## Citation for published version:

Bacik, K, Bacik, B \& Rogers, T 2023, 'Lane nucleation in complex active flows', Science.
https://doi.org/10.1126/science.add8091

DOI:
10.1126/science.add8091

Publication date:
2023

Document Version
Peer reviewed version

Link to publication

Publisher Rights
CC BY
This is the authors' accepted manuscript of the published article: Bacik, Karol, Bogdan Bacik, and Tim Rogers.
"Lane nucleation in complex active flows". Science. 2023. https://doi.org/10.1126/science.add8091

## University of Bath

## Alternative formats

If you require this document in an alternative format, please contact: openaccess@bath.ac.uk

[^0]
## Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Lane nucleation in complex active flows 

Karol A. Bacik, ${ }^{1}$ Bogdan S. Bacik, ${ }^{2}$ Tim Rogers, ${ }^{1 *}$<br>${ }^{1}$ Centre for Networks and Collective Behaviour, Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, United Kingdom<br>${ }^{2}$ Department of Human Motor Behavior, Institute of Sport Sciences, Academy of Physical Education in Katowice, ul. Mikołowska 72c, 40-065 Katowice, Poland<br>*To whom correspondence should be addressed; E-mail: t.c.rogers@bath.ac.uk.

One sentence summary An improved mathematical theory reveals universal patterns in active binary flows that are confirmed in experiments on human crowds.

Laning is a paradigmatic example of spontaneous organization in active twocomponent flows which has been observed in diverse contexts including pedestrian traffic, driven colloids, complex plasmas, and molecular transport. We introduce a kinetic theory which elucidates the physical origins of laning and quantifies the propensity for lane nucleation in a given physical system. Our theory is valid in the low-density regime and it makes different predictions about situations in which lanes may form that are not parallel with the direction of flow. We report on experiments with human crowds that verify two remarkable consequences of this phenomenon: tilting lanes under broken chiral symmetry and lane nucleation along elliptic, parabolic and hyperbolic curves in the presence of sources or sinks.

When two groups of pedestrians move past each other in opposite directions, the crowd
can spontaneously segregate into contra-flowing lanes (1-9). Moving in lanes reduces the risk of a collision and increases the efficiency of motion, but lane formation does not require conscious optimisation effort. Indeed, the spontaneous emergence of lanes is also observed for driven colloids $(10,11)$ and complex plasmas $(12,13)$, and this phenomenon is hypothesized as a key to facilitating bidirectional intercellular transport in elongated domains such as axons (14). Provided the crowd density is low enough to avoid jamming, lane formation is also robustly reproduced by cellular automata $(15,16)$, as well as both lattice $(17,18)$ and off-lattice (19-28) agent-based (AB) simulations. The ubiquity of lane formation across a broad class of physical systems as well as the fact that the emergence of lanes in numerical simulations is largely independent of their microscopic implementation suggests the existence of a universal mechanism underpinning lane formation. Nevertheless, notwithstanding substantial progress over many years, the scientific community has not yet reached consensus about the physical origin of lane formation (19, 22-24, 29).

The spontaneous emergence of an apparently regular pattern suggests a continuum fieldtheory approach to crowd modelling may be fruitful, where we might expect to observe a hydrodynamic instability around the homogenous mixed state. Calculations have previously been performed using $a d$ hoc partial differential equation (PDE) models based on heuristic reasoning, identifying at least three different plausible mechanisms of lane nucleation: gradient-induced drift $(19,22,29)$, undulation-induced diffusion $(11,23)$ and density-induced drift when chiral symmetry is broken (30). To settle this debate we need to make a direct connection between the microscopic interaction rules of a given system and the emergent PDE at the macroscale. Despite some notable advances for one-component systems (31,32), for active two-component flows this has only been performed for the repulsive torque model (33), as well as in the case of extremely soft and dense particles (24), which do not exhibit a laning instability.

We introduce a theoretical approach using temporal coarse-graining akin to the Einstein's
kinetic theory of Brownian motion (34). Our averaging scheme is valid in the case of nonjamming mixtures of hard particles, where the dynamics is dominated by pairwise interactions, which is a good approximation for typical pedestrian flows as well as dilute colloids. We recover and unify in a systematic manner the fundamental insights of Vicsek and Helbing (22) as well as Vissers et al. (11) and Klymko et al. (23), by showing that undulation-induced drift and diffusion can both contribute to lane nucleation. We also demonstrate that diffusive processes suppress the formation of very narrow lanes, thereby providing a dynamical selection mechanism favouring the nucleation of lanes of a particular width. We provide explicit formulae for the propensity of a given system to nucleate lanes and we present a simple approximate rule that lanes emerge at a rate proportional to the product of agent speed, density, and an effective parameter related to the average magnitude of lateral displacement in agent-agent collisions.

The transport equations for agent densities we derive are also readily extended to the case of systems exhibiting broken chiral symmetry, e.g. pedestrians with a preference for turning right (or left) when dodging each other (30,35,30), rotating robots (37), or non-spherical driven particles. We identify an additional density-induced mechanism (previously suggested in e.g. (30)) which is the dominant driver of laning in the case of strongly asymmetric interactions, and in even mildly chiral system is responsible for producing lanes which are no longer parallel to the direction of motion. This interesting prediction is confirmed in an experiment with human crowds.

Finally, we show that lanes which are not parallel to the direction of motion also occur in situations when the trajectories of the two agent types cross at an angle. This observation encompasses a range of more complex scenarios with point targets and sources, which are particularly germane to pedestrian flows - for example entrances and exits into open spaces. We demonstrate both theoretically and experimentally that in this case the laning instability is triggered along curves describing conic sections with different experimental scenarios achieving
parabolic, elliptic, or hyperbolic lane nucleation.
Our theory is designed to give insight into the fundamental mechanisms of lane formation in dilute two-component flows. Although we have tested some novel predictions of the theory in experiments with human crowds, it should be emphasised that our model does not incorporate a range of system-specific details which may have important effects in different settings. These include hydrodynamic effects in colloidal suspensions (25) or complex stimulus-response mechanisms (38), anticipation (39) and gait mechanics (40) in pedestrians. We also do not address multi-body interactions which arise in high-density regimes, so effects such as jamming (23, 41, 42) (discussed in more details in (43), Section 4B) are not captured, and we have not examined the role of boundaries which are important in many pedestrian scenarios $(3,44)$. Such features could be incorporated into more specialized versions of our model with a specific setting in mind, but here we present only the most simple form in order to elucidate the underpinning mathematics and show that many salient features of lane nucleation can be successfully explained by a simple and interpretable theory.

Kinetic description Consider two groups of agents (labelled '+' and '-') moving in opposite directions at equal speeds in a bounded two-dimensional domain. An agent's intended motion across the domain is interrupted by encounters with agents of the opposite type; we seek to compute the statistics of many such events accumulated over a period of time $\Delta t$. Tracking a focal agent $i$ of type ' + ' we obtain the approximate expression

$$
\begin{equation*}
\Delta \mathbf{r}_{i}^{+}=\mathbf{v}^{+} \Delta t+\sum_{\text {collisions }} \mathbf{G}\left(\mathbf{r}_{i}^{+}-\mathbf{r}_{j}^{-}\right) . \tag{1}
\end{equation*}
$$

The first term on the right hand side here describes the intended motion of an agent in free space, while the sum ranges over all agents $j$ of the opposite type who will collide with agent $i$ in the time period of interest. The interaction events are assumed to be much shorter than the characteristic time interval $\Delta t$, which is why we refer to them as collisions. In some systems
(e.g. pedestrians) the agents do not physically collide, but perform collision avoidance maneuvers; the distinction is not important for our theory as both effectively modify the position of interacting agents. The direction and magnitude of this perturbation is specified by the collisional operator G. We choose our coordinate system so that the preferred velocities are equal in magnitude and aligned with the $y$-axis $\left(\mathbf{v}^{ \pm}= \pm v \mathbf{e}_{y}\right)$, and so $\mathbf{G}$ can be treated as a function only of the lateral displacement $x=\left(\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{\mp}\right) \cdot \mathbf{e}_{x}($ Fig. 1C $)$.

The function $\mathbf{G}$ contains all necessary information about the microscopic interactions of the model and may be specified ab initio (for example in a simulation of hard spheres) or derived from data. For deterministic dynamics it is a univalued function, but in thermal or stochastic systems it should be conceptualised as a random variable (44,45). Data extracted from an experiment on human crowds is given as an example of what the empirical lateral displacement $G_{x}(x)$ may look like in a pedestrian system (Fig. 1D).

It should be emphasised that this kinetic description is valid in the flowing regime in which pairwise interactions dominate. For higher agent densities, the dynamics will be altered by multi-body interactions, eventually leading to jamming ( $23,41,42$ ). Propagation of chaos dictates a stochastic formulation of the evolution of agent position from a disordered initial condition. We consider the position vector of each agent as a random variable with probability density function (pdf) $\rho_{i}^{ \pm}(\mathbf{r}, t)$ and the corresponding agent (number) densities $\rho^{ \pm}=\sum_{i} \rho_{i}^{ \pm}$. Applying the central limit theorem to (1) we find that $\Delta \mathbf{r}_{i}^{ \pm}$is approximately normally distributed, with mean $\pm v \Delta t\left[\mathbf{e}_{y}+2 \mathbf{A}^{ \pm}\right]$and variance $2 v \Delta t B^{ \pm}$, where $\mathbf{A}$ and $B$ are obtained from the agent density via

$$
\begin{align*}
\mathbf{A}^{ \pm}\left[\mathbf{r}, \rho^{\mp}\right] & =\int \rho^{\mp}\left(\mathbf{r}+x \mathbf{e}_{\mathbf{x}}\right) \mathbb{E} \mathbf{G}( \pm x) d x \\
B^{ \pm}\left[\mathbf{r}, \rho^{\mp}\right] & =\int \rho^{\mp}\left(\mathbf{r}+x \mathbf{e}_{\mathbf{x}}\right) \mathbb{E}\left[\mathbf{G}^{T}( \pm x) \mathbf{G}( \pm x)\right] d x . \tag{2}
\end{align*}
$$

Put simply if the interval $\Delta t$ is sufficiently long to encompass several collisions, the motion of agents/particles behaves as a set of almost uncorrelated random walks (see (43), section 1 for
mathematical derivation). Summing eq. 1 and its '-'-type counterpart over $i$ then leads to a pair Fokker-Planck equations for the agent densities

$$
\begin{equation*}
\frac{\partial \rho^{ \pm}}{\partial t} \pm v \frac{\partial \rho^{ \pm}}{\partial y}=\mp 2 v \nabla \cdot\left(\rho^{ \pm} \mathbf{A}^{ \pm}\right)+v \nabla \cdot\left(\nabla^{T} \rho^{ \pm} B^{ \pm}\right) \tag{3}
\end{equation*}
$$

We can derive several interesting conclusions from this result. Examining equation 3, we find on the left hand side terms describing simple advection in the preferred direction of motion, while the right hand side describe the effects of inter-agent interactions: the first corresponding to interaction-induced drift (arising from the mean collisional displacement) and the second to interaction-induced diffusion (arising from the variance of collisional displacement). In the following we will disentangle the roles of these terms and explore their consequences for the emergent nucleation of lanes.

Instability of homogeneous flow The transport equation 3 admits a trivial homogenous solution $\rho^{ \pm}(\mathbf{r}, t) \equiv \rho_{0}$. We interrogate the stability of this solution with respect to wave-like perturbations. No generality is lost as every perturbation can be represented as a linear combination of Fourier modes, which initially evolve independently. Linearizing and taking a Fourier transform (denoted with ${ }^{\sim}$ ), we obtain

$$
\begin{align*}
i \omega \tilde{\rho}^{ \pm} \pm i l v \tilde{\rho}^{ \pm}= & \mp 2 i v \rho_{0} \tilde{\mathbf{A}}(0) \cdot \mathbf{k} \tilde{\rho}^{ \pm} \mp 2 v i \rho_{0} \tilde{\mathbf{A}}(\mp k) \cdot \mathbf{k} \tilde{\rho}^{\mp}  \tag{4}\\
& -v \rho_{0} \mathbf{k}^{T} \tilde{B}(0) \mathbf{k} \tilde{\rho}^{ \pm}-v \rho_{0} \mathbf{k}^{T} \tilde{B}(\mp k) \mathbf{k} \tilde{\rho}^{\mp},
\end{align*}
$$

where $\omega$ is temporal frequency and $\mathbf{k}=(k, l)^{T}$ is the wave vector, which can also be expressed via $\lambda \mathbf{k}=2 \pi(\cos \theta,-\sin \theta)^{T}$ in terms of the perturbation wave length $\lambda$ and pattern orientation $\theta$.

The four terms on the right of eq. 4 , which are derived from the two interaction terms in eq. 3 , allow for a more nuanced analysis of the different dynamical processes. The terms involving $\tilde{\mathbf{A}}(0)$ and $\tilde{B}(0)$ describe the density-induced processes and the terms involving their frequency
dependent counterparts $\tilde{\mathbf{A}}(k)$ and $\tilde{B}(k)$ control gradient sensing processes. For example, the first term on the right describes the density-induced drag which one group imposes on the other by its shear presence, whereas the second term describes drag occurring in response to spatial inhomogeneities.

For chirally symmetric interactions, the density-induced drag represents the retardation associated with each collision acting in the $y$-direction and the inhomogeneity-induced drag arises as a consequence of an imbalance between left and right 'turns' leading to a net drift in the $x$ direction. In this case, as expected, one can show that the most unstable perturbations are the lane-like undulations in the $x$-direction $\left(\mathbf{k}=(k, 0)^{T}\right)$ and the growth rate $\sigma=-\Im[\omega]$ satisfies the dispersion relation

$$
\begin{equation*}
\sigma(k)=v \rho_{0}\left[2 k\left|\tilde{A}_{x}(k)\right|+k^{2}\left|\tilde{B}_{x x}(k)\right|-k^{2} \tilde{B}_{x x}(0)\right] . \tag{5}
\end{equation*}
$$

Thus, we recover the two previously conjectured mechanisms of lane nucleation: gradientinduced drift proportional to $\left|\tilde{A}_{x}(k)\right|(22)$ and gradient-induced diffusion controlled by $\left|\tilde{B}_{x x}(k)\right|$ $(11,23)$. We further discover the regularizing role of density-induced diffusion proportional $\tilde{B}_{x x}(0)$, which prevents an 'ultra-violet catastrophe' in the shortwave limit $k \rightarrow \infty$, instead inducing a cut-off wavenumber $k_{\text {cut }}$ such that high frequency oscillations are fully suppressed.

Considering the prototypical case of excluded volume interactions of hard spheres with diameter $D$ is instructive. As calculated in (43), the hard sphere dispersion relation has analytic form

$$
\begin{equation*}
\sigma(k)=v \rho_{0}\left[-\frac{D^{3}}{6} k^{2}+3 D-\frac{3 \sin (D k)}{k}\right] . \tag{6}
\end{equation*}
$$

We plot the dispersion relation plotted as a function of wavelength $\lambda$ (Fig. 2A) and it obtains a unique maximum at characteristic wavelength $\lambda_{\max } \approx 2.07 D$, with a sharp cut off to the left at $\lambda_{\text {cut }} \approx 1.34 D$ and an algebraic tail to the right.

Numerical simulations reveal the same characteristic shape of the dispersion relation after
appropriate rescaling for hard ellipses and hard spheres with non-parallel preferred velocity $\mathbf{v}^{ \pm}$(Fig. 2A). The growth of lanes exhibit some strongly non-linear characteristics, in that the logarithmic derivative of a given average Fourier mode $\langle\sigma(k)\rangle=\left\langle\dot{\tilde{\rho}}^{ \pm}\right\rangle /\left\langle\tilde{\rho}^{ \pm}\right\rangle$is not constant in time ( (43), Section IV). Nevertheless, at the 'moment of maximal linear growth' $t^{*}$ such that $\langle\sigma\rangle\left(\left(k^{*}, 0\right), t^{*}\right)=\max _{k, t}\langle\sigma\rangle((k, 0), t)$ that dominates lane formation, we find close agreement between $\langle\sigma(k)\rangle$ and the theoretically predicted growth rate (eq. (5)). We confirm this relationship for excluded volume models of ellipses and spheres (Fig. 2A) but a similar agreement has been found for other deterministic and stochastic AB models (see (43), Section IV). These models include a data-driven event-based simulation, which bootstrap the experimental displacement data displayed (Fig. 1D) to generate a large number of 'in silico' pedestrian experiments with the same displacement statistics.

Dimensional analysis of eq. 5 shows that the $k$-dependent terms act to select a characteristic length scale of lane nucleation, itself emerging from the action of the displacement operator. This length scale can be closely approximated by the square root of the $L^{1}$ norm of the $x$ component, $\mu=\sqrt{2\left\|G_{x}\right\|_{1}}$. We show that the rate of lane nucleation measured from simulations is robustly predicted by the simple expression $\sigma_{\max } \approx v \rho_{0} \mu$ (Fig. 2B). The first two terms of this expression determine the rate of collisions, whilst $\mu$ controls the size the lateral displacement per collision, which might vary due to the softness of particles, or the angle of approach.

Chiral interactions We emphasise that the fact that the most unstable perturbations of the homogeneous state are aligned with the differential velocity is only the case for chirally symmetric interactions. If the collisional displacement $\mathbf{G}(x)$ is not symmetric about zero, one can show that the maximal growth rate is attained at

$$
\begin{equation*}
\theta_{\max } \approx \tan ^{-1}\left[2 \rho_{0} \tilde{A}_{x}(0)\right] . \tag{7}
\end{equation*}
$$

The inclination of incipient lanes is driven through density-induced drift, evident from this expression, which itself can provide an alternative instability mechanism (30). Indeed, if the interactions are chirally biased, the very presence of the other agents will induce drift and agent segregation. This should be contrasted with the symmetric case, where the drift is gradientinduced, i.e. it relies on the imbalance of left and right turns associated with encountering other more agents of the opposite type on left or right.

The qualitative prediction of tilting lanes has been confirmed in a controlled experiment with pedestrians moving across a square arena (Fig. 3). When the participants were divided in two groups and asked to cross to the other side, they formed lanes parallel to the direction of motion (Supplementary Video 1). However, when the experiment marshals instructed the participants to repeat the same exercise while obeying a 'pass on the right' traffic rule, the crowds spontaneously segregated into lanes with marked inclination (Supplementary Video 2). A non-trivial tilting angle has also been confirmed in a quantiative manner with a Fourier analysis of the trajectory data from repeated trials. The magnitude of the tilt is consistent with the properties of the empirical collisional operator ( (43), Section VIIF). Validity of the formula (7) has also been confirmed by agent-based simulations of 'frictional rotating disks' (see (43), Section IV). It is also worth noting that if the lane orientation and the direction of motion are misaligned, collisions do not cease once the lanes are formed and the agents are forced to increase the lengths of their paths. None of these effects are desired in pedestrian traffic, which suggest that in crowd management the 'pass on the right' rule should be used with caution.

