PNAS Nexus, 2024, 3, pgae120
https://doi.org/10.1093/pnasnexus/pgae120
Advance access publication 19 March 2024
Research Report

Dimensionless numbers reveal distinct regimes in the structure and dynamics of pedestrian crowds

Jakob Cordes [a,b,*], Andreas Schadschneider [b] and Alexandre Nicolas [c]

[a]Institute of Advanced Simulation, Forschungszentrum Jülich, 52428 Jülich, Germany
[b]Institut für Theoretische Physik, Universität zu Köln, 50937 Köln, Germany
[c]Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, UMR5306, F-69100 Villeurbanne, France
*To whom correspondence should be addressed: Email: IAS7-ped-dyn-modelling@fz-juelich.de
Edited By: Cristina Amon

Abstract

In fluid mechanics, dimensionless numbers like the Reynolds number help classify flows. We argue that such a classification is also relevant for crowd flows by putting forward the dimensionless Intrusion and Avoidance numbers, which quantify the intrusions into the pedestrians' personal spaces and the imminency of the collisions that they face, respectively. Using an extensive dataset, we show that these numbers delineate regimes where distinct variables characterize the crowd's arrangement, namely, Euclidean distances at low Avoidance number and times-to-collision at low Intrusion number. On the basis of these findings, a perturbative expansion of the individual pedestrian dynamics is carried out around the noninteracting state, in quite general terms. Simulations confirm that this expansion performs well in its expected regime of applicability.

Significance Statement

Pedestrian streams are ubiquitous, but very diverse. Classifying them is critical in practice for crowd management, but also for the validation of models. However, a robust way to do this was still missing. By introducing two dimensionless numbers, rooted in pedestrian psychology and rendering the ideas of preservation of personal space and anticipation of collisions, we show that different regimes can be delineated, corresponding to different types of arrangement of the crowd. The relevance of these quantities, demonstrated using an extensive empirical dataset, prompts a perturbative approach to pedestrian dynamics, not contingent on any further assumptions.

Introduction

Crowds often look like an ocean made of hundreds or thousands of heads, ruffled by ripples and waves (1), moving in synchrony or not; this impression struck poets (2) long before it inspired scientists (3, 4). Yet, even as of now, pedestrian dynamics as a discipline does not stand on the same footing as fluid mechanics. In the latter field, the classical motion of particles at the microscale is governed by an exact equation, Newton's law, whose homogenization yields the universal Navier–Stokes equation. In practice, modelers resort to a plethora of approximate schemes (e.g. Stokesian dynamics, lattice Boltzmann methods, Euler equations for inviscid flows) but the choice among these is guided, and theoretically bolstered, by the calculation of dimensionless numbers, such as the Reynolds number and the Mach number.

On the other hand, a zoo of models for pedestrian dynamics coexist (see, e.g. (5–8)) and the realm of applicability of each is ill-defined. The crowd's density is generally used to delineate different regimes, for instance the levels of service defined by Fruin for crowds (9, 10). Each level is marked by a dominant behavior: (un)avoidable contact, necessity to change gait, possibility to turn around, etc., and it has been argued that as the density changes crowd dynamics should be controlled by distinct laws (11). However, the watersheds between the regimes are arbitrary. Even from a practical standpoint, for safety assessments, crowds at similar densities may present contrasted characters and risk profiles. Consider the difference between a densely packed, but static audience in a concert hall and people vying for escape in an emergency evacuation (12). Recently, yearning for a better classification of these scenarios, it was proposed to gauge congestion on the basis of a dimensionless number related to the vorticity of the velocity field, instead of the density (13). This quantity is practically relevant, notably for safety issues, but gives no insight into the determinants of pedestrian dynamics at the microscale.

In this paper, we argue that in common scenarios, pedestrian dynamics are dominated by two variables, rendering the ideas

Competing Interest: The authors declare no competing interest.
Received: March 5, 2024. **Accepted:** March 7, 2024

of preservation of personal space (proxemics) and anticipation of collisions. Their averages over the crowd define dimensionless parameters that delineate regimes of crowd flows which display distinctive features. In the spirit of the Reynolds number in fluid mechanics, these dimensionless numbers help gauge the range of validity of pedestrian models and prompt specific perturbative expansions of the equations of motion in each regime.

Results

Psychological studies on proxemics indicate that people pay attention to their personal space, defined as "the area individuals maintain around themselves into which others cannot intrude without arousing discomfort" (14), more than to global density (15). We are therefore led to define an intrusion variable $\mathcal{I}n_i$ centered on each agent i, which vanishes for isolated pedestrians and, ideally, diverges at physical contact, so as to underscore the transition from no-contact dynamics to contacts and pushes. Here, to keep at bay ambiguities in the definition of a local density, we make use of the center-to-center distances r_{ij} to define $\mathcal{I}n_i$, viz.,

$$\mathcal{I}n_i = \sum_{j\in\mathcal{N}_i} \underbrace{\left(\frac{r_{soc}-\ell_{min}}{r_{ij}-\ell_{min}}\right)^{k_I}}_{\mathcal{I}n_{ij}}. \tag{1}$$

Taking $k_I = 2$, this represents the sum of *areal* encroachments of other agents j on i's personal space.

