Classification of Pedestrian Crowds by Dimensionless Numbers

Jakob Cordes\textsuperscript{1,2} · Alexandre Nicolas\textsuperscript{3} · Andreas Schadschneider\textsuperscript{2}

\textsuperscript{1} Institute of Advanced Simulation, Forschungszentrum Jülich GmbH, Jülich, Germany
E-mail: jakob_cordes@yahoo.de
\textsuperscript{2} Institut für Theoretische Physik, Universität zu Köln, Germany
\textsuperscript{3} Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, Villeurbanne, France

Received: 6 November 2023 / Last revision received: 19 March 2024 / Accepted: 30 April 2024
DOI: 10.17815/CD.2024.164

Abstract Classifications of pedestrian crowds primarily rely on density. This fails to encompass the diverse behaviours and risk profiles observed. We have introduced two dimensionless numbers, the Intrusion number $I_n$, based on the desire to maintain one’s personal space, and the Avoidance number $A_v$, based on the anticipation of collisions. These two numbers delineate different flow regimes, as we intuitively expect and as we have empirically demonstrated using an extensive dataset. Similarly to Fluid Mechanics, where dimensionless numbers guide the choice between different approximations, the dynamics of crowds can be approached in each regime by perturbative expansions of the individual pedestrian dynamics, which yield approximate equations of motion applicable in the corresponding regime (and only there).

Keywords Pedestrians · modelling · dimensionless numbers · collision avoidance · intrusion

1 Introduction

The enterprise of classifying pedestrian crowds goes back at least to Fruin \cite{Fruin1971}, who introduced the concept of level of service (LoS) in pedestrian dynamics. LoS is mainly based on density as classification criterion, following the idea that each level will be marked with a distinctive dominant behaviour of the individuals, e.g., (un)avoidable contact, necessity to change gait, possibility to turn around, etc. However, the boundaries between the classes are rather arbitrary. Furthermore, using just the density as classifier is not sufficient, as one has come to realise: crowds at similar densities can exhibit different beh-
haviours and risk profiles, e.g. when comparing a static audience at a concert and people vying for escape in an emergency evacuation [2].

This observation has led to alternative proposals for classifiers. Recently, it has been suggested that a dimensionless number related to the vorticity of the velocity field better discriminates between such scenarios [3]. Although this quantity may be very relevant for safety analysis, it does not provide much insight into the determinants of pedestrian dynamics at the microscale. Gaining such insight, however, would be highly desirable as it would also clarify the realm of applicability of the large zoo of models that have been proposed for pedestrian and crowd dynamics.

Our approach draws inspiration from fluid dynamics, where suitably defined dimensionless numbers help specify properties of fluid flows. A prominent example is the Reynolds number. In a similar spirit, we introduce two dimensionless numbers for crowd motion [4]. These numbers aim to quantitatively capture two essential factors for an individual’s motion amid a crowd, namely, the preservation of one’s personal space and the anticipation of collisions. In the following, we propose a pedagogical discussion that elaborates on the findings already reported in [4].

2 Concepts of Intrusion and Avoidance

People generally shun too close contacts with other people, especially if they are unrelated. Preferred interpersonal distances vary between countries and cultures and depending on the relationship between people [5], but the will to preserve some personal space is virtually universal [6] and naturally impacts the arrangement of crowds. To capture this effects in a quantitative fashion, we introduce an intrusion variable $I_{ni,j}$ for binary interactions between agents $i$ and $j$, viz.,

$$I_{ni,j} = \left( \frac{r_{soc} - \ell_{\text{min}}}{r_{ij} - \ell_{\text{min}}} \right)^2,$$

