CHAPTER 21

The Field Theory

In picoscopic modeling, drivers are modeled as an intelligent agent who is able to gather information from his or her driving environment and make a decision to achieve his or her goals—for example, traveling to the destination on the preferred route at the desired speed while avoiding hazards. The outputs of the driver model are driving decisions, including steering, accelerating, and braking, which, in turn, can be represented by acceleration in the longitudinal and lateral directions. To serve this purpose, this chapter introduces a generic modeling approach, called the field theory of traffic flow, that represents everything in the environment as a field perceived by the subject driver whose mission is to achieve his or her goals by navigating through the overall field.[1]

21.1 MOTIVATION

Research on highway traffic flow over the past half century has resulted in many follow-the-leader theories, each of which was proposed with its own motivation. For examples, in the General Motors family of models [55, 56], a driver's response (e.g., desired acceleration or deceleration) was the result of stimuli (e.g., spacing and relative speed) from his or her leader; the Pipes model [52], the Forbes model [53, 54, 66], and the Gipps model [57] were inspired by safe driving rules; in psychophysical models [64, 111], driver reactions were triggered by perception thresholds; rule-based models [67, 112] were motivated by the fuzzy logic in driver decision making. Though the motivation behind some other car-following models such as the Newell nonlinear model [58] and equilibrium traffic flow models such as those in Refs. [9–12] might not be clear, they were so formulated because of their reasonable performance. Two questions naturally arise. First, would it be possible to have a unifying framework that coherently interprets and relates these models? Second, would it be possible to root such a unifying framework in first principles so that traffic flow theory

[1] This chapter is reproduced from [110].

http://dx.doi.org/10.1016/B978-0-12-804134-5.00021-0
Copyright © 2016 Elsevier Inc.
All rights reserved.

is furnished with a solid foundation and connected to other branches of sciences and engineering?

This chapter and Chapters 23 and 24 are motivated by the above questions. This chapter attempts to address the second question; Chapter 23 is intended for the first question; Chapter 24 presents a multiscale modeling perspective. In this chapter, our attention shall be devoted to the modeling of driver operational control in a transportation system—that is, the motion and interaction of driver-vehicle units on a long homogeneous highway. From first principles (e.g., physical laws and social rules), a phenomenology is postulated which represents the driving environment perceived by a subject driver as an overall field. In this field, objects (e.g., roadways and vehicles) in the environment are each represented as a component field, and their superposition represents the overall hazard that the subject driver tries to avoid. Hence, the modeling of vehicle motion is simply seeks the least hazardous route by navigating through the overall field along its valley.

21.2 PHYSICAL BASIS OF TRAFFIC FLOW

Three systems are of particular interest: a physical system, a transportation system, and a social system, as illustrated in Figure 21.1. The physical system typically consists of nonliving objects whose motion and interaction are subject to physical laws such as Newton's laws of motion. In contrast, the social system involves living entities such as humans whose behaviors differ widely among the population but generally follow some loosely defined rules (e.g., seeking gains and avoiding losses). As such, physical science is recognized as "hard" since it is more objective, rigorous, and accurate, while social science is perceived as "soft" because of its subjectivity, vagueness, and inexactness. Straddling the above two systems is the transportation system, which involves both living entities (human drivers) and nonliving objects (roadways and vehicles). Hence, transportation science can be perceived as "firm" (for the lack of a proper word between "hard" and "soft") since it deals with both physical laws and social rules. Actually, it is close to the "soft" end when strategic planning is concerned, while it migrates toward the "hard" end if tactical decisions and particularly operational control are of interest.

Many traffic flow phenomena are similar to those in the physical system, yet the transportation system has something special to distinguish itself. Some examples of such similarities are given below.

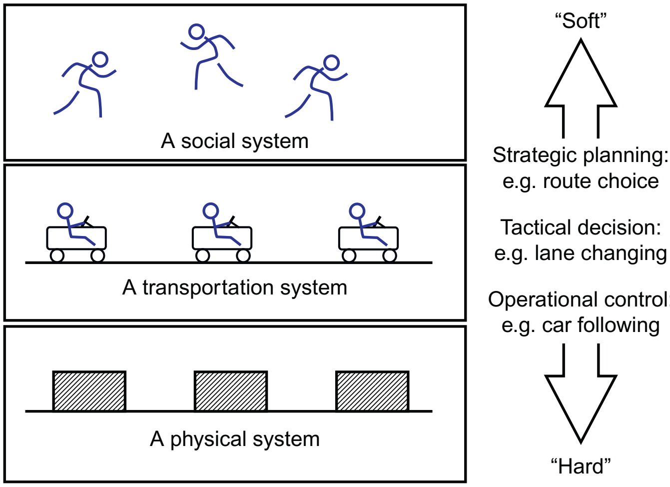

Figure 21.1 Three systems.