For simplicity, we overlook anisotropic effects and assume uniform circular shapes for the pedestrian bodies and personal spaces, of diameter $\ell_{min} = 0.2\,m$ and radius $r_{soc} = 0.8\,m$, respectively. The sum runs over the set $\mathcal{N}_i$ of all close neighbors j of i, here defined by $r_{ij} \leqslant 3\, r_{soc}$. That the intrusions of diverse neighbors should be added up makes sense for physical contacts (superposition of mechanical forces), but also for proxemic behavior (16, 17).

While this variable gives a sense of the level of crowding, it neither provides a full reflection of psychological experience (feeling of congestion) in the midst of the crowd (13, 18), nor fully controls the agent's dynamics: when two people i and j run toward each other, they will not behave as though they were isolated, even though they may still be separated by several meters, hence, $\mathcal{I}n_{ij} \to 0$. This anticipatory behavior is well captured by an anticipated time-to-collision (TTC) τ_{ij}, defined as the delay until the first collision if both i and j keep their current velocities ($\tau_{ij} = \infty$ if no collision is expected). Humans are indeed capable of identifying the most imminent collision between multiple objects and estimating TTCs (19), notably via purely optical quantities, namely, the optical angle divided by its derivative (20). Experiments showed that the TTC is instrumental in humans' decisions as to "when" to avoid an approaching pedestrian (21). Accordingly, the TTC can be used to define a nondimensional avoidance variable $\mathcal{A}v_i$ quantifying the risk of an imminent collision,

$$\mathcal{A}v_i = \sum_{j\in\mathcal{N}'_i} \underbrace{\left(\frac{\tau_0}{\tau_{ij}}\right)^{k_A}}_{\mathcal{A}v_{ij}}. \tag{2}$$

Here, $\tau_0 = 3\,s$ is a timescale above which collisions are hardly dreaded. The exponent k_A is simply set to 1 (see SI Appendix B for variations), and, in contrast with Eq. 1, the set of neighbors $\mathcal{N}'_i$ is restricted to the agent with the shortest τ_{ij}, i.e. the most imminent risk. Indeed, for collision avoidance, it has been ascertained that participants immersed in a virtual crowd tend to fixate a particular agent with a high risk of collision just before performing an avoidance maneuver around this person (22).

With these two variables in hand, one can hope for a finer delineation of pedestrian streams than with the traditional density-based levels of service. To this end, the foregoing agent-centered variables are averaged over the $N(t)$ agents observed in the crowd at time t, and then over time. This average defines the dimensionless Avoidance number $\mathcal{A}v$ and Intrusion number $\mathcal{I}n$. As $\mathcal{A}v$ should quantify the urgency of expected collisions, we only consider data points with a finite TTC in the average. Especially in the sparse datasets, this allows to focus on the parts where interactions occur.

Figure 1A illustrates the regimes of crowd flow that one would intuitively expect to find in a diagram parametrized by $\mathcal{A}v$ and $\mathcal{I}n$, using exemplary cases. The bottom left corner, $\mathcal{I}n$, $\mathcal{A}v \ll 1$, corresponds to very sparse crowds with hardly any interactions. As one moves up the $\mathcal{I}n$-axis, the setting gets more crowded, and pedestrians are eager to maintain a certain social distance with respect to others, as in a *unidirectional* flow. When $\mathcal{I}n \gg 1$, personal space can no longer be preserved and physical contact may eventually be unavoidable, as in a tightly packed static crowd (*Waiting* scenario). A very different way to depart from the noninteracting case is to consider people walking or running toward each other. This is well approximated by the beginning of an *Antipodal* experiment, in which participants initially positioned all along the circumference of a circle (with $\mathcal{I}n \ll 1$) are asked to reach the antipodal position. This induces conflicting moves, with risks of collision in the center of the circle, hence $\mathcal{A}v \gg 1$. Finally, competitive evacuations though a *bottleneck* exemplify the regime of large $\mathcal{I}n$ and $\mathcal{A}v$, which features contacts, pushes, as well as conflicting moves.

These are of course idealized expectations. To test them, we have collated an extensive dataset of pedestrian trajectories, including controlled experiments (single-file motion (23), bottleneck flows (24), corridor flows (25–27), antipodal scenarios (28)) and empirical observations in outdoor settings (29, 30); further details about these scenarios and the way we have smoothed out head sways from the trajectories can be found in SI Appendix A. For each scenario and each realization, we have computed $\mathcal{A}v(t)$ and $\mathcal{I}n(t)$ every 0.5 s, and averaged over the whole quasistationary state (unless otherwise stated).

Figure 1B shows that the idealized diagram worked out intuitively (Fig. 1A) is largely corroborated by the empirical datasets. Indeed, single files of amply spaced pedestrians are found in the bottom left corner, at small $\mathcal{I}n$ and $\mathcal{A}v$, whereas the top of the diagram, at large $\mathcal{I}n$, is occupied by situations in which physical contacts are almost inevitable. More interestingly, unidirectional flows and crossflows may have similar $\mathcal{I}n$ numbers, but they are distinguished by $\mathcal{A}v$, which takes larger values for crossflows, prone to more conflicts. In the same vein, antipodal maneuvers have intrusion numbers comparable to those of some typical outdoor scenarios, but larger avoidance numbers. The spread of points for a given type of scenario is expected and sensible, as a given geometry can give rise to flows of different natures, depending on the inflow and density, notably. Conversely, scenarios bearing different names may be similar on the whole and thus have comparable dimensionless numbers; for instance, a bidirectional flow with thick lanes is mostly made of unidirectional flows, apart from the few interactions between the lanes. Note that the $\mathcal{I}n$ and $\mathcal{A}v$ axes have been plotted orthogonally, whereas skewed axes should in principle be used if the variables exhibit some correlations; this does not alter the topology of the diagram, however. Nor do variations of the (somewhat arbitrary) precise definitions of $\mathcal{I}n$ and $\mathcal{A}v$, see SI Appendix B. In particular, $\mathcal{I}n$ is related to, but yet different from, the density, in that it is associated with