Complex flows Our theoretical framework is also easily extended to agents with generalized preferred velocity Indeed, this case can be reduced to the system with $\mathbf{v}^{ \pm}= \pm v \mathbf{e}_{y}$ by performing a local orthogonal coordinate transformation into a frame moving with the average speed $\frac{1}{2} \mathbf{v}^{p}=\frac{1}{2}\left[\mathbf{v}^{+}+\mathbf{v}^{-}\right]$and the $y$-axis is aligned with the differential velocity $\mathbf{v}^{m}=\mathbf{v}^{+}-\mathbf{v}^{-}$.

Consequently, in this case lanes emerge along the direction of $\mathbf{v}^{m}$, grow at a theoretical rate proportional to $v=\frac{1}{2}\left|\mathbf{v}^{m}\right|$ (c.f. eq. (5),), and shift with wave-speed $\frac{1}{2} \mathbf{v}^{p}$ (46). We show that this theoretical estimate of lane growth rate is consistent with hard spheres simulations as we vary the angle $\psi$ between the two preferred directions (Fig. 2A).

The formation of travelling laning patterns has also been observed in the experiment with two groups of pedestrians moving in perpendicular directions in which case the lanes are inclined at $45^{\circ}$ (Fig. 4A, Supplementary Video 3) (47). We confirm that even though the diagonal organisation appears more complex, it corresponds to a 'simple' stationary laning pattern in the appropriate frame of reference (Fig. 4C).

The observation that in symmetric systems lanes form along the direction of differential velocity can be used to analyze lane nucleation in more general cases where the preferred velocity varies in space. This may occur for example, when the agents proceed with equal speeds towards two distinct point targets $\mathbf{f}^{ \pm}$, where the lane formation is triggered along ellipses with foci in $\mathbf{f}^{ \pm}$(Fig. 4D). When $\mathbf{f}^{+}$is a target, but $\mathbf{f}^{-}$is a point of repulsion, lanes nucleate along hyperbolae with foci $\mathbf{f}^{ \pm}$(Fig. 4E). Finally, when the ${ }^{\prime}+^{\prime}$ group has a target direction $\mathbf{v}^{+}$, but the '-' group has a target point $\mathbf{f}^{-}$, the pattern curves are parabolae with focus in $\mathbf{f}^{-}$and a directrix perpendicular to $\mathbf{v}^{+}$(Fig. 4F). All of these facts can be easily proved using the reflective properties of conic sections. In more sophisticated scenarios, e.g. when the two speeds are not equal, computing the laning curves may require numerical integration. Importantly, these systems do not possess a stable steady state, so the local onset of lane nucleation does not progress to nonlinear growth and saturation. In particular, accumulation of agents near the attractive points will lead an increase in density and eventually to the breakdown of our theory as multi-body interactions dominate.

Despite these theoretical limitations, we have observed curved lane nucleation in a suite of controlled pedestrian experiments mimicking the three canonical scenarios (Fig. 4D-F Supple-
mentary Videos 4-6). In the experiment, the point targets are associated with narrow exit gates and the point repeller is approximately realized with a narrow entry gate. To avoid excessive crowding the entrance and exit gates were appropriately wide (1m) relative to the throughput of pedestrians; although higher pedestrian densities were observed near gates, encounters between agents of opposite types were distributed across the centre of the experimental domain ( (43), Fig. S25). In each case the directions of nucleating lanes were found to be consistent with the theoretical prediction. We also show that conic section patterns are observed in numerical simulations ( (43), section VI), where the average laning direction over a large number of replicates can be extracted in a quantitative manner.

Discussion In summary, we considered the complex flow of two groups of agents driven towards different targets. Most importantly, we have introduced a simple gas-kinetic framework which directly relates the individual behavior agents to their collective dynamics, thereby bridging the gap between the exact agent-based studies $(38,48,49)$ and heuristic continuum models ( $2,19,29,30,50$ ). Our analysis systematizes different previously proposed mechanisms of lane formation in dilute active mixtures $(11,22,23,30)$ and provides quantitative estimates for the propensity of lane nucleation induced by binary collision-like interactions. Our theory also generates predictions about the non-trivial orientation of lanes in chirally-biased systems, as well as curved lanes in complex flows with spatially varying direction of motion. All these have been confirmed both by numerical simulations, as well as in a controlled pedestrian experiment with human participants. The latter validation step is particularly important as it shows that our model correctly describes the characteristics of lane nucleation in real-world systems with complex interaction laws.

## References

1. M. Isobe, T. Adachi, T. Nagatani, Experiment and simulation of pedestrian counter flow, Physica A 336, 638-650 (2004).
2. D. Helbing, L. Buzna, A. Johansson, T. Werner, Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions, Transportation Science 39, 1-24 (2005).
3. J. Zhang, W. Klingsch, A. Schadschneider, A. Seyfried, Ordering in bidirectional pedestrian flows and its influence on the fundamental diagram, J. Stat. Mech. 2012 (2012).
4. J. Bamberger, et al., Crowd research at school: Crossing flows, Traffic and Granular Flow, 2013 (2015).
5. S. Cao, A. Seyfried, J. Zhang, S. Holl, W. Song, Fundamental diagrams for multidirectional pedestrian flows, J. Stat. Mech: Theory Exp. 3, 033404 (2017).
6. C. Feliciani, K. Nishinari, Empirical analysis of the lane formation process in bidirectional pedestrian flow, Phys. Rev. E 94, 032304 (2016).
7. N. Guo, Q. Y. Hao, R. Jiang, M. B. Hu, B. Jia, Uni- and bi-directional pedestrian flow in the view-limited condition: Experiments and modeling, Transportation Research Part C: Emerging Technologies 71, 63-85 (2016).
8. N. Guo, H. X. Liu, R. Jiang, B. Jia, M. B. Hu, Improving heuristics-based model to reproduce lane formation, International Journal of Modern Physics C 91, 1850069 (2018).
9. M. Moussaïd, et al., Traffic instabilities in self-organized pedestrian crowds, PLoS Comput. Biol. 8, e1002442 (2012).
10. M. E. Leunissen, et al., Ionic colloidal crystals of oppositely charged particles, Nature 437, 235-240 (2005).
11. T. Vissers, et al., Lane formation in driven mixtures of oppositely charged colloids, Soft Matter 7, 2352-2356 (2011).
12. K. R. Sütterlin, et al., Dynamics of lane formation in driven binary complex plasmas, Phys. Rev. Lett. 102, 085003 (2009).
13. C. R. Du, et al., Experimental investigation on lane formation in complex plasmas under microgravity conditions, New J. Phys. 14, 073058 (2012).
14. R. Jose, L. Santen, Self-organized lane formation in bidirectional transport by molecular motors, Phys. Rev. Lett. 124, 198103 (2020).
15. S. Nowak, A. Schadschneider, Quantitative analysis of pedestrian counterflow in a cellular automaton model, Phys. Rev. E 85, 066128 (2012).
16. Y. Sun, Kinetic monte carlo simulations of bi-direction pedestrian flow with different walk speeds, Physica A 549, 124295 (2020).
17. M. Burger, S. Hittmeir, H. Ranetbauer, M. T. Wolfram, Lane formation by side-stepping, SIAM J. Math. Anal. 48, 981-1005 (2016).
18. H. Ohta, Lane formation in a lattice model for oppositely driven binary particles, EPL 99, 40006 (2012).
19. J. Chakrabarti, J. Dzubiella, H. Löwen, Reentrance effect in the lane formation of driven colloids, Phys. Rev. E 70, 012401 (2004).
20. J. Cristín, V. Méndez, D. Campos, General scaling in bidirectional flows of self-avoiding agents, Sci. Rep. 9, 1-9 (2019).
21. J. Delhommelle, Should "lane formation" occur systematically in driven liquids and colloids?, Phys. Rev. E 71, 016705 (2005).
22. D. Helbing, T. Vicsek, Optimal self-organization, New Journal of Physics 1, 13 (1999).
23. K. Klymko, P. L. Geissler, S. Whitelam, Microscopic origin and macroscopic implications of lane formation in mixtures of oppositely driven particles, Phys. Rev. E 94, 022608 (2016).
24. A. Poncet, O. Bénichou, V. Démery, G. Oshanin, Universal long ranged correlations in driven binary mixtures, Phys. Rev. Lett. 118, 118002 (2017).
25. M. Rex, H. Löwen, Influence of hydrodynamic interactions on lane formation in oppositely charged driven colloids, European Physical Journal E 26, 143-150 (2008).
26. T. Xiong, P. Zhang, S. C. Wong, C. W. Shu, M. P. Zhang, A macroscopic approach to the lane formation phenomenon in pedestrian counterflow, Chin. Phys. Lett. 28, 108901 (2011).
27. H. Kuang, X. Li, T. Song, S. Dai, Analysis of pedestrian dynamics in counter flow via an extended lattice gas model, Phys. Rev. E 78, 066117 (2008).
28. C. Totzeok, An anisotropic interaction model with collision avoidance, Kinetic and Related Models 13, 1219-1242 (2020).
29. J. Chakrabarti, J. Dzubiella, H. Löwen, Dynamical instability in driven colloids, Europhys. Lett. 61.3, 415 (2003).
30. D. Helbing, A fluid-dynamic model for the movement of pedestrians, Complex Systems $\mathbf{6}$, 391-415 (1992).
31. A. Bricard, J.-B. Caussin, N. Desreumaux, O. Dauchot, D. Bartolo, Emergence of macroscopic directed motion in populations of motile colloids, Nature 503, 95-98 (2013).
32. N. Bain, D. Bartolo, Dynamic response and hydrodynamics of polarized crowds, Science 363, 46-49 (2019).
33. N. Bain, D. Bartolo, Critical mingling and universal correlations in model binary active liquids, Nat. Commun. 8 (2017).
34. A. Einstein, Investigations on the Theory of the Brownian Movement, Dover, New York (1956).
35. X. Jia, C. Feliciani, D. Yanagisawa, K. Nishinari, Experimental study on the evading behavior of individual pedestrians when confronting with an obstacle in a corridor, Physica A 531 (2019).
36. M. Moussaïd, et al., Experimental study of the behavioural mechanisms underlying selforganization in human crowds, Proceedings of the Royal Society B: Biological Sciences 276, 2755-2762 (2009).
37. C. Scholz, M. Engel, T. Pöschel, Rotating robots move collectively and self-organize, Nat. Comтип. 9, 1-8 (2018).
38. I. Karamouzas, B. Skinner, S. J. Guy, Universal power law governing pedestrian interactions, Phys. Rev. Lett. 113, 238701 (2014).
39. Q. Xu, M. Chraibi, A. Seyfried, Anticipation in a velocity-based model for pedestrian dynamics, Transportation Research Part C: Emerging Technologies 133 (2021).
40. J. Wang, et al., Linking pedestrian flow characteristics with stepping locomotion, Physica A: Statistical Mechanics and its Applications 500, 106-120 (2018).
41. M. Muramatsu, T. Irie, T. Nagatani, Jamming transition in pedestrian counter flow, Physica A 267, 487-498 (1999).
42. C. Reichhardt, C. J. O. Reichhardt, Velocity force curves, laning, and jamming for oppositely driven disk systems, Soft Matter 14, 490-498 (2018).
43. Materials and methods.
44. H. Murakami, C. Feliciani, Y. Nishiyama, K. Nishinari, Mutual anticipation can contribute to self-organization in human crowds, $S c i . A d v .7$, eabe7758 (2021).
45. A. Corbetta, J. A. Meeusen, C. M. Lee, R. Benzi, F. Toschi, Physics-based modeling and data representation of pairwise interactions among pedestrians, Phys. Rev. E 98, 062310 (2018).
46. J. Dzubiella, H. Löwen, Pattern formation in driven colloidal mixtures: tilted driving forces and re-entrant crystal freezing, J Phys: Condens Matter 14 (2002).
47. P. Mullick, et al., Analysis of emergent patterns in crossing flows of pedestrians reveals an invariant of stripe formation in human data, PLoS Comput Biol 18 (2022).
48. D. Helbing, P. Molnár, Social force model for pedestrian dynamics, Phys. Rev. E 51, 4282 (1995).
49. F. Dietrich, G. Köster, Gradient navigation model for pedestrian dynamics, Phys. Rev. E 89, 062801 (2014).
50. J. Toner, Y. Tu, Flocks, herds, and schools: A quantitative theory of flocking, Phys. Rev. E 58, 4828 (1998).
51. K. Bacik, B. Bacik, T. Rogers, Dataset for 'Lane nucleation in complex active flows", Bath: University of Bath Research Data Archive. (2023).
52. C. S. O'Hern, S. A. Langer, A. J. Liu, S. R. Nagel, Random packing of frictionless particles, Phys. Rev. Lett. 88 (2002).
53. E. Olson, Apriltag: A robust and flexible visual fiducial system, Proceedings - IEEE International Conference on Robotics and Automation (2011).

## Acknowledgments

BSB and all the authors would like to thank the colleagues from the Department of Human Motor Behaviour at the Academy of Physical Education in Katowice, in particular Grzegorz Sobota, Wojciech Marszałek, and Michał Pawłowski, for their help in conducting the experiment. KAB and TR thank Enrico Gavagnin and Johanna Meyer for their advice on experiments involving AprilTags and humans, respectively, as well as Becky Hayward for useful discussions. The authors would also like to thank Anton Souslov for proofreading the manuscript. The authors gratefully acknowledge the University of Bath's Research Computing Group (doi.org/10.15125/b6cd-s854) for their support in this work. Funding: KAB and TR acknowledge the support of Engineering and Physical Sciences Research Council grant number EP/V048228/1. The cost of the experiment was covered by the Academy of Physical Education in Katowice. Author contributions: KAB performed the theoretical analysis and numerical simulations, designed the experiments, and drafted the manuscript. BSB designed and conducted the experiments. TR conceived the study, performed the theoretical analysis, and drafted the manuscript. All authors approved the final text of the manuscript. Competing interests:
none. Data and materials availability: All data from the pedestrian experiment is available in a repository which can be accessed through the following link https://doi.org/10.15125/BATH01242 (51). The research protocol of the pedestrian experiment was approved by the Bioethics Committee for Scientific Research at the Academy of Physical Education in Katowice (No. 1/2022/6/23) and met the ethical standards of the Declaration of Helsinki.

## Supplementary materials

Materials and Methods (including Ref. (52, 53))
Supplementary Videos SV1-SV6 showing example experiments with human crowds (c.f. Materials and Methods, Sec. VII).

SV1: Symmetric lane formation. Paradigmatic experiment with two groups of pedestrians heading in opposite directions without any traffic rules imposed.

SV2: Chiral lane formation. An experiment with two groups heading in opposite directions with a 'pass on the right' rule imposed. Due to the broken chiral symmetry, nucleating lanes are no longer parallel to the direction of motion (c.f. Fig. 3).

SV3: Travelling lanes in perpendicular cross-streams. An experiment with two groups of pedestrians with perpendicular target directions, spontaneously organizing into diagonal travelling lanes (c.f. Fig. 4(a-c)).

SV4: Elliptical lane formation. Two groups of pedestrians heading towards two distinct gates (c.f. Fig. 4(d)).

SV5: Hyperbolic lane formation. One group of pedestrians (red) heads towards an exit gate (top). The members of the other group (blue) enter through a narrow gate on the left and exit through one of the preassigned wide gates on the right (c.f. Fig. 4(e)).

SV6: Parabolic lane formation. One group of pedestrians (red) crosses the experimental area in a prescribed direction, while the the other group (blue) targets a narrow exit gate on the right (c.f. Fig. 4(f)).


Figure 1: Lane formation examples. (A) Bidirectional pedestrian flow realised in a controlled experiment. (B) Agent-based simulation of driven hard spheres. Kinetic model. (C) Set-up sketch focusing on a ' + ' type agent moving with a preferred velocity $v \mathbf{e}_{y}$ (red). The presence of the ' - ' type agent (blue) moving with velocity $-v \mathbf{e}_{y}$ alters its trajectory, so instead of advancing by $v \Delta t \mathbf{e}_{y}$, its displacement within a $\Delta t$ time interval is given by $v \Delta t \mathbf{e}_{y}+\mathbf{G}(x)$, where $\mathbf{G}$ is the collisional operator and $x$ is the initial lateral offset between the two agents. (D) The $x$ component of the collisional displacement extracted from the pedestrian experiment (see (43) for inference procedure). Each gray dot corresponds to one interaction event, the red line is a running average of this data and the shading shows the corresponding standard deviation. (E) Diagram explaining the hydrodynamic equation (3). In the linearized model the evolution of agent density can be understood as a superposition of five processes: active migration in the preferred direction, density- and inhomogeneity-induced drift (which in a symmetric system act in orthogonal directions), as well homogeneous and inhomogeneous diffusion.


Figure 2: Stability analysis. (A) The growth rate of lane-like Fourier modes at the moment of maximal linear growth $t^{*}$ extracted through a Fourier transform from an ensemble of 5000 simulations of $2 N=300$ hard spheres moving in a doubly-periodic domain with preferred velocities subtending angle $\psi$ (large angles correspond to intense blue symbols, as explained by the colorbar in top-right), as well as hard ellipses with varying aspect ratio $\eta$ (red symbols of varying intensity, see inset). It shows an agreement with the theoretical dispersion relation with a global maximum for $\lambda_{\max } \approx 2 D$. The simulation details can be fund in (43), section IV. (B) The maximal growth rate $\sigma_{\max }$ can also be approximated using a simple heuristic scaling $\sigma_{\max } \propto v \rho_{0} \mu$.


Figure 3: Chiral intractions. (A) Experimental set-up with participants gathering behind the starting line of the square $5.8 \mathrm{~m} \times 5.8 \mathrm{~m}$ arena. The experiment was repeated in two sessions, one with 60 and one with 73 participants. In total, 10 benchmark trials and 7 trials with explicitly biased dodging maneuvers were recorded. (B) When a 'pass on the right' rule is imposed, pedestrians form tilting lanes. (C) The tilt can be systematically detected as a dominant pattern angle through Fourier analysis of the pedestrian trajectories. The tilting trend is reproducible across the repeated trials. (D) Example pedestrian trajectories for unbiased interactions. (E) Example pedestrian trajectories from the experiment with biased dodging maneuvers confirming the tilt.


Figure 4: Complex flows. (A) Two groups of agents with perpendicular target velocities form diagonal lanes migrating with the average velocity. The arrows highlight that the lanes are expected to form along the direction of differential velocity. The dots represent the position of pedestrians crossing the experimental arena (Fig. 3A) as two perpendicular streams. (B) Pedestrian trajectories in the lab frame cross at $90^{\circ}$. (C) The same trajectories viewed in an appropriately rotated reference frame moving with velocity $0.5\left(\mathbf{e}_{x}+\mathbf{e}_{y}\right) \mathrm{ms}^{-1}$ (approximating the average velocity of the two groups) reveal the laning pattern. (D) For a more complex scenario, where the two groups head towards two point targets the laning instability is triggered along ellipses. This can be proved by using the fact that lanes emerge along the direction of differential velocity (see arrows) as well as the reflective properties of an ellipse. As before, the dots show pedestrian positions observed in an experiment, showing segregation consistent with the theoretically predicted elliptical lines. (E) When one of the groups (blue) moves away from a particular point, the laning lines for a family of hyperbolas. In the experiment the 'repulsive point' is replaced with a narrow entry gate. (F) When one group (red) has a preferred direction and the other (blue) has a point target, the incipient lanes form along parabolas.

