

The Ergodic Theory of Traffic Jams

Lawrence Gray¹ and David Griffeath²

Received November 16, 2000; revised June 26, 2001

We introduce and analyze a simple probabilistic cellular automaton which emulates the flow of cars along a highway. Our Traffic CA captures the essential features of several more complicated algorithms, studied numerically by K. Nagel and others over the past decade as prototypes for the emergence of traffic jams. By simplifying the dynamics, we are able to identify and precisely formulate the self-organized critical evolution of our system. We focus here on the Cruise Control case, in which well-spaced cars move deterministically at maximal speed, and we obtain rigorous results for several special cases. Then we introduce a symmetry assumption that leads to a two-parameter model, described in terms of acceleration (α) and braking (β) probabilities. Based on the results of simulations, we map out the (α, β) phase diagram, identifying three qualitatively distinct varieties of traffic which arise, and we derive rigorous bounds to establish the existence of a phase transition from free flow to jams. Many other results and conjectures are presented. From a mathematical perspective, Traffic CA provides local, particle-conserving, one-dimensional dynamics which cluster, and converge to a mixture of two distinct equilibria.

KEY WORDS: Traffic jam; conservative flow; interacting particle system; phase separation; ergodic.

1. INTRODUCTION

This study is motivated primarily by the simulations of K. Nagel and his coworkers on the emergence of traffic jams. Over the past decade, first in Germany and then at Los Alamos National Labs, the foundations of a theory of traffic have been investigated in more than a dozen papers (e.g., [Kra], [Nag1], [NP], [SSNI], and the recent surveys [KNW] and [CSS])

¹ School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455; e-mail: gray@math.umn.edu

² Mathematics Department, University of Wisconsin, Madison, Wisconsin 53706; e-mail: griffeat@math.wisc.edu

which include many additional references). The first cellular automaton with realistic traffic jam behavior was introduced by Nagel and Schreckenberg [NS] in 1992. It was used to model the flow of a single lane of cars as a directed, conservative particle system. In the study of that model and its relatives, metaphors and methods of statistical physics were used to argue that jams represent a kind of self-organized state which arises when the traffic reaches a critical density. A consistent paradigm seems to govern the various algorithms, as summarized in [KNW]:

... different models of traffic flow display similar mechanisms that lead to traffic flow breakdown or recovery after breakdown, respectively. It can be assumed that the mechanisms are mathematically identical.

Subsequent work at Los Alamos has applied the same basic principles to increasingly complex traffic networks, from two-lane highways with passing [NWWS] all the way to car-level micro-simulations of Dallas and Portland [Nag2]. A colorful account of this research appeared in the August 5, 1999 *Washington Post*. Staff writer Alan Sipress [Sip] summarizes the most basic findings thus:

Scientists have identified “phase changes” in traffic, similar to the sudden transitions that occur when steam turns to water or water to ice. Understanding the timing and dynamics of phase changes in traffic, like those in nature, poses a challenge for physicists. ...

Phase 1. When traffic is light, motorists drive much as they like, moving at the speed they want and changing lanes easily. Motorists are comparable to steam particles with great freedom of movement.

Phase 2. As the road becomes crowded, motorists suddenly lose much of their freedom and are forced to drive as part of the overall traffic stream, moving at the speed of the general flow and often unable to change lanes. This phase, similar to water, has been called “synchronized” flow.

Phase 3. In heavy congestion, traffic is stop-and-go. Like water freezing into ice, the motorists are stuck in place.

Intrigued by the phenomenology of real-world traffic and the simulations of Nagel and others, we set out to find the simplest spatial interactions which give evidence of the three phases Sipress describes. We consider only one-lane systems; a recent study [RT] in *Nature* finds that the effects of passing and lane changes in congested conditions are relatively minor. From our background in interacting particle systems (cf. [Lig]), we are particularly interested in the *ergodic theory* of traffic, by which we mean the asymptotic distributions of these random streams, their domains of attraction, and their dependence on key system parameters. The natural setting for such foundational questions is an infinite highway, in which case one may investigate whether a model exhibits bona fide phase transitions

from “free flow” to “synchronized flow” to “heavy traffic” as the density of cars increases. Mathematics teaches us that the simpler the prototype, the better the prospects for rigorous results.

Traffic has been modeled for the past several decades, with either discrete or continuous state space, and in discrete or continuous time. See [LeV] for an account of one classical approach via partial differential equations. Nagel [Nag1] provides an overview of more recent fluid dynamic approaches, but then argues the merits of models discrete in both time and space: *probabilistic cellular automata* (PCA). Indeed, discrete space seems appropriate since cars are evidently discrete entities with emergent aggregate behavior on mesoscopic scales, while discrete time is of course well-suited to simulation. The PCA of [NS], [Nag1], [NP], [SSNI], and related work, typically involve occupancy of the sites of the one-dimensional integer lattice by a set of cars, with at most one car per site, each car having a velocity from a discrete set $\{0, \dots, v_{\max}\}$. Transitions are governed by additional parameters which control acceleration, braking, and random fluctuations. The resultant algorithm is easily implemented on a computer, but far too complicated for exact mathematical analysis.

After much experimentation we have identified a simpler PCA, which seems to capture all the essential features discovered by Nagel and his coworkers, but which offers brighter prospects for theorems and proofs. Our *Traffic CA*, which we denote hereafter as TCA, is as follows. At time $t = 0$, cars are situated at certain sites of the lattice \mathbb{Z} , at most one car to a site. We represent an occupied site by a 1, a vacant site by a 0. The TCA will typically start from the initial probability distribution

$$\mu_\rho = \text{Bernoulli product measure with density } \rho,$$

which assigns a car to each site independently with probability $\rho \in [0, 1]$. Then at each time $t \geq 0$, a car at x advances to site $x + 1$ at time $t + 1$ provided that two conditions hold: (i) there is no car occupying $x + 1$ at time t , and (ii) an independent coin flip succeeds, where the probability of success is determined by the occupancy of sites $x - 1$ and $x + 2$ according to Table I.

Table I. Traffic Cellular Automaton Update Rule

| transition type | $(x-1)$ | x | $(x+1)$ | $(x+2)$ | probability of advance |
|-----------------|---------|-----|---------|---------|------------------------|
| accelerating | 1 | 1 | 0 | 0 | α |
| braking | 0 | 1 | 0 | 1 | β |
| congested | 1 | 1 | 0 | 1 | γ |
| driving | 0 | 1 | 0 | 0 | δ |

Let ξ_t denote the state of the whole system, or *configuration*, at time t , and write $\xi_t(x) = 1$ if a car occupies the site x at time t , or $\xi_t(x) = 0$ if x is vacant at time t . We write $\xi_t = \xi_t^\rho$ when we need to make explicit the dependence on the initial probability distribution μ_ρ . Using this notation, the *accelerating* transition can be described as follows: $\xi_{t+1}(x+1) = 1$ with probability α provided $\xi_t(x-1) = \xi_t(x) = 1$ and $\xi_t(x+1) = \xi_t(x+2) = 0$. The idea is that this situation arises most often when a car has reached the front of a jam, sees plenty of room ahead, and so increases its speed. Similarly, the *braking* transition occurs most often when a car is about to join the back of a jam, and the *congested* transition applies to cars within a jam. The *driving* transition occurs when a car is surrounded by wide open space. To summarize: cars which can move do so *synchronously*, advancing independently according to whichever one of the four update probabilities applies. Informally, we refer to these update probabilities as “rates.” To avoid confusion about the interpretation given here, note that higher β means higher probability of moving, hence less of a “braking” effect in this terminology. We were surprised to learn recently that a continuous-time cousin of TCA with quite different system behavior was considered more than 15 years ago by Katz, Lebowitz and Spohn [KLS]. A version of that continuous-time model was studied recently in a traffic context by Antal and Schuetz [AS].

Let us illustrate some characteristics of TCA dynamics with a sample simulation. We approximate the infinite highway \mathbb{Z} by a finite lattice of $L = 500$ sites, with wrap-around at the edges. As plausible but rather ad hoc parameters, we choose $(\alpha, \beta, \gamma, \delta) = (0.5, 0.4, 0.3, 0.9)$. We run the system for 2,000 updates starting from a random distribution of cars μ_ρ , first with density $\rho = 0.2$, then with $\rho = 0.4, 0.6$, and 0.8 . The final 400 updates of each run are shown in the four space-time plots of Fig. 1, with time running from top to bottom.

For density $\rho = 0.2$, each car travels at close to its maximal velocity 0.9, delayed only by intermittent and short-lived local congestion. This is *Phase 1* of the traffic trichotomy. However, at $\rho = 0.4$ we see the emergence of persistent jams moving backwards along the highway, as is typical of real traffic above a critical “free flow” density. Note that individual cars seem to alternate between periods of free flow and jams in this regime. Next, when $\rho = 0.6$, the pattern is more suggestive of *Phase 2* synchronized flow, a congested steady state with rather long-range correlations. Finally, at $\rho = 0.8$, movement is *Phase 3*, which is to say that it is stop and go with extended regions of gridlock.

Of course, the progression observed here may only be qualitative, rather than a sequence of true *phase transitions* in the mathematical sense: abrupt quantitative changes in system behavior marked by singularities at

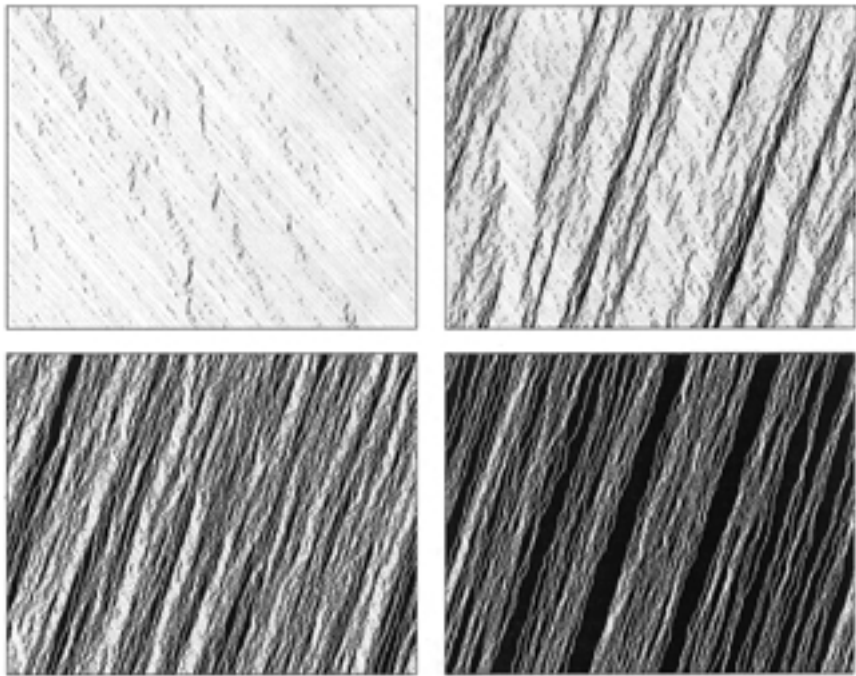


Fig. 1. The $(0.5, 0.4, 0.3, 0.9)$ Traffic Cellular Automaton on 500 sites, from 1600 to 1999 updates, with initial densities $\rho = 0.2, 0.4, 0.6, 0.8$ (top left, top right, bottom left, bottom right).

parameter values located along boundaries in the phase portrait. Our primary object in this paper is to argue that the TCA on \mathbb{Z} , with suitably chosen values of $\alpha, \beta, \gamma, \delta$, does indeed exhibit several phase transitions as the density of cars increases. In particular, the $\rho = 0.4$ case of Fig. 1 is suggestive of an intriguing *clustering* of critical free flow and a self-organized jam equilibrium which occurs for certain parameter values.