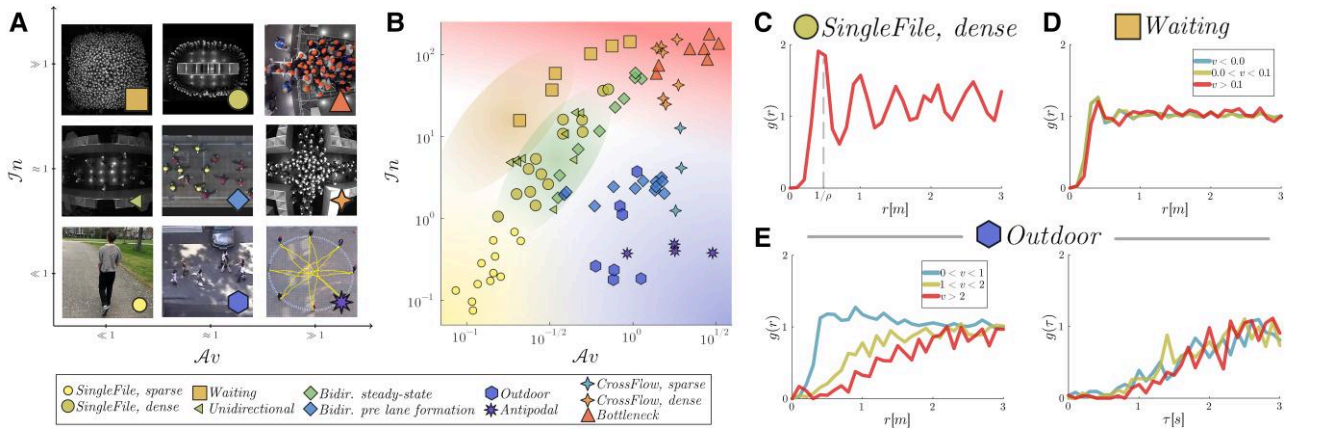

Fig. 1. Delineation of crowd flow regimes with the dimensionless numbers $\mathcal{I}n$ and $\mathcal{A}v$. A) Sketch of the expected diagram, illustrated with snapshots from the collated datasets (described in SI Appendix A). B) Empirical diagram obtained from various pedestrian datasets. Each datapoint corresponds to one experimental run or observational sequence. The experimental data from *SingleFile* and *CrossFlow* were split into a *sparse* and a *dense* dataset. Colored gradients are visual guides, to indicate different regimes. C) Pair-distribution function (pdf) for one run of the *SingleFile, dense* dataset. D) Pdf for the *Waiting* dataset and E) for the *Outdoor* dataset. The curves are binned according to the rate of approach $v = -\mathrm{d}r/\mathrm{d}t$ (given in m/s). As for D) the pdf is well parameterized by r and poorly by τ (see Fig. S3) and in E) vice versa.

individuals and captures both psychological and biomechanical aspects, as further discussed in SI Appendix B.

In practice, the visual delineation of regimes on the diagram of Fig. 1B appears sensible. But its physical relevance will only transpire if the delineated regimes exhibit constitutive differences. Remarkably, we find a major difference in the arrangement of the crowd, not in terms of static symmetry of the structure (which distinguishes, say, a liquid from a crystal), but in the nature of this self-organized "structure," i.e. more pragmatically, in the variables that characterize it. Drawing inspiration from condensed matter physics and following Ref. (31), we use as structural probe the pdf $g(x) = P(x)/P_{\mathrm{NI}}(x)$ between pedestrians, which quantifies the probability that two interacting pedestrians are found a given distance x apart, renormalized by the probability P_{NI} of measuring this distance for pedestrians that do not interact. This probability can be approximated by randomizing the time or space information (cf. SI Appendix C).

Starting from the origin ($\mathcal{I}n$, $\mathcal{A}v = 0$) and moving up along the $\mathcal{I}n$-axis while keeping $\mathcal{A}v \ll 1$, the crowd gets structured in real space, as evidenced by its radial pdf $g(r)$, where r is the Euclidean spacing between people. This is conspicuous for 1D configurations; indeed, the pdf of dense single files (Fig. 1C) develops a series of gradually decaying oscillations, with peaks positioned at multiples of the mean spacing, resembling the pdf of a liquid or a dense suspension of active colloids (32). But structural features are also visible in 2D settings, notably the dense static waiting crowd (Fig. 1D). Its pdf displays a strong dip at short distances, below 0.3–0.4 m, reflecting strong short-range repulsion, due to hard-core impenetrability and the reluctance for intrusion into the intimate space; the dip is followed by a peak at the nearest-neighbor distance. These features in real space are insensitive to dynamic variables such as the rate of approach v (i.e. the rate at which the distance between two pedestrians declines): the radial pdf exhibit the very same trend (Fig. 1D), quite independently of v.