21.2.1 Mechanics Phenomena

In physics, forces are the cause of a change of motion. In addition, they are measurable and their effects are reproducible. For example, Newton's second law of motion stipulates that the velocity of an object changes if it is subject to a nonzero external force; Newton's third law says that for every action there is an equal and opposite reaction. Similarly, "forces" exist in traffic flow, but such forces are subjective matters. Consequently, they are nonmeasurable, and their effects do not repeat precisely. For example, a fast driver feels a "force" (a stress in the driver's mind) when he or she approaches a slow vehicle, and hence needs to slow down or change lane. In return, the slow driver may or may not be subject to the reaction "force" depending on whether or not the driver pays attention and responds to the force. If the driver does so, he or she speeds up or gives way in response. Otherwise, Newton's third law does not take effect in this case. More examples of mechanics phenomena are provided below:

M1: Directional flow

Traffic always flows in a predetermined direction much like free objects always fall to the ground. Free objects fall because they are constantly subject to Earth's gravity. Similarly, it is reasonable to imagine that vehicles in the traffic are subject to a "gravity" along the roadway. Such a roadway gravity is, again, a subjective matter since it exists in the mind of drivers and is not measurable, but it is recognized that the gravity is related to factors such as driver personalities (e.g., aggressiveness), vehicle properties (e.g., engine power), and road conditions (e.g., freeways versus streets).

M2: Free flow

An object in free fall will accelerate to an equilibrium speed because of air resistance, and so does a vehicle in free flow. In this case, the "resistance" comes from the driver's willingness to comply with traffic rules (e.g., speed limits) as opposed to rolling, grade, and air resistances. Unlike the free-fall speed, which is deterministic and replicable given the same condition, the free speed of a vehicle is, once again, a subjective matter because it is largely the driver's choice. Given the same conditions, the choice may differ for different drivers and, for the same driver, at different times. In addition, different roadways support different free speeds. To avoid confusion, the free speed chosen by a driver is termed his or her "desired speed," whereas the free speed aggregated over a group of vehicles on a particular road is called the "free-flow speed" supported by the road. Generally, the desired speed is related to driver personalities and road conditions, while the free-flow speed is affected by road conditions and the driver population.

M3: Stopping at a red light

Much like a moving object being slowed to a stop behind a wall, a vehicle decelerates to a stop in front of a red light. The analogous "repelling force" in the latter case resides in the driver in that if he or she ignores the red light, the consequence is costly (e.g., an accident or a ticket). Unlike the moving object, which always stops in the same fashion in repeated experiments, drivers are entitled to decelerate at a comfortable rate to a stop and, in some extreme cases, drivers may forget to stop.

M4: Road barriers

Vehicles moving in the same direction on a roadway are separated by lane lines. To avoid colliding with vehicles in adjacent lanes, a driver must keep in his or her lane as if he or she were guided by barriers at both edges of the lane. If, however, the driver unconsciously departs from the current lane, he or she will perceive some stress, which motivates him or her to steer back into the lane as if a correction "force" from the barrier acts on the vehicle and pushes it toward the center of the lane. If the driver is blocked by a slow vehicle, the desire for mobility will motivate the driver to change lane as if he or she were energized or elevated above the barrier so he or she can cross it and land on the adjacent lane. Running off the road is discouraged,