The first step toward precise characterization of TCA phases is to introduce *throughput*

$$\theta = \theta(\rho) = \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{s=0}^{t-1} P(\xi_s(0) = 1, \xi_{s+1}(0) = 0), \quad (1.1)$$

i.e., the asymptotic rate at which cars pass a given location along the highway. In physics contexts (cf. [KS]) the quantity θ is known as *flux*. For our TCA, $\theta \leq \rho$ since cars can advance at speed at most 1. Of course, there is no a priori guarantee that the above limit exists, but there is every

indication that the system converges to *equilibrium* for all initial densities. That is, we believe that for each initial probability distribution μ_ρ (introduced earlier), there exists a probability measure ν_ρ , possibly a mixture, such that

$$P(\xi_t^\rho \in \cdot) \Rightarrow \nu_\rho \quad \text{as } t \rightarrow \infty \quad (1.2)$$

in the sense of weak convergence, so that the flow $\xi_t^{\nu_\rho}$ started from ν_ρ is stationary, and

$$\theta = P(\xi_t^{\nu_\rho}(0) = 1, \xi_{t+1}^{\nu_\rho}(0) = 0) \quad \text{for all } t \geq 0. \quad (1.3)$$

Indeed, there seems to be no simple example of a local homogeneous interaction on $\{0, 1\}^Z$ for which the convergence (1.2) does not take place. Throughout the remainder of the paper we will take care to make clear when the existence of θ is provable, and when we must rely on the Ansatz that limit (1.1) exists.

Traffic researchers capture basic system statistics by means of the *fundamental diagram*, which plots throughput as a function of density. According to Nagel and others (cf. [KNS], [CSS]), a typical real-world graph looks something like the “reverse lambda” of Fig. 2, where the gray lines mark approximate boundaries between Phases 1, 2, and 3 of the trichotomy, from left to right. The short dotted line indicates a metastable free flow phase at certain densities. For these values of ρ , random fluctuations in the traffic eventually cause jams to form, and then the throughput drops to the lower value indicated by the solid line. Levels of θ in the intermediate case of synchronized flow are less than the maximal free flow, so in mixed patterns like the second frame of Fig. 1 the jams gradually

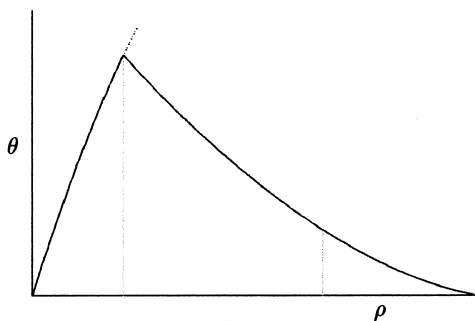


Fig. 2. The “reverse lambda” form of a realistic fundamental diagram for traffic flow.

move backward along the highway. In general, as we shall see later in the paper, whether θ is increasing or decreasing at a given density depends very subtly on system parameters.

Another natural quantity to measure is the average speed σ of a car. Since cars cannot pass one another, translation invariance and (1.2) imply that every car has asymptotic speed σ , where σ is the conditional expected speed of a car at the origin under ν_ρ , given that there is a car at the origin. Since ρ is the probability that there is a car at the origin, it follows that $\theta = \rho \cdot \sigma$, or

$$\sigma = \frac{\theta}{\rho}, \quad (1.4)$$

so the graph of speed vs. density is easily determined from the fundamental diagram.

But the question remains: *Precisely what are jams in a Traffic CA?* The remainder of this paper is devoted to a mathematical formulation of the answer to this question, and to explaining how jams emerge from free flow and related matters. The organization is as follows. We begin in Section 2 by identifying various special cases of TCA which have been studied previously or which have structural properties that make them particularly amenable to rigorous analysis. After Section 2 we begin the main focus of this study: the ergodic theory of *Cruise Control*, which is the case with $\delta = 1$.

An important aspect of Cruise Control is the ease with which *free flow* can be identified (both formally, and visually in computer experiments). If $\delta = 1$ then well-spaced cars advance deterministically with speed 1, so free flow is characterized by the condition $\sigma = 1$, which is equivalent to $\theta = \rho$ and the *maximal free flow density* ρ_* can be defined as

$$\rho_* = \sup\{\rho : \theta(\rho) = \rho\}. \quad (1.5)$$

In Section 3, we consider three extreme cases for which ρ_* and the fundamental diagram can be exactly determined: (a) $\alpha = 0$; (b) $\beta = 1$; and (c) $\beta = 0$. We also give some results for the case (d) $\alpha = 1$. Section 4 presents a proof of the lower and upper bounds

$$\frac{\alpha}{1+2\alpha} \leq \rho_* \leq \frac{1}{3} \quad (0 < \alpha, \beta < 1) \quad (1.6)$$

for Cruise Control, establishing a nontrivial phase transition from free flow to traffic with speed less than 1. Our lower bound gives a rough sense of the critical value's dependence on the acceleration α .

Section 5 begins our empirical investigation of the *Symmetric Cruise Control* case, for which $\gamma = \delta = 1$. By setting γ equal to δ , we impose a "particle/hole symmetry" that simplifies the mathematics somewhat. The remaining two parameters live in the unit square $\mathcal{S} = \{(\alpha, \beta) : 0 \leq \alpha, \beta \leq 1\}$. We find that this square is divided into three regimes with qualitatively different ergodic behavior. In one region the traffic exhibits only free flow (Phase 1), "anti-free flow" (Phase 3), or a mixture of the two. In another region, backward-moving synchronized jams (Phase 2) emerge at suitable traffic densities, while the third region gives rise to forward-moving jams (again Phase 2). We offer a complete description of asymptotic system behavior depending on which of the three regions includes (α, β) , and where the traffic intensity ρ falls in relation to ρ_* and another more mysterious critical value called ρ^* . Our understanding is based on extensive computer simulation, together with analysis of the interface dynamics between free flow, jams, and anti-free flow. Experimentalists will likely be convinced by the evidence we offer, whereas purists may prefer to interpret our findings as a series of challenging conjectures worthy of verification.

Of particular interest is Symmetric Cruise Control behavior at densities just above ρ_* , as jams begin to form. Perhaps the most significant discovery of our computer simulations is the *self-organized critical clustering* which occurs for such densities throughout the interior of \mathcal{S} . The TCA evolves into patches of ever-increasing length, alternating between stretches of *maximal free flow* with distribution ν_{ρ_*} and stretches of a *minimal ergodic jammed state* with distribution ν_{ρ^*} . In particular, the law of ξ_t^ρ converges to a proper mixture:

$$P(\xi_t^\rho \in \cdot) \Rightarrow c\nu_{\rho_*} + (1-c)\nu_{\rho^*} \quad \text{as } t \rightarrow \infty, \quad (1.7)$$

where c , of course, satisfies $c\rho_* + (1-c)\rho^* = \rho$. This clustering is quite subtle close to the various phase boundaries in \mathcal{S} and also near the edges of \mathcal{S} , so in Section 6 we present extensive numerical data for one of the clearest examples of the phenomenon: $\alpha = \beta = 0.6$. As one manifestation of (1.7), our simulated fundamental diagram for this particular case supports the presence of a *flat piece*, i.e., linearly interpolating throughputs over an interval of traffic densities which includes the interval (0.32, 0.40).

Next, in Section 7 we offer a preliminary description of Symmetric Cruise Control near the right ($\alpha = 1$), lower ($\beta = 0$), and upper ($\beta = 1$) edges of \mathcal{S} . The ultimate goal is to prove rigorous results for perturbations of the exact cases analyzed in Section 3. Near the right edge, traffic is

described by diffusing particles and ballistic antiparticles on a background ether of ideal free flow. Near the bottom edge, and in the top left corner, there are related particle/antiparticle representations. We hope to exploit these perturbation schemes in the future to obtain theorems which establish clustering and other basic features of “slow-to-start” Cruise Control.

Section 8 contains the most speculative, but also the most intriguing case study of the paper. We identify a slow-to-start TCA ($\alpha < \delta$) with all four parameters strictly between 0 and 1, in particular not Cruise Control, for which clustering and convergence to a mixture (1.7) seem to occur. This is apparently an example of a local interacting system, with particle conservation but otherwise irreducible in a natural sense, for which the limiting measure ν_ρ (corresponding to a suitable choice of density ρ) is not extreme. The mechanism whereby this system achieves arbitrarily large length scales is still poorly understood, but certainly warrants further investigation. Finally, Section 9 provides pointers to software and Web resources without which this research would have been unthinkable, and which the diligent reader will find indispensable for understanding the subtleties of emergent traffic jams.

2. SPECIAL TRAFFIC CELLULAR AUTOMATA AND THEIR BASIC PROPERTIES

Here we catalogue and discuss some examples of TCA, using the notation $(\alpha, \beta, \gamma, \delta)$ to designate their parameters.

(1, 1, 1, 1) Rule 184

In perhaps the simplest of all nontrivial discrete directed flows, every particle moves one step to the right provided that cell is vacant. According to Wolfram’s unenlightening taxonomy [Wol], this cellular automaton is called Rule 184. The model was first discussed in the context of traffic in [BML]. Even though the dynamics are deterministic, one can study the limit distributions starting from μ_ρ and other random initial states, explore the dependence of throughput (flux) on density, and so forth. This program has been carried out in a recent paper by Belitsky and Ferrari [BF]. They give a complete description of the invariant measures for Rule 184, and show that the throughput is

$$\theta(\rho) = \begin{cases} \rho, & \rho \leq \frac{1}{2} \\ 1 - \rho, & \rho > \frac{1}{2}. \end{cases} \quad (2.1)$$

See [BKNS] for a simple two-lane variant of Rule 184 in which lane changes increase the throughput at some intermediate traffic densities, but not in free flow or heavy congestion.

($\alpha, \alpha, \alpha, \alpha$) Synchronous Totally Asymmetric Simple Exclusion Process (STASEP)

This is the discrete-time version of one of the most widely studied interacting particle systems, the Asymmetric Simple Exclusion Process (ASEP). See [Lig] for basic theory of this and related exclusion models in continuous time. Some beautiful recent work on the ASEP appears in [Rez] and [Sep]. The discrete-time STASEP is both a generalization of Rule 184, and a special case of the original Nagel-Schreckenberg model with maximum velocity 1 (see [Nag1], Section IV). The equilibria for the STASEP were first characterized in [Yag]. (See also [SSNI], where these same results were obtained). Under these dynamics there are no persistent jams of the type that interest us. That is, clustering (1.7) cannot arise in the STASEP. Its throughput is given by

$$\theta(\rho) = \frac{1 - \sqrt{1 - 4\alpha\rho(1 - \rho)}}{2},$$

which matches (2.1) when $\alpha = 1$.

$\alpha < \delta$ Slow-to-Start

Perhaps the most significant discovery of Kai Nagel and his colleagues during their extensive empirical and simulation studies of traffic flow was the role of a simple principle, as summarized in [KNW]:

In order for jams to be stable, the reaction time and thus the minimum time headway needs to be smaller than the “escape” time. ... A way to obtain models that represent this aspect of the dynamics correctly is the use of so-called *slow-to-start* rules. One simply slows down acceleration from low speeds.

Whether mechanical or psychological in origin, this dynamical property seems key to the emergence of traffic jams. For TCA, the condition becomes $\alpha < \delta$, i.e., that the driving rate exceed the acceleration rate. Note that our previous example was not slow-to-start. Throughout the remainder of this paper we will treat only slow-to-start systems, since these are the ones which give rise to realistic jams such as those depicted in Fig. 1.

($\alpha, \beta, \gamma, 1$) Cruise Control

An important simplifying feature already mentioned in the Introduction is Cruise Control: the case $\delta = 1$ in which well-spaced cars travel deterministically at maximum velocity (1 cell per update). The terminology here is

due to Nagel and Schreckenberg [NS], who considered such deterministic free flow dynamics as limits of their more elaborate traffic models. Indeed, the $(\alpha, 1, \alpha, 1)$ TCA, like STASEP, may be viewed as a special case of the NS model with maximum velocity 1. As explained earlier, Cruise Control has the advantage that free flow and the maximal free flow density ρ_* are easily defined. Our approach to the complexity of TCA dynamics is to first understand the $\delta = 1$ case as fully as possible, and then to study models with $\delta < 1$ as perturbations of Cruise Control.

(\alpha, 1, \gamma, 1) Slow-to-Start TCA with Computable Throughput

In this case, isolated cars advance deterministically. Since $\beta = 1$, it can be shown that the system either enters free flow, or it clusters into alternating stretches of free flow and solid blocks. Because of this simple structure, we have been able to compute the throughput exactly:

$$\theta(\rho) = \begin{cases} \rho, & 0 \leq \rho \leq \rho_* \\ (1 - \rho) \frac{\alpha}{1 + \alpha - \gamma}, & \rho_* < \rho \leq 1, \end{cases} \quad \text{where } \rho_* = \frac{\alpha}{1 + 2\alpha - \gamma}. \quad (2.2)$$

In a follow-up paper, we will give the proof of (2.2) and other facts about this case. As indicated above, this is a special case of the NS model when $\alpha = \gamma$.