(1)
where \( r_{ij} \) is the centre-to-centre distance of pedestrians \( i \) and \( j \). It quantifies the areal encroachment of other agents \( j \) on the personal space of agent \( i \). As illustrated in Fig. 1 (Left), \( I_{nj} \) decays with the distance between people; it vanishes for isolated pedestrians, but diverges when people come into physical contact. For simplicity, we assume that both agents and their personal space are circular (with diameter \( \ell_{\text{min}} = 0.2 \) m and radius \( r_{\text{soc}} = 0.8 \) m, respectively), i.e. effects of anisotropy are neglected. As several agents \( j \) may surround agent \( i \), the latter experiences an intrusion

\[
I_{ni} = \sum_{j \in N_i} I_{nj}.
\]  

For the data presented below, the sum in (Eq. 2) has been taken to run over the set \( N_i \) of all close neighbours \( j \) of \( i \), here defined by \( r_{ij} \leq 3 r_{\text{soc}} \). The exact definition of a number \( I_{nj} \) that fulfils the above criteria can of course be varied, for instance changing the power exponent in Eq. 1, but the variations we have tested do not qualitatively alter the results, even though they may not delineate regimes as well as Eq. 1-Eq. 2 (see below, but also [4]).

Short interpersonal distances are not the only factor that pedestrians are keen to avoid. Indeed, they also baulk at standing in a collision path with another pedestrian, all the more so as the possible collision is imminent, because this implies that they will have to abruptly swerve to avoid it. We quantify the risk of an imminent collision between two agents \( i \) and \( j \) with the avoidance variable

\[
A_{vij} = \frac{\tau_0}{\tau_{ij}},
\]

where \( \tau_{ij} \) is the anticipated time-to-collision (TTC) and \( \tau_0 = 3 \) s denotes the timescale above which potential collisions are hardly dreaded. TTC is defined as the time until the first collision if both agents \( i \) and \( j \) keep their current velocities. It is set to \( \tau_{ij} = \infty \) (hence, \( A_{vij} = 0 \)) if no collision is expected. Importantly, as sketched in Fig. 1 (Right), the avoidance variable may take large values even if there is strictly no intrusion at present, and conversely it may be vanishing even if the personal space is strongly violated. In contrast to Eq. 2, we assume that each agent is mostly concerned with only the most imminent risk of collision, in the light of experimental evidence that pedestrians tend to gaze mostly at one risk of imminent collision [7], so that in the definition of the agent-centred variable

\[
A_{vi} = \sum_{j \in N'_i} A_{vij},
\]

we restrict the set \( N'_i \) to the agent \( j \) with the shortest \( \tau_{ij} \), hence \( A_{vi} = \max_j A_{vij} \). Again, alternative definitions of \( A_{vi} \) have been put to the probe, for instance taking the square of \( \tau_0/\tau_{ij} \) in Eq. 3, with little incidence on our qualitative results [4].

### 3 Delineation of Crowd Regimes

To classify crowd flows with many pedestrians, the foregoing agent-centred variables are averaged over the \( N(t) \) agents observed in the crowd at time \( t \), and then over time. This
Figure 2 Classification of pedestrian crowds by the dimensionless numbers In and Av. **Left:** Schematic diagram drawn according to the intuition and illustrated with exemplary snapshots. Adapted from [4]. **Right:** Empirical diagram obtained from the extensive pedestrian dataset. Each data-point corresponds to one experimental run or observational sequence.

average value defines two dimensionless numbers, the Avoidance number (Av) and the Intrusion number (In). For the calculation of Av, we only consider data points with a finite Time to Collision (TTC), because in sparse situations many pedestrians are observed with $Av_i = 0$, which would lead to a very low Av although the pedestrians whose trajectories are actually influenced by others may have significant $Av_i$; in other words, pedestrians with $Av_i = 0$ are supposed to be non-interacting and are discarded to obtain a neater classification.

