{1/2} \Delta t \frac{k(\sigma^2 + \boldsymbol{\overline{\nu}} \|x(t^*)\|^2)^\beta}{\boldsymbol{\nu} \sigma^{2\beta}} \|v(0)\|$$

which implies

$$(\sigma^2 + \boldsymbol{\overline{\nu}} \|x(t^*)\|^2)^{1/2} \leq \left(\boldsymbol{\overline{\nu}}^{1/2} \|x(0)\| + \sigma \right) + \boldsymbol{\overline{\nu}}^{1/2} \Delta t \frac{k(\sigma^2 + \boldsymbol{\overline{\nu}} \|x(t^*)\|^2)^\beta}{\boldsymbol{\nu} \sigma^{2\beta}} \|v(0)\|. \quad (14)$$

Let $z = (\sigma^2 + \boldsymbol{\overline{\nu}} \|x(t^*)\|^2)^{1/2}$,

$$\boldsymbol{a} = \frac{k \boldsymbol{\overline{\nu}}^{1/2} \Delta t}{\boldsymbol{\nu} \sigma^{2\beta}} \|v(0)\|, \qquad \text{and} \qquad \boldsymbol{b} = \boldsymbol{\overline{\nu}}^{1/2} \|x(0)\| + \sigma.$$

Then (14) can be rewritten as $F(z) \leq 0$ with

$$F(z) = z - \boldsymbol{a} z^{2\beta} - \boldsymbol{b}.$$

(i) Assume $\beta < 1/2$. By Lemma 2, $F(z) \leq 0$ implies that $(\sigma^2 + \boldsymbol{\overline{\nu}} \|x(t^*)\|^2) \leq U_0^2$ with

$$U_0 = \max \left\{ \left(\frac{2k \boldsymbol{\overline{\nu}}^{1/2} \Delta t}{\boldsymbol{\nu} \sigma^{2\beta}} \|v(0)\| \right)^{\frac{1}{1-2\beta}}, 2 \left(\boldsymbol{\overline{\nu}}^{1/2} \|x(0)\| + \sigma \right) \right\}.$$

Since U_0 is independent of t we deduce that, for all $t \in \mathbb{N}$,

$$\|x(t)\|^2 \leq B_0^2 := \frac{U_0^2 - \sigma^2}{\overline{\boldsymbol{\nu}}}$$

and therefore, by Lemma 3,

$$\kappa(t) \leq \frac{k(\sigma^2 + \Gamma(t))^\beta}{\boldsymbol{\nu}\sigma^{2\beta}} \leq \frac{k(\sigma^2 + \overline{\boldsymbol{\nu}}\|x(t)\|^2)^\beta}{\boldsymbol{\nu}\sigma^{2\beta}} \leq \kappa_* := \frac{kU_0^{2\beta}}{\boldsymbol{\nu}\sigma^{2\beta}}.$$

By Corollary 1, for $t \in \mathbb{N}$,

$$\|v(t)\| \leq \prod_{i=0}^{t-1}\left(1 - \frac{1}{\kappa(i)}\right)\|v(0)\| \leq \left(1 - \frac{1}{\kappa_*}\right)^t \|v(0)\|$$

and this expression tends to zero when $t \to \infty$.

Finally, for $T > t$, reasoning as above, we have

$$\begin{aligned} \|x(T) - x(t)\| &\leq \sum_{j=t}^{T-1} \|x(j+1) - x(j)\| \leq \Delta t \sum_{j=t}^{T-1} \|v(j)\| \\ &\leq \Delta t \sum_{j=t}^{T-1} \left(1 - \frac{1}{\kappa_*}\right)^j \|v(t)\| \leq \Delta t \kappa_* \|v(t)\|. \end{aligned}$$

Since $\|v(t)\|$ tends to zero, we deduce that $\{x(t)\}_{t\in\mathbb{N}}$ is a Cauchy sequence and there exists $\widehat{x} \in X$ such that $x(t) \to \widehat{x}$.