In order to understand the remaining examples, it is useful to think of empty spaces as *anticars* that move to the left. Modulo this change of direction, it is easy to check that the dynamics of the anticar process are exactly the same as for cars, except that the roles of parameters γ and δ are reversed.

(\alpha, 1, 1, \delta) Convoys

This somewhat strange model is obtained from the $(\alpha, 1, \gamma, 1)$ model by switching the cars and anticars and also switching left and right directions. The cars in the new model cluster into longer and longer “convoys,” where a convoy is defined to be a cluster of cars in which there is at most one space between any two successive cars. The lead car of each such cluster moves forward (at speed α or δ , depending on the local circumstances) until it arrives within one unit of the back of the cluster ahead of it, at which time the two convoys have coalesced. While reminiscent of certain East European highways where travel is regulated by a smattering of Trabants which max out at 45mph, this is not a terribly realistic traffic flow model. In reality, one expects cars to be able to escape from the front of a

jam. Obviously, the throughput for this case can be obtained from (2.2): simply replace γ by δ and ρ by $(1 - \rho)$.

$(\alpha, \beta, \delta, \delta)$ Symmetric TCA

We can impose a “car/anticar symmetry” (also known as “particle/hole symmetry”) by assuming $\gamma = \delta$. Admittedly this seems unrealistic, since one would normally expect cars to move more slowly in congestion than in the open. Note, however, that under this assumption cars within a jam still move more slowly than those in free flow because congestion affords less chances to advance. Our motivation here is purely mathematical: the car/anticar symmetry implies that the fundamental diagram is symmetric about $\rho = \frac{1}{2}$. Extensive simulation suggests that while this reduction alters quantitative details, the essential qualitative aspects of TCA system behavior remain intact, at least for $\rho \leq \frac{1}{2}$. Over that range of traffic densities we shall see that symmetric fundamental diagrams remain true to the spirit of Fig. 2.

$(\alpha, \beta, 1, 1)$ Symmetric Cruise Control

This is of course the intersection of the Cruise Control and Symmetric cases. Our empirical investigations in Sections 5 through 8 are primarily devoted to this case.

We close this section by introducing an *interface representation* for TCA dynamics which will prove useful later on, and which is interesting in its own right. This sort of construction goes back to Rost [Ros], and Krug and Spohn [KS] for particle systems. See also [EG] and [Sep] for more recent treatments. At each discrete time t , the interface is the graph interpolating between points $\{x, \eta_t(x)\}$ of a map on \mathbb{Z} such that adjacent sites are mapped to integer values which differ by ± 1 :

$$|\eta_t(y) - \eta_t(x)| = 1 \quad \text{whenever} \quad |y - x| = 1.$$

In other words, the interface consists of a doubly infinite sequence of diagonal bonds of slope ± 1 linked end to end (see Fig. 3). Cars are identified with the heads of downward pointing arrows, empty spaces with the heads of upward pointing arrows. To normalize the initial position of the interface, set $\eta_0(0) = 0$. The dynamics are directed up—relative minima advance two units, thereby becoming relative maxima, with probabilities determined by the slopes of the neighboring bonds, as summarized in Fig. 3 (light gray being the new position). With the identification between cars, spaces, and arrows given above, we have

$$\xi_t(x) = \frac{1}{2} (1 + \eta_t(x) - \eta_t(x - 1)),$$

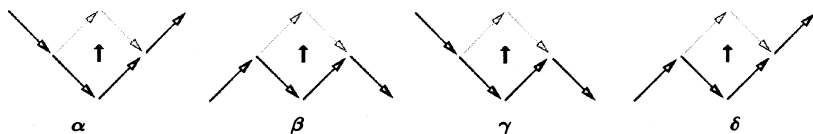


Fig. 3. Interface representation of TCA dynamics.

and it is easy to check that the transitions of Fig. 3 correspond exactly to the TCA update rule of Table I. Observe that the throughput θ of (1.2) is half the interface's speed of advance at $x = 0$,

$$\theta = \lim_{t \rightarrow \infty} \frac{1}{2t} \eta_t(0)$$

(again, assuming this limit exists). The interface representation also clearly shows the relationship between the motion of cars to the right and spaces to the left, since interchanging the values of γ and δ is equivalent to interchanging upward and downward arrows together with the left and right directions.

3. CRUISE CONTROL: ANALYSIS OF EXTREME CASES

In this section we prove a theorem which identifies the exact ergodic behavior of the Cruise Control model ($\delta = 1$) in the various special cases when either α or β equals 0 or 1. An informal description of our results is as follows:

(a) Not surprisingly, when the acceleration α is 0, then all cars eventually grind to a halt, so the throughput is identically equal to 0.

(b) The case $\alpha > 0$ and $\beta = 1$ was already mentioned in the previous section ("Slow-to-Start TCA with Computable Throughput"). The throughput can be computed exactly, and clustering occurs. In this paper, we will only give the proof under the additional assumption that $\gamma = 1$ (Symmetric Cruise Control); the proof for more general γ will appear elsewhere.

(c) For $\alpha > 0$ and $\beta = 0$, the critical value ρ_* and throughput θ are independent of $\alpha > 0$. We give a fairly complete description of the systems' limiting configurations. In the asymmetric case ($\gamma < 1$), the ergodic behavior exhibits an interesting transition as the density crosses $\rho = \frac{1}{2}$.

(d) Asymptotics for the case $\alpha = 1$ are somewhat involved, depending on β in a way which seems to defy exact calculation. Nevertheless, it can be shown that the maximal free flow density is $\rho_* = \frac{1}{3}$ when $\gamma = 1$.

Here is some terminology that will be used in the statement and proof of the theorem, as well as subsequent discussion. We will assume that each car is labeled in some unique way, so that the cars retain their identity throughout the time evolution. We say that a given car is *permanently stuck* at time t if the car has speed 0 at all times $\geq t$. Similarly, a car is in *permanent free flow* at time t if the car has speed 1 at all times $\geq t$, and a car is *permanently isolated* (respectively *permanently double-isolated*) at time t if the car is separated by at least one (respectively two) vacant site(s) from all other cars at all times $\geq t$. Becoming permanently isolated is a necessary condition for entering permanent free flow. When $\beta = 1$, it is also sufficient. When $\beta < 1$, it is almost surely the case that a car enters permanent free flow if and only if it becomes permanently double-isolated. Note that if one car becomes permanently stuck, then eventually all cars to the left of it also become permanently stuck. Since the initial configuration is always translation invariant and ergodic in our model, it follows from standard arguments that there is a 0–1 law: either all cars become permanently stuck with probability 1, or the probability is 0 that any car becomes permanently stuck. A similar argument shows that a 0–1 law also applies to achieving permanent free flow: either all cars enter permanent free flow with probability 1, or the probability is 0 that any car enters permanent free flow. It is natural to make the following conjecture, which we believe applies to the general Cruise Control case:

Conjecture. If $\rho < \rho_*$, then all cars enter permanent free flow with probability 1.

If two cars are not separated by another car, then we say they are *consecutive*. The (possibly empty) set of vacant sites between two consecutive cars is called the *gap* between them, and the number of such sites is the *size* of the gap. If the gap between two cars has size 0, then we say that there is a *bond* between the cars. Suppose there are three consecutive cars, with a bond between the two on the left, and a gap of size 1 between the two on the right, giving the configuration 1101. If the middle car moves and the car on the right does not, then we say that the bond between the two cars on the left *moves* one step to the right. Bonds can only move to the right. Bonds can also be *created* or *destroyed* when cars move. However, when $\beta = 0$, it is easily seen that bonds cannot be created. For anticars (vacant spaces), we define *antigaps* and *antibonds* analogously. Of course, an antigap is simply an interval of consecutive cars. An antigap of size at least 2 is sometimes called a *solid block*. This redundant terminology allows us to efficiently use symmetry in the arguments below.

Here are some consequence of our first theorem which the reader may find surprising. First, from parts (a) and (c), we see that the throughput θ does not always depend continuously on the parameters of the model. In particular, there is a discontinuity at $\alpha = \beta = 0$. Second, from parts (b) and (c), we learn that the dependence of θ on the parameter β is not monotonic. Namely, (3.2) shows that $\theta(\frac{1}{3}) = \frac{1}{3}$ when $\beta = 0$, whereas (2.2) shows that when $\beta = 1$, the value of $\theta(\frac{1}{3})$ can be either strictly less than or strictly greater than $\frac{1}{3}$, depending on the value of γ . Our explicit formulas for the throughput are nondecreasing in the parameter α , and we conjecture that this monotone dependence on α holds in general for the Cruise Control model. Finally, note that the fundamental diagram is not unimodal in case (c) for any $\gamma > 0$.

Theorem 1. Let ξ_t^p be Cruise Control with acceleration α , braking β , congestion γ , and traffic density ρ .

(a) If $\alpha = 0$ then every car becomes permanently stuck with probability 1, and $\theta(\rho) \equiv 0$. (This conclusion also holds for general δ .)

(b) Let $\alpha > 0, \beta = 1$. If $\gamma < 1$ then the critical density ρ_* and throughput $\theta(\rho)$ are given by (2.2). If $\gamma = 1$ (Symmetric Cruise Control), then the following statements hold. If $0 \leq \rho < \frac{1}{2}$, then every car becomes permanently isolated and enters permanent free flow, with probability 1; a corresponding result for $\frac{1}{2} < \rho \leq 1$ is obtained by using car/anticar symmetry. If $\rho = \frac{1}{2}$, then isolation occurs in the following weaker sense:

$$P(\xi_t(x) = \xi_t(x+1) = 1) \rightarrow 0 \quad \text{for all } x \quad \text{as } t \rightarrow \infty.$$

The throughput θ has the form (cf. the first graph of Fig. 4)

$$\theta(\rho) = \begin{cases} \rho, & 0 \leq \rho \leq \frac{1}{2} \\ 1 - \rho, & \frac{1}{2} < \rho \leq 1. \end{cases}$$

(c) Let $\alpha > 0, \beta = 0$. If $0 \leq \rho < \frac{1}{2}$, then every car becomes permanently isolated with probability 1, while if $\rho = \frac{1}{2}$, then

$$P(\xi_t(x) = \xi_t(x+1)) \rightarrow 0 \quad \text{for all } x \quad \text{as } t \rightarrow \infty.$$

Furthermore, if $0 \leq \rho < \frac{1}{3}$, then every car becomes permanently double-isolated and enters permanent free flow, with probability 1. On the other hand, if $\frac{1}{3} \leq \rho \leq \frac{1}{2}$, then for all x ,

$$\begin{aligned} P(\xi_t(x-1) = 0, \xi_t(x) = 1, \xi_t(x+1) = \xi_t(x+2) = 0, \xi_t(x+3) = 1) \\ \rightarrow 1 - 2\rho \quad \text{as } t \rightarrow \infty, \end{aligned} \tag{3.1}$$

and

$$P(\xi_t(x-1) = 0, \xi_t(x) = 1, \xi_t(x+1) = 0, \xi_t(x+2) = 1) \rightarrow 3\rho - 1 \quad \text{as } t \rightarrow \infty.$$

For $\rho > \frac{1}{2}$, all of the anticars (vacant spaces) eventually become permanently isolated, after which they behave in a manner equivalent to the STASEP introduced in Section 2, with $q = \gamma$. (See the proof of Theorem 1 for a more precise description of this behavior.) The throughput θ is given by (cf. the second graph of Fig. 4 for the case $\gamma = 1$)

$$\theta(\rho) = \begin{cases} \rho & 0 \leq \rho \leq \frac{1}{3} \\ 1 - 2\rho & \frac{1}{3} < \rho \leq \frac{1}{2} \\ \frac{\rho - \sqrt{\rho^2 - 4\gamma(2\rho - 1)(1 - \rho)}}{2} & \frac{1}{2} < \rho \leq 1. \end{cases} \quad (3.2)$$

(d) Let $\alpha = 1$, $0 < \beta < 1$, and $\gamma = 1$. For each ρ , the throughput $\theta(\rho)$ exists and is given by (1.1). When $\rho < \frac{1}{3}$, all cars enter permanent free flow almost surely, and when $\rho = \frac{1}{3}$, (3.1) holds. When $\frac{1}{3} < \rho \leq \frac{1}{2}$, the cars almost surely do not enter permanent free flow, and all gaps and all solid blocks eventually have size no greater than 2. In particular, $\rho_* = \frac{1}{3}$.