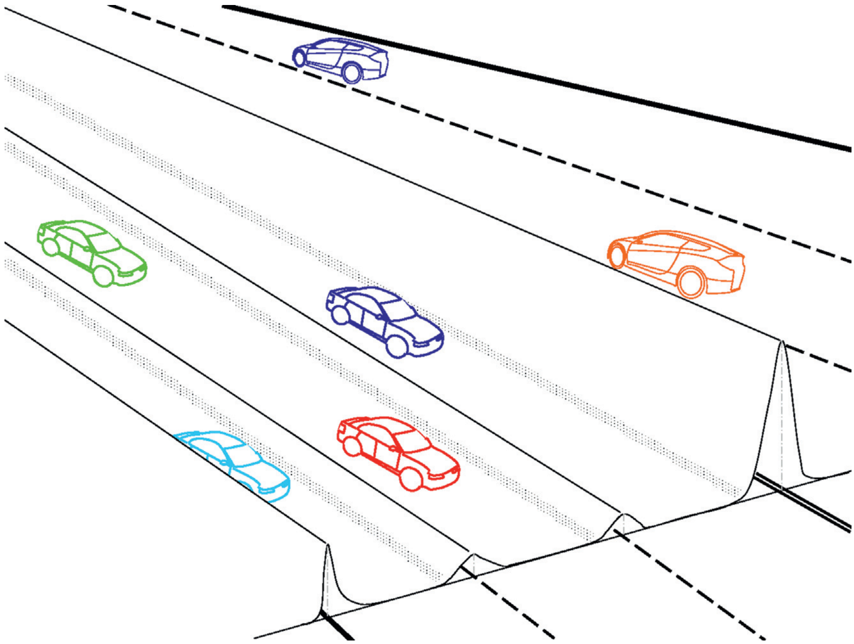

Figure 21.2 Road barriers.

so barriers at road edges are typically higher than lane barriers. Encroaching into the opposite direction of travel is so dangerous that the barrier at the center line is very high (see Figure 21.2 for an illustration of the barriers). These barriers are not real objects, but are only imaginary in drivers' minds.

21.2.2 Electromagnetic Phenomena

An object can exert a force on another object in either of the following ways: collision and action at a distance. For example, hitting a ball with a bat is an example of the former and finding a needle with use of a magnet is an example of the latter. Though collisions are not uncommon on highways, action at a distance is how vehicles normally interact with each other, and examples of this kind include some of the above-mentioned mechanics phenomena as well as the following:

E1: Car following

When a fast vehicle catches up with a slow vehicle, the fast driver perceives an imminent collision if he or she keeps driving at the high speed. The cost and fear of the collision motivates the fast driver to take actions in advance.

If a lane change is not an option and the slow driver does not speed up, the fast driver has to decelerate when he or she approaches the slow vehicle, and then adopts the slow vehicle's speed separated by a safe following distance. This is analogous to moving a charge A toward a like charge B. According to Coulomb's law, the electric force between them is directly proportional to the product of their charges and inversely proportional to the square of their distance. Similarly in car following, the "force" (stress) acting on the fast driver is larger if he or she runs into the slow vehicle faster and their separation is shorter. However, the same opposite force may or may not act on the slow driver as he or she may or may not notice the vehicle approaching from behind.

E2: Tailgating

We continuing with the above example and assume that the fast vehicle tailgates (i.e., follows at a dangerously short distance). Then, it is likely that the opposite force is perceived by the slow driver, who may respond by speeding up or giving way to the fast follower. We return to the analogy, but charge B is now driven (or driven away) by charge A and Newton's third law holds in this case. In general, a "force" must be perceived by a driver before the force has an effect on the person. In addition, a driver's ability to perceive something depends on where he or she scans and how frequently this happens.

E3: Shying away

If two vehicles happen to run in parallel, one or both drivers may feel intimidated. The fear of a side collision motivates them to spread out in space (longitudinally or laterally). Such a shying-away effect becomes more evident when one of the vehicles is a heavy truck.

21.2.3 Wave Phenomena

W1: Harmonic wave

A platoon of vehicles on a roadway is like a harmonic wave. The platoon is characterized by flow (in vehicles per hour), traffic speed (in kilometers per hour), and density (in vehicles per kilometer), while the wave is determined by frequency (in hertz, or cycles per second), wave speed (in meters per second), and wave length (in meters). One immediately recognizes that flow is equivalent to frequency, traffic speed is equivalent to wave speed, and the spacing (the inverse of density) is equivalent to wave length. The upper part of Figure 21.3 shows a platoon of vehicles as a harmonic wave.

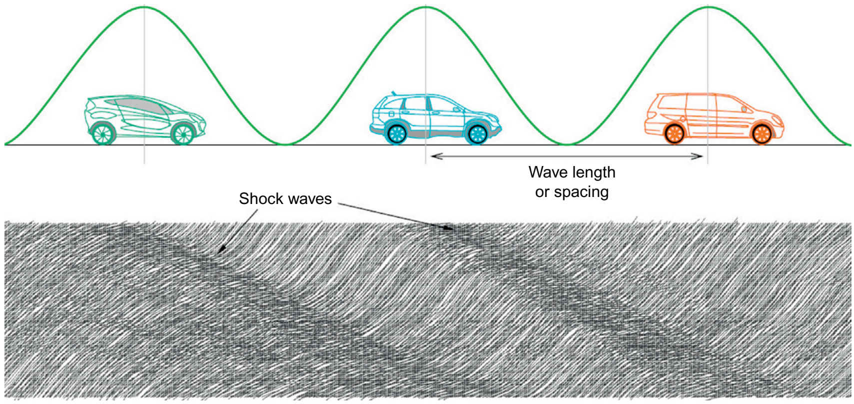

Figure 21.3 Traffic and waves.