Fig. 2 (**Left**) shows the crowd flow patterns that one would tentatively expect to see on a diagram parametrized by Av and In, using exemplary cases. Close to the origin ($In, Av \ll 1$), the agents can freely pursue their goals, which corresponds to very sparse crowds with little interactions. Moving up along the In-axis, the scenery gets more crowded and pedestrians are eager to keep a certain distance to each other. The simplest example is that of a (mostly static) waiting line, where people halt to preserve their mutual personal spaces. This regime can also be exemplified by a sparse crowd walking in the same direction, i.e., a unidirectional flow. If more and more people are present in the scene, then $In \gg 1$, i.e. intrusions into the personal or intimate space are unavoidable and mechanical contacts may occur, e.g. in a tightly packed static crowd (Waiting scenario). However, one can also leave the origin by moving to the right, along the Av-axis. This is exemplified by the beginning of the Antipodal experiment, where the participants are well separated but face an anticipated collision in the centre of the circle. The top corner ($Av, In \gg 1$) represents e.g. Bottleneck experiments with high motivation.

The foregoing discussion is guided by the intuition, but it can be made concrete by calculating Av and In from real pedestrian trajectories. Therefore, we have collected a large dataset consisting of controlled experiments (single-file motion [8], bottleneck flows [9], corridor flows [10–12], antipodal scenarios [13], the intruder scenario [14], a static queue [15]). Further details about the data sets, the method to smooth trajectories, and the calculation of the Avoidance and Intrusion number (every 0.5 s over the whole
Figure 3 Alternative classification of pedestrian crowds, where the *In* number on the vertical axis has been substituted by the global density $\rho$.

(quasi-)stationary state) can be found in the Appendix of [4].

Indeed, the empirical values of $Av$ and *In* shown in Fig. 2 (Right) comforts our intuition from Fig. 2 (Left). Low density single-file experiments have small *In* and $Av$. In comparison, a static *Queue* has a very low Avoidance number as well, but displays a higher Intrusion number. A very densely packed static crowd (*Waiting*) can be found at the top left. While these could be delineated by the density (or the Level of Service) alone, many scenarios can only be distinguished by $Av$. For example, unidirectional from cross flows, the beginning of an Antipodal experiment from typical sparse outdoor settings, or a static crowd with and without an intruder passing through it. More subtle differences are also apparent, for example, the difference between the same bottleneck runs with low and high motivation, where the latter shows both higher *In* and $Av$ numbers.

We chose to have an Intrusion number based on distances instead of using the local density. This is partly justified by the ambiguity in the definition of a local density. However, we acknowledge that the averaged *In* is still closely related to the density, which certainly is the quantity most commonly used to classify crowds. In Fig. 3, we have substituted the Intrusion number with the global density $\rho$, calculated as the number of pedestrians divided by the available space. The delineation of different regimes is still clearly visible. Only the Single-File data strongly deviate from the original diagram, where, relative to the rest of the diagram, moderate intrusions seem to correspond to high densities. In the Single-file scenario, the pedestrians do not have lateral neighbours. Therefore, the deviation might actually be reflected in the subjective feeling of the pedestrians. Alternatively, the deviation might also be explained by the presence of obstacles, e.g. walls, which are very close to all of the agents in the single-file scenario and have not been taken into account. Besides, the Cross scenario and the Bottleneck scenario are at lower densities, but higher *In* values, compared to the Waiting scenario. In the latter, people are distributed

---

1To enable us to plot the Single-File data along with the rest, the one-dimensional density, calculated as the number of people divided by the length of the track, was rescaled according to [16], where we assumed a width of 0.3 m.
very homogeneously, whereas inhomogeneities are conspicuous in the former, including regions of tight packing. The Intrusion number puts more weight on these tightly packed neighbourhoods.

We emphasise that the spread of points for a given geometry is expected, as the nature of the flow in one geometry can change with ‘boundary conditions’ (e.g. inflows or density). Conversely, two different scenarios can lead to similar numbers. It seems to us that these points are in support of the chosen approach, rather than shortcomings, in that this shifts the focus on the flow properties experienced by pedestrians. For example, crossing flows at high densities generate effective ‘bottlenecks’, so much so that the properties of crossing flows at higher densities are thought to be controlled by emergent bottlenecks; finding similar dimensionless numbers for crossing flows and actual bottlenecks is therefore not so surprising.