The situation is widely different if one departs from the noninteracting regime by turning up $\mathcal{A}v$, i.e. considering very sparse crowds ($\mathcal{I}n \ll 1$) with more and more conflicting moves, as in the antipodal scenario or sparse outdoor crowds. This is the regime analyzed in Ref. (31). Strikingly, the radial pdfs do not collapse onto a single curve in this case; binned by rates of approach v, their pdfs display different shapes (left of Fig. 1E). In particular, the faster pedestrians approach each other, the larger is the Euclidean spacing at which they begin to interact.

Instead, if the TTC τ is substituted for r as the argument of the pdf, then a master curve is recovered, as shown in Fig. 1E (right) for the *Outdoor* dataset. In particular, the pdf gets more and more strongly depleted as τ becomes shorter, signaling the risk of an imminent collision. Thus, crowds in this regime also have some structure, but this is mostly hidden in real space and only becomes apparent in TTC space. This major finding of (31) is here contextualized by ascribing it to a particular regime of crowd flow: it does not hold for the waiting room (finite $\mathcal{I}n$, small $\mathcal{A}v$) (Fig. S3).

Discussion

To what extent can these observations be rationalized theoretically? Formally, the dynamics of a pedestrian i (or any other entity) is a function of their perceived surroundings, more precisely, the set $\mathcal{R}(t) = (\boldsymbol{r}_1(t), \ldots, \boldsymbol{r}_N(t))$ of all positions of the N agents (and, if need be, body orientations) observed so far, the agents' shapes $\mathcal{S}$, and some variable ξ_i gathering all unobserved features, which (in the worst case) may vary from realization to realization. Without loss of generality, it is possible to recast this functional dependence as a minimization, by *designing* a suitable mathematical function $\mathcal{C}_i$ (hereafter called cost function to follow the common terminology (33–35), but with no implication on its nature or properties), viz.,

$$\boldsymbol{v}_i^\star = \underset{\boldsymbol{v}\in\mathbb{R}^2}{\arg\min}\, \mathcal{C}_i[\boldsymbol{v}, \{\mathcal{R}(t'),\, t' \leqslant t\}, \mathcal{S}, \xi_i], \tag{3}$$

where $\boldsymbol{v}_i^\star$ denotes the decision of agent i which serves as an input to a mechanical layer which yields the actual velocity $\boldsymbol{v}_i(t)$. Unfortunately, neither the cost function $\mathcal{C}_i$ nor the hidden variables ξ_i are known. Nevertheless, in the physical sciences, generic perturbative expansions often afford ground for the study of systems near their critical states, leveraging symmetries to compensate blurred microscopic insights. Here, one cannot rely on conventional symmetry considerations, but the empirical classification of crowd regimes performed above has confirmed the prominent role of the Intrusion and Avoidance number.

This justifies the assumption that the agents' responses are mostly controlled by $\mathcal{I}n_i$ and $\mathcal{A}v_i$, so that the second part of Eq. 3 reduces to

$$\boldsymbol{v}_i^\star = \underset{\boldsymbol{v}\in\mathbb{R}^2}{\arg\min}\ \mathcal{C}_i\big[\boldsymbol{v}, \mathcal{I}n_i\big(\boldsymbol{r}_i(t)+\boldsymbol{v}\delta t\big), \mathcal{A}v_i(\boldsymbol{v})\big], \tag{4}$$

where $\mathcal{A}v_i$ (i.e. the TTC) is evaluated with the test velocity $\boldsymbol{v}$ and $\mathcal{I}n_i$ at the associated position $\boldsymbol{r}_i(t)+\boldsymbol{v}\delta t$, where δt is a time step. In the noninteracting scenario ($\mathcal{I}n_i, \mathcal{A}v_i = 0$), the agent freely pursues her goal at velocity $\boldsymbol{v}_{\mathrm{des},i}$, hence, to leading order,

$$\mathcal{C}_i(\boldsymbol{v}) \approx (\boldsymbol{v}_{\mathrm{des},i}-\boldsymbol{v})^2. \tag{5}$$

Expanding Eq. 4 around this reference situation, as detailed in SI Appendix D, yields the $\mathcal{A}v{*}\mathcal{I}n$-model,

$$\mathcal{C}_i\big[\boldsymbol{v}, \mathcal{I}n_i\big(\boldsymbol{r}_i(t)+\boldsymbol{v}\delta t\big), \mathcal{A}v_i(\boldsymbol{v})\big] \approx \big[\boldsymbol{v}_{\mathrm{des},i}-\boldsymbol{v}+\beta\nabla\mathcal{I}n_i\big(\boldsymbol{r}_i(t)\big)\big]^2+\alpha\,\mathcal{A}v_i(\boldsymbol{v}), \tag{6}$$

with $\alpha, \beta \geq 0$. We will refer to the case $\alpha = 0$ as the $\mathcal{I}n$-model and $\beta = 0$ as the $\mathcal{A}v$-model. One should bear in mind that these models were derived as generic asymptotic expansions of Eq. 4, rather than designed in an ad hoc way; their purpose is not so much to be realistic for the widest possible range of scenarios, as it is to shed light on crowd dynamics in limiting cases. In Eq. 6, we have neglected all mechanical interactions between the agents and the actual velocity relaxes toward the optimum $\boldsymbol{v}_i^\star$ over a timescale τ_R.