## Materials and Methods

## CONTENTS

I. Derivation of the kinetic equations ..... 1
II. Linear stability analysis ..... 3
A. Most unstable direction ..... 3
B. Symmetric interactions ..... 4

1. Hard sphere limit ..... 4
C. Long-wave limit ..... 5
III. Differential AB models ..... 6
2. Collisional operator ..... 6
A. Soft spheres ..... 7
B. Soft Ellipses ..... 8
3. Soft Ellipses vs. Social Force Model ..... 10
C. Rotating spheres ..... 10
IV. Numerical simulations ..... 12
A. Simulation details ..... 13
B. Empirical Stability Inference ..... 14
4. Results - symmetric interactions ..... 15
5. Results - chiral interactions ..... 17
6. Results - density effects ..... 18
V. Data-driven AB model ..... 20
A. Even-based simulation ..... 20
B. Lane nucleation in the surrogate simulation ..... 21
VI. Numerical simulations of complex flows ..... 22
A. Simulation details ..... 22
7. Two point targets (ellipse) ..... 23
8. A target and a repeller (hyperbola) ..... 23
9. A point target and a direction target (parabola) ..... 24
B. Automated detection of incipient lanes ..... 24
VII. Pedestrian experiments ..... 24
A. Data capture ..... 25
B. Experimental protocols ..... 25
C. Empirical collisional operator ..... 27
D. Pedestrian density Scenarios 1 and 2 ..... 28
E. Statistical analysis of tilting lanes ..... 28
F. A priori estimate of the tilting angle ..... 29
G. Density estimates in complex flows ..... 30
H. Data used in figures and videos ..... 31
References ..... 31

## I. DERIVATION OF THE KINETIC EQUATIONS

As described in the article, we start our theoretical considerations from an agent-based (AB) description, where after time interval $\Delta t$, the agent positions are updated according to

$$
\begin{equation*}
\Delta \mathbf{r}_{i}^{ \pm}= \pm v \mathbf{e}_{y} \Delta t+\sum_{j=1}^{N} \mathbf{F}_{\Delta t}^{ \pm}\left(\mathbf{r}_{i}^{ \pm}(t)-\mathbf{r}_{j}^{\mp}(t)\right) \tag{S.8}
\end{equation*}
$$

where $\mathbf{F}_{\Delta t}^{ \pm}$quantifies the effect of interactions. In general, $\mathbf{F}$ is a random variable, but for the differential AB models (see Sec. III) it becomes deterministic.

To begin with, it is instructive to consider isotropic interactions with finite range $D$. In this case, agent $i_{ \pm}$will interact with agent $j_{\mp}$ within time interval $\Delta t$ if and only if $\mathbf{r}_{j}^{-}(t) \in \mathcal{A}^{ \pm}$, where

$$
\begin{equation*}
\mathcal{A}^{ \pm}=\left\{\mathbf{r}_{i}^{ \pm}(t)+k \tilde{y}+D \cos \theta \tilde{x}+D \sin \theta \tilde{y}: k \in[0, \pm 2 v \Delta t], \theta \in[0, \pm \pi]\right\} \tag{S.9}
\end{equation*}
$$

and $D$ is the maximal interaction range (Fig. S1). If the time interval $\Delta t \gg t_{I}$, where $t_{I}$ is the duration of a typical interaction, we can approximate the interaction function $\mathbf{F}$ with

$$
\begin{equation*}
\mathbf{F}_{\Delta t}^{ \pm}(\mathbf{r})= \pm \mathbb{1}[0 \leq \mp y \leq 2 v \Delta t,|x| \leq D] \mathbf{G}( \pm x) \tag{S.10}
\end{equation*}
$$

even if the collision range is not isotropic. This motivates the approximation presented in the article

$$
\begin{equation*}
\Delta \mathbf{r}_{i}^{ \pm}= \pm v \mathbf{e}_{y} \Delta t \pm \sum_{\text {collisions }} \mathbf{G}\left( \pm \mathbf{r}_{i}^{ \pm}(t) \mp \mathbf{r}_{j}^{\mp}(t)\right) \tag{S.11}
\end{equation*}
$$

where $\mathbf{G}(x): \mathbb{R} \rightarrow \mathbb{R}^{2}$ is the collision operator, which describes the collisional displacement as a function of the initial lateral offset between the agents.


Figure S1. Schematic explaining the theoretical considerations. On the left hand side, the centre of the white circle is the location of agent $i_{+}$moving 'upwards' with speed $v$. This agent will collide with a 'downwards' moving agent $j_{i}$ within time interval $\delta t$ as long as $\mathbf{r}_{j}^{-}(t) \in \mathcal{A}^{+}$. The right hand side panel shows an analogous region $\mathcal{A}^{-}$, such that $\mathbf{r}_{j}^{+}(t) \in \mathcal{A}^{-}$implies that agents $i_{-}$and $j_{+}$interact within time interval $\Delta t$.

Assuming that the position of agent $i_{ \pm}$follows probability distribution $\rho_{i}^{ \pm}(\mathbf{r}, t)$, we can use eq. (S.11) to calculate the expected value of the displacement of agent $i$ after $\Delta t$, conditional on its current position $\mathbf{r}_{i}^{ \pm}(t)$.

$$
\begin{align*}
\mathbb{E}[\Delta \mathbf{r}] & \approx \pm v \mathbf{e}_{\mathbf{y}} \Delta t \pm \sum_{j} \int_{\mathcal{A}^{ \pm}} \rho_{j}^{\mp}(\mathbf{r}, t) \mathbb{E} \mathbf{G}\left(x_{i}^{ \pm}(t)-x\right) d \mathbf{r} \\
& = \pm v \mathbf{e}_{\mathbf{y}} \Delta t \pm 2 v \Delta t \int_{-\infty}^{\infty} \rho^{\mp}\left(\mathbf{r}_{i}^{ \pm}+x \mathbf{e}_{\mathbf{x}}, t\right) \mathbb{E} \mathbf{G}(x) d x+O\left((\Delta t)^{2}\right)  \tag{S.12}\\
& = \pm v \Delta t\left[\mathbf{e}_{y}+2 \mathbf{A}^{ \pm}\right]+O\left((\Delta t)^{2}\right)
\end{align*}
$$

where

$$
\begin{equation*}
\rho^{ \pm}=\sum_{i=1}^{N} \rho_{i}^{ \pm} \tag{S.13}
\end{equation*}
$$

and

$$
\begin{equation*}
\mathbf{A}^{ \pm}\left[\mathbf{r}, \rho^{\mp}\right]=\int_{-\infty}^{\infty} \rho^{\mp}\left(\mathbf{r}_{i}^{ \pm}+x \mathbf{e}_{\mathbf{x}}, t\right) \mathbb{E} \mathbf{G}( \pm x) d x \tag{S.14}
\end{equation*}
$$

In a similar way, we can compute the variance matrix

$$
\begin{align*}
\operatorname{Var}[\Delta \mathbf{r}] & \approx 2 v \Delta t \int_{-D}^{D} \rho^{\mp}\left(\mathbf{r}_{i}^{ \pm}+x \mathbf{e}_{\mathbf{x}}, t\right)\left[(\mathbb{E} \mathbf{G}(x)-\mathbb{E}[\Delta \mathbf{r}])^{T}(\mathbb{E} \mathbf{G}(x)-\mathbb{E}[\Delta \mathbf{r}])+\operatorname{Var}[\mathbf{G}(x)]\right] d x+O\left((\Delta t)^{2}\right) \\
& =2 v \Delta t \int_{-D}^{D} \rho^{\mp}\left(\mathbf{r}_{i}^{ \pm}+x \mathbf{e}_{\mathbf{x}}, t\right)\left[\mathbb{E} \mathbf{G}^{T}(x) \mathbb{E} \mathbf{G}(x)+\operatorname{Var}[\mathbf{G}(x)]\right] d x+O\left((\Delta t)^{2}\right)  \tag{S.15}\\
& =2 v \Delta t \int_{-D}^{D} \rho^{\mp}\left(\mathbf{r}_{i}^{ \pm}+x \mathbf{e}_{\mathbf{x}}, t\right) \mathbb{E}\left[\mathbf{G}^{T}(x) \mathbf{G}(x)\right] d x+O\left((\Delta t)^{2}\right) \\
& =2 v \Delta t B^{ \pm}+O\left((\Delta t)^{2}\right)
\end{align*}
$$

where

$$
\begin{equation*}
B^{ \pm}\left[\mathbf{r}, \rho^{\mp}\right]=\int_{-\infty}^{\infty} \rho^{\mp}\left(\mathbf{r}_{i}^{ \pm} \pm x \mathbf{e}_{\mathbf{x}}, t\right) \mathbb{E}\left[\mathbf{G}^{T}(x) \mathbf{G}(x)\right] d x \tag{S.16}
\end{equation*}
$$

Note that in general we consider $\mathbf{G}(x)$ as a random variable with mean $\mathbb{E} \mathbf{G}(x)$ and variance $\operatorname{Var}[\mathbf{G}(x)]$. Thus, the calculations in eq. (S.12) and eq. (S.15) can be understood as the law of iterated expectation and the law of total variance, respectively. In deterministic systems $\operatorname{Var}[\mathbf{G}(x)]=0$.

Having calculated the expected value of the agent displacement, we can make an approximation for the evolution of agent position, given in equation (S.11). By assuming that the agents are indistinguishable, so that they have identical probability distributions $\rho_{i}^{ \pm}=\frac{1}{N} \rho^{ \pm}$, for large $N$ and sufficiently large $\Delta t$ we can approximate the sum $\sum_{\text {collisions }} \mathbf{G}\left(\mathbf{r}_{i}^{ \pm}(t)-\mathbf{r}_{j}^{\mp}(t)\right)$ by appealing to the central limit theorem, so that

$$
\begin{equation*}
\mathbf{r}_{i}^{ \pm}(t+\Delta t) \approx \mathbf{r}_{i}^{ \pm}(t) \pm v \Delta t\left[\mathbf{e}_{y}+2 \mathbf{A}^{ \pm}\right]+\mathcal{N}\left(0,2 v \Delta t B^{ \pm}\right) \tag{S.17}
\end{equation*}
$$

where $\mathcal{N}$ is the standard normal distribution. In the limit of $\Delta t \rightarrow 0$, equation (S.17) could be realized by a stochastic differential equation

$$
\begin{equation*}
d \mathbf{r}_{i}^{ \pm}= \pm v\left[\mathbf{e}_{y}+2 \mathbf{A}^{ \pm}\right] d t+2 v B^{ \pm} d \mathbf{W} \tag{S.18}
\end{equation*}
$$

where $d \mathbf{W}$ is a two-dimensional Brownian motion. Note that the our arguments implicitly assume a time scale separation, so that $\Delta t$ is short compared to the global evolution of density, but long compared to the timescale of individual collision.

For each agent obeying equation (S.17) it is possible to write a corresponding Fokker-Planck equation for $\rho_{i}^{ \pm}$, By summing these equations over all the agents, we obtain a closed system of PDEs for the agent densities $\rho^{ \pm}$.

$$
\begin{equation*}
\frac{\partial \rho^{ \pm}}{\partial t} \pm v \frac{\partial \rho^{ \pm}}{\partial y} \pm 2 v \nabla \cdot\left[\rho^{ \pm} \mathbf{A}^{ \pm}\right]=v \nabla \cdot\left[\nabla^{T} \rho^{ \pm} B^{ \pm}\right] \tag{S.19}
\end{equation*}
$$

## II. LINEAR STABILITY ANALYSIS

Equation (19) is trivially solved by time-independent homogeneous solution

$$
\begin{equation*}
\rho^{ \pm}(x, y, t)=\rho_{0}=\frac{N}{|\mathcal{D}|} \tag{S.20}
\end{equation*}
$$

where $|\mathcal{D}|$ is the area of our domain. We will now analyse its linear stability by assuming a wave-like perturbation

$$
\begin{equation*}
\rho^{ \pm}:=\rho_{0}+\epsilon \rho^{ \pm} e^{i(\omega+\mathbf{k} . \mathbf{r})}+\text { c.c. } \tag{S.21}
\end{equation*}
$$

where $\epsilon \ll 0, \omega$ is the temporal frequency, and $\mathbf{k}=(k, l)$ is the wave-vector. The wave-vector can also be rewritten in terms of wavelength $\lambda$, and the orientation angle $\theta$, so that

$$
\begin{equation*}
\mathbf{k}=\frac{2 \pi}{\lambda}(\cos \theta,-\sin \theta) \tag{S.22}
\end{equation*}
$$

By noting that $\mathbf{A}^{ \pm}\left[\mathbf{r}, \rho^{\mp}\right]$ and $B^{ \pm}\left[\mathbf{r}, \rho^{\mp}\right]$ are linear in density, we may linearize equation (S.19) around this state to obtain

$$
\begin{equation*}
\frac{\partial \rho^{ \pm}}{\partial t} \pm v \frac{\partial \rho^{ \pm}}{\partial y} \pm 2 v \rho_{0} \nabla \cdot \mathbf{A}^{ \pm}\left[\mathbf{r}, \rho^{\mp}\right] \pm 2 v \mathbf{A}^{ \pm}\left[\mathbf{r}, \rho_{0}\right] \nabla \cdot \rho^{ \pm}=v \rho_{0} \nabla \cdot\left[\nabla^{T} B^{ \pm}\left[\mathbf{r}, \rho^{\mp}\right]\right]+v \nabla \cdot\left[\nabla^{T} \rho^{ \pm} B^{ \pm}\left[\mathbf{r}, \rho_{0}\right]\right] \tag{S.23}
\end{equation*}
$$

By taking a Fourier transform o equation (S.23), we get

$$
\begin{equation*}
i \omega \tilde{\rho}^{ \pm} \pm i l v \tilde{\rho}^{ \pm} \pm 2 i v \rho_{0} \tilde{\mathbf{A}}(0) \cdot \mathbf{k} \tilde{\rho}^{ \pm} \pm 2 v i \rho_{0} \tilde{\mathbf{A}}(\mp k) \cdot \mathbf{k} \tilde{\rho}^{\mp}+v \rho_{0} \mathbf{k}^{T} \tilde{B}(0) \mathbf{k} \tilde{\rho}^{ \pm}+v \rho_{0} \mathbf{k}^{T} \tilde{B}(\mp k) \mathbf{k} \tilde{\rho}^{\mp}=0 \tag{S.24}
\end{equation*}
$$

where

$$
\begin{equation*}
\tilde{\mathbf{A}}(k)=\int_{-D}^{D} e^{-i k x} \mathbb{E} \mathbf{G}(x) d x, \quad \tilde{B}(k)=\int_{-D}^{D} e^{-i k x} \mathbb{E}\left[\mathbf{G}^{T}(x) \mathbf{G}(x)\right] d x \tag{S.25}
\end{equation*}
$$

The solvability condition for equation (S.24) gives rise to a quadratic equation for $\omega$ with a solution

$$
\begin{align*}
\omega= & v \rho_{0} \mathbf{k}^{T} \tilde{B}(0) \mathbf{k} i \\
& \pm v \sqrt{\left(l+2 \rho_{0} \tilde{\mathbf{A}}(0) \cdot \mathbf{k}\right)^{2}-4 \rho_{0}^{2} \Im\left[\mathbf{k}^{T} \tilde{B}(-k) \mathbf{k} \tilde{\mathbf{A}}(k) \cdot \mathbf{k}\right]-4 \rho_{0}^{2}|\tilde{\mathbf{A}}(k) \cdot \mathbf{k}|^{2}-\rho_{0}^{2}\left|\mathbf{k}^{T} \tilde{B}(k) \mathbf{k}\right|^{2}} . \tag{S.26}
\end{align*}
$$

## A. Most unstable direction

In general the dispersion relation (26) is rather complicated, but we can make some analytical progress by considering the limit $\rho_{0} \ll 1$. Note that if $l=O(1)$, we can approximate equation (S.26) with

$$
\begin{equation*}
\omega= \pm v l+v \rho_{0}\left(\mathbf{k}^{T} \tilde{B}(0) \mathbf{k} i \pm 2 \tilde{\mathbf{A}}(0) . \mathbf{k}\right)+O\left(\rho_{0}^{2}\right) \tag{S.27}
\end{equation*}
$$

which (given that $\tilde{B}(0)$ is positive definite and $\tilde{A}(0) \in \mathbb{R}^{2}$ ) implies a negative growth rate

$$
\begin{equation*}
\sigma=-\Im \omega=-v \rho_{0} \mathbf{k}^{T} \tilde{B}(0) \mathbf{k}<0 \tag{S.28}
\end{equation*}
$$

On the other hand, for $l=l^{\prime} \rho_{0}$, we have

$$
\begin{align*}
\omega= & v \rho_{0} k^{2} \mathbf{e}_{\mathbf{x}}^{T} \tilde{B}(0) \mathbf{e}_{\mathbf{k}} i \\
& \pm v \rho_{0} \sqrt{\left(l^{\prime}+2 k \tilde{\mathbf{A}}(0) \cdot \mathbf{e}_{\mathbf{x}}\right)^{2}-4 k^{3} \Im\left[\mathbf{e}_{\mathbf{k}}^{T} \tilde{B}(-k) \mathbf{e}_{\mathbf{k}} \tilde{\mathbf{A}}(k) \cdot \mathbf{e}_{\mathbf{k}}\right]-4 k^{2}\left|\tilde{\mathbf{A}}(k) \cdot \mathbf{e}_{\mathbf{k}}\right|^{2}-k^{4}\left|\mathbf{e}_{\mathbf{k}}^{T} \tilde{B}(k) \mathbf{e}_{\mathbf{k}}\right|^{2}}+O\left(\rho_{0}^{3 / 2}\right) \tag{S.29}
\end{align*}
$$

where $\mathbf{e}_{\mathbf{k}}=(1,0)^{T}$. Thus, for fixed $|\mathbf{k}|^{2}=k^{2}+l^{2}$, the most unstable direction can be approximated as

$$
\begin{equation*}
\frac{l}{k}=-2 \rho_{0} \tilde{A}_{x}(0) \tag{S.30}
\end{equation*}
$$

## B. Symmetric interactions

When the interaction is chirally symmetric $G_{x}$ is odd and $G_{y}$ is even, so $\tilde{A}_{x}(0)=0$ and the maximum is achieved for $l=0$ and the (maximal) growth rate is given by

$$
\begin{equation*}
\sigma(k)=2 v \rho_{0} \Im\left[\tilde{A}_{x}(-k)\right] k-v \rho_{0} k^{2}\left(\tilde{B}_{x x}(0)-\left|\tilde{B}_{x x}(k)\right|\right), \tag{S.31}
\end{equation*}
$$

where we have introduced notation

$$
\tilde{\mathbf{A}}(k)=\binom{\tilde{A}_{x}(k)}{\tilde{A}_{y}(k)}, \quad \tilde{B}(k)=\left(\begin{array}{cc}
\tilde{B}_{x x}(k) & \tilde{B}_{x y}(k)  \tag{S.32}\\
\tilde{B}_{x y}(k) & \tilde{B}_{y y}(k)
\end{array}\right)
$$

for the Fourier parameters. Noting that for symmetric interactions $\tilde{A}_{x}$ is pure imaginary and typically for $k>0$, $\Im \tilde{A}_{x}(-k)>0$, we can rewrite eq. (31) as

$$
\begin{equation*}
\sigma(k)=2 v \rho_{0}\left|\tilde{A}_{x}(k)\right| k-v \rho_{0} k^{2}\left(\tilde{B}_{x x}(0)-\left|\tilde{B}_{x x}(k)\right|\right) \tag{S.33}
\end{equation*}
$$

## 1. Hard sphere limit

For hard spheres with diameter $D$ (see Section III A),

$$
\begin{equation*}
G_{x}(x)=\max \left[\frac{1}{2}(-x+D \operatorname{sign}(x)), 0\right] \tag{S.34}
\end{equation*}
$$

so for $k \neq 0$

$$
\begin{align*}
\left|\tilde{A}_{x}(k)\right| & =\frac{(D k-\sin (D k))}{k^{2}}  \tag{S.35}\\
\left|\tilde{B}_{x x}(k)\right| & =\frac{(D k-\sin (D k))}{k^{3}} \tag{S.36}
\end{align*}
$$

and

$$
\begin{equation*}
\tilde{B}_{x x}(0)=\frac{D^{3}}{6} \tag{S.37}
\end{equation*}
$$

Thus,

$$
\begin{equation*}
\sigma(k)=v \rho_{0}\left[-\frac{D^{3}}{6} k^{2}+3 D-\frac{3 \sin (D k)}{k}\right] \tag{S.38}
\end{equation*}
$$

Although, these cannot be written in an analytical form, numerically we can find that the growth rate attains its maximum

$$
\begin{equation*}
\sigma_{\max }=\max _{k} \sigma(k)=\sigma\left(k_{\max }\right) \approx 1.36 v \rho_{0} D \tag{S.39}
\end{equation*}
$$

for

$$
\begin{equation*}
k_{\max } \approx 3.04 D^{-1} \tag{S.40}
\end{equation*}
$$

and that $\sigma\left(k_{\text {cut }}\right)=0$ for

$$
\begin{equation*}
k_{\mathrm{cut}} \approx 4.67 D^{-1} \tag{S.41}
\end{equation*}
$$

which correspond to the most unstable wavelength

$$
\begin{equation*}
\lambda_{\max }=\frac{2 \pi}{k_{\max }} \approx 2.07 D \tag{S.42}
\end{equation*}
$$

and the cut-off wavelength

$$
\begin{equation*}
\lambda_{\mathrm{cut}}=\frac{2 \pi}{k_{\mathrm{cut}}} \approx 1.34 D \tag{S.43}
\end{equation*}
$$



Figure S2. (a) Perturbation growth rate plotted as a function of the wave number $k$ for the hard sphere limit. (b) The same relationship plotted as a function of the wavelength.