That being said, to some extent there is a link between the flow properties and the typical self-organisation of the pedestrians. To get insight into the relation between the structure and the dimensionless numbers, and following [17], we have studied the crowds’ structure in asymptotic regimes using the pair-distribution function [4]. This has led us to the observation that Euclidean distances are relevant to describe the crowd’s structure at low Avoidance number, whereas the TTCs are relevant at low Intrusion number. In the first case, the interactions governing the dynamics can be considered as "spatially controlled" whereas in the second case they are "temporally controlled".

4 Consequences for Modeling

These distinct types of arrangements, depending on the regime, contribute to making the delineation of regimes useful, especially from a modelling perspective. Quite generally, the velocity adopted by a pedestrian \(i\) at the next time step \(t + \delta t\) can be modelled as the result of a minimisation of a suitably defined cost function \(C_i(\vec{v}, \ldots)\) [18–20], viz.,

\[
\vec{v}_i(t + \delta t) = \arg\min_{\vec{v} \in \mathbb{R}^2} C_i(\vec{v}, \ldots).
\]  

In principle, \(C_i\) can be a very complex function of all environmental variables \(E_i := \{\vec{r}_j(t), \vec{v}_j(t), \ldots\}\), not to mention the agent’s desired velocity, \(\vec{v}_{\text{des},i}\), i.e., the velocity that would be taken were the agent alone. However, in the spirit of Landau’s theory of phase transitions in condensed matter, we aspire to get a simple perturbative expansion of \(C_i\), drawing on the symmetries of the system. By definition, if there exists such a symmetry \(\mathcal{S}\), then the cost is identical for all symmetric configurations, i.e., \(C_i(\mathcal{S} \cdot (\vec{v}, E_i)) = C_i(\vec{v}, E_i)\). Here, there are no conventional symmetries, but the above considerations strongly suggest that \((\vec{v}, E_i)\) is dominated by the Intrusion and Avoidance variables. Accordingly, all \((\vec{v}, E_i)\) sharing the same \(\text{In}_i(\vec{v}, E_i)\) and \(\text{Av}_i(\vec{v}, E_i)\) can be regarded as ‘symmetric’ in some loose sense. To compute \(\text{In}_i(\vec{v}, E_i)\) and \(\text{Av}_i(\vec{v}, E_i)\) with the test velocity \(\vec{v}\), it is sensible to evaluate \(\text{In}_i\) in the current environment \(E_i\), except that agent \(i\) is shifted from \(\vec{r}_i\) to \(\vec{r}_i + \vec{v} \delta t\), and to evaluate \(\text{Av}_i\) in the current environment, except
that agent $i$ assumes velocity $\vec{v}$ instead of $\vec{v}_i(t)$ (positional changes at $t + \delta t$ are of higher order). Therefore, we arrive at

$$\vec{v}_i(t + \delta t) \simeq \arg\min_{\vec{v} \in \mathbb{R}^2} C_i\left[ In_i(\vec{r}_i + \vec{v} \delta t), Av_i(\vec{v}) \right]. \quad (6)$$

In [4], we showed that a perturbative expansion of Eq. 6 around the non-interacting situation ($In, Av = 0$) leads to

$$C_i(\vec{v}) \approx Av_i(\vec{v}) + \frac{1}{\alpha} \left[ \vec{v}_{des,i} + \beta \nabla In_i(\vec{r}_i(t)) - \vec{v} \right]^2, \quad (7)$$

with $\alpha > 0$, $\beta \geq 0$, which is a first-order model in the newly introduced variables that we will call the $Av \times In$-model. We will refer to the case $\alpha \to 0$ as the $In$-model and $\beta = 0$ as the $Av$-model. The actual velocity of agent $i$ relaxes towards the minimum of $C_i$ on a time-scale $\tau_R$.\footnote{We chose the following parameters $\alpha = 2/3 \ s^2/m^2$, $\beta = 0.02 /s$, $v_{max} = 1.7 \ m/s$, $v_{des} = 1.4 \ m/s$, and $\tau_R = 0.1 \ s$. We have increased the agents’ size in the $Av$-part of the model to a social radius $\ell_{soc} = 0.4 \ m$. A small scalar is subtracted from $In_{ij}$ to make it continuous across the cut-off radius. Note that, in Eq. 7, we omitted the dependencies of $Av_i$ and $In_i$ on the positions and velocities of the other agents.}