W2: Signal propagation

The signal here does not mean a traffic signal, rather it refers to any quantity that clearly defines the location and speed of a perturbation in a medium. When the leading vehicle of a compact platoon brakes briefly, a kinematic wave forms and propagates against the platoon, where the signal here is the brief speed reduction. When a platoon of fast vehicles catches up with a platoon of slow vehicles, a shock wave is generated and propagates against the traffic, where the signal here is the interface between fast and slow vehicles. The bottom part of Figure 21.3 illustrates a few shock waves observed in vehicle trajectories.

W3: Wave-particle duality

All matter, particularly small-scale objects, exhibits both wavelike and particle-like properties. The latter is prominent when individual objects are concerned (e.g., the photoelectric effect), while the former becomes significant when the behavior of many objects is viewed collectively (e.g., diffraction of waves). In traffic flow, individual vehicles act like particles (e.g., car following and lane changing), while a platoon or platoons of vehicles act like waves (e.g., kinematic waves and shock waves).

21.2.4 Statistical Mechanics Phenomena

Traffic flow has been modeled by many authors as a one-dimensional compressible fluid, such as a gas. In gases, the speeds of gas molecules follow a Maxwell-Boltzmann distribution (see Figure 21.4), top left, for an illustration. Remarkable in the distribution is the increase in average