Proof. In each case, the proof is trivial if $\rho = 0$, so let us assume $\rho > 0$ throughout.

(a) With probability 1, there is a configuration of the form 1100 somewhere in the initial state. Since $\alpha = 0$, the two cars in this configuration

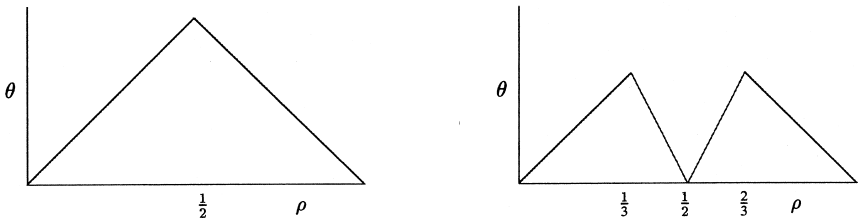


Fig. 4. Fundamental diagrams for the top and bottom edges of the SCC phase portrait.

are permanently stuck. The 0–1 law mentioned above now implies that all cars become permanently stuck with probability 1. (Clearly, this argument also works for general δ).

(b) As mentioned earlier, we only present the case where $\gamma = 1$ here. The more general case will be treated in a separate paper. Consider a gap of size at least 2 in the initial configuration. Since $\alpha > 0$, the car immediately to the left of this gap eventually moves forward, with probability 1, thereby reducing the length of the gap by 1 at its left end. Since δ equals 1, the car continues to move forward at speed 1, at least until all of the antibonds in the gap have disappeared. Furthermore, since β and γ both equal 1, a gap cannot grow if it has size 1. In other words, once the car immediately to the left of a gap has moved at least once, the number of antibonds in that gap cannot increase.

It is easy to see from this argument that if the car immediately to the right of a gap does not enter permanent free flow, then from some point on, the gap contains no antibonds. Because of the 0–1 law for permanent free flow, we conclude that if any car does not eventually enter permanent free flow, then with probability 1, all gaps eventually have size bounded above by 1. If $\rho < \frac{1}{2}$, the average gap size is greater than 1, so it must be the case that all cars enter permanent free flow with probability 1. In the case that $\rho = \frac{1}{2}$, suppose that all cars do not enter permanent free flow with probability 1, so that all gaps eventually have size bounded above by 1. The car/anticar symmetry implies in this case that, with probability 1, all antigaps eventually have size bounded above by 1, or in other words, that all solid blocks almost surely eventually disappear. The weaker conclusion for the case $\rho = \frac{1}{2}$ now follows easily.

(c) Since $\beta = 0$, it is easy to see that it is not possible for new bonds or antibonds to be created. Bonds move to the right with speed γ , unless they are part of a 1100 configuration. Similarly, antibonds move left with speed $\delta = 1$ unless they are part of such a configuration. When a bond and an antibond meet in a 1100 configuration, they destroy one another once the indicated acceleration transition occurs.

Now consider a particular site x . If $\rho < \frac{1}{2}$, it follows from the Strong Law of Large Numbers that every sufficiently large interval containing x has more antibonds than bonds in the initial state. Hence every bond is eventually annihilated with probability 1. Furthermore, if a car starts at x , it follows that only finitely many of the bonds to the left of x will catch the car before being annihilated. A car cannot catch up with a bond on its right since $\beta = 0$, so a given car can only be a part of finitely many different bonds during the time evolution. That is, every car eventually becomes permanently isolated with probability 1. The corresponding weaker conclusion

for the case $\rho = \frac{1}{2}$ follows from the Weak Law. The assertion about the throughput for the case $\rho = \frac{1}{2}$ also follows.

We now prove the remaining conclusions in part (c) for $\rho < \frac{1}{2}$, in which case every car becomes permanently isolated with probability 1. Since $\beta = 0$, permanently isolated cars have only two speeds, 1 or 0, depending on whether or not the two sites ahead of them are vacant. Consider two consecutive permanently isolated cars. It is easy to see that if the gap between them has size 2, then that size cannot increase. It follows that at least one of the following conditions holds: either the size of the gap between the two cars is eventually bounded above by 2 forever, or both cars eventually enter permanent free flow. The 0–1 law for permanent free flow now implies that there are only two possibilities: (i) all of the cars enter permanent free flow, and hence they also all become permanently double-isolated; or (ii) no car enters permanent free flow, and it is then easy to see that, with probability 1, every car eventually “catches up” within two units of the car ahead of it, so that the gap between them eventually stays bounded above by size 2 forever. If $\rho < \frac{1}{3}$ there is too much vacant space for the second of the two possibilities to occur so every car enters permanent free flow almost surely. If $\rho > \frac{1}{3}$, there is not enough vacant space for the first possibility to occur, so every gap is eventually bounded above by size 2. Since all cars become permanently isolated with probability 1, it is also true that every gap is eventually bounded below by size 1.

To complete the case in which $\rho < \frac{1}{2}$, we need to show that the proportion of cars which have a gap of size 2 in front of them converges to $(1-2\rho)/\rho$ as $t \rightarrow \infty$. Consider the distribution at time t of the “typical” gap size, defined in the usual manner by the Ergodic Theorem, and denote this distribution by Q_t . Because cars do not appear or disappear, the expected value of this distribution equals its value at time 0, which is $(1-\rho)/\rho$. A standard formula for expected value now gives

$$\sum_{k=1}^{\infty} Q_t([k, \infty)) = \frac{(1-\rho)}{\rho} \quad \text{for all } t \geq 0.$$

Since gaps of size 2 or larger cannot grow and since all gaps eventually have size 1 or 2, $Q_t([k, \infty))$ decreases to 0 as t tends to ∞ for all $k > 2$. Since $Q_t([1, \infty)) = 1$ and $Q_t(1) = 1 - Q_t([2, \infty))$, it follows from the Dominated Convergence Theorem that $\lim_{t \rightarrow \infty} Q_t(1) = (1-2\rho)/\rho$. This is the asymptotic proportion of cars with speed 1. In the limit, all remaining cars have speed 0. The desired form of the throughput function $\theta(\rho)$ follows immediately.

We now consider the case in which $\rho > \frac{1}{2}$. By using the Strong Law as before, we see that every antibond eventually disappears. A standard large

deviations estimate for the initial state can be used to strengthen this result as follows: given any finite set of sites A , there exists a time, with probability 1, after which there are no antibonds associated with any of the sites in A . Consequently, we can determine the asymptotic behavior of the system by restricting our attention to configurations in which there are no antibonds. That is, we may assume that all of the anticars are already permanently isolated at time 0. It is not hard to check that when the state of the system satisfies this restriction, then the anticars move according to STASEP dynamics, with $q = \gamma$, except for one detail: anticars are forced to remain two units apart, rather than a single unit apart. But this is a trivial difference which is easily handled by simply suppressing one car between each pair of anticars at each time step. Such a transformation changes the density of anticars from $(1 - \rho)$ to $(1 - \rho)/\rho$. The formula for the throughput in this transformed process is obtained by substituting this density for ρ and γ for q in the throughput formula for STASEP given in Section 2. To recover the throughput of our original model, we multiply by the scaling factor ρ , thereby obtaining the desired result.

(d) The assertion about the existence of throughput in this case follows from discrete-time versions of results from [EG]. In order to prove the remaining assertions, we consider any antigap of size at least 2. Since $\alpha = 1$, cars leave the front of this antigap at rate 1. So an antigap cannot grow if it has size at least 2. By symmetry, a gap also cannot grow if it has size at least 2. Now consider two consecutive cars, separated by a gap of size greater than 2. Because $\alpha = 1$, the car to the left of the gap has speed 1. Therefore, unless the car to the right of the gap enters permanent free flow, the gap will decrease in size until it has size 1. It follows that we have the same two alternatives mentioned at the end of the second paragraph in the proof of part (c). Therefore, when $\rho < \frac{1}{3}$ all cars enter permanent free flow almost surely. We further conclude that if $\rho > \frac{1}{3}$, it is almost surely the case that no cars enter permanent free flow and the sizes of all gaps are eventually bounded above by 2. The weaker result for $\rho = \frac{1}{3}$ also follows from this argument. Finally, symmetry implies that the sizes of all antigaps are almost surely eventually bounded above by 2 if $\rho < \frac{2}{3}$, and the proof is complete. ■

We end this section with a few additional remarks about the TCA models with $\alpha = \gamma = \delta = 1$ (Symmetric Cruise Control, but not slow-to-start). As part (d) of Theorem 1 suggests, for densities just above $\rho_* = \frac{1}{3}$, these systems evolve toward a background “ether” of *ideal free flow*, namely the periodic configuration

$$\dots 100100100\dots, \tag{3.3}$$

interrupted by occasional pairs of cars separated by a gap of size 1. Call such a pair a *virtual particle*. It is easy to check that any virtual particle well separated from the others performs a random walk within the ether, advancing 1 unit with probability β (when its trailing car advances), or falling back 2 units with probability $1-\beta$ (when its trailing car stays put, and so becomes the virtual particle's lead car). Thus the expected displacement of these mini-jams is

$$1 \cdot \beta + (-2) \cdot (1 - \beta) = 3\beta - 2,$$

and the motion has positive drift for $\beta > \frac{2}{3}$, negative drift for $\beta < \frac{2}{3}$, and mean 0 for $\beta = \frac{2}{3}$. Of course, more complicated interactions play a role when virtual particles wander close together, but we will see in subsequent sections that the point $(1, \frac{2}{3})$ has special significance in the Symmetric Cruise Control phase portrait. The following conjecture is well-supported by simulations:

Conjecture. If $\alpha = \gamma = \delta = 1$ and if the initial condition for TCA is a left half-space of cars, then the border between ether on the right and antiether on the left remains tight if $\beta < \frac{2}{3}$, but not if $\beta > \frac{2}{3}$.

Letting $\theta_{\alpha, \beta}(\rho)$ denote the throughput of Symmetric Cruise Control with density ρ and parameters (α, β) , a simple coupling of interface representations yields the inequality

$$\theta_{\alpha, \beta}(\rho) \leq \theta_{1, \beta}(\rho) \tag{3.4}$$

for all α, β , and ρ . Indeed, if the (α, β) and $(1, \beta)$ systems evolve simultaneously as upward moving interfaces in the manner of Fig. 3, using the same “ β -coin,” and if the (α, β) and $(1, \beta)$ interfaces initially agree everywhere, then it is easy to check that under no circumstance can any portion of the former graph ever lie above the corresponding portion of the latter. In particular, the interface positions at 0 must maintain this ordering, and (3.4) follows.

In light of (3.4), it would be especially useful to obtain a good, explicit upper bound on $\theta_{1, \beta}(\rho)$, so we pose this as an open problem. Simulations indicate that $\theta_{1, \beta}(\frac{1}{3}) = \frac{1}{3}$ is the maximal throughput of the $(1, \beta)$ system for any $\beta < \frac{2}{3}$, which would imply that for any density ρ the throughput is at most $\frac{1}{3}$ throughout the “lower two-thirds” of \mathcal{S} . For β close to 1, presumably $\theta_{1, \beta}(\frac{1}{2})$ gives a uniform upper bound on $\theta_{\alpha, \beta}(\rho)$.

4. CRUISE CONTROL: BOUNDS ON THE MAXIMAL FREE FLOW DENSITY

We now present simple upper and lower bounds on the maximal free flow density ρ_* . These apply to the general Cruise Control model, under the assumption that $\beta < 1$. Our upper bound is trivial, but the lower bound captures something of the cutoff's dependence on acceleration. Possible improvements will be mentioned after the proof.

Theorem 2. For $\beta < 1$ and $\delta = 1$,

$$\frac{\alpha}{1 + 2\alpha} \leq \rho_* \leq \frac{1}{3}.$$

Proof. Since $\beta < 1$, we have already seen that permanent double-isolation is equivalent to eventual free flow. In other words, if $\rho < \rho_*$ then eventually there can be at most one car every three spaces along the road. But traffic density is preserved, so this implies $\rho \leq \frac{1}{3}$, which yields the upper bound.