## C. Long-wave limit

Helbing and Vicsek in Ref. [22] propose a heuristic continuous (PDE) model which shows how in chirally symmetric systems lane formation could be initiated by gradient-induced instability. We will now demonstrate how this fundamental insight is confirmed by our model by analysing equation (S.19) in the long wave limit.

The long wave limit could be defined as a situation where the the agent densities vary slowly in space. In this case, the agent densities can be locally approximated with a Taylor series. In particular, $\mathbf{A}^{ \pm}$and $B^{ \pm}$which have been defined in eq. (S.14) and (S.16), respectively, can be approximated as follows.

$$
\begin{align*}
\mathbf{A}^{ \pm} & =\int_{-D}^{D} \rho^{\mp}\left(\mathbf{r}_{i}^{ \pm}+x \mathbf{e}_{\mathbf{x}}, t\right) \mathbb{E} \mathbf{G}( \pm x) d x \\
& \approx \int_{-D}^{D}\left[\rho^{\mp}\left(\mathbf{r}_{i}^{ \pm}, t\right)+x \frac{\partial \rho^{\mp}\left(\mathbf{r}_{i}^{ \pm}, t\right)}{\partial x}\right] \mathbb{E} \mathbf{G}( \pm x) d x  \tag{S.44}\\
& = \pm \rho^{\mp}\left(\mathbf{r}_{i}^{ \pm}, t\right) \tilde{\mathbf{A}}(0)+\frac{\partial \rho^{\mp}\left(\mathbf{r}_{i}^{ \pm}, t\right)}{\partial x} \boldsymbol{\Delta} \tilde{\mathbf{A}}, \\
B^{ \pm}= & \int_{-D}^{D} \rho^{\mp}\left(\mathbf{r}_{i}^{ \pm} \pm x \mathbf{e}_{\mathbf{x}}, t\right) \mathbb{E}\left[\mathbf{G}^{T}(x) \mathbf{G}(x)\right] d x \\
\approx & \int_{-D}^{D}\left[\rho^{\mp}\left(\mathbf{r}_{i}^{ \pm}, t\right)+x \frac{\partial \rho^{\mp}\left(\mathbf{r}_{i}^{ \pm}, t\right)}{\partial x}\right] \mathbb{E}\left[\mathbf{G}^{T}(x) \mathbf{G}(x)\right] d x  \tag{S.45}\\
= & \rho^{\mp}\left(\mathbf{r}_{i}^{ \pm}, t\right) \tilde{B}(0)+\frac{\partial \rho^{\mp}\left(\mathbf{r}_{i}^{ \pm}, t\right)}{\partial x} \Delta \tilde{B},
\end{align*}
$$

where $\tilde{\mathbf{A}}(0), \tilde{B}(0)$ have already been defined in eq. (S.25) and

$$
\begin{equation*}
\Delta \tilde{\mathbf{A}}=\int_{-D}^{D} x \mathbb{E} \mathbf{G}( \pm x) d x, \quad \Delta \tilde{B}=\int_{-D}^{D} x \mathbb{E}\left[\mathbf{G}^{T}(x) \mathbf{G}(x)\right] d x \tag{S.46}
\end{equation*}
$$

With these approximations, equation (S.19) becomes

$$
\begin{equation*}
\frac{\partial \rho^{ \pm}}{\partial t} \pm v \frac{\partial \rho^{ \pm}}{\partial y} \pm 2 v \nabla \cdot\left[\rho^{ \pm}\left( \pm \rho^{\mp} \tilde{\mathbf{A}}(0)+\frac{\partial \rho^{\mp}}{\partial x} \Delta \tilde{\mathbf{A}}\right)\right]=v \nabla \cdot\left[\nabla^{T} \rho^{ \pm}\left(\rho^{\mp} \tilde{B}(0)+\frac{\partial \rho^{\mp}}{\partial x} \Delta \tilde{B}\right)\right] . \tag{S.47}
\end{equation*}
$$

and the dispersion relation (S.26) becomes

$$
\begin{align*}
\omega= & v \rho_{0} \mathbf{k}^{T} \tilde{B}(0) \mathbf{k} i \\
& \pm v \sqrt{l^{2}+\rho_{0} \tilde{\mathbf{A}}(0) \cdot \mathbf{k} l-\rho_{0}^{2}\left(2 k \Delta \tilde{\mathbf{A}} \cdot \mathbf{k}+\mathbf{k}^{T} \tilde{B}(0) \mathbf{k}\right)^{2}-i \rho_{0}^{2} k^{2} \mathbf{k}^{T} \Delta \tilde{B} \mathbf{k}\left(2 \Delta \tilde{\mathbf{A}} \cdot \mathbf{k}+\mathbf{k}^{T} \Delta \tilde{B} \mathbf{k}\right)+\left(\rho_{0} k \mathbf{k}^{T} \Delta \tilde{B} \mathbf{k}\right)^{2}} \tag{S.48}
\end{align*}
$$

The analysis of Sec. II A is still valid, implying the same most unstable direction $\frac{l}{k}=-2 \rho_{0} \tilde{A}_{x}(0)$. For symmetric interactions $\tilde{A}_{x}(0)=0$, so the most unstable perturbations are parallel to the direction of motion and their growth rate is given by

$$
\begin{equation*}
\sigma(k)=2 v \rho_{0} k^{2}[\boldsymbol{\Delta} \mathbf{A}]_{x} \tag{S.49}
\end{equation*}
$$

The same expression can also be derived by taking a limit $k \rightarrow 0$ in equation (S.31) and noting that for $k \ll 1$

$$
\begin{equation*}
\Im \tilde{A}_{x}(-k)=\Im\left[\int_{-D}^{D} e^{i k x} \mathbb{E} G_{x}(x) d x\right] \approx k \int_{-D}^{D} x \mathbb{E} G_{x}(x) d x=k[\boldsymbol{\Delta} \mathbf{A}]_{x} \tag{S.50}
\end{equation*}
$$

## III. DIFFERENTIAL AB MODELS

Agent-based ( AB ) models of binary interacting mixtures often take form of first-order (over-damped) differential equations

$$
\begin{equation*}
\frac{d \mathbf{r}_{i}^{ \pm}}{d t}= \pm v \mathbf{e}_{y}+\sum_{j} \mathbf{f}_{\mp}\left(\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{\mp}, \frac{d \mathbf{r}_{j}^{\mp}}{d t}\right)+\sum_{j \neq i} \mathbf{f}_{ \pm}\left(\mathbf{r}_{i}^{ \pm}-\mathbf{r}^{ \pm}, \frac{d \mathbf{r}_{j}^{ \pm}}{d t}\right) \tag{S.51}
\end{equation*}
$$

where $\mathbf{f}_{ \pm}(\mathbf{r}, \mathbf{u})$ represents the interaction force. For physical systems, such as colloids, $\mathbf{f}_{ \pm}(\mathbf{r}, \mathbf{u})=\mathbf{f}(\mathbf{r})$, but the velocity dependence and type dependence $\left(\mathbf{f}_{+} \neq \mathbf{f}_{-}\right)$is often introduced in pedestrian systems to accounts for anticipatory effects.

## 1. Collisional operator

For systems governed by equation (S.51), the collisional operator may be computed by integrating the governing equations for two agents of opposite type $\left(r^{ \pm}\right)$subject to initial condition $\mathbf{r}^{ \pm}(t=0)=\mathbf{r}_{0}^{ \pm}$. For two agents satisfying equation (S.51), the governing equations are

$$
\begin{equation*}
\frac{d \mathbf{r}_{i}^{ \pm}}{d t}= \pm v \mathbf{e}_{y} \pm f_{\mp}\left(\mathbf{r}^{+}-\mathbf{r}^{-}, \frac{d \mathbf{r}^{\mp}}{d t}\right) \tag{S.52}
\end{equation*}
$$

By symmetry, we expect

$$
\begin{equation*}
\frac{d}{d t}\left(\mathbf{r}^{+}+\mathbf{r}^{-}\right)=0 \tag{S.53}
\end{equation*}
$$

which allows us to rewrite equation (S.52) as as a single equation for the differential position $\mathbf{r}=\mathbf{r}^{+}-\mathbf{r}^{-}$

$$
\begin{equation*}
\frac{d \mathbf{r}}{d t}=2 v \mathbf{e}_{y}+\mathbf{f}_{-}\left(\mathbf{r},-\frac{1}{2} \frac{d \mathbf{r}}{d t}\right)-\mathbf{f}_{+}\left(\mathbf{r}, \frac{1}{2} \frac{d \mathbf{r}}{d t}\right) \tag{S.54}
\end{equation*}
$$

By integrating (S.54), we obtain the collisional operator

$$
\begin{equation*}
\mathbf{F}_{\Delta t}^{+}\left(\mathbf{r}_{0}\right)=\frac{1}{4} \int_{t=0}^{\Delta t} \mathbf{f}_{-}\left(\mathbf{r},-\frac{1}{2} \frac{d \mathbf{r}}{d t}\right)-\mathbf{f}_{+}\left(\mathbf{r}, \frac{1}{2} \frac{d \mathbf{r}}{d t}\right) d t \tag{S.55}
\end{equation*}
$$

If at time $t$, the two agents are outside each other's sphere of influence and $\Delta t$ is greater than the duration of their interaction $t_{I}$, then $\mathbf{F}_{\Delta t}$ equals the collisional operator $\mathbf{G}\left(x_{0}\right)$

$$
\begin{equation*}
\mathbf{F}_{\Delta t}^{+}\left(\mathbf{r}_{0}\right)=\frac{1}{2}\left[\mathbf{r}(\Delta t)-\left(\mathbf{r}_{0}+v \mathbf{e}_{y} \Delta t\right)\right]=\mathbf{G}\left(x_{0}\right) \tag{S.56}
\end{equation*}
$$

Note that the collisional displacement is independent of $\Delta t$ and it is only a function of the initial lateral offset $x_{0}$, which can be illustrated visually by plotting a phase space diagram for $\mathbf{r}$ (Fig. S3).

We shall now discuss the properties of the collisional displacement and the associated Fourier transforms defined in equation (S.32) for the example models used in the article: soft and hard spheres, hard ellipses, and the rotating spheres model.


Figure S3. The phase space diagram of equation (S.54). Blue lines are example trajectories and the black dashed circle corresponds to the interaction range $D$. If at $t=0,|\mathbf{r}(0)|=\left|\mathbf{r}_{0}\right|>D$, and $\Delta t$ is longer than the interaction time $t_{I}$, the displacement is only a function of the initial lateral offset $x_{0}$. The interaction time $t_{I}\left(x_{0}\right)$ is finite for $x_{0} \neq 0$, and for $x_{0}=0$ a saddle-type fixed point may appear, depending on the details of the interaction function $\mathbf{f}$.

## A. Soft spheres

The first family of interaction functions we consider in the article is a simple isotropic soft sphere model, with

$$
\begin{equation*}
\mathbf{f}_{ \pm}(\mathbf{r}, \mathbf{u})=\mathbf{f}(\mathbf{r})=\alpha \max [D-|\mathbf{r}|, 0] \frac{\mathbf{r}}{|\mathbf{r}|} \tag{S.57}
\end{equation*}
$$

parametrized by the 'hardness parameter' $\alpha$. The key non-dimensional parameter of this model, which compares the magnitude of the interaction term and the driving term is

$$
\begin{equation*}
\gamma=\frac{\alpha D}{v} \tag{S.58}
\end{equation*}
$$

Figure S4 shows the effect of $\gamma$ on the phase space topology. For $\gamma \geq 1$ the dynamical system of equation (S.54) has a saddle-type equilibrium located at $x=0, y=D\left(-1+\gamma^{-1}\right)$. As $\gamma \rightarrow \infty$, the dynamics converges to a hard sphere limit, where the agents maintain constant distance throughout interaction.





Figure S4. Numerically computed phase space diagrams for family of interaction functions given by equation (S.57). Parameter $\gamma=\frac{\alpha R}{v}$ controls the relative strength of the repulsion and determines the shape of the trajectories.

Figure S 5 shows numerically computed collisional displacement $\mathbf{G}(x)=\left(G_{x}(x), G_{y}(x)\right)$ as we vary $\gamma$. As $\gamma$ increases, the collisional displacement generally increases in amplitude, and in the limit $\gamma \rightarrow \infty, \mathbf{G}$ approaches the analytical
hard sphere limit

$$
\begin{align*}
& G_{x}^{\text {hard }}(x)=\max \left[\frac{1}{2}(D \operatorname{sign}(x)-x), 0\right] \\
& G_{y}^{\text {hard }}(x)=\max \left[\frac{1}{2}\left(\sqrt{D^{2}-x^{2}}-D \ln \left(\frac{D+\sqrt{D^{2}-x^{2}}}{|x|}\right)\right), 0\right] \tag{S.59}
\end{align*}
$$

The limiting lateral displacement $G_{x}^{\mathrm{hard}}$ can be inferred by noting that in the hard sphere limit all trajectories with $0<|x|<D$ go through one of the points $x= \pm D, y=0$. The longitudinal displacement $G_{y}^{\text {hard }}$ can be computed by assuming that at $|\mathbf{r}|=D$, the radial component of the freestream differential velocity is balanced by the interaction force. Thus, the evolution of $\mathbf{r}$ for $|\mathbf{r}|=D$ is equivalent to a point mass moving on a circle of radius $D$ with angular speed

$$
\begin{equation*}
\dot{\theta}= \pm \frac{2 v}{D} \sin \theta \tag{S.60}
\end{equation*}
$$

which can be integrated analytically to yield the result.


Figure S5. (a) Numerically computed lateral component of the collisional displacement $G_{x}(x)$ for the soft sphere model with parameter $\gamma$. (b) The corresponding longitudinal displacement $G_{y}(x)$. As $\gamma$ increases, the collisional displacement $\mathbf{f}$ approaches the hard sphere limit given by equation (S.59). Note that $G_{y}(0)=-\infty$ for $\gamma \geq 1$, when the fixed point appears (c.f. Fig. S4).

As discussed in Section II B, in the hard sphere limit the Fourier transforms of the mean collisional displacement can be calculated analytically as well. Figure S6, which uses the notation of equation (S.32), shows how this hard sphere limit is approached as $\gamma \rightarrow \infty$. Of particular interest is Figure $\mathrm{S} 6(\mathrm{f})$, which shows the theoretical growth rate $\sigma(\lambda)$, It shows that the maximal growth rate $\sigma_{\max }$, as well as the most unstable wavelength $\lambda_{\max }$ both increase with $\alpha$.

## B. Soft Ellipses

Another AB model which we consider are soft ellipses, motivated by the Social Force Model (SFM) of pedestrian motion [22]. The SFM aims to incorporate the idea of anticipation into the avoidance maneuvers by assuming that the social force imposed by the agents is normal to an ellipse with a semi-major axis aligned with their preferred direction of motion. The said pedestrian is one of the foci of the ellipse and the other focus is located at a distance $d$ in their preferred direction of motion (Fig. S7). After the foci are specified, the magnitude of the social force is assumed to be constant along the the corresponding family of ellipses.

In the soft ellipse mode, the distance between the two foci is fixed at

$$
\begin{equation*}
d=D\left(1-\eta^{2}\right) \tag{S.61}
\end{equation*}
$$

and the corresponding interaction function (c.f. eq. (S.51)) can be concisely written as

$$
\begin{equation*}
\mathbf{f}_{ \pm}^{\text {ellipse }}(\mathbf{r})=\alpha \max \left[\left(\eta D-\frac{1}{2} \sqrt{\left(|\mathbf{r}|+\left|\mathbf{r} \pm d \mathbf{e}_{y}\right|\right)^{2}-d^{2}}, 0\right] \frac{\mathbf{n}_{ \pm}}{\left|\mathbf{n}_{ \pm}\right|},\right. \tag{S.62}
\end{equation*}
$$



Figure S6. Fourier parameters of the collisional displacement for the soft sphere model approaching the hard sphere limit as the parameter $\gamma \rightarrow \infty$. (a) Longitudinal density-induced drift (retardation) parameter. (b) Lateral gradient-induced drift parameter. (c) Density-induced diffusion lateral diffusion parameter. (d) Density-induced diffusion longitudinal diffusion parameter. (e) Gradient-induced cross-diffusion parameter.
where

$$
\begin{equation*}
\mathbf{n}_{ \pm}=\frac{\mathbf{r}}{|\mathbf{r}|}+\frac{\mathbf{r} \mp d \mathbf{e}_{y}}{\left|\mathbf{r} \mp d \mathbf{e}_{y}\right|} \tag{S.63}
\end{equation*}
$$

Similarly to the soft sphere model,

$$
\begin{equation*}
\gamma=\frac{\alpha D}{v} \tag{S.64}
\end{equation*}
$$

is the non-dimensional softness parameter and the additional parameter $\eta \in(0,1]$ controls the eccentricity of the ellipse

$$
\begin{equation*}
e=\frac{1-\eta^{2}}{1+\eta^{2}} \tag{S.65}
\end{equation*}
$$

Note that the largest ellipse for which $\mathbf{f}_{ \pm}^{\text {ellipse }} \neq 0$ has a semi-minor axis $\eta D$ and the maximal distance between this ellipse and its focus is $D$ (Fig. S7).