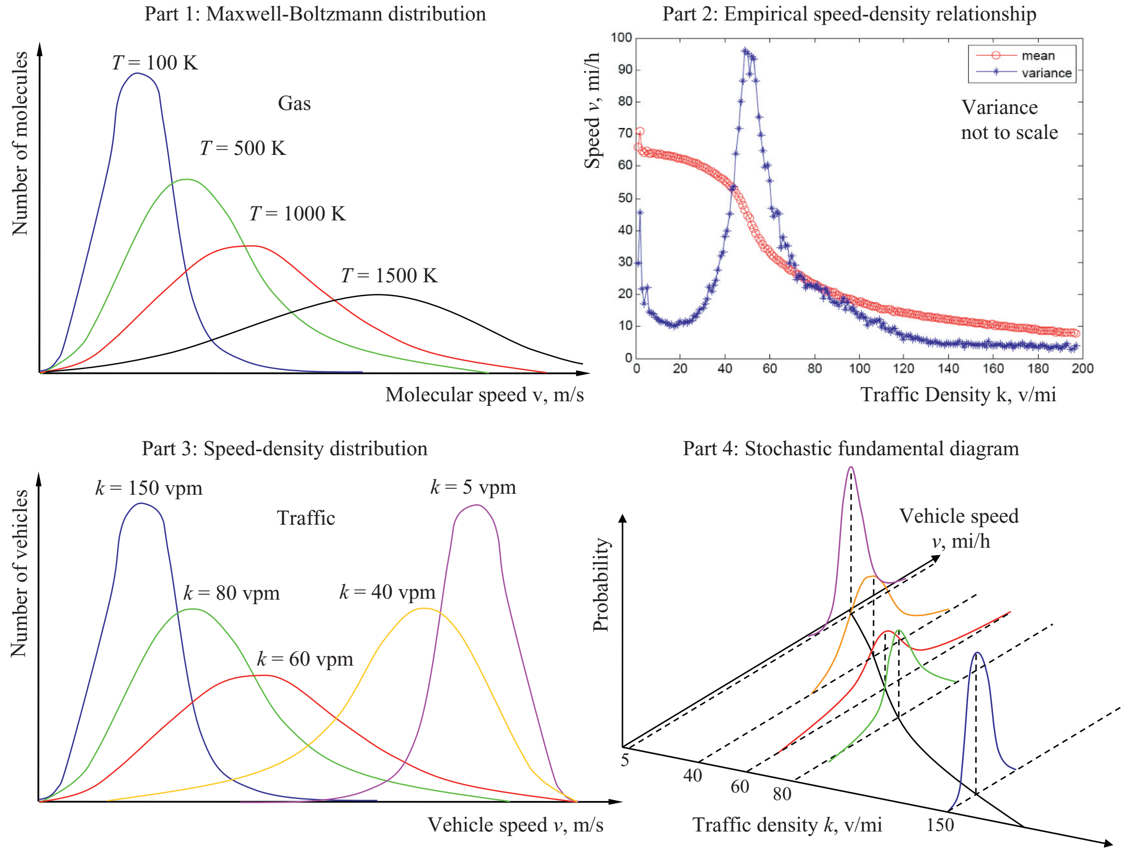

Figure 21.4 Maxwell-Boltzmann distribution in traffic flow.

speed and speed variance as temperature increases. In contrast, traffic flow exhibits a different trend. Empirical observations (Figure 21.4, top right) show that the variance of traffic speed peaks around the optimal density (where capacity flow occurs) and drops at both ends. Such a distribution is illustrated Figure 21.4, bottom left, in contrast to the Maxwell-Boltzmann distribution, and is further elaborated in a three-dimensional model in Figure 21.4, bottom right, which forms the basis of a stochastic fundamental diagram.

21.3 THE FIELD THEORY

The above-mentioned similarities between the transportation system and the physical system provide motivation for a phenomenology of traffic flow (i.e., the field theory) which aims to describe traffic phenomena in a way that is consistent with first principles but is not directly derived from them. Since the transportation system involves both living entities (e.g., human drivers) and nonliving objects (e.g., roadways and vehicles), it is subject to both physical laws and social rules. As such, the phenomenology is formulated progressively on the basis of a set of postulates, two of which (Postulates 1 and 3) are physical and two of which (Postulates 2 and 4) are social.

21.3.1 Postulate 1: A Road is a Physical Field

Postulate 1 is motivated by phenomena M1, M2, and M4 in Section 21.2. In the longitudinal (x) direction, a driver-vehicle unit is subject to a gravity along the road:

$$G_i = m_i \times g_i, \tag{21.1}$$

where i denotes the unit's ID, G_i is the roadway gravity acting on the unit, m_i is the mass of the unit, and g_i is the acceleration of roadway gravity perceived by driver i. As discussed in M1, g_i is a function of driver personalities Θ, vehicle properties Λ, and road conditions Ξ—that is, $g_i = g_i(\Theta, \Lambda, \Xi)$.

Meanwhile, the unit is also subject to a resistance R_i perceived by the driver due to his or her willingness to observe traffic rules (e.g., speed limits). As discussed in M2, R_i is related to the driver's perceived difference between his or her actual speed $\dot{x}_i$ and the desired speed v_i, which in turn is related to the free-flow speed of the road v_f—that is, $R_i = R_i(\dot{x}_i, v_i, v_f)$. Therefore, the net force acting on unit i in the longitudinal direction can be expressed as

$$m_i\ddot{x}_i = G_i - R_i, \tag{21.2}$$

where $\ddot{x}_i$ is the acceleration of unit i. Since the right-hand side represents the amount of net force that can be used to accelerate the unit, it can be interpreted as the driver's *unsatisfied desire for mobility*. As the unit speeds up, the right-hand side decreases (because R_i increases). Eventually, the right-hand side varnishes, at which time the unit reaches its desired speed v_i. If, somehow, a random disturbance brings the unit's speed above v_i, the right-hand side becomes negative. In this case, the unit decelerates and finally settles back to v_i.

In the lateral direction of the road, there are lane lines, road edges, and center lines to guide and separate traffic. As discussed in M4 in Section 21.2, these cross-section elements of the road can be mapped into a roadway potential field U_i^{R} perceived by the driver. When the unit deviates from its lane, the unit is subject to a correction force N_i, which can be interpreted as the stress on the driver to keep in his or her lane (see Figures 21.2 and 21.6). The effect of such a force is to push the unit back to the center of the current lane. On the basis of physical principles, the force can be determined as the derivative of the roadway field, U_i^{R}, with respect to the unit's lateral displacement y_i:

$$N_i = -\frac{\partial U_i^{\mathrm{R}}}{\partial y_i}. \tag{21.3}$$

21.3.2 Postulate 2: A Driver Responds to His or Her Surroundings Anisotropically

The interaction between two driver-vehicle units differs from the collision of two objects in two ways: one pertains to Newton's third law of motion, which is discussed below, and the other concerns noncontact forces, which are the subject of the next postulate.