Turning to the lower bound, we will assume $\rho > \rho_*$, and show that $\rho \geq \frac{\alpha}{1+2\alpha}$. Condition on the event that a car starts at the origin at time 0, let X_t be the position of that car at time $t \geq 0$, and let Y_t be the position of the next car to the right of X_t . Standard theory implies that $E(Y_t - X_t) = \frac{1}{\rho}$ for all $t \geq 0$. Define the following events for $0 \leq s < t$:

$$A_t = \{Y_t - X_t \leq 2\}$$

$$B_{s,t} = \{Y_s - X_s = 2 \text{ and } Y_u - X_u > 2 \text{ for } s < u \leq t\}$$

$$C_t = \{Y_u - X_u > 2 \text{ for } 0 \leq u \leq t\}.$$

Under our assumption that $\rho > \rho_*$, we have $\lim_{t \rightarrow \infty} P(C_t) = 0$. Therefore,

$$\frac{1}{\rho} \leq \lim_{t \rightarrow \infty} \left[2P(A_t) + \sum_{s=0}^t P(B_{s,t}) E(Y_t - X_t | B_{s,t}) \right].$$

Since $2 < (1 + 2\alpha)/\alpha$ for $\alpha > 0$, it follows from this last inequality that we will have the desired lower bound if we can show that

$$E(Y_t - X_t | B_{s,t}) \leq \frac{1 + 2\alpha}{\alpha} \quad \text{for all } 0 \leq s < t.$$

Actually, we will show a little more; namely that

$$E(Y_t - X_t | B_{s,t}, Y_t - Y_s, \mathcal{F}_s) \leq \frac{1+2\alpha}{\alpha} \quad \text{for all } 0 \leq s < t$$

where \mathcal{F}_s is the σ -field generated by the traffic process up to time s .

If there is a vacant site immediately to the left of X_{s+1} , then it is easy to see that $Y_t - X_t = 3$ if the event $B_{s,t}$ occurs. Since $\frac{1+2\alpha}{\alpha} > 3$ when $\alpha < 1$, we therefore only need to consider the case in which there is a car at the site immediately to the left of X_{s+1} . In this case, we can imagine that a coin is tossed at each time unit, starting at time $u = s+1$. The probability of Heads is α , and the car at X_u does not move until Heads appears. The coin tosses are independent of the σ -field \mathcal{F}_s , but they are not independent of the event $B_{s,t}$ or of the quantity $Y_t - Y_s$. We will now consider the nature of the dependence.

Let $K = (t-s) - (Y_t - Y_s)$. If $B_{s,t}$ occurs, then a little thought reveals that the results of the coin tosses at the times $s+1, \dots, s+K$ must all be Tails, since otherwise we would have $Y_u - X_u = 2$ for some u in the interval $(s, t]$. A little more thought reveals that the coin tosses after time $s+K$ are independent of $B_{s,t}$ and $Y_t - Y_s$, and that $Y_t - X_t = 3 + N$ where N equals the number of consecutive Tails obtained starting at time $s+K+1$.

Thus, we have generally that if the event $B_{s,t}$ occurs, then $Y_t - X_t \leq 3 + N$, and N is independent of $B_{s,t}$, $Y_t - Y_s$, and \mathcal{F}_s . Since $E(N) = \frac{1-\alpha}{\alpha}$ it follows that

$$E(Y_t - X_t | B_{s,t}, Y_t - Y_s, \mathcal{F}_s) \leq 3 + \frac{1-\alpha}{\alpha} = \frac{1+2\alpha}{\alpha},$$

as desired. ■

Even though cars arranged as density $\frac{1}{3}$ ideal free flow (3.3) all advance with maximal speed 1, simulations strongly suggest that $\rho_* < \frac{1}{3}$ whenever $\alpha, \beta, \gamma < 1$. Confirmation of this strict inequality remains an open problem for Cruise Control dynamics. By (1.4), such a result would imply that random initial distributions with certain densities below $\frac{1}{3}$ have cars traveling at less than maximal speed due to jamming. The timed on-ramps in Los Angeles, Minneapolis [Bla], and elsewhere attempt to inject order into an otherwise chaotic traffic stream, an ambitious control strategy in light of the inherent instability of ideal free flow.

5. SYMMETRIC CRUISE CONTROL: EMPIRICAL STUDY OF THE PHASE PORTRAIT

In Symmetric Cruise Control (SCC), with $\gamma = \delta = 1$, we have what may be the simplest prototype for the complexities of traffic flow. This model is defined by only three parameters: acceleration α , braking β , and the traffic intensity ρ . Thus, for each point (α, β) in the unit square $\mathcal{S} = [0, 1] \times [0, 1]$ there is a family of SCC processes ξ_t^ρ , $\rho \in (0, 1)$, with throughput function $\theta(\rho)$ as well as more detailed density-dependent ergodic behavior. Most of the rest of this paper is devoted to a mathematically precise description of this *phase portrait*. A related, but inherently qualitative look at traffic behavior in terms of acceleration and braking was carried out by Krauss [Kra] in his dissertation, using algorithms of Nagel *et al.* For other intriguing examples of two-dimensional phase portraits for nonlinear interacting stochastic systems, see [DN] and [MDDGL].

To begin our study of the SCC phase portrait we consider its left, top, bottom, and right edges

$$\begin{aligned} \mathcal{E}_L &= \{\alpha = 0\}, \\ \mathcal{E}_T &= \{\alpha > 0, \beta = 1\}, \\ \mathcal{E}_B &= \{\alpha > 0, \beta = 0\}, \\ \mathcal{E}_R &= \{\alpha = 1, 0 < \beta < 1\}. \end{aligned}$$

We have seen in parts (a), (b), and (c) of Theorem 1 that the processes ξ_t^ρ have quite simple limiting behavior along the first three of these edges, and that their fundamental diagrams are exactly computable. SCC dynamics on \mathcal{E}_R are more involved, and not slow-to-start, but have regularity properties which allowed us to establish the existence of $\theta(\rho)$ along this edge as well. By contrast, we are presently able to prove existence of throughputs in $\mathcal{S}^\circ = (0, 1) \times (0, 1)$ only in cases where the system goes into free flow.

More detailed rigorous results on Symmetric Cruise Control seem hard to come by, so we turn to simulation for a preliminary understanding of traffic patterns in the interior \mathcal{S}° of the phase diagram. We start with the following experimental observations, which relate to the results in Theorems 1 and 2. For the mathematician, these are also conjectures, some cases of which may not be too difficult to prove.

$$\rho_* \rightarrow \frac{1}{3} \text{ as } (\alpha, \beta) \text{ approaches } \mathcal{E}_B \cup \mathcal{E}_R. \tag{5.1}$$

$$\begin{aligned} \rho_* \rightarrow \varphi(\alpha) \text{ as } (\alpha, \beta) \rightarrow (\alpha, 1) \text{ from within } \mathcal{S}^\circ, \text{ where } \varphi \text{ is a continuous} \\ \text{function that equals 0 at } \alpha = 0 \text{ and } \frac{1}{3} \text{ at } \alpha = 1. \end{aligned} \tag{5.2}$$

Further experimental work is facilitated by *border analysis*—the study of the dynamics of the region that divides traffic jams and free flow. Such analysis provides a rather complete picture of SCC ergodic behavior. Our basic *experiment* with density ρ runs on a large torus of N sites, with cars initially in a solid block of size $\lfloor \rho N \rfloor$, and the remaining space vacant. As it turns out, three qualitatively different outcomes of the basic experiment give rise to distinct phases for the system and partition \mathcal{S}° into regimes called \mathcal{T} , \mathcal{U}^- , and \mathcal{U}^+ , as we will now explain.

Region \mathcal{T} (tight borders; stationary jams)

In the simplest scenario, our basic experiment with density $\frac{1}{2}$ produces a stable border between a single block of maximal free flow and a single block of minimal anti-free flow. Within this regime, cars gradually leave the initial solid block at the front, sending corresponding anticars backward through the jam, but the border between the front of the jam and exiting free flow stays *tight*. Formally, this means that the diameter of the intermediate region containing cars which are neither in free flow nor in antifree flow stays stochastically bounded by an a.s. finite random variable, uniformly in time and the size of the torus N . (Alternatively, one can require a tight border in an infinite experiment with cars initially on the negative half-line and empty space on the positive side.)

Of course, on a finite torus cars in free flow eventually wrap around and meet the back border of the anti-free flow jam. By symmetry, that border must also be tight, so the basic experiment with density $\frac{1}{2}$ is inherently stable, maintaining one block of each phase for eons if N is at all large, until some exceedingly rare event causes a transient instability. In contrast, if the initial state is product measure μ_ρ with density ρ , the system finds free flow if $\rho \leq \rho_*$, and settles into anti-free flow if $\rho \geq 1 - \rho_*$. But for intermediate densities $\rho \in (\rho_*, 1 - \rho_*)$, traffic self-organizes into alternating regions of maximal free flow and minimal anti-free flow. These regions move diffusively and cluster in a manner reminiscent of the basic one-dimensional voter model (cf. [Lig]). Consequently, (1.7) holds with $\rho^* = 1 - \rho_*$. Figure 5 shows SCC clustering with parameters (0.2, 0.6), a representative phase point from region \mathcal{T} . The traffic has formed two regions of heavy congestion (anti-free flow) which will eventually merge into one as their borders fluctuate somewhat like random walks.

Region \mathcal{U}^- (unstable back border, stable front; jams move backward)

For other phase points, such as (0.6, 0.6), the basic experiment with density $\frac{1}{2}$ evolves quite differently. As illustrated by the four-panel time trace of Fig. 6, free flow exits the initial solid block with a tight border, but after cars wrap around the torus their interaction with the back border of

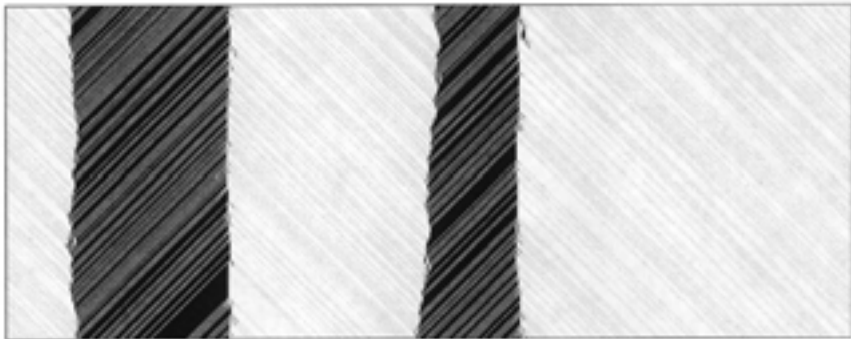


Fig. 5. (0.2, 0.6) Symmetric Cruise Control on 1000 sites, from 19,600 to 19,999 updates, with initial density $\rho = 0.37$.

the jam is *unstable*. A new *synchronized flow* emerges, spreads at linear speed, and eventually pervades the entire lattice. Evidently there is a stable, ergodic, density $\frac{1}{2}$ measure $\nu_{1/2}$ for traffic with these parameters which is neither free flow, nor anti-free flow, but rather an instance of “Phase 2” in the Introduction. Not surprisingly, traffic started from $\mu_{1/2}$ appears to converge to the same distribution $\nu_{1/2}$ whereas the system attains free flow if $\rho \leq \rho_*$.

But what happens for ρ between ρ_* and $\frac{1}{2}$? At densities just below $\frac{1}{2}$, traffic converges to an ergodic equilibrium ν_ρ as one would expect. (The emergence of a new phase, as shown in Fig. 6, continues to take place in the basic experiment, albeit at a slower linear rate.) But for substantially

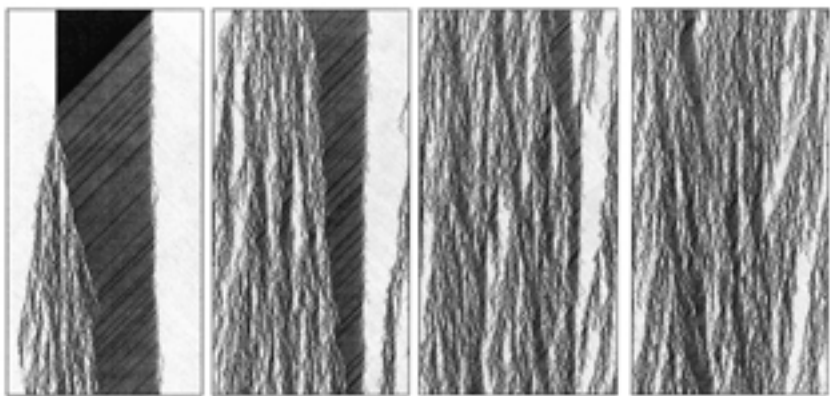


Fig. 6. The density $\frac{1}{2}$ basic experiment for (0.6, 0.6) Symmetric Cruise Control: emergence at the synchronized traffic phase at the back border of a solid jam.