Figure S7. A sketch of the soft ellipse model showing the social force $\mathbf{f}_{-}$imposed by the downwards moving blue agent on the upwards moving red agent. The social force is normal to the red ellipse with one focus coinciding with the position of the red agent and another focus at distance $d$ away from it, in its direction of motion. The magnitude of $\mathbf{f}_{-}$is a function of the semi-minor axis of the blue ellipse. In the Soft Ellipse Model, $d$ is fixed, but in the Social Force Model it depends on the speed of the red agent. For soft ellipses, the interaction range of the red agent is delimited by the black ellipse where the maximal longitudinal interaction distance (ahead) is $D$ and the maximal lateral interaction distance (semi-minor axis of the black ellipse) is $\eta D$.

Crucially, in the hard ellipse limit $(\gamma \rightarrow \infty)$, the lateral displacement approaches the limit

$$
\begin{equation*}
G_{x}^{\mathrm{hard}}(x)=\max \left[\frac{1}{2}(\eta D \operatorname{sign}(x)-x), 0\right] \tag{S.66}
\end{equation*}
$$

(c.f. Equation (S.59)). Thus, the growth rate takes the analytical expression analogous to equation (S.67)

$$
\begin{equation*}
\sigma(k)=v \rho_{0}\left[-\frac{\eta^{3} D^{3}}{6} k^{2}+3 \eta D-\frac{3 \sin (\eta D k)}{k}\right] \tag{S.67}
\end{equation*}
$$

In other words, as far as lane nucleation is concerned, the hard ellipse model is equivalent to hard spheres with diameter $\eta D$.

## 1. Soft Ellipses vs. Social Force Model

The soft ellipse interaction model is a simplification of the Social Force Model proposed by Helbing et al, where the interaction function is speed-dependent

$$
\begin{equation*}
\mathbf{f}_{ \pm}^{S F M}(\mathbf{r}, \mathbf{u})=g\left(\frac{1}{2} \sqrt{\left(|\mathbf{r}|+|\mathbf{r} \pm \kappa| \mathbf{u}\left|\mathbf{e}_{y}\right|\right)^{2}-(\kappa|\mathbf{u}|)^{2}}\right) \frac{\mathbf{n}_{ \pm}}{\left|\mathbf{n}_{ \pm}\right|} \tag{S.68}
\end{equation*}
$$

so that the distance between the two foci $d=\kappa|\mathbf{u}|$ which is supposed to represent the anticipated step size and $g$ is the interaction function, e.g. $g(x)=\max (\eta D-x, 0)$.

Mindful of our limited computational resources we use the speed-independent model, which does not require such a fine temporal resolution of a numerical solver. Nevertheless, it is worth pointing out that the simplification is only justified when the agents' preferred directions of motions are close to parallel. Indeed, the soft ellipse model may produce some undesired coalescence effects when the target directions of the two colliding agents are divergent.

In this, more general, case the soft ellipse potential would be

$$
\begin{equation*}
\mathbf{f}_{ \pm}^{\text {ellipse }}(\mathbf{r})=\alpha \max \left[\left(\eta D-\frac{1}{2} \sqrt{\left(|\mathbf{r}|+\left|\mathbf{r}+d \mathbf{e}_{ \pm}\right|\right)^{2}-d^{2}}, 0\right] \frac{\mathbf{n}_{ \pm}}{\left|\mathbf{n}_{ \pm}\right|}\right. \tag{S.69}
\end{equation*}
$$

where $\mathbf{e}_{ \pm}$is the target direction of a ( $\pm$)-type agent and

$$
\begin{equation*}
\mathbf{n}_{ \pm}=\frac{\mathbf{r}}{|\mathbf{r}|}+\frac{\mathbf{r}-d \mathbf{e}_{ \pm}}{\left|\mathbf{r}-d \mathbf{e}_{ \pm}\right|} \tag{S.70}
\end{equation*}
$$

As described in the article, by choosing an appropriate travelling frame of reference, the dynamics of two agents crossing at an angle may be reduced to form (S.54), so that

$$
\begin{equation*}
\mathbf{e}_{+} \cdot \mathbf{e}_{x}=\mathbf{e}_{-} . \mathbf{e}_{x}, \quad \mathbf{e}_{+} \cdot \mathbf{e}_{y}=-\mathbf{e}_{-} \cdot \mathbf{e}_{y} \tag{S.71}
\end{equation*}
$$

Nevertheless, this transformation does not change the relative orientation of the elliptical interaction shells

$$
\begin{equation*}
\mathbf{e}_{+} \cdot \mathbf{e}_{-}=\cos \psi \tag{S.72}
\end{equation*}
$$

Interestingly, Figure S 8 shows that for large $\gamma$, as $\psi$ decreases from $\pi$ (head-on collision), the unstable fixed point of the two-agent dynamical system (c.f. Figures S3, S4) undergoes a subcritical pitchfork bifurcation and becomes stable. The existence of a stable fixed point with a sizable basin of attraction implies that instead of dodging each other, the agents may 'cling' to each other and come to a halt. By modulating the shape of the interaction zone, the speed-dependence of the SFM (eq. (S.68)) automatically removes this undesired effect. Indeed, as the agents slow down, their 'shape' becomes more circular and a pair of driven soft spheres cannot form a stable doublet. In summary, because of the 'clinging bifurcation', the soft ellipse simplification of the SFM should be used with care for agent cross-streams, but in this work we only apply it to cases where $\psi=\pi$.

## C. Rotating spheres

The rotating sphere model is introduced as a toy example of an $A B$ model with broken chiral symmetry. One way of interpreting this model is to think of rotating discs with a rough surface which impose frictional force on each other.


Figure S8. (a) Bifurcation diagram showing a subcritical pitchfork bifurcation in the soft ellipse model ( $\gamma=100, \eta=0.5$ ). Open circles show numerically computed boundaries of the basin of attraction of the fixed point at $x=0$. (b) Phase space diagram for an oblique collision angle $\psi=0.8 \pi$, where three fixed points coexist (l.h.s. of panel (a)). For sufficiently small $\left|x_{0}\right|$ the agents are attracted to a stable doublet configuration (magenta trajectories), but for larger $\left|x_{0}\right|$ they dodge each other (blue trajectories). (c) Phase space diagram for a head-on collision, $(\psi=\pi)$ where the basin of attraction of the fixed point at $x=0$ is one-dimensional, so the agents almost surely pass each other.




Figure S9. (a) Phase space diagram of the rotating sphere collision for $\gamma=100$ and $\beta / \alpha=1$. Critically, the phase space portrait is not symmetric about $x=0$. If the agents were to be interpreted as pedestrians, red trajectories would correspond to right turns ans blue to left turns. (b) This corresponding lateral collisional displacement $G_{x}$ with a singularity at $x \approx-0.7 D$ corresponding to the dividing line between left and right turns. (c) The longitudinal collisional displacement.

Here, the interaction force $f_{ \pm}(\mathbf{r})=f^{\text {chiral }}(\mathbf{r})$ is a superposition of soft sphere repulsion parallel to $\mathbf{r}$ and a 'frictional' force perpendicular to it, so that

$$
\begin{equation*}
f^{\text {chiral }}(\mathbf{r})=\max [D-|\mathbf{r}|, 0]\left(\alpha \frac{\mathbf{r}}{|\mathbf{r}|}+\beta \frac{\mathbf{t}}{|\mathbf{r}|}\right) \tag{S.73}
\end{equation*}
$$

where $\mathbf{r} . \mathbf{t}=0, \mathbf{r} \times \mathbf{t}=|\mathbf{r}|^{2} \mathbf{e}_{z}$. The relative strength of the swirling term and the repulsion is controlled by $\beta / \alpha$, where $\beta>0$ correspond to spheres rotating counter-clockwise and $\beta<0$ corresponds to spheres rotating clockwise.

Another, less direct, way of interpreting this model associates agents with pedestrians with biased dodging manoeuvres. Indeed, by looking at the phase space diagrams in Figure S9, we can infer that $\beta>0$ would correspond to agents having a right-turn preference (and $\beta<0$ would correspond to agents who prefer to turn left) when dodging each other. This symmetry breaking shifts the location of the unstable fixed point away from $x=0$ and the collision displacements are no longer symmetric about $x=0$.

As discussed in Section II A, if the chiral symmetry is broken, lanes no longer are expected to nucleate along the direction of motion. The growth rate for every wave number can be calculated out of the full dispersion relation (S.26). Figure S10 shows how it changes as we increase the chirality parameter $\beta / \alpha$.


Figure S10. Stability analysis of the rotating spheres model for $\rho_{0}=0.375$. Panels (a-d) show the stability diagrams for different values of the chirality parameter $\beta / \alpha$. Every possible sinusoidal perturbation is uniquely described by the orientation $\theta$ and the wavelength $\lambda$. Thus, each of them corresponds to one point in the $\lambda-\theta$ space. The black curve is the neutral stability contour, which separates growing and decaying modes. The growth rate of the amplified modes $\sigma$ is illustrated with the colour scheme explained in a legend on the right hand side, and the white cross shows the location of the fastest growing mode. For clarity, the negative values of $\sigma$ are not shown in the figure. (a) For $\beta / \alpha=0$, the dynamics is chirally symmetric, so the most unstable perturbation is parallel to the direction of motion $\left(\theta_{\max }=0\right)$ (c.f. Fig. S2). (b-d) As the chirality strength increases, the stability region shifts.(e) The most unstable angle increases with $\beta / \alpha$ and it is well approximated by the theoretical expression (S.74). (f) Interestingly, the maximal growth rate is not monotone in $\beta / \alpha$ and it is maximal for the symmetric interaction. (f) As the chirality intensifies, the most unstable wavelength increases.

For $\beta / \alpha=0$ (Fig. $\mathrm{S} 10(\mathrm{a}))$ the system is chirally symmetric so, as expected, the most unstable perturbation is parallel to the direction of motion $\left(\theta_{\max }=0\right)$. As the results were computed in the hard sphere regime, the most unstable wavelength is given by equation (S.42). In Section IIB where that expression was introduced, we also introduced the notion of a cut-off wavelength, which corresponds to a zero growth rate. In Figure S10, this cut-off wavelength corresponds to the intersection of the neutral stability curve (black) with the x-axis. The neutral stability curve separates the amplified and damped perturbations in the $\lambda-\theta$ space.

As the chirality parameter $\beta / \alpha$ increases, the instability region (bounded by the neutral stability curve) expands and shifts (Fig. S10(b-d)) towards larger values of $\theta$ and $\lambda$. Figure $\mathrm{S} 10(\mathrm{e})$ shows that the most unstable orientation agrees with the approximation given by equation (S.30), so that

$$
\begin{equation*}
\theta_{\max }=\tan ^{-1}\left(2 \rho_{0} \tilde{A}_{x}(0)\right) \tag{S.74}
\end{equation*}
$$

As far as the maximal growth rate $\sigma_{\max }$ is concerned, it turns that for the rotating spheres model, it is a nonmonotonic function of the chirality parameter $\beta / \alpha$ with a maximum at $\beta / \alpha=0$ and a local minimum at $\beta / \alpha \approx \pm 0.6$ (Fig. S10(f)). Finally, Figure $\mathrm{S} 10(\mathrm{~g})$ demonstrates that for the rotating spheres model, the most unstable wavelength $\lambda_{\text {max }}$ increases with $|\beta / \alpha|$.

It is worth pointing out that even though Figure S10 explores $\beta>0$ (right turn preference), we can easily infer the most unstable mode for $\beta<0$ (left turn preference) as well. By symmetry,

$$
\begin{equation*}
\theta_{\max }(-\beta)=-\theta_{\max }(\beta), \quad \lambda_{\max }(-\beta)=\lambda_{\max }(\beta), \quad \sigma_{\max }(-\beta)=\sigma_{\max }(\beta) \tag{S.75}
\end{equation*}
$$

## IV. NUMERICAL SIMULATIONS

This section is dedicated to the numerical solutions of the differential AB models introduced in Section III. In Subsection IV A we describe the finite difference scheme we use and we list the parameters of our simulations. In Subsection IV B we explain how the empirical pattern growth rate in inferred from the simulation data.

## A. Simulation details

To solve any of the models from Section III (soft spheres, soft ellipses, rotating spheres) we use a simple forward Euler finite difference scheme with time step $\Delta t=0.05$ and the total simulation time $T=100$.. The simulation domain is chosen to be a doubly-periodic square with side $L=20$ and we fix the number of agents in each group to $N=150$, as well as the preferred speed $(v=0.1)$ for all the simulations. In total, we performed five simulation campaigns to explore different aspects of the dynamics. In Campaigns $1-4, D=0.3$ is fixed, and the variable parameters are softness $\alpha$ (Eq. (S.58)), the slenderness parameter $\eta$ in case of ellipses (Eq. (S.65)), or the chirality parameter $\beta$ in case of rotating spheres (eq. (S.73)). In Campaign 5, we vary the diameter $D$ to investigate jamming effects.

## 1. Soft spheres simulation with variable softness.

Here for each agent we solve the differential equation

$$
\begin{equation*}
\frac{d \mathbf{r}_{i}^{ \pm}}{d t}= \pm 0.1 \mathbf{e}_{y}+\sum_{ \pm} \sum_{j=1}^{N=150} \mathbb{1}\left[\left|\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right|>0\right] \alpha \max \left[0.3-\left|\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right|, 0\right] \frac{\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}}{\left|\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right|} \tag{S.76}
\end{equation*}
$$

where the softness parameter is varied and we set it to one of the ten test values $\alpha=\frac{1}{10}, \frac{2}{10}, \ldots, 1$. As far as the stability results are concerned, each of these ten values constitutes one data points. However, to achieve statistical convergence for each of these 10 values of $\alpha$ we repeat the simulation $N_{\text {repeat }}=5000$ times, each time starting from a different random initial condition with independently and uniformly distributed agents.

## 2. 'Hard' spheres simulation with variable cross-stream angle.

Although perfectly hard spheres would require softness $\alpha=\infty$, we can closely approximate this limit without abandoning the simple finite difference scheme by choosing a sufficiently setting $\alpha$. Here we set $\alpha=10$, which in the notation of Equation (S.58) corresponds to $\gamma=30$ (c.f. Fig. S5).
Specifically, for each agent we solve the following differential equation

$$
\begin{equation*}
\frac{d \mathbf{r}_{i}^{ \pm}}{d t}=0.1 \mathbf{e}_{ \pm \psi / 2}+\sum_{ \pm} \sum_{j=1}^{N=150} \mathbb{1}\left[\left|\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right|>0\right] 10 \max \left[0.3-\left|\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right|, 0\right] \frac{\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}}{\left|\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right|} \tag{S.77}
\end{equation*}
$$

where $\mathbf{e}_{ \pm \psi / 2}$ is a unit vector in a given direction.
Note that the preferred speed is always fixed and the differential velocity between the two agent types is always parallel to $\mathbf{e}_{y}$. Nevertheless, the differential speed increases with the cross-stream angle $\psi$. In the archetypal set-up with head-on directions $\psi=\pi$. In our simulations we explored eleven different crossing angles equispaced between 0 and $\pi$, i.e. $\psi=0, \frac{\pi}{10}, \frac{2 \pi}{10}, \ldots \pi$. As in the case of soft spheres, for each of the eleven values we repeated the simulation $N_{\text {repeat }}=5000$ times.

## 3. 'Hard' ellipse simulation with variable slenderness.

The details of the elliptic potential model can be found in Section III B. The governing equations are given by

$$
\begin{equation*}
\frac{d \mathbf{r}_{i}^{ \pm}}{d t}= \pm 0.1 \mathbf{e}_{y}+\sum_{ \pm} \sum_{j=1}^{N=150} \mathbb{1}\left[\left|\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right|>0\right] \mathbf{f}_{ \pm}^{\mathrm{ellipse}}\left(\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right) \tag{S.78}
\end{equation*}
$$

where the two-parameter interaction function is given by equation (S.62). The first parameter is softness $\alpha$, analogous to the same quantity in the sphere model. Here, by fixing $\alpha=10$ we ensure that the agent-agent interactions closely approximate those of 'hard ellipses'. The second parameter is slenderness $\eta \in[0,1]$, which characterizes the eccentricity in a particular way ( $\eta=1$ corresponds to a circle). In our simulations we use 11 test values: $\eta=\frac{2}{12}, \frac{3}{12}, \ldots 1$, and as before in each setting the simulation is repeated $N_{\text {repeat }}=5000$ times.

## 4. 'Hard' rotating spheres with variable swirl.

This model is described in Section III C and the equations are given by

$$
\begin{equation*}
\frac{d \mathbf{r}_{i}^{ \pm}}{d t}= \pm 0.1 \mathbf{e}_{y}+\sum_{ \pm} \sum_{j=1}^{N=150} \mathbb{1}\left[\left|\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right|>0\right] \mathbf{f}^{\text {chiral }}\left(\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right) \tag{S.79}
\end{equation*}
$$

where $\mathbf{f}^{\text {chiral }}$ is defined in equation (S.73). This interaction force is a superposition of a radial soft sphere interaction with softness $\alpha$ and an azimuthal component whose strength is parametrized by $\beta$. The ratio of $\beta / \alpha$ is a measure of chirality of the dynamics. Here, we fix $\alpha=10$ (asymptotic hard regime) and we use 11 different values for $\beta=0,2, \ldots, 20$, noting that $\beta=0$ reduces to simple hard spheres. Again, each simulation setting has been repeated $N_{\text {repeat }}=5000$ times starting from a randomized initial condition.
5. 'Hard spheres' and 'hard rotating spheres' with variable radius. To test the influence of density, and the associated jamming effects, on lane nucleation, we also simulated 'hard spheres' with variable $D=$ $0.1,0.2, \ldots, 1.2$. The number of particles is kept the same $(N=150$ in each group) and the domain sized is fixed $(L=20)$. Thus, the corresponding packing fraction

$$
\begin{equation*}
\Phi=\frac{N \pi D^{2}}{2 L^{2}} \tag{S.80}
\end{equation*}
$$

ranges from 0.006 to 0.84 . To keep the dynamics invariant and in the 'hard sphere' regime we set $(\gamma=30)$ (c.f. Eq. (S.58)) and vary the interaction strength

$$
\begin{equation*}
\alpha=\frac{\gamma v}{D}=\frac{3}{D} \tag{S.81}
\end{equation*}
$$

First set of simulations was performed for chirally symmetric interactions, and then for rotating spheres with $\beta=\alpha$.

It is worth emphasising that, even though we neglect them in our theoretical analysis, the interactions between agents of the same type are retained in the simulation. Similarly, nothing in the simulation prevents interactions involving multiple agents at the same time.