In classical physics, Newton's third law of motion holds when two objects collide with each other. However, the law generally does not hold in the interaction between two driver-vehicle units. For example, when a fast vehicle catches up with a slow vehicle, the fast driver perceives a "repelling force" (i.e., stress) as the gap closes. The smaller the gap, the greater the force. Conversely, the reaction force may or may not be perceived by the slow driver depending on whether he or she notices the approaching fast vehicle and his or her willingness to respond. Since drivers all sit facing the front, it is the driver behind who is responsible for watching fir safe distances and who is held liable for a rear-end collision should it happen. Therefore,

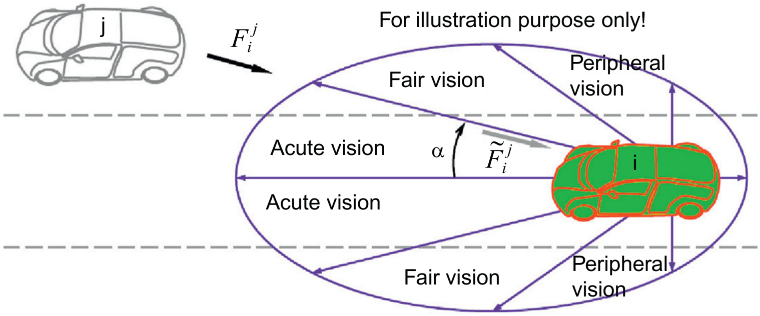

Figure 21.5 Distribution of driver's attention.

it is not uncommon that a leading driver does not respond to situations happening behind such as an approaching fast vehicle.

In general, a driver's responsiveness to his or her surroundings varies with his or her viewing angle and scanning frequency (see Figure 21.5). For example, the area immediately in front of the driver, especially in the same lane, falls into the driver's acute vision zone. The driver is responsible for watching this area constantly and responding to a situation promptly. Roughly in the driver's fair vision zones, the frontal areas in side lanes receive a considerable amount of the driver's attention since vehicles in the side lanes may change to the subject lane and the driver needs to watch this area when changing lanes. In comparison, the driver scans less frequently at both sides of his or her vehicle (roughly the driver's peripheral vision zones) unless the driver needs to change lanes or avoid parallel running. The last and least attended area is the rear of the vehicle, not only because it is difficult to access (indirectly by means of side or rear mirrors), but also because liability rests with drivers behind. Therefore, it is reasonable to assume that the driver's directional response to his or her surroundings, γ_i, is a function of his or her viewing angle α_i—that is, $\gamma_i = \gamma(\alpha_i)$. Consequently, the force that actually acts on the unit, $\tilde{F}_i$, is the product of the force that might have been perceived by the driver if he or she had paid full attention to it, F_i, and his or her directional response γ_i—that is,

$$\tilde{F}_i = F_i \times \gamma(\alpha_i), \tag{21.4}$$

where $\alpha_i \in [-\pi, \pi]$ is the viewing angle. For example, if one chooses $\gamma(0) = 1$ and $\gamma(\pi) = 0$, the driver responds to F_i in full when it comes from a leading vehicle (i.e., $\alpha_i = 0$) and ignores F_i when it comes from a trailing vehicle (i.e., $\alpha_i = \pi$).

21.3.3 Postulate 3: A Driver Interacts with Others by Action at a Distance

As described in E1, E2, and E3 in Section 21.2, a driver is able to sense the presence of other vehicles and obstacles in his or her vicinity and take preventive actions to avoid a collision. It is postulated that such an action at a distance is mediated by a field which is perceived by the driver as the danger of a collision. One may imagine the field as a hill; the higher and steeper the hill is, the more difficult it is to climb. The base of the hill/field delineates a region, outside of which the driver is not influenced by the field. For example, the dash-dotted oval (labeled as "Base j") in Figure 21.6 represents the base of the field perceived by driver i due to unit j. One may also interpret the field as the personal space of unit j, into which intrusion is discouraged. The deeper unit i intrudes, the stronger the repelling force it receives. The longitudinal section of the field is illustrated as the curve above the x-axis.