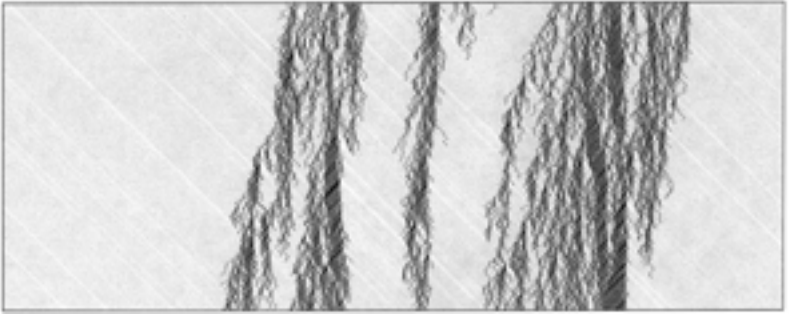


Fig. 7. (0.6, 0.6) Symmetric Cruise Control on 1000 sites, from 19,600 to 19,999 updates, with initial density $\rho = 0.37$.

lower densities a new phenomenon arises, as suggested by Fig. 7. Namely, the traffic pattern appears to converge to a *mixture* of maximal free flow and a minimal ergodic synchronized jam state v_{ρ^*} . Density ρ^* is characterized as the infimum over all ρ for which homogeneous synchronous flow emerges in the basic experiment with density ρ . Within \mathcal{U}^- , it turns out that ρ^* is always strictly greater than ρ_* ; in between, (1.7) holds and traffic clusters into alternating stretches of maximal free flow and self-organized critical synchronized jams which grow without bound on the infinite lattice. The next two sections will present additional evidence for this hypothesis. Since the front of heavy congestion does not advance in this regime, the borders between synchronized flow and maximal free flow retreat, and the jams move backward, as is apparently typical of real-world traffic.

Region \mathcal{U}^+ (unstable front border, stable or unstable back; jams move forward)

Within a third region of \mathcal{S}° our basic experiment with density $\frac{1}{2}$ is unstable at the front border of the initial solid jam. A representative phase point for this case is (0.6, 0.8). The resultant ergodic behavior is analogous to that in \mathcal{U}^- : from μ_ρ , the system converges to an ergodic equilibrium v_ρ if ρ is sufficiently close to $\frac{1}{2}$, to free flow if $\rho \leq \rho_*$, and (1.7) holds for $\rho \in (\rho_*, \rho^*)$, where ρ^* is the infimum of densities for which the synchronized flow emerging from the front border takes over the entire lattice. Figure 8 provides an illustration of convergence to a mixture in this setting. The distinguishing feature is forward motion of the jams.

A subtle issue arises: can it happen that both the front and back borders of the basic experiment are unstable, and if so, how does the traffic self-organize in this case? Indeed, careful simulation reveals a small region of phase emergence at both the front and rear of a congested jam. Thus, there



Fig. 8. (0.6, 0.8) Symmetric Cruise Control on 1000 sites, from 19,600 to 19,999 updates, with initial density $\rho = 0.37$.

can be two distinct ergodic equilibria in a neighborhood of density $\frac{1}{2}$. Further experimentation indicates that the synchronized jam produced at the initial front border always dominates that produced at the back, in that the former displaces the latter at a linear rate once they meet. In other words, v_ρ emerges from the front, and is stable, whereas the measure v'_ρ that emerges from the back is unstable. Consequently, critical jams move forward in this case as well, and it is effectively a portion of \mathcal{U}^+ .

By systematically simulating our basic experiment at a large number of phase points throughout \mathcal{S}° , and also studying the corresponding evolutions from various μ_ρ , we have arrived at the sketch of the SCC phase portrait shown in Fig. 9. Behavior near the phase boundaries is difficult to interpret,

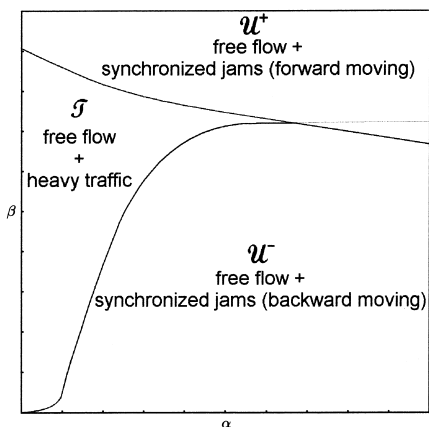


Fig. 9. The Symmetric Cruise Control Phase Portrait.

either visually or numerically, so it is important to stress that this diagram is only approximate. Some qualitative features are worth mentioning, however. First, note that the boundary between \mathcal{T} and \mathcal{U}^+ intersects the left edge near $\beta = 0.9$ rather than in the upper left corner. Thus, congested flow is unstable for models with β sufficiently close to 1 no matter how small the acceleration α . At the right edge, the boundary between \mathcal{U}^+ and \mathcal{U}^- seems to end at exactly $\beta = \frac{2}{3}$, a feature which can be understood in terms of the virtual particles introduced in Section 3 and used in Section 7 below to elucidate SCC dynamics near \mathcal{E}_R . The light gray curve in our sketch marks the upper boundary of a small “triangular” region where both the front and back borders of a congested jam are unstable, so that in addition to v_ρ there is an unstable ergodic equilibrium v'_ρ for traffic densities in a neighborhood of $\frac{1}{2}$. Finally, as β decreases, the boundary between \mathcal{T} and \mathcal{U}^+ at first glance seems headed for \mathcal{E}_B , but upon closer inspection ends in the lower left corner. This is the most speculative portion of the sketch since accurate simulation requires enormous array size and run time when both parameters are small and traffic is increasingly sluggish.

The apparent behavior of ρ^* for parameters approaching the edges of \mathcal{S} or the boundary between tight and unstable borders is summarized as follows:

$$\rho^* \rightarrow \frac{1}{3} \text{ as } (\alpha, \beta) \text{ approaches } \mathcal{E}_B \cup \mathcal{E}_R. \quad (5.3)$$

$$\rho^* \rightarrow \varphi(\alpha) \text{ as } (\alpha, \beta) \rightarrow (\alpha, 1) \text{ from within } \mathcal{S}^\circ, \text{ where } \varphi \text{ is the function} \\ \text{given in (5.2).} \quad (5.4)$$

$$\rho^* \rightarrow \frac{1}{2} \text{ as } (\alpha, \beta) \text{ approaches the boundary between } \mathcal{T} \text{ and } \mathcal{U}^+ \cup \mathcal{U}^-. \quad (5.5)$$

$$\rho^* > \rho_* \text{ throughout } \mathcal{S}^\circ. \quad (5.6)$$

We conclude this section with a few remarks about the shape of SCC fundamental diagrams. Of course $\theta(\rho)$ is linear up to ρ_* by definition (1.5), while (1.7) implies that θ interpolates linearly between (ρ_*, ρ_*) and $(\rho^*, \theta(\rho^*))$. On \mathcal{T} , $\rho^* = 1 - \rho_*$ and $\theta(\rho^*) = \rho_*$ by symmetry, so θ consists of 3 linear pieces with the middle one constant. On \mathcal{U}^- the jams move backward, which implies that $\theta(\rho^*) < \rho_*$. Thus the linear piece on the interval (ρ_*, ρ^*) has negative slope. Presumably the graph of θ is a smooth curve over the interval of densities $(\rho^*, 1 - \rho^*)$ corresponding to ergodic synchronous traffic. Similarly, on \mathcal{U}^+ the linear piece from ρ_* to ρ^* has positive slope since the jams move forward. Figure 10 shows caricatures of fundamental diagrams in each of the three basic regions of Fig. 9. As a final indication of throughput subtleties, we note that along the boundary between \mathcal{U}^- and \mathcal{U}^+ , with α near 1 and $\beta \approx 0.705$, there would appear to be

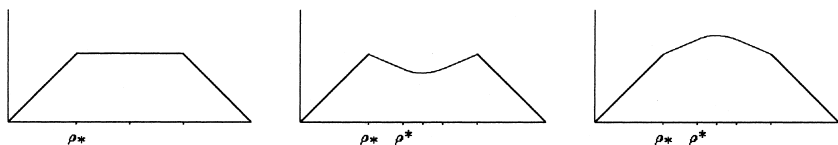


Fig. 10. Qualitative Shape of the Fundamental Diagram on \mathcal{T} , \mathcal{U}^- , and \mathcal{U}^+ (left to right).

cases where $\rho_* < \rho^*$ and $\theta(\rho_*) = \theta(\rho^*) \approx \frac{1}{3}$. For such parameters, it is pretty inconceivable that θ is constant over $[\rho^*, 1 - \rho^*]$. So the fundamental diagrams in a neighborhood of such (α, β) are presumably neither convex nor concave over $[\rho_*, 1 - \rho_*]$.

6. SCC: AN INSTANCE OF SELF-ORGANIZED CLUSTERING

Our description of the Symmetric Cruise Control phase portrait in the previous section relied mainly on computer visualization: of border dynamics, and of clustering from μ_p . To bolster the evidence for convergence to a mixture (1.7) at suitable traffic densities, we have carried out extensive Monte Carlo numerical estimation of the throughput in the representative case of Figs. 6 and 7: $\alpha = \beta = 0.6$. For densities ρ ranging from 0.28 to 0.50 in increments of 0.005, we performed 10 simulations each, up to time 100,000 on an array of 4,000 sites. In each simulation, the time-averaged throughput at 0 was computed from $t=20,000$ to $t=100,000$ in order to diminish the impact of initial transients. Variance reduction was achieved by populating the lattice with exactly $\lfloor 4000\rho \rfloor$ cars, randomly placed. Then $\hat{\theta}(\rho)$ was set as the average of the 10 simulated throughputs. A sample of the data, to four decimal places, is shown in Table II.

The missing entry for $\rho = 0.31$ reflects the presence of observable finite-size effects this close to ρ_* . (By doing analogous simulations on larger arrays we were able to obtain an estimate of θ at that density consistent with the other entries of the table.) Plots of our data for the throughput and speed (as determined by (1.4)) are shown in Fig. 11. We have superimposed a least squares fit of the throughput data from $\rho = 0.32$ to $\rho = 0.40$ on the estimated fundamental diagram. Our results are reasonably consistent with critical values $\rho_* \approx 0.306$ and $\rho^* \approx 0.43$, although this latter value is still quite tentative. In any case, there is rather compelling evidence for the existence of a linear piece in the diagram just above the maximal free flow density, but stopping short of density $\frac{1}{2}$. This scenario is best explained by asymptotic behavior (1.7), in which case the throughput within this regime is a mixture of the maximal free flow throughput and the minimal synchronized traffic throughput, as postulated in Section 5.

Table II. Monte Carlo Estimation of Throughput θ vs. Density ρ for (0.6, 0.6) Symmetric Cruise Control

| ρ | $\hat{\theta}$ |
|--------|----------------|
| 0.30 | 0.3000 |
| 0.31 | — |
| 0.32 | 0.3031 |
| 0.33 | 0.3016 |
| 0.34 | 0.3001 |
| 0.35 | 0.2987 |
| 0.36 | 0.2973 |
| 0.37 | 0.2962 |
| 0.38 | 0.2950 |
| 0.39 | 0.2940 |
| 0.40 | 0.2926 |
| 0.41 | 0.2910 |
| 0.42 | 0.2907 |
| 0.43 | 0.2893 |
| 0.44 | 0.2883 |
| 0.45 | 0.2876 |
| 0.46 | 0.2867 |
| 0.47 | 0.2859 |
| 0.48 | 0.2854 |
| 0.49 | 0.2849 |
| 0.50 | 0.2849 |

7. SCC: VIRTUAL PARTICLES NEAR THE EDGES OF THE PHASE DIAGRAM

The best prospects for rigorous mathematical confirmation of convergence to a mixture (1.7) and clustering in slow-to-start Traffic CA dynamics would seem to reside close to the top, bottom, and right edges of the phase portrait, i.e., within \mathcal{S}° but near the boundaries \mathcal{E}_T , \mathcal{E}_B , or \mathcal{E}_R , where the behavior is known precisely from Theorem 1. In keeping with many familiar models from statistical mechanics and cellular automata, these perturbations of exactly solvable dynamics admit simplified descriptions in terms of virtual particles which track the evolution of “dislocations” from the unperturbed system. Our objective in this section is to sketch three types of virtual dynamics which arise, in reverse order: right, then bottom, then top, and then to present a fourth, unified representation which is exact throughout \mathcal{S} . Systematic analysis of this approach to Cruise Control will be relegated to a subsequent paper.