## B. Empirical Stability Inference

To quantify the evolution of the distribution of agents, we focus on the density of one of the groups

$$
\begin{equation*}
\rho^{+}(\mathbf{r}, t)=\sum_{i=1}^{N} \delta\left(\mathbf{r}-\mathbf{r}_{i}^{+}(t)\right) \tag{S.82}
\end{equation*}
$$

where $\delta$ is the Dirac delta function. In order to assess the growth of a particular Fourier mode characterized by a wave vector $\mathbf{k}$, we compute the Fourier coefficient

$$
\begin{equation*}
\tilde{\rho}^{+}(\mathbf{k}, t)=\frac{1}{L^{2}} \int_{0}^{L} \int_{0}^{L} \rho^{+}(\mathbf{r}, t) e^{-i \mathbf{r} \cdot \mathbf{k}} d x d y=\frac{1}{L^{2}} \sum_{i=1}^{N} e^{-i \mathbf{r}_{i}^{+}(t) \cdot \mathbf{k}} \tag{S.83}
\end{equation*}
$$

In the periodic domain, the density can be then represented as a Fourier series

$$
\begin{equation*}
\rho^{+}(\mathbf{r}, t)=\sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \tilde{\rho}^{+}\left(\left(\frac{2 \pi i n}{L}, \frac{2 \pi i m}{L}\right), t\right) e^{\frac{2 \pi i}{L}(n x+m y)} \tag{S.84}
\end{equation*}
$$

Thus, there is an argument that we should only consider wave vectors of the form $\mathbf{k}=\left(\frac{2 \pi i n}{L}, \frac{2 \pi i m}{L}\right)$. Nevertheless, $\tilde{\rho}^{+}(\mathbf{k}, t)$ introduced in equation (S.83) is well defined for any $k$ and for sufficiently short wavelength it can always be approximated by a periodic perturbation.

Crucially, rather than analysing the evolution of Fourier coefficients $\tilde{\rho}^{+}(\mathbf{k}, t)_{n}$ for each individual simulation (which tend to be quite noisy), for each setting we repeat the numerical experiment $N_{\text {repeat }}=5000$ times (see Section IV A) and we compute the ensemble averaged quantity

$$
\begin{equation*}
\langle\tilde{\rho}\rangle(\mathbf{k}, t)=\frac{1}{N_{\text {repeat }}} \sum_{n=1}^{N_{\text {repeat }}}\left|\tilde{\rho}^{+}(\mathbf{k}, t)_{n}\right| . \tag{S.85}
\end{equation*}
$$

To quantify their growth, we need to compute the time derivative of $\langle\tilde{\rho}\rangle$. For robustness, $\langle\dot{\tilde{\rho}}\rangle(t)$ is estimated by fitting a straight line to $\langle\tilde{\rho}\rangle(t)$ over the interval $[t-\tau, t+\tau]$, where $\tau=10$ is the smoothing parameter. In the context of linear stability analysis, it is useful to introduce the logarithmic derivative

$$
\begin{equation*}
\langle\sigma\rangle=\frac{\langle\dot{\tilde{\rho}}\rangle}{\langle\tilde{\rho}\rangle} \tag{S.86}
\end{equation*}
$$



Figure S11. (a) The growth of Fourier coefficients associated with the modes parallel to the direction of motion for different wavelengths $\lambda$ (see colorbar on the r.h.s for scale). (b) Logarithmic derivative of the Fourier coefficients. The maximal peak value $\sigma_{\text {max }}$ is attained at time $t^{*}$, which is interpreted as the 'moment of maximal growth'. The lines appear truncated as the derivative is calculated based on a centered approximation. Thus, data is not available at the ends.

Linear dynamics is then equivalent to constant logarithmic derivative, i.e.

$$
\begin{equation*}
\frac{d\langle\sigma\rangle(\mathbf{k}, t)}{d t}=0 . \tag{S.87}
\end{equation*}
$$

## 1. Results - symmetric interactions

In campaigns 1-3, where the interactions are chirally symmetric, we confine our attention to the coefficients in the $x$-direction

$$
\begin{equation*}
\tilde{\rho}^{+}((k, 0), t)=\sum_{i=1}^{N} e^{-i x_{i}^{+}(t) k} . \tag{S.88}
\end{equation*}
$$

A growth of these modes can be associated with lanes nucleating in along the direction of differential velocity. Figure S11(a) shows the time growth of $\tilde{\rho}^{+}((k, 0), t)$ for different values of $k$ for 'hard' spheres in head-on motion (from the experimental campaign 1).

Figure S11(b) demonstrates, however, that in the hard sphere simulations the lane formation process is far from linear, in that the logarithmic derivative of Fourier coefficients is not constant. For long wavelengths $\lambda=\frac{2 \pi}{k} \gg D$, the logarithmic time derivative initially increases, but then it attains a local maximum and starts to decrease. Interestingly, within the time span of our simulation, we also observe a superlinear growth $\left(\frac{d\langle\sigma\rangle}{d t}>0\right)$ of modes with wavelength $\lambda \approx D$.

Despite the manifest nonlinear properties, the linear stability theory still encodes some highly relevant aspects of the early time dynamics $\left(t<10 v D^{-1}\right)$. At this stage, the most energetic (and the fastest growing) modes are those associated with wavelength $\lambda \approx 2 D$, which coincides with the most unstable wavelength $\lambda_{\max }=2.07 D$ predicted by the linear stability theory (c.f. eq. (S.42)). A more direct comparison with the linear stability results can by found


Figure S12. Logarithmic derivative of the Fourier coefficients at $t^{*}$ - the moment of maximal linear growth for head-on collisions of driven 'soft spheres' with parameter softness $\alpha$. The different panels correspond to different data points of the simulation campaign 1. The simulation-inferred $\sigma\left(k, t^{*}\right)$ is plotted in magenta and the black lines are the theoretically predicted growth rate $\sigma(k)$.
by defining

$$
\begin{equation*}
\left\langle\sigma_{\max }\right\rangle=\langle\sigma\rangle\left(\left(k^{*}, 0\right), t^{*}\right)=\max _{k, t<T}\langle\sigma\rangle((k, 0), t) \tag{S.89}
\end{equation*}
$$

Wavenumber $k^{*}$ can be plausibly associated with the most unstable pattern and time $t^{*}$ can be interpreted as a moment when the pattern growth is at its premium. If we focus on this particular moment in time, hereafter referred to as 'the moment of maximal growth', we find a close relationship between $\langle\sigma\rangle\left((k, 0), t^{*}\right)$ and the theoretical growth rate $\sigma(k)$ from equation (S.67). Figure 2(a) of the article shows that this relationship holds nor only for head-on collisions of 'hard spheres', but also for oblique counterflows, as well as 'hard ellipses'. As discussed in the article, $\langle\sigma\rangle\left((k, 0), t^{*}\right)$ follows the same trend as $\sigma(k)$, but it is consistently lower. Figure S 12 shows that the same is true for soft spheres, where the theoretical growth rate $\sigma(k)$ is not given by an analytical formula, but it can be computed numerically.


Figure S13. Empirical stability results for campaign 4 (rotating spheres). (a) The circles show the maximal momentary growth rate, as defined in eq. (S.90), measured in the simulation for different values of the swirling strength $\beta$ ( $\alpha=10$ is fixed). The dashed line shows the maximal growth rate predicted by the linear stability analysis (c.f. Fig. S10). (b) The corresponding 'moment of maximal growth' $t^{*}$ when the maximum of the logarithmic growth is attained is of order 10 turnover times $D v^{-1}$. (c) The orientation of the empirically most unstable mode (circles) is consistent with analytical estimate (dashed line). (d) The wavelength of the empirically most unstable mode (circles) agrees with the theoretical prediction (dashed line) only for $\beta / \alpha<0.5$.

## 2. Results - chiral interactions

An analogous stability inference procedure is used in campaign 4, where we simulate the 'rotating spheres' with broken chiral symmetry. Nevertheless, this time we expect a non-trivial pattern orientation, which is why we consider a range of wave vectors $\mathbf{k}$ with different magnitude and direction. Based on exploratory computations, for the final computations we consider wave vectors of the form $\mathbf{k}=\frac{2 \pi}{\lambda}(\cos \theta,-\sin \theta)$ for $\lambda \in[0,20 D]$ and $\theta \in[-.05, .2]$ and in each simulation we compute $\tilde{\rho}^{+}(\mathbf{k}, t)$, as defined in eq. (S.83), for 100 values of $\lambda$ and 100 values of $\theta$. The results are then ensemble averaged over the 5000 numerical realisations and we search for

$$
\begin{equation*}
\left\langle\sigma_{\max }\right\rangle=\langle\sigma\rangle\left(\frac{2 \pi}{\lambda^{*}}\left(\cos \theta^{*},-\sin \theta^{*}\right), t^{*}\right)=\max _{\mathbf{k}, t<T}\langle\sigma\rangle(\mathbf{k}, t) \tag{S.90}
\end{equation*}
$$

This peak value of the logarithmic growth rate $\sigma_{\max }$ for different values of the chirality measure $\beta / \alpha$ is shown in Figure S13(a). This maximum is attained at $t^{*} \in[6,13] D v^{-1}$ without any systematic dependency on $\beta / \alpha$ (Figure $\mathrm{S} 13(\mathrm{~b})$ ). The non-monotonic relationship between $\beta / \alpha$ (Fig. S13(a)) and the empirically measured growth rate $\sigma_{\max }$ is qualitatively similar to the trend which we found for the theoretical maximal growth rate (dashed line, c.f. Fig. S10(f)). Nevertheless, the results are quantitatively different, particularly for strongly chiral systems where the empirically observed maximum of the growth rate is much larger than the theoretical estimate.

Two robust consequences of broken chiral symmetry predicted by the linear stability analysis, which are manifest in the simulation are the non-trivial pattern orientation (Fig. S13(c)), as well as an increased width of nucleating lanes (Fig. S13(d)). However, even here the quantitative agreement is only partial.

More clues about this apparent disagreement can be found in Figure S14, where we show the temporal growth rate at the moment of maximal growth $t^{*}$ for a range of wave vectors $\mathbf{k}$. These heat maps are a 2 D analogue of Fig. S12 for symmetric interactions, where a single parameter sufficed to parameterise all the perturbations of interest. They can also be directly compared with the linear stability results shown in Fig. S10(a-d) (note that in our simulations $\alpha=10$, so the relative swirling strength ranges between 0 and 2 ).

The linear stability diagrams and the empirical findings of Figure S14 differ quite significantly. First of all, even though in the simulation we observe strong amplification of certain Fourier modes, we rarely measure strong damping.


Figure S14. Logarithmic derivative of the Fourier coefficients at $t^{*}$ (the moment of maximal linear growth) for perturbations of varied orientation $\theta$ and wavelength $\lambda$. The different panels correspond to different data points of the simulation campaign 4. The most unstable perturbation is indicated with a with cross. As the 'swirling strength' $\beta$ is increased and the asymmetry increases, the most unstable direction departs from $\theta=0$ and the most unstable wavelength increases.

Consequently, the neutral stability curve $(\sigma(\lambda, \theta)=0)$ is not easy to infer in Figure S14. Having said that, for sufficiently large $\sigma\left(t^{*}\right)>0$ we could still distinguish contours of $\sigma$ which could locate the 'highly unstable regions' in the $\lambda-\theta$ space. In Fig. S14, as well as Fig. S10(a-d), these regions manifest themselves as intensely red patches. Their shape, however, is markedly different in theory and in the simulations. The linear stability diagrams of Fig. S10(a-d) show highly unstable regions which are approximately symmetric in $\theta=\theta^{*}$, which is clearly not the case in Fig. S14 where the highly unstable region appears 'slanted'. One could hypothesise that this particular shape is a consequence of the fact that at the linear level $\theta_{\max }$ is proportional to the agent density (c.f. Section below). Thus, as the density inhomogeneities grow, we may observe a lanes at different orientation, which visually indeed appears to be the case in the simulations. The fact that the most unstable mode is density-dependent could also explain why, despite robust qualitative predictions, the linear stability theory is less successful in capturing quantitatively the dynamics of systems with broken chiral symmetry.

## 3. Results - density effects

The kinetic model presented in Sec. I assumes that the dynamics is dominated by binary collisions. While the assumption holds true for dilute systems, as the agent density increases, multi-body interactions may inhibit lane formation and ultimately lead to jamming. Such effects are not captured by our theory, by we can estimate the range


Figure S15. The density-induced inhibition of lane nucleation. (a) The circles in red show the maximal momentary growth rate, as defined in eq. (S.90), measured in the simulation of hard spheres for different values of packing fraction $\Phi$ (Campaign 5 ). For sparse mixtures, the growth rate increases linearly with low values of $\Phi$ as predicted by the kinetic theory (dashed line, c.f. eq. (S.93)), but as the density increases multi-body interactions and jamming effects suppress lane formation. (b) The orientation of the empirically most unstable mode $\theta^{*}$ (c.f. eq. (S.90)) aligns with the direction of differential velocity up for $\Phi<0.6$, in agreement with the results of Reichhardt and Reichhardt [42] who in the hard sphere limit identified a transition from a laned state with multiple lanes to a 'phase separated' state with two groups at $\Phi=0.55$, and a jamming transition at $\Phi=0.7$ (as indicated by the background colouring). (c) For rotating hard spheres, the most unstable angle is density-dependent. The theoretical estimate (dashed line, c.f. eq. (S.94)) agrees with the numerical results for $\Phi<0.2$, but as the density increases it underpredicts the lane tilt.
of validity of our kinetic equations, by virtue of numerical simulations. For hard spheres, the most natural measure of density is the packing fraction

$$
\begin{equation*}
\Phi=\frac{\pi D^{2}}{2} \rho \tag{S.91}
\end{equation*}
$$

where $D$ is the particle diameter, and $\rho$ is the (single-species) number density of one agent group. By convention, $\Phi$ accounts for both groups, which are assumed to be equal in size. Thus, $\Phi=1$ corresponds to complete coverage, $\Phi=\frac{\pi \sqrt{3}}{6} \approx 0.9$ to the hexagonal close packing, $\Phi \approx 0.84$ to the random close pack [52]. Reichhardt and Reichhardt, who studied this system numerically, identified three possible states in the hard sphere limit: lanes for $\Phi<0.55$, 'phase separated' state with two groups for $0.55<\Phi<0.7$, and a jammed state form $\Phi>0.7$ [42].

In the main article, we found an approximate expression for the maximal growth rate

$$
\begin{equation*}
\sigma_{\max } \approx v \rho_{0} \mu \tag{S.92}
\end{equation*}
$$

where $\mu=\sqrt{2\left\|G_{x}\right\|_{1}}$, which for hard spheres simplifies to

$$
\begin{equation*}
\sigma_{\max } \approx v \rho_{0} D=\frac{2 v \Phi}{\pi D} \tag{S.93}
\end{equation*}
$$

i.e. the linear theory predicts that for fixed particle size the lane nucleation rate increases linearly with packing fraction. Figure $\operatorname{S15}$ (a) shows that this is indeed the case for sufficiently small $\Phi$, but for $\Phi>0.2$, the empirically measured lane nucleation rate is dampened. We associate this effect with jamming and multi-body interactions. It is important to emphasise, that for $\Phi>0.2$ lanes still form, but slower that predicted by the simple kinetic theory. Indeed, Fig. S15(b) shows that $\theta=0$, which corresponds to lanes parallel to the direction of motion, remains the dominant pattern orientation up to $\Phi \approx 0.6$, in agreement with the finding of Reichhardt and Reichhard [42] who identified $\Phi=[0,0.55]$ as the interval where the 'laned state' is observed in the hard sphere limit. For $\Phi>0.6$ we observe that the most unstable mode is not the one parallel, but perpendicular to the direction of motion. While our theory does not offer any analytical insights in the jammed regime, this bifurcation is strongly reminiscent of the transition between 'dynamic lanes' and 'jammed bands' observed by Leunissen et al. in an experiment on driven colloids [10].

For the chirally biased systems, we predict a density-dependent pattern orientation

$$
\begin{equation*}
\theta_{\max }=\tan ^{-1}\left(\frac{4 \tilde{A}_{x}(0) \Phi}{\pi D^{2}}\right) \tag{S.94}
\end{equation*}
$$

(c.f. eq.(30)). Numerical simulations show that for an example chiral system ('rotating spheres model', with $\gamma=30$ and $\alpha=\beta$, c.f. Sec. III C) this expression approximately predicts the most unstable pattern orientation for $\Phi \leq 0.6$, with a particularly close agreement for $\Phi<0.2$. Thus, the validity range of equation (S.94) for the most unstable patter orientation in systems is similar to the validity range for the pattern growth rate (eq. (S.93)) in non-chiral hard spheres systems, indicating that $\Phi=0.2$ is the packing fraction where the multi-body interactions start to have notable dynamical effects for driven hard spheres.

## V. DATA-DRIVEN AB MODEL

Apart from the differential schemes, we also simulated a data-driven AB pedestrian model, which has been constructed as follows. First, we compiled two libraries of pedestrian-pedestrian interactions based on the experimental data. Then, we used them to bootstrap the empirical data to perform a surrogate simulation starting from a homogeneous state, even thought the experiment itself followed the 'release' protocol. The technical details of the empirical collision inference can be found in Section VII C and the details of the surrogate simulations are described here below.

## A. Even-based simulation

The pseudocode of the surrogate data-driven AB simulation is presented in a box called Algorithm 1. It is important to emphasise that it does not use the averaged collisional displacement presented, but a full dataset comprising empirically observed initial offsets

$$
\begin{equation*}
\operatorname{model} . x=\left(x^{1}, x^{2}, \ldots, x^{N_{m}}\right) \tag{S.95}
\end{equation*}
$$

and the corresponding lateral displacements

$$
\begin{equation*}
\operatorname{model} . G_{x}=\left(G_{x}^{1}, G_{x}^{1}, \ldots, G_{x}^{N_{m}}\right) \tag{S.96}
\end{equation*}
$$

Here we test the algorithm with two different empirical libraries of collisions - one based on our own lane formation experiments (see Scenario 1 in Sec. VIIB) and the other based on the publicly available datasets of Murakami et al. [44] (Fig. S16). The experimental protocols of both experiments are very similar; two groups of pedestrians (30-36 people each in our experiment and 26 people each in Ref. [44]) are released from the 'waiting zones' and head in the opposite directions. One notable difference is that in our experiments the experimental arena is a square $(5.8 \mathrm{~m} \times$ 5.8 m ) and in Ref. [44] it is a rectangle elongated in the direction of motion ( $10 \mathrm{~m} \times 3 \mathrm{~m}$ ). Our experiments included 10 trials and in Murakami et al. the same experiment was repeated 12 times and we use the data from all these trials.

It is also worth emphasising that we do not model collisional displacement in the $y$-direction, which is because it is difficult to extract this quantity from the data. Nevertheless, Algorithm 1 can be adapted to account for the collision-induced retardation as well.