Let us test this perturbative expansion in the corresponding (asymptotic) regimes. First, we simulate the *Waiting* scenario: In the $\mathcal{I}n$-model, the agents make use of the available space to keep social distances to the others, which results in comparable averaged $\mathcal{I}n$ numbers ($\mathcal{I}n = 16$ for the experiments vs. $\mathcal{I}n = 14$ with the $\mathcal{I}n$-model). In contrast, the $\mathcal{A}v$-model fails to capture these features: the system remains frozen in its initial state as no collision is expected. The central role of $\mathcal{I}n$ is also readily understood in the case of a waiting line, where people halt to preserve each other's personal space. As a consequence, in a macroscopic model of the crowd, the local flow will depend solely on the density field, echoing the finding of a density-based hydrodynamic response of the crowd at the start of a marathon (1). In the opposite regime, the basic features of the sparse *CrossFlow*, notably successful collision avoidance, are well replicated by the $\mathcal{A}v$-model, contrary to the $\mathcal{I}n$-model in which the agents bump into each other. They are unable to maintain reasonable spacings (in TTC or in real space) with respect to each other, as also testified by the values of the dimensionless numbers (we find $\mathcal{A}v =$ 1.8/1.1/6.8 and $\mathcal{I}n = 1.2/2.9/5.3$ in the experiments/$\mathcal{A}v$-model/$\mathcal{I}n$-model, respectively).

The deficiency of models premised solely on $\mathcal{I}n$ or $\mathcal{A}v$ is even more manifest in scenarios which are not confined to the vicinity of the axes of the ($\mathcal{A}v$, $\mathcal{I}n$) plane. For example, let us pay attention to the temporal evolution of a bidirectional flow, using as input the experimental data of Ref. (27) and averaging over multiple similar realizations. The process of lane formation and then disappearance of the lanes after the two groups have passed each other entails a loop in the phase space, as represented in Fig. 2. Shortly after pedestrians enter the measurement area, in panel A, the limited space for each crowd leads to moderate values of $\mathcal{I}n$, but $\mathcal{A}v$ gets relatively high as the groups are walking toward each other, until they form lanes in panel B, thus lowering $\mathcal{A}v$, while $\mathcal{I}n$ is large because space is limited; finally, in panel C, the crowds have passed each other (low $\mathcal{A}v$) and the pedestrians make use of the available space by dissolving the lanes (moderate $\mathcal{I}n$), marking a return to the origin. Even though all models reproduce the formation of lanes, only the $\mathcal{A}v{*}\mathcal{I}n$ -model produces a loop comparable to the empirical one. While the $\mathcal{I}n$-model is unable to keep in check the growth of $\mathcal{A}v$ prior to lane formation, the $\mathcal{A}v$-model fails to ensure sufficient space between people when lanes have formed, leading to very high $\mathcal{I}n$ values. The dynamics of all scenarios are shown in the Movie S1.

While the focus was here put on the asymptotic $\mathcal{I}n$ and $\mathcal{A}v$-models, the discussion has bearing on the broader category of agent-based models: their equations of motion often hinge on variants of either the $\mathcal{I}n_i$ variable (36, 37) or the $\mathcal{A}v_i$ variable (31, 38, 39) (for instance, the *interactions* in the models of Refs. (31, 38) boil down to those of the $\mathcal{A}v$-model with $k_A = 1$ or 2 in Eq. 2, respectively), thereby limiting their range of applicability to the associated regime; a detailed inspection of this broader model category is deferred to a future publication.

Finally, in all regimes discussed so far, contacts between pedestrians were at most scarce. The situation is different in the high-$\mathcal{I}n$ region, which is highlighted in red in Fig. 1B and notably includes competitive bottleneck flows; in that case, more realistic (e.g. elliptic) shapes and mechanical contacts should be considered.

Conclusion

In summary, we have shown that the desire to preserve one's personal space from intrusions and the anticipation of collisions, quantified by the dimensionless numbers $\mathcal{I}n$ and $\mathcal{A}v$, delineate different regimes at the crowd's scale. These are marked by specific dynamics and "structural" arrangements. The importance of taking into account these factors to model the dynamics of individual agents depends on the regime under study.

At present, only collisions between the hard cores have been taken into account, in the absence of which ($\mathcal{A}v_i = 0$), agents are deemed isolated and have thus been left aside in the averaged $\mathcal{A}v$. In reality, the "softer" collisions, i.e. the anticipated intrusions into the private or intimate space, are also avoided. A more sophisticated definition of $\mathcal{A}v$ should be able to capture these.

Beyond $\mathcal{I}n$ and $\mathcal{A}v$, other dimensionless numbers can, and certainly should, be introduced to describe specific features of crowd dynamics such as an analog of the Mach number for the propagation of waves in crowds or some variant of the Péclet number (ratio between diffusion and advection rates) to account for the variability in the outcome of nominally similar experiments, due to the hidden variables ξ_i in Eq. 3. Interestingly, such a series of dimensionless numbers would mark successive departures from the conservation laws and invariance principles traditionally encountered in physical systems: while in the $\mathcal{I}n$-regime agents do not differ from particles subjected to distance-based interactions, $\mathcal{A}v$ introduces a velocity-based component to the interactions and a marginal violation of the reciprocity of forces. Better capturing the asymmetry of perception between pedestrians would make the violation of reciprocity more acute, with all its implications in active systems (40). Eventually, the violation of Galilean invariance in crowds would be mirrored by paying attention not only to TTC, but also to absolute time gaps, should the neighbors suddenly come to a halt. By gradually relaxing the symmetries applicable in physical systems, the way is thus paved for a general theoretical study of the statistical physics of pedestrian assemblies. In particular, it may be a good strategy to first focus on regimes where many of the aforementioned numbers are zero in order to derive a macroscopic flow theory starting from the Boltzmann equation (41, 42) (along the lines propounded by Chapman and Enskog for