(ii) Assume now $\beta = 1/2$. Then (14) takes the form

$$(\sigma^2 + \overline{\boldsymbol{\nu}}\|x(t^*)\|^2)^{1/2}\left(1 - \frac{k\overline{\boldsymbol{\nu}}^{1/2}\Delta t}{\boldsymbol{\nu}\sigma^{2\beta}}\|v(0)\|\right) - \left(\overline{\boldsymbol{\nu}}^{1/2}\|x(0)\| + \sigma\right) \leq 0$$

which implies that

$$\|x(t^*)\|^2 \leq B_0 := \frac{1}{\overline{\boldsymbol{\nu}}}\left(\left(\frac{\overline{\boldsymbol{\nu}}^{1/2}\|x(0)\| + \sigma}{1 - \frac{k\overline{\boldsymbol{\nu}}^{1/2}\Delta t}{\boldsymbol{\nu}\sigma^{2\beta}}\|v(0)\|}\right)^2 - \sigma^2\right)$$

which is positive since, by hypothesis, $0 < 1 - \frac{k\overline{\boldsymbol{\nu}}^{1/2}\Delta t}{\boldsymbol{\nu}\sigma^{2\beta}}\|v(0)\| \leq 1$. We now proceed as in case (i).

(iii) Assume finally $\beta > 1/2$. Letting $\alpha = 2\beta$ as in the proof of Theorem 2, the arguments therein show that the derivative $F'(z) = 1 - \alpha\boldsymbol{a}z^{\alpha-1}$ has a unique zero at $z_* = \left(\frac{1}{\alpha\boldsymbol{a}}\right)^{\frac{1}{\alpha-1}}$ and $F(z_*) = \left(\frac{1}{\boldsymbol{a}}\right)^{\frac{1}{\alpha-1}}\left[\left(\frac{1}{\alpha}\right)^{\frac{1}{\alpha-1}} - \left(\frac{1}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}\right] - \boldsymbol{b}$. Our hypothesis then implies that $F(z_*) \geq 0$. This shows that the graph of F is as in Figure 1.

For $t \in \mathbb{N}$ let $z(t) = (\sigma^2 + \overline{\boldsymbol{\nu}}\|x(t^*)\|^2)^{1/2}$. When $t = 0$ we have $t^* = 0$ as well and

$$z(0) \le \overline{\boldsymbol{\nu}}^{1/2}\|x(0)\| + \sigma = \boldsymbol{b} < \left(\frac{1}{\boldsymbol{a}}\right)^{\frac{1}{\alpha-1}} \left(\frac{1}{\alpha}\right)^{\frac{1}{\alpha-1}} = z_*.$$

This actually implies that $z(0) \le z_\ell$. Assume that there exists $t \in \mathbb{N}$ such that $z(t) \ge z_u$ and let T be the first such t. Then $T = T^* \ge 1$ and, for all $t < T$

$$(\sigma^2 + \overline{\boldsymbol{\nu}}\|x(t)\|^2)^{1/2} \le z(T-1) \le z_\ell.$$

This shows that, for all $t < T$,

$$\|x(t)\| \le \left(\frac{z_\ell^2 - \sigma^2}{\overline{\boldsymbol{\nu}}}\right)^{1/2} \le B_0 := \left(\frac{z_*^2 - \sigma^2}{\overline{\boldsymbol{\nu}}}\right)^{1/2}.$$

In particular,

$$\|x(T-1)\|^2 \le \frac{z_\ell^2 - \sigma^2}{\overline{\boldsymbol{\nu}}}$$

For T instead, we have

$$\|x(T)\|^2 \ge \frac{z_u^2 - \sigma^2}{\overline{\boldsymbol{\nu}}}.$$

This implies

$$\|x(T)\|^2 - \|x(T-1)\|^2 \ge \frac{z_u^2 - z_\ell^2}{\overline{\boldsymbol{\nu}}} \ge \frac{z_*^2 - z_\ell^2}{\overline{\boldsymbol{\nu}}} \ge \frac{(z_* - z_\ell)z_*}{\overline{\boldsymbol{\nu}}}. \tag{15}$$

From the intermediate value theorem, there is $\xi \in [z_\ell, z_*]$ such that $F(z_*) = F'(\xi)(z_* - z_\ell)$. But $F'(\xi) \ge 0$ and $F'(\xi) = 1 - a\alpha\xi^{\alpha-1} \le 1$. Therefore,

$$z_* - z_\ell \ge F(z_*)$$

and it follows from (15) that

$$\|x(T)\|^2 - \|x(T-1)\|^2 \ge \frac{z_* F(z_*)}{\overline{\boldsymbol{\nu}}}. \tag{16}$$