Figure S16. Two different libraries of pedestrian interactions which we use as input to our data-driven simulations. Each panel shows the lateral component of the collisional operator $G_{x}$ extracted from a different set of human crowd experiments. Each collision corresponds to one of the scattered data points (gray), the inferred mean $\mathbb{E} G_{x}(x)$ is presented as a solid line, and the shading corresponds to the standard deviation. (a) The library of collisions constructed based on our own experiments (Scenario 1) by fixing the collision inference parameter $D_{\text {coll }}=1 \mathrm{~m}$ (Sec. VII C). (b) The library of collisions constructed with the same method based on the dataset accompanying Ref. [44].


## B. Lane nucleation in the surrogate simulation

As expected, simulations based on the surrogate event-based algorithm also lead to lane formation. To allow for comparison with the numerical simulations of the differential AB models (Sec. IV), we performed data-driven simulations with the same number of agents in each species $N=150$. This time the collisional displacement is measured in meters and we can ascribe the same units to the computational domain, which is a doubly-periodic square domain with side length $L=20 \mathrm{~m}$. Notably, in an event-based simulation the base velocity of agents $v$ does not change the sequence of agent-agent interactions and by increasing $v$ we merely increase their apparent rate of occurrence. The time-step of an even-based simulation is variable and it is dictated by the density of agents and their relative position. Thus, the only relevant temporal parameter of our simulation is the total simulation time in relation to the characteristic time scale of agent motion, here set to $10 \mathrm{v} D^{-1}$, where $D=D_{\text {coll }}=1 \mathrm{~m}$.

As in the case of differential models, for each of the two data-driven models we repeated the simulation 5000 times from a randomized initial condition (initial position of each agent is uniform and independent of the other agents) and we used the same methodology as in Sec. IV B to quantify the growth rate of nucleating lanes. Figure S17, which is exactly analogous to Figure S11, shows the time evolution of the ensemble averaged Fourier coefficients and the corresponding logarithmic derivative which quantifies their growth for the simulation for the the empirical collisional operators (Fig. S16). By searching for a peak in the logarithmic derivative of the dominant wavelength of $\lambda \approx 1.2 \mathrm{~m}$, we can compare the empirical measure of growth rate at the time of maximal growth $t^{*}(\langle\sigma(t *)\rangle$ to the theoretically


Figure S17. Lane formation in data-driven simulations. (a) The average growth of Fourier coefficients associated with the modes parallel to the direction of motion for different wavelengths $\lambda$ (see colorbar in panel (b) for scale) for the data-driven simulations based on our own dataset. (b) The average growth of Fourier coefficients in the simulations based on the dataset extracted from Murakami et al. [44] is slightly lower. (c) Logarithmic derivative of the Fourier coefficients for the data-driven simulations based on our own dataset. The maximal peak value $\sigma_{\max }$ is attained at time $t^{*}$, which is interpreted as the moment of maximal growth. (d) The same quantity for the data-driven simulations using the library of collisions extracted from the dataset of Murakami et al. [44]. The lines in panels (c) and (d) appear truncated as the derivative is calculated based on a centered approximation. Thus, data is no available at the ends.
predicted growth rate $\sigma(k)$.
Figure S18 shows that the dominant wavelength $\lambda \approx 1.2 \mathrm{~m}$ is similar in both of our data-driven models, pointing to some reproducible characteristics of pedestrian motion. The perturbation growth rate is larger in the simulations based on our own data, which is probably related to a larger amplitude of the collisional displacement (Fig. S16). Nevertheless, in both cases the maximal momentary growth measured in the simulation is in good agreement with the theoretical instability rate computed from the statistics of the empirical collisional operator through equation (S.31).

## VI. NUMERICAL SIMULATIONS OF COMPLEX FLOWS

In the final part of the article we consider generalized scenarios with pedestrians heading towards (or away from) gates and we showed that in this case lanes nucleate along curved conic sections. To illustrate this phenomenon we used numerical simulations of appropriate differential AB models. In this Section we discuss the set-up simulations (Sec. VI A) and we describe the network-based technique of detecting the germs of nucleating lanes (Sec. VIB).

## A. Simulation details

Simulating pedestrian flows with point targets is more challenging than simulating pedestrian flows with direction targets for at least two reasons. First, point attractors and repellers are most naturaly defined in unbounded, rather than periodic, domains. Second, agents accumulate and jam near their target points. To avoid spurious results


Figure S18. Logarithmic derivative of the Fourier coefficients at $t^{*}$ - the moment of maximal linear growth for the two data driven models compared to the theoretical growth rate $\sigma(k)$ (solid lines) computed according to eq .(31) using the statistics of the empirically-inferred collisional operators.
associated with these effects, we limit the duration of each simulation and we implement some additional features, such as agent removal in the vicinity of the gates.

We will now describe the details of the simulations for the three canonical scenarios. In all of these simulations we use $N=200$ agents of each type with preferred speed $v=0.1$, interacting through a 'hard' sphere interaction force given by equation (S.57) with $\alpha=10$ and $D=0.3$. Thus, the average density of agents is now larger than in the simulations of Section IV, but we use the same parameters for the hard sphere inter-agent force as in 'Campaign 2'. We also use the same time-step $\Delta t=0.05$ and the same total simulation time $T=100$ and all the simulations are initialized with a random (uniform and independent) distribution of agents. Similarly to the previous simulations, the computational domain is a square of side $L=20$ with coordinates $(x, y) \in[0, L]^{2}$, but the boundary conditions require some modifications for each one of the scenarios.

## 1. Two point targets (ellipse)

The two attractive points (gates) are located at $\mathbf{f}^{+}=(x=5, y=10)$ and $\mathbf{f}^{-}=(x=15, y=10)$, respectively. Thus, given the assumed hard sphere potential, the agents obey the following equation of motion

$$
\begin{equation*}
\frac{d \mathbf{r}_{i}^{ \pm}}{d t}=0.1 \frac{\mathbf{f}^{ \pm}-\mathbf{r}_{i}^{ \pm}}{\left|\mathbf{f}^{ \pm}-\mathbf{r}_{i}^{ \pm}\right|}+\sum_{ \pm} \sum_{j=1}^{N=200} \mathbb{1}\left[\left|\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right|>0\right] 10 \max \left[0.3-\left|\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right|, 0\right] \frac{\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}}{\left|\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right|} \tag{S.97}
\end{equation*}
$$

To alleviate jamming near the gates, we implement circular exclusion zones of unit radius near each one of the gates. Thus, if at some point in the simulation $\left|\mathbf{r}_{i}^{ \pm}(t)-\mathbf{f}^{+}\right|<1$ or $\left|\mathbf{r}_{i}^{ \pm}(t)-\mathbf{f}^{-}\right|<1$, then instead of evolving $\mathbf{r}_{i}^{ \pm}(t)$ according to equation (S.97), the corresponding agent will be 'teleported' to a random location, in that $\mathbf{r}_{i}^{ \pm}(t+\Delta t)$ will be chosen uniformly at random. Note that in this simulation the distances between agents are computed in a toroidal sense, but the distance towards the gate is not. This could potentially lead to an anomalous behaviour near the boundaries, but due to the attraction, the agents almost never cross the boundary.

## 2. A target and a repeller (hyperbola)

The two distinguished points are located at $\mathbf{f}^{ \pm}$as before, but this time $\mathbf{f}^{-}$is a point of repulsion, in that the agents satisfy

$$
\begin{equation*}
\frac{d \mathbf{r}_{i}^{ \pm}}{d t}= \pm 0.1 \frac{\mathbf{f}^{ \pm}-\mathbf{r}_{i}^{ \pm}}{\left|\mathbf{f}^{ \pm}-\mathbf{r}_{i}^{ \pm}\right|}+\sum_{ \pm} \sum_{j=1}^{N=200} \mathbb{1}\left[\left|\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right|>0\right] 10 \max \left[0.3-\left|\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right|, 0\right] \frac{\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}}{\left|\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right|} \tag{S.98}
\end{equation*}
$$

As before, we impose and exclusion zone and a teleportation mechanism within a unit circle of the attractive gate $\mathbf{f}^{+}$. Moreover, to prevent spurious accumulation of the (-)-type agents near the boundaries, we implement an additional exclusion margin of unit width near the boundaries of the simulation domain.

## 3. A point target and a direction target (parabola)

In this simulation the $(+)$ agents move have a preferred velocity along the x -axis, i.e.

$$
\begin{equation*}
\frac{d \mathbf{r}_{i}^{+}}{d t}=-0.1 \mathbf{e}_{x}+\sum_{ \pm} \sum_{j=1}^{N=200} \mathbb{1}\left[\left|\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right|>0\right] 10 \max \left[0.3-\left|\mathbf{r}_{i}^{+}-\mathbf{r}_{j}^{ \pm}\right|, 0\right] \frac{\mathbf{r}_{i}^{+}-\mathbf{r}_{j}^{ \pm}}{\left|\mathbf{r}_{i}^{+}-\mathbf{r}_{j}^{ \pm}\right|} \tag{S.99}
\end{equation*}
$$

and the $(-)$ agents target a gate located at $\mathbf{f}^{-}=(x=15, y=10)$

$$
\begin{equation*}
\frac{d \mathbf{r}_{i}^{-}}{d t}=0.1 \frac{\mathbf{f}^{-}-\mathbf{r}_{i}^{-}}{\left|\mathbf{f}^{-}-\mathbf{r}_{i}^{-}\right|}+\sum_{ \pm} \sum_{j=1}^{N=200} \mathbb{1}\left[\left|\mathbf{r}_{i}^{ \pm}-\mathbf{r}_{j}^{ \pm}\right|>0\right] 10 \max \left[0.3-\left|\mathbf{r}_{i}^{-}-\mathbf{r}_{j}^{ \pm}\right|, 0\right] \frac{\mathbf{r}_{i}^{-}-\mathbf{r}_{j}^{ \pm}}{\left|\mathbf{r}_{i}^{-}-\mathbf{r}_{j}^{ \pm}\right|} . \tag{S.100}
\end{equation*}
$$

Near this gate we use the same teleportation mechanism are in scenario 1.

## B. Automated detection of incipient lanes

Incipient lanes nucleating along curved pattern lines manifest themselves as small spatially elongated chains of agents aligned with the expected direction of laning instability. To detect such chains in a simulation for each agent type we construct a time-evolving proximity network $\mathcal{G}^{ \pm}$with $N$ nodes corresponding to the $N$ agents. Let $D_{i j}^{ \pm}$be a matrix of pairwise distances between the agents, then the adjacency matrix $A^{ \pm}$of our proximity network is given by

$$
A_{i j}^{ \pm}= \begin{cases}1 & D_{i j}^{ \pm} \leq d_{0}  \tag{S.101}\\ 0 & D_{i j}^{ \pm}>d_{0}\end{cases}
$$

where $d_{0}$ is a threshold parameter that we set to $d_{0}=0.45=\frac{3}{2} D$.
An example proximity network is shown in Figure S19. It consists of several connected components and each of them is associated with one cluster of agents. Let $\mathcal{C}_{1}, \ldots, \mathcal{C}_{k(t)}$ be the clusters with at least two members from both proximity networks $\left(\mathcal{G}^{+}\right.$and $\left.\mathcal{G}^{-}\right)$. Note that the number of clusters $k(t)$ depends on the current positions of all the agents and it can vary in time. The clusters often assume a chain-like spatial arrangement, whose orientation can be quantified by treating all the agents in one cluster as a cloud of points and computing its first principal component. Mathematically, this first principal component $\mathbf{o}_{i}$ is the leading eigenvector of the covariance matrix $\operatorname{cov}\left(X_{i}, Y_{i}\right)$, where $X_{i}$ is the vector of x-coordinates of all the members of $\mathcal{C}_{i}$ and $Y_{i}$ is the vector of their y-coordinates. The corresponding eigenvalue $e_{i}$ is a measure of slenderness of the corresponding cluster.

The procedure of finding clusters is repeated repeated periodically at $t=1,2, \ldots, 100$, which provides a large data set of clusters observed at different locations, here the location of a cluster $\mathbf{r}_{i}$ is simply the average position of all its members. As a final step of our analysis we average our data by introducing a $20 \times 20$ grid tesselating the simulation into $1 \times 1$ squares and we use it to compute the average orientation at a lattice point $\left\langle\mathbf{o}_{n m}\right\rangle$. This average orientation is based on all the orientation vectors $\mathbf{o}_{i}$ (recorded at different times) such that the corresponding cluster position $\mathbf{r}_{i}$ falls within the appropriate grid domain. Specifically, $\left\langle\mathbf{o}_{n m}\right\rangle$ is the first principal component of the augmented dataset comprising each $\mathbf{o}_{i}$, as well as $-\mathbf{o}_{i}$. This way way of defining $\left\langle\mathbf{o}_{n m}\right\rangle$ makes it independent of the arbitrary sign of $\mathbf{o}_{i}$. The orientation of the $\left\langle\mathbf{o}_{n m}\right\rangle$ for the three simulation kinds of this section (each repeated 50 times) is presented in Fig. S20 with the help of red 'needles'. This length of each needle is a monotonic function of the grid-averaged slenderness $\langle e\rangle$ to highlight the locations where the lane nucleation is more prominent.

## VII. PEDESTRIAN EXPERIMENTS

The pedestrian experiments were conducted in the sports hall at the Academy of Physical Education in Katowice and the participants were volunteers recruited among students and staff. The experiments were conducted in two separate sessions - one with 60 and one with 73 participants, randomly assigned into two groups. In Session 1 the groups were equal; in session two one of them comprised 36 , and the other 37 people. All participants have signed a written participation consent and they have been informed about the nature of the experiments. Nevertheless, they have not been informed about the specific scientific objective and the phenomenon of lane formation has not been explicitly mentioned or suggested.


Figure S19. A snapshot from a simulation with two point targets showing the links in the proximity network. The connected components with more than two elements (full circles) frequently form chains aligned with the dominant direction of lane nucleation.

## A. Data capture

The experiments were recorded with a GoPro HERO 9 camera suspended 9.5 m above the centre of the experimental arena. The camera recorded image with resolution $3540 \times 2160$ at a frequency of 30 Hz .

The participants wore square hats (mortarboards) adorned with distinct Apriltag [53] codes (selected from the $36 h 11$ family). The fiducial system was then used for automated tracking of pedestrians utilising an in-built Matlab algorithm. With appropriate calibration, the inferred head position was then translated into the real-world floor coordinates.

## B. Experimental protocols

The experimental arena was a $5.8 \mathrm{~m} \times 5.8 \mathrm{~m}$ square with corners marked out by plastic poles (Fig. S21(a)). To reduced possible 'wall effects', the boundary of the arena was only an indication and if an avoidance manoeuvre required moving outside the participants could cross the boundary. In all trials the participants were instructed to


Figure S20. The mean grid-averaged orientation of agent chains $\left\langle\mathbf{o}_{n m}\right\rangle$ measured for the different simulation scenarios of this section is shown with the red 'needles'. Their length is an affine function (the same for all panels) of the grid-averaged slenderness $\langle e\rangle$. (a) Two targets. (b) A target and a repeller.(c) A point target and a direction target. In all cases the average orientation measured in the simulation is in qualitative agreement with the theoretically predicted conic section pattern lines, here shown in black.
move at a normal walking speed while maintaining safe distance and avoiding collisions. Ahead of each trial, the participants would gather in an unstructured manner (no pre-defined lanes) behind an appropriate starting line and they would start walking towards a specified target when prompted by the experiment marshal.

Each experimental session included 5-6 repetitions of the following exercises

1. Simple head-on cross-streams. The two groups started at the opposite ends of the arena and they were instructed to cross to the other side (Fig. S21(a)). This experiment reproduced the formation of archetypal lanes aligned with the dominant direction of motion.
2. Chirally biased cross-streams. The two groups started at adjacent sides of the arena (as in (Fig. S21(a)) and they were instructed to cross to the other side while maintain 'pass on the right' traffic rule. This experiment reproduced the formation of 'tilted lanes'.
3. Simple perpendicular cross-streams. The two groups started at adjacent sides of the arena and they were instructed to cross to the other side (Fig. S21(b)). This experiment led to the formation of diagonal travelling lanes.
4. Two exit gates. The two groups started at adjacent sides of the arena (spanning their entire length) and they


Figure S21. Experimental set-up. (a) Two groups gathering before the start of a head-on cross-stream experiment. The boundary central experimental arena is marked with a dashed line. (b) The beginning of a perpendicular cross-stream experiment. (c) An experiment with two exit gates marked out with plastic poles. (d) An experiment with one exit (red) gate and one entry gate (blue on the left). The members of the blue group are divided into two subgroups targeting the two wide exit gates on the opposite side (blue on the right). (e) An experiment with one exit gate (blue group) and a stream (red group).
were instructed to enter 'gate' located at the centre of the other side (Fig. S21(c)). The gate was marked with two plastic poles approximately 1 m from each other. This experiment led to elliptical lane formation.
5. Entry gate + exit gate. One group gathered along one of the boundary lines of the square arena and was instructed to move towards an exit gate on the opposite side, as in the previous scenario (Fig. S21(d)). The other group gathered behind an entry gate located at an adjacent side and its members were divided into two subgroups. Both subgroups were instructed to go through the narrow entry gate, but once they passed it, they were supposed to move towards one of the two wide (approximately 2 m ) exit gates located on the opposite side. This scenario was only an approximate realisation of a repulsive point, but it provided the proof-of-concept of hyperbolic lane formation in a realistic traffic scenario.
6. Exit gate + stream. The two groups started at adjacent sides of the arena (Fig. S21(e)). One was instructed to cross to the other side (as in the 'simple cross-stream' experiment) and the other was moving towards a narrow exit gate on the other side (as in the 'two exit gates' experiment). This experiment led to parabolic lane formation.

## C. Empirical collisional operator

The trajectory data from Scenarios 1 and 2 can be used to infer what the collisional operator may look like in a pedestrian experiment (see Sec. I for the definition).

Even though in our experiments participants do not come into physical contact, as explained in the manuscript, we refer to their interactions as 'collisions' in keeping with the theoretical framework, which assumes that the interactions are short compared to the timescale of macroscopic evolution. We identify the start of a collision with a moment in time $t_{0}$ when two pedestrians with opposite target directions come within distance $D_{\text {coll }}$ of each other, i.e.

$$
\begin{equation*}
t_{0}=\min _{t}\left\{t:\left|\mathbf{r}^{+}(t)-\mathbf{r}^{-}(t)\right| \leq D_{\text {coll }}\right\} \tag{S.102}
\end{equation*}
$$

and the end of a collision with a moment in time when they pass each other shoulder to shoulder, i.e.

$$
\begin{equation*}
t_{1}=\min _{t}\left\{t: y^{+}(t)=y^{-}(t)\right\} \tag{S.103}
\end{equation*}
$$

(see Fig. S22). We then compare their initial lateral offset $x_{0}=x^{+}\left(t_{0}\right)-x^{-}\left(t_{0}\right)$ and with their final lateral offset $x_{1}=x^{+}\left(t_{1}\right)-x^{+}\left(t_{1}\right)$ and we define the empirical (lateral) collisional displacement as

$$
\begin{equation*}
\bar{G}_{x}\left(x_{0}\right)=\frac{1}{2}\left(x_{1}-x_{0}\right) . \tag{S.104}
\end{equation*}
$$

The interaction distance $D_{\text {coll }}$ is a parameter of our data mining technique and it is a priori unclear what value is appropriate, so in Fig. S 23 we compare the estimated collisional operator for $D_{\text {coll }}$ ranging from 0.5 m to 1.5 m . As we


Figure S22. Example pedestrian trajectories (blue) extracted from one particular repetition of Scenario 1 with 73 participants, plotted in the frame of reference of a pedestrian localized at the origin. The pedestrians must alter their motion in order to avoid a collision and the collisional displacement can be computed by comparing their lateral offset $x_{1}$ when they pass each other shoulder-to-shoulder and their lateral offset $x_{0}$ when they were $D_{\text {coll }}$ away (the red trajectory is distinguished only for demonstration). The trajectories in this Figure correspond comes from one trial and the density of trajectories reflects the lane structure which emerged in this particular realisation. Our complete data set is approximately 10 times larger, but only a subset of the trajectories is plotted for the sake of clarity.
are interested in isolated binary interactions, arguably $D_{\text {coll }}$ should not be longer than a few steps, but for small $D_{\text {coll }}$ we may not have enough data to correctly infer $G_{x}(x)$ for small $|x|$ (Fig. S23(a-b)). On the other hand, for large $D_{\text {coll }}$ the collective dynamics of nucleating lanes has spurious influence on our estimate of the collisional operator. These are particularly prominent for the chiral experiments (Scenario 2), where $G_{x}(x)$ seems to be positive for all values of $x$ due to the global rightward drift (Fig. S23(f)).