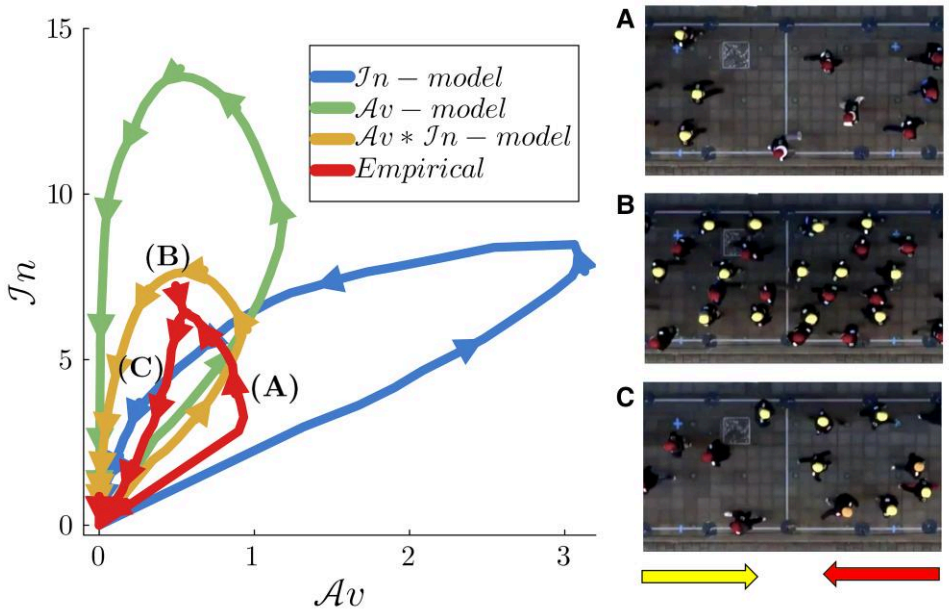

Fig. 2. Phase space trajectory during and after the formation of lanes using empirical results (27) and simulations of the proposed models. The arrows in the $\mathcal{I}n$–$\mathcal{A}v$ plot are spaced by 2.5 s. The temporal values of $\mathcal{A}v$ and $\mathcal{I}n$ are averaged over multiple realizations of the experiment as well as the simulations, with random initial conditions. The temporal evolution of the experiment is illustrated by three snapshots. Panel A shows the crowd before lane formation, in panel B the lanes have formed, and in panel C the two crowds have passed each other and the lanes have dissolved.

fluids), before facing the specific challenges raised by nonzero dimensionless numbers (e.g. the nonlocality inherent in the $\mathcal{A}v \neq 0$ regime).

Materials and methods

Empirical Datasets

Each dataset contains trajectories expressed in real-world coordinates as a function of time. For the *Waiting* scenario, we extracted trajectories by ourselves using the semiautomatic tracking mode of the *PeTrack* software (43), from existing videos of a controlled experiment conducted in Germany in 2013 in the frame the *BasiGo* project.

The *SingleFile* and the *Cross* datasets were split into a *dense* part and a *sparse* part (comprising the 6 and 3 runs with the lowest global densities). The initial transients were discarded in most scenarios, but not for the bidirectional flows, where they were used to probe the situation before lane formation.

Data processing

Pedestrian trajectories typically feature oscillations due to head sways and empirical noise, which both affect the calculation of $\mathcal{I}n$ and above all $\mathcal{A}v$. To smooth out the head sways, a 4th order Butterworth filter with critical frequency 0.5 Hz was applied to the trajectories. Velocities were then computed as the distance covered in approximately 1 s. Then, the TTC was computed by assuming that each pedestrian is a disk of diameter $\ell = 0.2$ m. This may generate apparent overlaps; to mitigate these artifacts, we set an upper bound $\mathcal{I}n_{ij}^{\max} = 400$ and $\mathcal{A}v_{ij}^{\max} = 60$ on all variables for the computation of the $\mathcal{I}n$ and $\mathcal{A}v$ numbers.

Some scenarios, particularly the *Outdoor* one, involve a large number of pedestrians that actually walk in isolation. To focus on actual interactions in the assessment of $\mathcal{A}v$, pedestrians with $\mathcal{A}v_i = 0$ were excluded.

Pair-distribution functions

For some variable x ($x = r$ or $x = \tau$ in the following), the pdf is given by the probability that two pedestrians are separated by x normalized by the probability $P_{NI}(x)$ that two noninteracting pedestrians are separated by x, viz., $g(x) = P(x)/P_{NI}(x)$. This normalization is aimed at correcting the lack of translational invariance in crowd observations. In practice, $P(x)$ is directly estimated from the frequency of occurrence of "separation" x in the dataset. $P_{NI}(x)$ is unknown in principle, but can be estimated by randomizing either the spatial or the temporal information (31), i.e. reshuffling pedestrians and frames. Finally, the pdf is obtained by binning the data into bins of size 0.1 m (for $x = r$) or 0.1 s (for $x = \tau$).