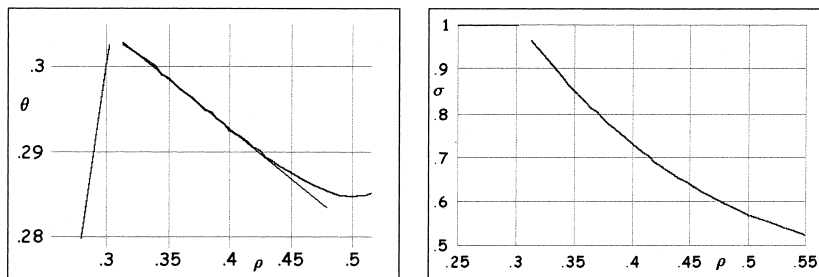


Fig. 11. Empirical estimation of the throughput θ (left) and velocity σ (right) for $(0.6, 0.6)$ Symmetric Cruise Control as a function of the density ρ .

(a) *Ballistic-Diffusive Representation Near \mathcal{E}_R*

As explained at the end of Section 4, on \mathcal{E}_R isolated virtual particles of the form 101 perform mean $3\beta - 2$ random walks on a background ether of density $\frac{1}{3}$ ideal free flow. Near \mathcal{E}_R , these walks are augmented by ballistic virtual antiparticles—10001 configurations which move to the right deterministically at speed 1. The virtual antiparticles are created when random walks collide to form a $\cdots 01011010 \cdots$ local configuration and then the second car fails to advance for two successive time steps (an event with probability $(1 - \alpha)^2$). Once that car manages to accelerate, the net result is a ballistic virtual antiparticle/particle pair adjacent to the colliding pair. Each virtual antiparticle continues to travel at maximal velocity until it encounters a dislocation, at which time the two annihilate. The points of contact in Fig. 12 between diagonal white lines (00) and dark gray random walks (101) signal such annihilations.

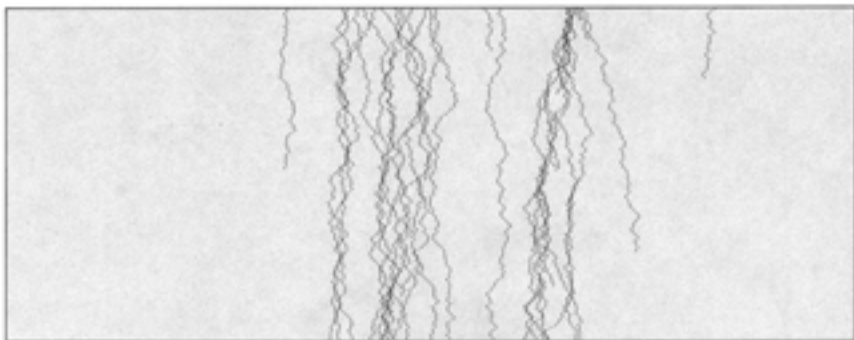


Fig. 12. $(0.95, 0.67)$ Symmetric Cruise Control on 1000 sites, from 19,600 to 19,999 updates, with density $\rho = 0.34$.

Since ballistic motion is effectively instantaneous on the diffusive scale of random motion, a simpler prototype for these dynamics consists of randomly walking particles such that, whenever a pair collides, there is simultaneous birth of a new particle at the point of collision, and annihilation of the nearest particle to the right. It can be shown that such an interaction among any finite number N of particles remains tight on a large torus, i.e., the diameter of the particle cluster remains stochastically bounded above, uniformly in time and in the size of the torus, by an a.s. finite random variable D_N . Within the cluster an equilibrium distribution of particles is attained as $N \rightarrow \infty$. This mechanism provides an illuminating cartoon for the self-organized critical traffic jams of (1.7).

(b) Virtual Particle/Antiparticle Representation Near \mathcal{E}_B

Along the bottom edge of the SCC phase diagram a related virtual particle scheme arises. There are two ethers: 100100... (ideal free flow), and 10101010... (ideal traffic jam). Virtual antiparticles travel deterministically through the free flow ether with speed 1, while virtual particles do a random walk in the free flow ether with speed $3\beta - 2$, just as they do on the right edge. But when β is small, the virtual particles move essentially deterministically.

It is possible to show that the traffic jams move with average speed $-2 + C\sqrt{\beta}$ in this region, so the jams are slightly faster than isolated virtual particles. This means that the front of a jam is tight, so that when $\alpha < 1$, virtual antiparticles are emitted at a positive linear rate. These virtual antiparticles clean up the backs of subsequent traffic jams, keeping them tight at the rear border.

Within the jam ether, additional virtual particle/antiparticle pairs are created everywhere at rate β . These virtual particles move deterministically to the right at speed 2 while the virtual antiparticles move almost deterministically to the left at about speed 2. When virtual particles of opposite type collide, they annihilate after a short delay which depends on α . Intra-jam dynamics and a few escaping virtual antiparticles are shown in Fig. 13.

Since traffic jams move parallel to virtual antiparticles, many jam virtual particles escape from the front and are incorporated in the free flow ether, or occasionally create free flow virtual antiparticles, but jam virtual antiparticles rarely escape. So the jams have an excess of virtual antiparticles, each incorporating a little empty space. This gives the jams a density slightly less than $\frac{1}{2}$, and helps explain why there can be ergodic equilibria with $\rho < \frac{1}{2}$. When ρ is just below $\frac{1}{2}$ there are more jam virtual antiparticles than jam virtual particles, and the extra virtual antiparticles "absorb" stripes of free flow. Similarly, when ρ is just above $\frac{1}{2}$ there are more jam

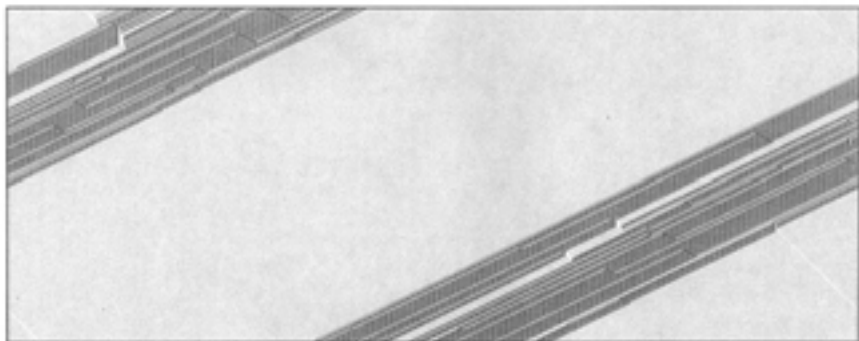


Fig. 13. $(0.2, 0.001)$ Symmetric Cruise Control on 1000 sites, from 19,600 to 19,999 updates, with density $\rho = 0.37$.

virtual particles than jam virtual antiparticles, with the virtual particles absorbing stripes of anti-free flow.

At first glance, all of this seems rather more complicated than our previous description along the right edge. However, the local interactions of virtual particles and antiparticles when β is small are actually simpler, so we suspect that a neighborhood of $(1, 0)$, with β much smaller than $1 - \alpha$, may offer the best hope for a rigorous proof of (1.7). As suggested by our simulations and Fig. 9, once α gets small enough the jam virtual particles have difficulty escaping from the front, start to back up, and create solid blocks of cars at the front of the jams. There is a battle at the border between the block in the front of the jam and the $1010\cdots$ ether at the rear, with the block winning out, so the size of the $1010\cdots$ ether stays the same on average, no matter how big the jam gets. Thus the density of a large jam is close to $1 - \rho_*$, and it is no longer possible to have an ergodic equilibrium for any density.

(c) Multiparticle Representation Near $\alpha=0$, $\beta=1$

Next, we mention an intriguing asymptotic regime in the upper left corner of \mathcal{S} . Here the acceleration is low, so cars at the front of a jam pile up in a solid block, and then leave at only rate α . But since there is almost no braking (β near 1), a $\cdots 10101$ trail follows each breakaway car until a " β -coin" finally fails to advance the next car, after a geometric number of successes with success probability β . This effect creates a space-time cone of cars and empty spaces, spreading to both the left and right at speed 1. But the ideal jam ether, a checkerboard in space-time, is vulnerable from within whenever a car fails to advance with (small) probability $1 - \beta$, in which case a 1100 local configuration initiates another backlog of slow-to-start

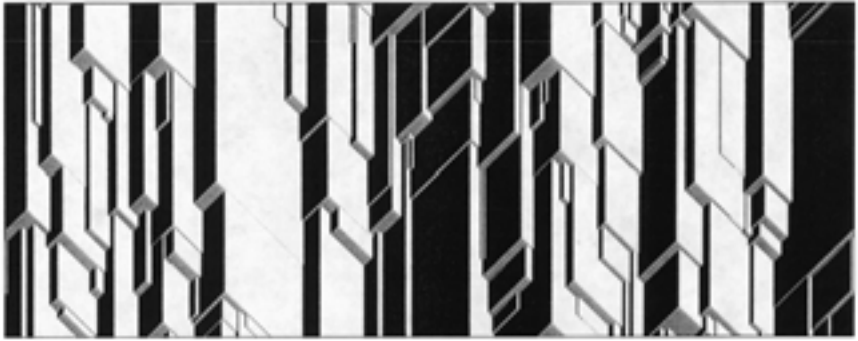


Fig. 14. $(0.01, 0.99)$ Symmetric Cruise Control on 1000 sites, from 19,600 to 19,999 updates, with density $\rho = 0.34$.

cars. Since ideal jam cones are created along each vertical slice between solid and empty regions but disrupted anywhere within the ensuing two-dimensional cone, a stable limiting random field is attained as $\alpha \rightarrow 0$, $\beta \rightarrow 1$, with $\alpha \sim C \sqrt{1 - \beta}$. The appropriate scaling, which creates a curious three-dimensional effect, is shown in Fig. 14. We conjecture that asymptotics for the traffic equilibria ν_ρ under this scaling can be described by means of a planar system of static and ballistic particles and antiparticles, with suitable interactions, exponential holding times, and Poisson initial condition.

(d) Gap CA—A Unified Particle–Antiparticle Representation for TCA

Associate each gap between successive cars in the TCA with a site in a new model, which we call the *Gap CA*, or GCA. If a given gap in the original model contains $n \geq 0$ empty spaces, then put $2 - n$ particles at the corresponding site in the GCA. Positive particle numbers are called *particles* in the GCA, negative particle numbers are called *antiparticles*. Thus, at a given site in the new model, there can be either one or two particles, or there can be arbitrarily many antiparticles, or the site can be vacant. The dynamics for the GCA model are now easy to sort out. Note that in this version, the negative particles don't move, while positive particles either stay still or move left. To recover the setting of Figs. 12 and 13, in which antiparticles move to the right at speed 1, simply make a space-time change of variables. For simplicity, we restrict our attention to the Cruise Control case ($\delta = 1$). It is straightforward to modify this model to accommodate the more general model.

The rules for GCA (with any parameters α , β , γ , and with $\delta = 1$), applied independently and simultaneously at the various sites during each time step, are as follows:

(i) If there are two particles at a site x , then one of the two particles moves to $x-1$ and the other stays at x . In addition, if $x+1$ is vacant or contains antiparticles, then with probability $1-\alpha$, a new particle is created at x and an antiparticle is created at $x+1$. (In the latter event, x will still contain 2 particles.)

(ii) If there is one particle at a site x , and if $x-1$ does not contain two particles, then the particle at x moves to $x-1$ with probability $1-\beta$ and stays at x with probability β .

(iii) If there is one particle at a site x , and if $x-1$ does contain two particles, then the particle at x moves to $x-1$ with probability $1-\gamma$ and stays at x with probability γ .