But

$$\begin{aligned} \|x(T)\| - \|x(T-1)\| &\le \|x(T) - x(T-1)\| \\ &= \Delta t\|v(T-1)\| \\ &\le \Delta t\|v(0)\| \end{aligned}$$

the last since $\|v\|$ is decreasing. Therefore,

$$\begin{aligned} \|x(T)\|^2 - \|x(T-1)\|^2 &\le (\Delta t)^2\|v(0)\|^2 + 2\Delta t\|v(0)\|\|x(T-1)\| \\ &\le (\Delta t)^2\|v(0)\|^2 + 2\Delta t\|v(0)\|B_0. \end{aligned}$$

Putting this inequality together with (16) shows that

$$z_* F(z_*) \le \overline{\boldsymbol{\nu}} \left((\Delta t)^2 \|v(0)\|^2 + 2\Delta t \|v(0)\| B_0\right)$$

or equivalently,

$$\left(\frac{1}{\boldsymbol{a}}\right)^{\frac{2}{\alpha-1}} \left[\left(\frac{1}{\alpha}\right)^{\frac{2}{\alpha-1}} - \left(\frac{1}{\alpha}\right)^{\frac{\alpha+1}{\alpha-1}}\right] - \boldsymbol{b} \le \overline{\boldsymbol{\nu}} \left((\Delta t)^2 \|v(0)\|^2 + 2\Delta t \|v(0)\| B_0\right)$$

which contradicts our hypothesis.

We conclude that, for all $t \in \mathbb{N}$, $z(t) \le z_\ell$ and hence, $\|x(t)\| \le B_0$. We now proceed as in case (i). □

Remark 5 **(i)** In the system (4) we could have replaced $\|L_x\|$ by $\frac{Kk}{\sigma^{2\beta}}$ and Theorem 3 would hold as well (with the same B_0 proved therein). Interpreting the latter choice would require less computational capabilities on the birds.

(ii) In the proof of Theorem 3 we could have used the bounds for $\overline{\boldsymbol{\nu}}$ and $\boldsymbol{\nu}$ exhibited in Lemma 1 and, in case (iii) the trivial bound $(\alpha \boldsymbol{a})^{-\frac{2}{\alpha-1}} - \sigma^2 \le (\alpha \boldsymbol{a})^{-\frac{2}{\alpha-1}}$. Recall, $V_0 := \Delta t \|v(0)\|$. Denoting as well $X_0 := \|x(0)\|$ and

$$g(\alpha) := \left[\left(\frac{1}{\alpha}\right)^{\frac{2}{\alpha-1}} - \left(\frac{1}{\alpha}\right)^{\frac{\alpha+1}{\alpha-1}}\right]$$

the sufficiency condition for convergence in case (iii) becomes

$$g(\alpha)\left(\frac{\sigma^\alpha}{3\sqrt{2}k^3}\right)^{\frac{1}{\alpha-1}} V_0^{-\frac{1}{\alpha-1}} \ge \sqrt{2}kX_0 + \sigma + 2\left(k^2 V_0^2 + V_0^{\frac{\alpha-2}{\alpha-1}} \left(\frac{\sigma^\alpha}{3\sqrt{2}k^3}\right)^{\frac{1}{\alpha-1}}\right).$$

It is apparent from the expression above that this condition is satisfied when V_0 is sufficiently small. It is also apparent that the larger k is, the smaller V_0 needs to be to satisfy the condition.

We note also that, for $\alpha > 1$, we have $0 < g(\alpha) < 1$ and that $g(a) \to 0$ when $\alpha \to 1$.

6 Language evolution

We now consider a linguistic population with k agents evolving with time. At time t, the state of the population is given by $(x(t), f(t)) \in (\mathbb{E}^3)^k \times \mathcal{H}^k$. Here $\mathbb{E}^3$ is interpreted as the space of positions and $\mathcal{H}$ as the space of languages of [2]. Thus, unlike the development in Section 3, the functions x and f do not belong to the same space.

We model the evolution of the population with the system of differential equations

$$\begin{aligned} x' &= -L_f x \\ f' &= -L_x f. \end{aligned}$$

Again, L_x is the Laplacian of the matrix A_x given by $a_{ij} = \eta_X(\|x_i - x_j\|^2)$ for some function $\eta_X : \mathbb{R}_+ \to \mathbb{R}_+$. Similarly with L_f for some function $\eta_{\mathcal{H}} : \mathbb{R}_+ \to \mathbb{R}_+$. The distance between languages in $\mathcal{H}$ is defined as in [2].