Asymptotic models

The $\mathcal{A}v$, $\mathcal{I}n$, and $\mathcal{A}v*\mathcal{I}n$ -models defined in the main text were simulated in *Julia*, by taking uniform and constant model parameters: $\alpha = 1.5\,\mathrm{m}^2/\mathrm{s}^2$, $\beta = 0.02\,\mathrm{m}^2/\mathrm{s}$, $v_{des} = 1.4$ m/s, and $\tau_R = 0.1$ s. Speeds were capped at $v_{max} = 1.7$ m/s. Agents are modeled as hard disks of diameter $\ell_{min} = 0.2$ m. Nevertheless, to account for the fact that people shun collisions not only between their hard cores but also between their private spaces, the diameter was increased to $\ell_{soc} = 0.4$ m for the computation of $\mathcal{A}v_i$. Also note that a small scalar $\varepsilon > 0$ is subtracted from $\mathcal{I}n_{ij}$ in Eq. 1 to make $\mathcal{I}n_i$ continuous across the cutoff distance.

Acknowledgments

The authors are grateful to Maik Boltes for giving them access to the Waiting Room experimental data and helping them in the process of extracting the trajectories. The authors also thank Antoine Tordeux and Mohcine Chraibi for insightful discussions. We thank Yao Xiao for granting permission to use the snapshot of the Antipodal scenario in Figure 1A.

Supplementary Material

Supplementary material is available at *PNAS Nexus* online.

Funding

The authors acknowledge financial support from the German Research Foundation (Deutsche Forschungsgemeinschaft DFG, grant number 446168800) and the French National Research Agency (Agence Nationale de la Recherche, grant number ANR-20-CE92-0033), in the frame of the French–German research project MADRAS.

Author Contributions

J.C., A.N., and A.S. conceptualized the study, and wrote and revised the manuscript. J.C. extracted and analyzed the data. A.N. performed the perturbative expansion. J.C. carried out numerical modeling experiments.

Preprints

This manuscript was posted on a preprint server: 10.48550/arXiv.2307.12786.

Data Availability

The data supporting the findings of this study are available in the supplementary material for the *Waiting* scenario. The data from the *Antipodal* scenario have been made available by Yao Xiao (xiaoyao9@mail.sysu.edu.cn) the corresponding author of Ref. (28). The remaining data are openly available in the repositories listed in SI Appendix A.

References

1 Bain N, Bartolo D. 2019. Dynamic response and hydrodynamics of polarized crowds. *Science*. 363(6422):46–49.
2 Hugo V. 1872. *L'Année terrible/les 7,500,000 oui*. Paris: Michel Lévy, Frères.
3 Henderson LF. 1974. On the fluid mechanics of human crowd motion. *Transp Res*. 8(6):509–515.
4 Hughes RL. 2003. The flow of human crowds. *Annu Rev Fluid Mech*. 35(1):169–182.
5 Martinez-Gil F, Lozano M, Garcia-Fernandez I, Fernandez F. 2017. Modeling, evaluation, and scale on artificial pedestrians: a literature review. *ACM Comput Surv*. 50:72.
6 Schadschneider A, Chraibi M, Seyfried A, Tordeux A, Zhang J. 2018. Pedestrian dynamics: from empirical results to modeling. In: Gibelli L, Bellomo N, editors. *Crowd dynamics: theory, models, and safety problems. Vol. 1*. Cham: Birkhäuser. p. 63–102.
7 Chraibi M, Tordeux A, Schadschneider A, Seyfried A. 2018. Modelling of pedestrian and evacuation dynamics. In: Meyers R, editor. *Encyclopedia of complexity and systems science*. Berlin, Heidelberg: Springer. p. 1–22.
8 Maury B, Faure S. 2018. *Crowds in equations: an introduction to the microscopic modeling of crowds*. Hackensack, NJ: World Scientific Publishing.
9 Fruin J. 1971. *Pedestrian planning and design*. New York (NY): Metropolitan Association of Urban Designers and Environmental Planners.
10 Schadschneider A, Chowdhury D, Nishinari K. 2010. *Stochastic transport in complex systems: from molecules to vehicles*. Amsterdam: Elsevier.
11 Best A, Narang S, Curtis S, Manocha D. 2014. DenseSense: interactive crowd simulation using density-dependent filters. In: Pettré J, Erleben K, editors. *SCA '14: Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*. Goslar: Eurographics Association. p. 97–102.
12 Feliciani C, Shimura K, Nishinari K. 2021. *Introduction to crowd management*. Cham (Switzerland): Springer.
13 Zanlungo F, *et al.* 2023. A pure number to assess "congestion" in pedestrian crowds. *Transp Res Part C Emerg Technol*. 148:104041.
14 Hayduk LA. 1978. Personal space: an evaluative and orienting overview. *Psychol Bull*. 85(1):117–134.
15 Evans GW, Wener RE. 2007. Crowding and personal space invasion on the train: please don't make me sit in the middle. *J Environ Psychol*. 27(1):90–94.
16 Lian L, Song W, Yuen KKR, Telesca L. 2018. Analysis of repulsion states among pedestrians inflowing into a room. *Phy Lett A*. 382(35):2424–2430.
17 Knowles ES, Kreuser B, Haas S, Hyde M, Schuchart GE. 1976. Group size and the extension of social space boundaries. *J Pers Soc Psychol*. 33(5):647–654.
18 Jia X, *et al.* 2022. Revisiting the level-of-service framework for pedestrian comfortability: velocity depicts more accurate perceived congestion than local density. *Transp Res Part F Traffic Psychol Behav*. 87:403–425.
19 Delucia PR, Novak JB. 1997. Judgments of relative time-to-contact of more than two approaching objects: toward a method. *Percept Psychophys*. 59(6):913–928.
20 Lee DN. 1976. A theory of visual control of braking based on information about time-to-collision. *Perception*. 5(4):437–459.
21 Pfaff LM, Cinelli ME. 2018. Avoidance behaviours of young adults during a head-on collision course with an approaching person. *Exp Brain Res*. 236(12):3169–3179.
22 Meerhoff LA, Bruneau J, Vu A, Olivier A-H, Pettré J. 2018. Guided by gaze: prioritization strategy when navigating through a virtual crowd can be assessed through gaze activity. *Acta Psychol (Amst)*. 190:248–257.
23 Seyfried A, Portz A, Schadschneider A. 2010. Phase coexistence in congested states of pedestrian dynamics. In: Bandini S, Manzoni S, Umeo H, Vizzari G, editors. Cellular automata. Vol. 6350. Berlin, Heidelberg: Springer Berlin Heidelberg. p. 496–505.
24 Adrian J, Seyfried A, Sieben A. 2020. Crowds in front of bottlenecks at entrances from the perspective of physics and social psychology. *J R Soc Interface*. 17(165):20190871.
25 Cao S, Seyfried A, Zhang J, Holl S, Song W. 2017. Fundamental diagrams for multidirectional pedestrian flows. *J Stat Mech Theory Exp*. 2017(3):033404.
26 Feliciani C, Murakami H, Nishinari K. 2018. A universal function for capacity of bidirectional pedestrian streams: filling the gaps in the literature. *PLoS One*. 13(12):e0208496.
27 Murakami H, Feliciani C, Nishiyama Y, Nishinari K. 2021. Mutual anticipation can contribute to self-organization in human crowds. *Sci Adv*. 7(12):eabe7758.
28 Xiao Y, *et al.* 2019. Investigation of pedestrian dynamics in circle antipode experiments: analysis and model evaluation with macroscopic indexes. *Transp Res Part C Emerg Technol*. 103:174–193.
29 Pellegrini S, Ess A, Schindler K, van Gool L. 2009. *2009 IEEE 12th International Conference on Computer Vision*. New York: IEEE. p. 261–268.