The purpose of these models is not to describe pedestrians crowds in every conceivable situation. Rather, the models should give some additional insights into the relevance of the characteristic numbers introduced, especially through their limitations! These limitations become conspicuous if we study the models in the asymptotic regimes.

As an example for the asymptotic $In$-regime ($Av \ll 1$), we simulate the Waiting scenario. In the $In$-model, the agents avoid intrusions and thus the crowd self-organises into a spatially homogeneous state. On the contrary, in the $Av$-model, the crowd will stay in its (possibly very inhomogeneous) initial condition, given that all $Av_i$ are vanishing. On the other hand, in the asymptotic $Av$-regime ($In \ll 1$), here illustrated with a sparse CrossFlow, the $In$-model fails to reproduce its basic features, namely, successful collision avoidance: the agents keep bumping into each other. In contrast, the $Av$-model solves these conflicts convincingly.

The limitations of models based solely on either $Av$ or $In$ become even more apparent when we consider scenarios that depart from the axes of the ($Av, In$) plane. For instance, all models can replicate the formation of lanes. However the $In$-model struggles to deal with the impending collisions before lane formation, and the $Av$-model fails to maintain sufficient spacing between individuals within each lane. Besides, the lanes will not dissolve even after the two crowds have passed each other. In [4], we show that only the $Av \times In$-model displays these features correctly, resulting in a very similar temporal evolution of $In$ and $Av$ if compared to the experimental data from [12]. All described features of the models can be seen in the videos uploaded at [21].

5 Conclusion

In order to delineate different crowd flow regimes, we have introduced the dimensionless Intrusion number $In$, based on the desire to preserve one’s personal space, and the (equally
Avoidance number $AV_i$, grounded in the anticipation of collisions. Classifying flows according to these numbers succeeds in generating a neat diagram of crowd scenarios. The way in which the crowd self-organizes was found to be best described by ‘distances’ in time (TTCs) in the low-$In_i$ regime and by distances in space in the low-$AV_i$ regime. In the future, studying distributions or (cross) correlations between $AV_i$ and $In_i$ might give complementary insight into the nature of the different crowd regimes.

Based on $AV_i$, $In_i$, or on these two numbers, perturbative models have been put forward. Similarly to Fluid Mechanics, where dimensionless numbers guide the choice between different approximations, the dynamics of crowds can be approached in each regime by perturbative expansions, which yield pedestrian models applicable in the corresponding regime (and only there). This finding has a much broader impact on agent-based pedestrian models, as $AV_i$ or $In_i$ appear in many of these [17, 22–25]. The study presented here should therefore be extended to organize the plethora of agent-based models and to connect their realm of applicability to specific crowd regimes.

To conclude with some limitations of the present work, it makes no doubt that $AV_i$ and $In_i$ cannot capture all possible crowd behaviours. For instance, the proposed models cannot directly reproduce the fundamental diagram. This hints at additional mechanisms that have been overlooked so far, such as asymmetric interactions e.g. due to perception, absolute time-gaps, and more complex shapes that might depend on the velocity. A more realistic description of pedestrian shapes and mechanical interactions would also be needed when $In_i$ reaches larger values. Further dimensionless numbers could be introduced to capture more features.

**Acknowledgements** We are grateful to Maik Boltes for giving us access to the Waiting Room experimental data and helping us in the process of extracting the trajectories. We also thank Mohcine Chraibi and Antoine Tordeux for insightful discussions.

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.

**References**