Independent of the parameter choice, we observe qualitative differences between Scenario 1 (canonical laning experiment) and Scenario 2 (chiral dodging). As anticipated, in the first case $G_{x}(x)$ is odd about zero, but for the biased dodging manoeuvres the symmetry is broken, demonstrating that marshals' instructions have been effective in changing the nature of pedestrian interactions.

## D. Pedestrian density Scenarios 1 and 2

Due to the transient nature of the experiment the number of people visible within the experimental arena ( $5.8 \mathrm{~m} \times$ $5.8 \mathrm{~m})$ changes in time. Figure S24 shows how it evolves in time for Scenarios 1 and 2. For both experimental sessions during the maximum number of people visible within is about 40 , suggesting that with more participants the crowd does not spontaneously increase density, but simply forms longer lanes.

In Session 2 the time necessary to complete the trial was markedly longer for the chiral experiments (Fig. S24(b)). The same trend was also observed in Session 1, with fewer participants, although the effect is not that prominent (Fig. S24(a)). This is in agreement with the fact that misalignment between the direction of motion and the orientation of the nucleating lanes may lead to individual frustration and an lengthening of pedestrian trajectories, thereby reducing the speed of motion. Nevertheless, a thorough experimental investigation with larger groups of people would be necessary to test the generality of our hypothesis and eliminate other effects, such as fatigue or learning.

## E. Statistical analysis of tilting lanes

To quantify and compare the orientation of lanes in Scenarios 1 and 2 we used the following automated procedure.
For each trial we computed pedestrian trajectories $\mathbf{r}_{i}^{ \pm}\left(t_{j}\right)$ by sampling time with frequency 10 Hz and confining our attention to the central square area of the experiment. We then computed the Fourier coefficients

$$
\begin{equation*}
\tilde{\rho}^{ \pm}(\lambda, \theta)=\sum_{i, j} \exp \left[\frac{2 \pi i}{\lambda} \mathbf{e}_{\theta} \cdot \mathbf{r}_{i}^{ \pm}\left(t_{j}\right)\right] \tag{S.105}
\end{equation*}
$$

where $\mathbf{e}_{\theta}=(\cos \theta, \sin \theta)$ for a range of orientation angles $\theta \in[0, \pi]$ and wavelengths $\lambda \in(0,2.9) \mathrm{m}$. We then searched


Figure S23. Empirical collisional operator inferred from the experimental data for different values of $D_{\text {coll }}$ ranging from 0.5 m to 1.5 m . The panels on the left present complied data from all the Scenario 1 experiments (classical laning) and the panels on the right present data from the 7 repetitions of Scenario 2 (chirally biased dodging). For Scenario 1 the data comes from 10 trials. Each black dot corresponds to one pairwise interaction, the solid red lines present the mean $\mathbb{E} G_{x}(x)$ which was computed by dividing the data into 71 equisized bins, and the shading corresponds to one standard deviation.
for the dominant orientation defined as

$$
\begin{equation*}
\theta_{\max }=\max _{\theta} \sum_{\lambda}\left[\tilde{\rho}^{+}(\lambda, \theta)-\tilde{\rho}^{-}(\lambda, \theta)\right]^{2} . \tag{S.106}
\end{equation*}
$$

In total, we computed $\theta_{\text {max }}$ for 10 trials of scenario 1 ( 5 from involving 60 participant and 5 involving 73 participants) and 7 trials of scenario 2 ( 2 involving 60 participants and 5 involving 73 participants) and its statistics are shown in Fig. 3(c) of the article. As we did not observe strong dependencies on the number of participants, we collate the results of the two sessions in one graph.

## F. A priori estimate of the tilting angle

One may ask whether the tilting angles we observe are consistent with the theoretical prediction of eq. (30). To this end, we need to compute the average of the collisional operator

$$
\begin{equation*}
\tilde{A}_{x}(0)=\int \mathbb{E} G_{x}(x) d x \tag{S.107}
\end{equation*}
$$

as well as the characteristic pedestrian density.
Unfortunately, the value of $\tilde{A}_{x}(0)$ depends on the way the empirical collisional operator is constructed. However, to get a ballpark estimate we use the methodology of Sec. VII C with parameter value $D_{\text {coll }}=1 \mathrm{~m}$ (Fig. S23(d)), which yields $\tilde{A}_{x}(0) \approx 0.1 \mathrm{~m}^{2}$. As discussed in Sec. VIID characteristic agent density $\rho_{0}$ is also not straightforward to define due to the transient character of the experiment. Nevertheless, as discussed in Sec. VIID the maximal densities we observe are at the order of

$$
\begin{equation*}
\rho_{\max }=\frac{20}{5.8 \times 5.8} \mathrm{~m}^{-2} \approx 0.6 \mathrm{~m}^{-2} . \tag{S.108}
\end{equation*}
$$



Figure S24. Number of people within the experimental arena as a function of time. Solid lines correspond to the mean across trials of the same type and the shadings represent standard deviation. In order to compare different trials, in each of them $t=0$ is defined as the time when more the number of visible participants exceeds 20 . (a) First experimental session with 60 participants. (b) Second experimental session with 73 participants. Note that when the number of participants is larger, the time taken to complete the experiments is larger, but the maximal density of people does not seem to increase.

Thus, by using eq. (7) of the manuscript we can find an approximate bound for the expected tilting angle as

$$
\begin{equation*}
\theta_{\max } \approx \tan ^{-1}\left[2 \rho_{0} \tilde{A}_{x}(0)\right] \approx 6^{\circ} \tag{S.109}
\end{equation*}
$$

According to our convention, the positive sign corresponds to a clockwise tilt and the numerical value is remarkably close to the experimentally observed mean tilt $4^{\circ}$.

## G. Density estimates in complex flows

For complex flows (Scenarios 4-6) the density of pedestrians is affected by the bottleneck effects near the gates. We quantified the spatial variations of density in the following way.

First, in each trial we approximated local density by counting all participants within radius $r_{a v}=0.5 \mathrm{~m}$

$$
\begin{equation*}
\rho^{ \pm}(\mathbf{r}, t)=\frac{\sum_{i} \mathbb{1}\left[\left|\mathbf{r}_{i}^{ \pm}(t)-\mathbf{r}\right|<r_{a v}\right]}{\pi r_{a v}^{2}} \tag{S.110}
\end{equation*}
$$

Then, for each trial we computed time-averaged density $\bar{\rho}^{ \pm}(\mathbf{r})$ and in order to make the results comparable between different trials, we confined our attention to the times when at least 20 participants were present within the experimental arena.

$$
\begin{equation*}
\bar{\rho}^{ \pm}(\mathbf{r})=\frac{\sum_{t} \rho^{ \pm}(\mathbf{r}, t) \mathbb{1}\left[\int \rho^{+}(\mathbf{r}, t)+\rho^{-}(\mathbf{r}, t) d \mathbf{r} \geq 20\right]}{\sum_{t} \mathbb{1}\left[\int \rho^{+}(\mathbf{r}, t)+\rho^{-}(\mathbf{r}, t) d \mathbf{r} \geq 20\right]} \tag{S.111}
\end{equation*}
$$

Finally, $\bar{\rho}^{ \pm}(\mathbf{r})$ was ensemble-averaged for all 11 trials of a given scenario ( 6 in Session 1 and 5 in Session 2) and this quantity is plotted in the first two columns of Figure S25. In all cases, we see a local increase of average density near the exit gates. Interestingly, the same is true for the entry gate as well (Fig. S25(e)). Moreover, for scenario 4 (two exit gates) we also found another local peak in average density which can be attributed to the slow-down due to pedestrian interactions (Fig. S25(a-b)).

To determine if these density variations have an important role in lane nucleation, we estimated frequency of close pedestrian encounters at any given location. To this end, we compute the encounter rate,

$$
\begin{equation*}
f(\mathbf{r})=\frac{\sum_{t} \mathbb{1}\left[\rho^{+}(\mathbf{r}, t)>1\right] \mathbb{1}\left[\rho^{-}(\mathbf{r}, t)>1\right] \mathbb{1}\left[\int \rho^{+}(\mathbf{r}, t)+\rho^{-}(\mathbf{r}, t) d \mathbf{r} \geq 20\right]}{\sum_{t} \mathbb{1}\left[\int \rho^{+}(\mathbf{r}, t)+\rho^{-}(\mathbf{r}, t) d \mathbf{r} \geq 20\right]} \tag{S.112}
\end{equation*}
$$

which is defined as a ratio of times when agents of both types were present inside radius $r_{a v}$. This quantity is a proxy for the frequency of interactions between the agents of opposite type and it shows that in all scenarios they are most likely to occur in the central part of the experimental arena (the right-most column of Fig. S25). The fact that the interactions between the two groups occurs relatively far away from the bottlenecks, suggesting that the gate-specific dynamics most likely do not affects the lane nucleation phenomena that we are interested in at the leading order.


Figure S25. Density analysis for complex flow experiments. Precise mathematical definitions of the quantities presented by the contour plots are given in eq. (S.111) and (S.112). Scenario 4 (two exit gates): (a) Time-averaged density $\left\langle\bar{\rho}^{+}\right\rangle\left[\mathrm{m}^{-2}\right]$ of the group moving towards the exit gate at the top. (b) Time-averaged density $\left\langle\bar{\rho}^{-}\right\rangle\left[\mathrm{m}^{-2}\right]$ of the group moving towards the exit gate on the right. (c) Ensemble-averaged encounter rate $\langle f\rangle$ showing that inter-type interactions are most likely to occur in the centre. Scenario 5 (entry gate + exit gate): (d) Time-averaged density $\left\langle\bar{\rho}^{+}\right\rangle\left[\mathrm{m}^{-2}\right]$ of the group moving towards the exit gate at the top. (e) Time-averaged density $\left\langle\bar{\rho}^{-}\right\rangle\left[\mathrm{m}^{-2}\right]$ of the group emerging from the entry gate on the left. (f) Ensemble-averaged encounter rate $\langle f\rangle$. Scenario 6 (exit gate + stream): (g) Time-averaged density $\left\langle\bar{\rho}^{+}\right\rangle\left[\mathrm{m}^{-2}\right]$ of the pedestrian stream moving along the $y$-axis. (h) Time-averaged density $\left\langle\bar{\rho}^{-}\right\rangle\left[\mathrm{m}^{-2}\right]$ of the group moving towards the exit gate on the right. (i) Ensemble-averaged encounter rate $\langle f\rangle$.

## H. Data used in figures and videos

All the Supplementary Videos come from the session with 60 participants and each them shows a selected trial from one of the six scenarios.

Photographs in Figure 3 of the article were taken during session with 60 participants.
Fig. 3D-E shows example trajectories from experiments with 73 pedestrians.
Figure 4 of the article is based on selected trials from the session involving 60 participants.
[1] M. Isobe, T. Adachi, and T. Nagatani, Physica A 336, 638 (2004).
[2] D. Helbing, L. Buzna, A. Johansson, and T. Werner, Transportation Science 39, 1 (2005).
[3] J. Zhang, W. Klingsch, A. Schadschneider, and A. Seyfried, J. Stat. Mech. 2012 (2012).
[4] J. Bamberger, A. L. Gessler, P. Heitzelmann, S. Korn, R. Kahlmeyer, X. H. Lu, Q. H. Sang, A. Jiewang, G. A. Yuan, M. Gauss, and T. Kretz, Traffic and Granular Flow, 2013 (2015).
[5] S. Cao, A. Seyfried, J. Zhang, S. Holl, and W. Song, J. Stat. Mech: Theory Exp. 3, 033404 (2017).
[6] C. Feliciani and K. Nishinari, Phys. Rev. E 94, 032304 (2016).
[7] N. Guo, Q. Y. Hao, R. Jiang, M. B. Hu, and B. Jia, Transportation Research Part C: Emerging Technologies 71, 63 (2016).
[8] N. Guo, H. X. Liu, R. Jiang, B. Jia, and M. B. Hu, International Journal of Modern Physics C 91, 1850069 (2018).
[9] M. Moussaïd, E. G. Guillot, M. Moreau, J. Fehrenbach, O. Chabiron, S. Lemercier, J. Pettré, C. Appert-Rolland, P. Degond, and G. Theraulaz, PLoS Comput. Biol. 8, e1002442 (2012).
[10] M. E. Leunissen, C. G. Christova, A. P. Hynninen, C. P. Royall, A. I. Campbell, A. Imhof, M. Dijkstra, R. V. Roij, and A. V. Blaaderen, Nature 437, 235 (2005).
[11] T. Vissers, A. Wysocki, M. Rex, H. Löwen, C. P. Royall, A. Imhof, and A. V. Blaaderen, Soft Matter 7, 2352 (2011).
[12] K. R. Sütterlin, A. Wysocki, A. V. Ivlev, C. Räth, H. M. Thomas, M. Rubin-Zuzic, W. J. Goedheer, V. E. Fortov, A. M. Lipaev, V. I. Molotkov, O. F. Petrov, G. E. Morfill, and H. Löwen, Phys. Rev. Lett. 102, 085003 (2009).
[13] C. R. Du, K. R. Sütterlin, K. Jiang, C. Räth, A. V. Ivlev, S. Khrapak, M. Schwabe, H. M. Thomas, V. E. Fortov, A. M. Lipaev, V. I. Molotkov, O. F. Petrov, Y. Malentschenko, F. Yurtschichin, Y. Lonchakov, and G. E. Morfill, New J. Phys. 14, 073058 (2012).
[14] R. Jose and L. Santen, Phys. Rev. Lett. 124, 198103 (2020).
[15] S. Nowak and A. Schadschneider, Phys. Rev. E 85, 066128 (2012).
[16] Y. Sun, Physica A 549, 124295 (2020).
[17] M. Burger, S. Hittmeir, H. Ranetbauer, and M. T. Wolfram, SIAM J. Math. Anal. 48, 981 (2016).
[18] H. Ohta, EPL 99, 40006 (2012).
[19] J. Chakrabarti, J. Dzubiella, and H. Löwen, Phys. Rev. E 70, 012401 (2004).
[20] J. Cristín, V. Méndez, and D. Campos, Sci. Rep. 9, 1 (2019).
[21] J. Delhommelle, Phys. Rev. E 71, 016705 (2005).
[22] D. Helbing and T. Vicsek, New Journal of Physics 1, 13 (1999).
[23] K. Klymko, P. L. Geissler, and S. Whitelam, Phys. Rev. E 94, 022608 (2016).
[24] A. Poncet, O. Bénichou, V. Démery, and G. Oshanin, Phys. Rev. Lett. 118, 118002 (2017).
[25] M. Rex and H. Löwen, European Physical Journal E 26, 143 (2008).
[26] T. Xiong, P. Zhang, S. C. Wong, C. W. Shu, and M. P. Zhang, Chin. Phys. Lett. 28, 108901 (2011).
[27] H. Kuang, X. Li, T. Song, and S. Dai, Phys. Rev. E 78, 066117 (2008).
[28] C. Totzeok, Kinetic and Related Models 13, 1219 (2020).
[29] J. Chakrabarti, J. Dzubiella, and H. Löwen, Europhys. Lett. 61.3, 415 (2003).
[30] D. Helbing, Complex Systems 6, 391 (1992).
[31] A. Bricard, J.-B. Caussin, N. Desreumaux, O. Dauchot, and D. Bartolo, Nature 503, 95 (2013).
[32] N. Brain and D. Bartolo, Science 363, 46 (2019).
[33] N. Bain and D. Bartolo, Nat. Commun. 8 (2017).
[34] A. Einstein, Dover, New York (1956).
[35] X. Jia, C. Feliciani, D. Yanagisawa, and K. Nishinari, Physica A 531 (2019).
[36] M. Moussaïd, D. Helbing, S. Garnier, A. Johansson, M. Combe, and G. Theraulaz, Proceedings of the Royal Society B: Biological Sciences 276, 2755 (2009).
[37] C. Scholz, M. Engel, and T. Pöschel, Nat. Commun. 9, 1 (2018).
[38] I. Karamouzas, B. Skinner, and S. J. Guy, Phys. Rev. Lett. 113, 238701 (2014).
[39] Q. Xu, M. Chraibi, and A. Seyfried, Transportation Research Part C: Emerging Technologies 133 (2021).
[40] J. Wang, M. Boltes, A. Seyfried, J. Zhang, V. Ziemer, and W. Weng, Physica A: Statistical Mechanics and its Applications 500, 106 (2018).
[41] M. Muramatsu, T. Irie, and T. Nagatani, Physica A 267, 487 (1999).
[42] C. Reichhardt and C. J. O. Reichhardt, Soft Matter 14, 490 (2018).
[43] "Materials and methods,".
[44] H. Murakami, C. Feliciani, Y. Nishiyama, and K. Nishinari, Sci. Adv. 7, eabe7758 (2021).
[45] A. Corbetta, J. A. Meeusen, C. M. Lee, R. Benzi, and F. Toschi, Phys. Rev. E 98, 062310 (2018).
[46] J. Dzubiella and H. Löwen, J Phys: Condens Matter 14 (2002).
[47] P. Mullick, S. Fontaine, C. Appert-Rolland, A.-H. Olivier, W. Warren, and J. Pettré, PLoS Comput Biol 18 (2022).
[48] D. Helbing and P. Molnár, Phys. Rev. E 51, 4282 (1995).
[49] F. Dietrich and G. Köster, Phys. Rev. E 89, 062801 (2014).
[50] J. Toner and Y. Tu, Phys. Rev. E 58, 4828 (1998).
[51] K. Bacik, B. Bacik, and T. Rogers, Bath: University of Bath Research Data Archive. (2023).
[52] C. S. O’Hern, S. A. Langer, A. J. Liu, and S. R. Nagel, Phys. Rev. Lett. 88 (2002).
[53] E. Olson, Proceedings - IEEE International Conference on Robotics and Automation (2011).


[^0]:    General rights
    Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