30 Lerner A, Chrysanthou Y, Lischinski D. 2007. Crowds by example. *Comput Graph Forum*. 26(3):655–664.
31 Karamouzas I, Skinner B, Guy SJ. 2014. A universal power law governing pedestrian interactions. *Phys Rev Lett*. 113(23):238701.
32 Klongvessa N. 2020. Study of dense assemblies of active colloids: collective behavior and rheological properties [PhD thesis]. School Université de Lyon.
33 van Toll W, *et al.* 2020. Generalized microscropic crowd simulation using costs in velocity space. In: Symposium on Interactive 3D Graphics and Games. San Francisco (CA): ACM. p. 1–9.
34 Hoogendoorn S, Bovy P. 2003. Simulation of pedestrian flows by optimal control and differential games. *Optim Control Appl Meth*. 24(3):153–172.
35 Guy SJ, Curtis S, Lin MC, Manocha D. 2012. Least-effort trajectories lead to emergent crowd behaviors. *Phys Rev E*. 85(1):016110.
36 Helbing D, Farkas I, Vicsek T. 2000. Simulating dynamical features of escape panic. *Nature*. 407(6803):487–490.
37 Tordeux A, Chraibi M, Seyfried A. 2016. Collision-free speed model for pedestrian dynamics. In: Knoop VL, Daamen W, editors. *Traffic and granular flow '15*. Cham: Springer. p. 225–232.
38 van den Berg J, Lin M, Manocha D. 2008. Reciprocal velocity obstacles for real-time multi-agent navigation. In: ICRA 2008—2008 IEEE International Conference on Robotics and Automation, May 19–23, 2008, Pasadena, CA.
39 van den Berg J, Guy S, Lin M, Manocha D. 2011. Reciprocal n-body collision avoidance. In: Pradalier C, Siegwart R, Hirzinger G, editors. *Springer tracts in advanced robotics*. Vol. 70. Berlin, Heidelberg: Springer. p. 3–19.
40 Fruchart M, Hanai R, Littlewood PB, Vitelli V. 2021. Non-reciprocal phase transitions. *Nature*. 592(7854):363–369.
41 Bertin E, Droz M, Grégoire G. 2006. Boltzmann and hydrodynamic description for self-propelled particles. *Phys Rev E*. 74(2):022101.
42 Arnault P, Guisset S. 2022. Chapman–Enskog derivation of multicomponent Navier–Stokes equations. *Phys Plasmas*. 29(9):090901.
43 Boltes M, Seyfried A, Steffen B, Schadschneider A. 2010. Automatic extraction of pedestrian trajectories from video recordings. In: Klingsch WWF, Rogsch C, Schadschneider A, Schreckenberg M, editors. Pedestrian and evacuation dynamics 2008. Berlin, Heidelberg: Springer. p. 43–54.