Similarly, unit k represents another field (whose base is labeled as "Base k") which also exerts an influence on unit i. Since unit k is in the lane at

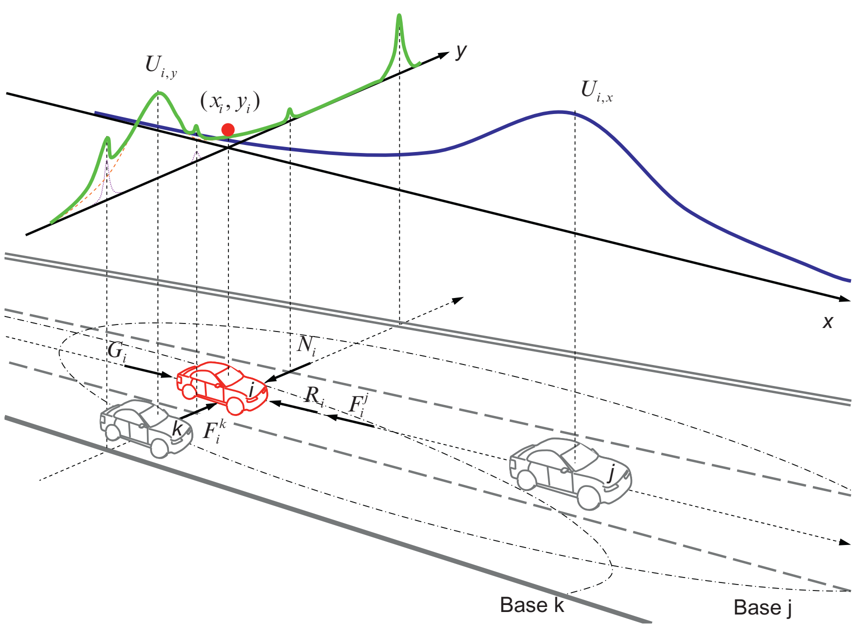

Figure 21.6 The illustration of a field.

the side of unit i, the influence is not in the longitudinal x direction but is in the lateral y direction—that is, the field results in a repelling force, F_i^k, on unit i, which motivates it to shy away from unit k.

The above fields and, consequently, forces are related to driver personalities Θ and vehicle dynamics Γ—that is, $U = U(\Theta, \Gamma)$ and $F = F(\Theta, \Gamma)$. For example, since an aggressive driver accepts shorter car-following distances, the field perceived by such a driver covers a smaller base. On the other hand, the faster a unit moves, the more hazard it imposes on neighboring vehicles, and thus the larger and steeper is the field it creates.

21.3.4 Postulate 4: A Driver Tries to Achieve Gains and Avoid Losses

A driver's strategy of moving on roadways is to achieve mobility and safety (gains) while avoiding collisions and violation of traffic rules (losses). Such a strategy can be represented with use of an overall potential field U_i which consists of component fields such as those due to moving units U_i^{B}, roadways U_i^{R}, and traffic control devices U_i^{C}—that is,

$$U_i = U_i^{\mathrm{B}} + U_i^{\mathrm{R}} + U_i^{\mathrm{C}}. \tag{21.5}$$

If U_i is viewed as a mountain range whose elevation denotes the risk of losses, the driver's strategy is to navigate through the mountain range along its valley—that is, the least stressful route. For example, Figure 21.6 illustrates two sections of such a field. Perceived by driver i, the longitudinal x section of the field, $U_{i,x}$, is dominated by unit j since it is the only neighboring vehicle in the center lane. Unit i is represented as a ball which rides on the tail of curve $U_{i,x}$ since the vehicle is within unit j's field. Therefore, unit i is subject to a repelling force F_i^j which is derived from $U_{i,x}$ as

$$F_i^j = -\frac{\partial U_{i,x}}{\partial x}. \tag{21.6}$$

The effect of F_i^j is to push unit i back to keep a safe distance. By incorporating the driver's unsatisfied desire for mobility $(G_i - R_i)$, we can determine the net force in the x direction as:

$$m_i \ddot{x}_i = \sum F_{i,x} = G_i - R_i - F_i^j = (m_i g_i - R_i) + \frac{\partial U_{i,x}}{\partial x}. \tag{21.7}$$

The section of U_i in the lateral y direction, $U_{i,y}$ (the bold curve), is the sum of two components: the cross section of the field due to unit k (the dashed curve) and that due to the roadway field (the dotted curve).