A rationale for this model could be the following. Agents tend to move towards other agents using languages close to theirs (and therefore, communicating better). Hence, the first equation. Also, languages evolve by the influence from other agents' languages and this influence decrease with distance (for instance, because of a decrease in the frequency of linguistic encounters). Hence, the second equation.

Theorem 4 *Let $\eta_X, \eta_{\mathcal{H}} : \mathbb{R}_+ \to (0,\infty)$ be non-increasing. Then, when $t \to \infty$, the state (x, f) tends to a point in the diagonal of $(\mathbb{E}^3 \times \mathcal{H})^k$.*

Proof. We use the ideas and notations from Section 3. Reasoning as in Proposition 3 we obtain, for all $t > 0$,

$$\Lambda'(f(t)) \le -2\phi_t \Lambda(f(t)) \qquad \text{and} \qquad \Gamma'(x(t)) \le -2\phi_{f(t)}\Gamma(x(t)).$$

This shows that both Λ and Γ are decreasing and satisfy

$$\Lambda(f(t)) \le \Lambda_0 e^{-2\int_0^t \phi_\tau d\tau} \qquad \text{and} \qquad \Gamma(f(t)) \le \Gamma_0 e^{-2\int_0^t \phi_{f(\tau)} d\tau}.$$

But since both η and Γ are non-increasing, by Proposition 2,

$$\phi_\tau \ge \boldsymbol{\nu}\eta(\max_{i \ne j} \|x_i(\tau) - x_j(\tau)\|^2) \ge \boldsymbol{\nu}\eta(\Gamma_\tau) \ge \boldsymbol{\nu}\eta(\Gamma_0).$$

Thus,

$$\int_0^t \phi_\tau d\tau \ge t\boldsymbol{\nu}\eta(\Gamma_0)$$

and $\Lambda(f(t)) \le \Lambda_0 e^{-2t\boldsymbol{\nu}\eta(\Gamma_0)}$. This shows the convergence to 0 of $\Lambda(t)$. That of $\Gamma(t)$ is similar. □

Remark 6 **(i)** We interpret the convergence of $x(t)$ to a fixed $\boldsymbol{x} \in \Delta_X$ as the formation of a tribe and the convergence of $f(t)$ to a fixed $\boldsymbol{f} \in \Delta_{\mathcal{H}}$ as the emergence of a common language as in Examples 2 and 3 of [2]. The first such example is taken from [5] were models are proposed (and studied via simulation) for the origins of language. The second, is a modification of it proposed in [2] for the emergence of common vowel sounds.

(ii) The assumption of symmetry is plausible in contexts where (unlike the Mother/Baby case discussed in [2, Example 4]) there are no leaders in the liguistic population.

(iii) Detailed learning mechanisms could be introduced by first deriving a result akin to Proposition 4 for discrete time and then follow [2].

(iv) We have not used any argument as those in the proof of Proposition 4. These arguments involved expressions like $\langle x, f \rangle$ which, in the situation at hand, would be meaningless.

References

[1] F. Cucker and S. Smale. Best choices for regularization parameters in learning theory. *Found. Comput. Math.*, 2:413–428, 2002.

[2] F. Cucker, S. Smale, and D.-X. Zhou. Modeling language evolution. *Found. Comput. Math.*, 4:315–343, 2004.

[3] G. Flierl, D. Grünbaum, S. Levin, and D. Olson. From individuals to aggregations: the interplay between behavior and physics. *J. Theor. Biol.*, 196:397–454, 1999.

[4] A. Jadbabaie, J. Lin, and A.S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Trans. on Autom. Control*, 48:988–1001, 2003.

[5] J. Ke, J. Minett, C.-P. Au, and W.S.-Y. Wang. Self-organization and selection in the emergence of vocabulary. *Complexity*, 7:41–54, 2002.

[6] B. Mohar. The Laplacian spectrum of graphs. In Y. Alavi, G. Chartrand, O.R. Oellermann, and A.J. Schwenk, editors, *Graph Theory, Combinatorics, and Applications*, volume 2, pages 871–898. John Wiley & Sons, 1991.

[7] T. Vicsek, A. Czirók, E. Ben-Jacob, and O. Shochet. Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Letters*, 75:1226–1229, 1995.