(iv) Antiparticles do not move. However, if one or more particles and one or more antiparticles exist at the same site, then they annihilate in pairs. (This situation arises from particle movement or from the creation of particle/antiparticle pairs.)

In (iv), it is easy to see that at most one annihilation can occur at a site during any one time unit. If $\beta = \gamma$, then (ii) and (iii) can be combined into a single rule.

Here are some remarks about the GCA, which we expect will prove helpful for future analysis of TCA:

When $\rho = \frac{1}{3}$, the average particle number per site is 0 (counting antiparticles as negative), and when $\rho = \frac{1}{2}$, the average particle number per site is 1. (7.1)

The system goes into free flow if and only if all particles eventually die. (7.2)

Clustering means that the system produces larger and larger intervals which contain no particles, without going into free flow. (7.3)

The simplest statement of our main goal for future research is: Find $(\alpha, \beta) \in \mathcal{S}^\circ$ such that clustering can be proved at some $\rho < \frac{1}{2}$, and such that it can be shown that there is no clustering at $\rho = \frac{1}{2}$. When ρ is small, or α is close to 1, particle/antiparticle creation events are rare, so the particles and antiparticles annihilate faster than they are created, at least until either particles or antiparticles become scarce enough. If the antiparticles outnumber the particles by a significant amount, such as when ρ is significantly less than $\frac{1}{3}$, this leads to free flow. But when $\rho > \frac{1}{3}$ or α is not close to 1, then there is a minimum rate at which creation events occur. Such events

favor clustering, because particles that try to enter a large region with no particles typically do not get very far before they get annihilated by anti-particles.

For ρ close to $\frac{1}{2}$, however, there is a counterbalancing effect which is best seen when α is close to 1 and β is close to 0. In this case, the average particle number is 1 per site, particle/antiparticle creation events are rare, and single particles move most of the time, but when there are double particles, only one of them moves, and a particle to the right of a double particle also doesn't move. The result is that particles spread out, with very few double particles. Moreover, the number of sites with double particles equals (the number of sites without particles) + (the number of antiparticles). So vacant sites and antiparticles are necessarily rare. And whenever intervals without particles do appear, they contain relatively few antiparticles, so they are easily invaded as a result of the activity of sites with double particles. Thus, there is no clustering. As suggested earlier, the best place to look for a rigorous instance of (1.7) would seem to be near the bottom right corner of phase space.

8. EVIDENCE FOR CLUSTERING IN AN IRREDUCIBLE TCA

One of our original motivations for the study of traffic prototypes concerned the possibility of clustering and convergence to a nontrivial mixture (1.7) in an interacting system with particle conservation, but with no other deterministic constraints. For our purposes, say that ξ_t^ρ is *irreducible* if each car stays put with positive probability, and advances with positive probability unless blocked by a car directly ahead of it. Of course this amounts to the assumption that $\alpha, \beta, \gamma,$ and δ are all strictly between 0 and 1, and implies that any invariant ν_ρ for ξ_t^ρ gives strictly positive probability to every possible configuration on each finite set $A \subset \mathbb{Z}$. (By contrast, Cruise Control with $\rho = \frac{1}{3}$ has an equilibrium which attaches no weight to an interval of 3 successive 0s.) Is there an irreducible TCA which separates into two distinct phases starting from a suitable μ_ρ ? Pictures similar to Fig. 1 from the experiments of Nagel *et al.* suggested the possibility of this exotic behavior among slow-to-start traffic models, but do such simulations merely indicate equilibria with long length scales, close in parameter space to clustering Cruise Control dynamics, yet stable nevertheless?

Based on an extensive, albeit preliminary empirical study, we have found compelling evidence for the possibility of phase separation (1.7) even in some irreducible cases. From a phenomenological point of view, this is probably the most intriguing discovery of the present study. The clustering mechanism for these examples is very poorly understood, so our purpose



Fig. 15. The (0.11, 0.05, 0.9, 0.9) Traffic Cellular Automaton on 1,000 sites, from 79,600 to 79,999 updates, with density $\rho = 0.3$.

here is merely to identify a promising parameter set and describe very briefly the observed behavior. Since we know of no other particle conserving irreducible lattice interaction which clusters, and since various of Nagel's simulations and the recent study [RSS] warn of subtle finite-size effects, a much more thorough analysis is planned.

The clearest indication of irreducible TCA clustering we have found occurs in a small neighborhood of the phase point (0.11, 0.05, 0.9, 0.9), and with traffic densities around $\rho = 0.3$. Figure 15 shows one simulation with those parameter values on an array of 1,000 sites. Note that traffic is about a 50–50 mix of “quasi free flow” and an exotic synchronized jam state. Occasional mini-jams can be seen attempting to nucleate within the low-density phase, a consequence of irreducibility, but these are evidently unstable. Also, the high-density phase is characterized by thick “quasi anti-free flow” regions which seem to somehow play a crucial role in the clustering process. We have run a great many simulations with these and nearby parameters, on array sizes up to $L=20,000$, for hundreds of thousands or even millions of updates. In all cases, the traffic has slowly separated in the manner of Fig. 14, and then retained clear and stable evidence of the distinct phases thereafter. Thus, it appears quite possible that even irreducible slow-to-start dynamics can give rise to traffic jams of arbitrarily large size.

9. ELECTRONIC RESOURCES

The analysis of Traffic CA system behavior described in this paper would have been impossible without an efficient and flexible simulation engine. Our first experiments, vital for “proof of concept,” were carried out

using an in-house, dedicated program written by Janko Gravner. Once a more systematic study was indicated, we were fortunate to make the (virtual) acquaintance of Mirek Wójtowicz, whose state-of-the-art general cellular automaton program MCell [Wój] supports hundreds of fascinating one-dimensional and two-dimensional CA rules. Over the past year we have worked with Mirek on a daily basis to develop a tool set suitable for the visual and numerical study of TCA, but also applicable to his entire library of interacting systems. The latest version of MCell includes a ready-to-run Traffic CA experiment.

We urge our readers to explore the resources and links of a companion Web page:

<http://psoup.math.wisc.edu/traffic/>

which provides additional graphics, instructions for downloading MCell, and a library of Mcell experiments illustrating many of the features of TCA dynamics discussed here. Indeed, a far better understanding of the subtleties of emergent traffic jams will be achieved by parsing those computer-based materials in parallel with this account.

ACKNOWLEDGMENTS

We would like to thank Kai Nagel, Mirek Wójtowicz, Janko Gravner, and Ann Scheels for their help with this research.

REFERENCES

- [AS] T. Antal and G. Schuetz, Asymmetric exclusion process with next-nearest-neighbor interaction: Some comments on traffic flow and a nonequilibrium reentrance transition, *Phys. Rev. E* **62**:83–93 (2000).
- [BF] V. Belitsky and P. Ferrari, *Invariant Measures and Convergence for Cellular Automaton 184 and Related Processes*. Preprint (1999).
- [BML] O. Biham, A. Middleton, and D. Levine, Self-organization and a dynamical transition in traffic-flow models, *Phys. Rev. A*, 6124–6127 (1992).
- [Bla] L. Blake, *Experts Around the Nation Interested in Meter Study Results* (Star Tribune, Minneapolis–St. Paul, November 9, 2000).
- [BKNS] V. Belitsky, J. Krug, E. Neves, and G. Schutz, *A Cellular Automaton for Two-Lane Traffic*. Preprint (2000).
- [CSS] D. Chowdhury, L. Santen, and A. Schadschneider, Statistical physics of vehicular traffic and some related systems, *Physics Reports (in Physics Letters)* **329**:199–329 (2000).
- [DN] R. Durrett and C. Neuhauser, Coexistence results for some competition models, *Ann. Appl. Probab.* **7**(1):10–45 (1997).
- [EG] M. Ekhaus and L. Gray, *Convergence to Equilibrium and a Strong Law for the Motion of Restricted Interfaces*. In preparation.

- [Fi] R. Fisch, Clustering in the one-dimensional three-color cyclic cellular automaton, *Ann. Probab.* **20**(3):1528–1548 (1992).
- [GK] M. Gerwinski and J. Krug, Analytic approach to the critical density in cellular automata for traffic flow, *Phys. Rev. E* **60**:188–196 (1999).
- [Gri] D. Griffeath, Primordial Soup Kitchen, <http://psoup.math.wisc.edu/kitchen.html>.
- [KLS] S. Katz, J. Lebowitz, and H. Spohn, Stationary nonequilibrium states for stochastic lattice gas models of ionic superconductors, *J. Statist. Phys.* **34**:497–537 (1984).
- [KNW] S. Krauss, K. Nagel, and P. Wagner, *The Mechanism of Flow Breakdown in Traffic Flow Models* (Los Alamos National Laboratory). Preprint LA-UR 98-3161 (1998).
- [Kra] S. Krauss, *Microscopic Modeling of Traffic Flow: Investigation of Collision Free Vehicle Dynamics*, Ph.D. thesis (University of Cologne, Cologne, Germany, 1997).
- [KS] J. Krug and H. Spohn, Universality classes for deterministic surface growth, *Phys. Rev. A* **43**:4271–4283 (1988).
- [LeV] R. LeVecque, *Numerical Methods for Conservation Laws*, Lectures in Mathematics ETH Zurich (Birkhauser, Basel, 1990).
- [Lig] T. Liggett, *Interacting Particle Systems* (Springer-Verlag, 1985).
- [MDDGL] J. Molofsky, R. Durrett, J. Dushoff, D. Griffeath, and S. Levin, Local frequency dependence and global coexistence, *Theoretical Population Biology* **55**:270–282 (1999).
- [Nag1] K. Nagel, Particle hopping models and traffic flow theory, *Phys. Rev. E* **53**:4655–4672 (1996).
- [Nag2] K. Nagel, Experiences with iterated traffic microsimulations in Dallas, in *Traffic and Granular Flow II*, Wolf and Schreckenberg, eds. In press.
- [NWWS] K. Nagel, D. Wolf, P. Wagner, and P. Simon, Two-lane traffic rules for cellular automata: A systematic approach, *Phys. Rev. E* **58**:1425–1437 (1995).
- [NP] K. Nagel and M. Paczuski, Emergent traffic jams, *Phys. Rev. E* **51**:2909–2918 (1995).
- [NS] K. Nagel and M. Schreckenberg, A cellular automaton model for freeway traffic, *J. Physique I France* **2**:2221 (1992).
- [RT] D. Redelmeier and R. Tibshirani, Why cars in the next lane seem to go faster, *Nature* (September 2, 1999).
- [Rez] F. Rezakhanlou, Hydrodynamic limit for attractive particle systems on \mathbb{Z}^d , *Comm. Math. Phys.* **140**:417–448 (1991).
- [Ros] H. Rost, Non-equilibrium behaviour of a many particle process: Density profile and local equilibrium, *Z. Wahrsch. Verw. Gebiete* **58**:41–53 (1981).
- [RSS] N. Rajewsky, T. Sasamoto, and E. R. Speer, Spatial particle condensation for an exclusion process on a ring, *Physica A* **279**:123–142 (2000).
- [SSNI] M. Schreckenberg, A. Schadschneider, K. Nagel, and N. Ito, Discrete stochastic models for traffic flow, *Phys. Rev. E* **51**:2939–2949 (1995).
- [Sep] T. Seppäläinen, Existence of hydrodynamics for the totally asymmetric simple K -exclusion process, *Ann. Probab.* **27**:361–415 (1999).
- [Sip] A. Sipress, *Studying the Ebb and Flow of Stop-and-Go. Los Alamos Lab Using Cold War Tools to Scrutinize Traffic Patterns* (Washington Post, August 5, 1999), p. A01.
- [Wol] S. Wolfram, *Cellular Automata and Complexity* (Addison–Wesley, Menlo Park, California, 1994).

- [WS] D. Wolf and M. Schreckenberg, eds., *Traffic and Granular Flow II* (Springer, Heidelberg, 1998).
- [Wój] M. Wójtowicz, *Mirek's Celebration: A 1D and 2D Cellular Automata Explorer*. Windows software, available for free download from <http://psoup.math.wisc.edu/mcell/>.
- [Yag] H. Yaguchi, Stationary measures for an exclusion process on one-dimensional lattices with infinitely many hopping sites, *Hiroshima Math. J.* **16**:449–475 (1986).