The former results in a repelling force F_i^k which makes unit i to shy away from unit k and the latter generates a correction force N_i should unit i depart its lane center. Therefore, the net effect can be expressed as:

$$m_i \ddot{y}_i = \sum F_{i,y} = F_i^k - N_i = -\frac{\partial U_{i,y}}{\partial y} \tag{21.8}$$

By incorporating time t, driver i's perception-reaction time τ_i, and driver i's directional response γ_i, we can express Equations 21.7 and 21.8 as

$$m_i \ddot{x}_i(t+\tau_i) = \sum \tilde{F}_{i,x}(t) = \gamma_i^0 [G_i(t) - R_i(t)] + \gamma(\alpha_i^j) \frac{\partial U_{i,x}}{\partial x}, \tag{21.9a}$$

$$m_i \ddot{y}_i(t+\tau_i) = \sum \tilde{F}_{i,y}(t) = -\gamma(\alpha_i^k) \frac{\partial U_{i,y}}{\partial y}, \tag{21.9b}$$

where $\gamma_i^0 \in [0,1]$ represents the unit's attention to its unsatisfied desire for mobility (typically $\gamma_i^0 = 1$), and α_i^j and α_i^k are viewing angles, which are also functions of time. The above system of equations summarizes the field theory in generic terms and constitutes the basic law governing a unit's motion on a planar surface.

21.4 SIMPLIFICATION OF THE FIELD THEORY

Though the generic form of the field theory is able to explain some traffic phenomena qualitatively, rigorous modeling of traffic flow requires a specific form, which is the focus of this section and the next chapter. In the generic theory, the functional forms of the field U_i, roadway gravity G_i, and resistance R_i are undetermined. It appears that the generic theory can take many specific forms, and it is impractical to enumerate all of them. In choosing a specific form, we find that Occam's razor turns out to be a good rule of thumb, and basically says that "entities should not be multiplied unnecessarily." Hence, the razor gives rise to the following considerations: (1) the chosen specific form should make physical sense, for which empirical observations are good tests, (2) it should take a simple functional form that involves physically meaningful parameters but not calibration coefficients, and (3) it should provide a sound microscopic basis for aggregated behavior—that is macroscopic equilibrium models. With these considerations in mind, some simplifications are made to the generic theory as the first step in the formulation of a specific form.

Simplification 1

Rather than formulating the field itself, the specific form formulates forces resulting from the field directly.

Simplification 2

The specific form decouples equations in the longitudinal x direction and lateral y direction—that is, the longitudinal equation is used to model driver's longitudinal control (e.g., car-following behavior) and the lateral equation is used only when a lane change is to be considered.

Simplification 3

Directional response $\gamma(\alpha)$ is treated as follows: for car following, the subject driver responds only to his or her leader; for lane changing, the subject driver responds to the leading and trailing vehicles in the current lane and the target lane.

21.4.1 Motion in a Longitudinal Direction

With these simplifications, vehicle motion in the longitudinal direction is formulated as follows. Note that time t and response delay (i.e., perception-reaction time τ) are dropped for convenience.

The term $(G_i - R_i)$ explains a driver-vehicle unit's unsatisfied desire for mobility. Intuitively, when a unit starts from standstill—that is, $\dot{x}_i = 0$—its unsatisfied desire for mobility is the greatest. As the unit speeds up, $(G_i - R_i)$ decreases accordingly, but is still positive—that is, it still accelerates the unit to higher speeds. When the unit achieves its desired speed—that is, $\dot{x}_i = v_i$—its desire for mobility has been fully satisfied and, hence, $G_i - R_i = 0$, which means that the unit settles at v_i if no other forces act on it. If a random perturbation brings $\dot{x}_i$ over v_i, the unit's desire for mobility is oversatisfied and $G_i - R_i$ becomes negative, which decelerates the unit back to v_i. With the above understanding, a specific form of the unsatisfied desire for mobility can be formulated as

$$G_i - R_i = m_i g_i \left[1 - \left(\frac{\dot{x}_i}{v_i} \right)^{\delta} \right], \tag{21.10}$$

where δ is a calibration parameter.

When a fast unit i (with displacement x_i, speed $\dot{x}_i$, and acceleration $\ddot{x}_i$) catches up with a slow unit j (with x_j, $\dot{x}_j$, and $\ddot{x}_j$), the former is subject to a noncontact force, F_i^j, from the latter. Such a noncontact force varies with the spacing between the two units, $s_{ij} = x_j - x_i$. For example, the force virtually has no effect on unit i when it is distant, but has an effect when unit i becomes close (e.g., within the range of its desired spacing s_{ij}^*), increases as the spacing becomes even shorter ($s_{ij} \downarrow$), and it goes to a maximum when $s_{ij} \to l_j$, where l_j represents the minimum "safety room" required by unit j, an extreme case of which is the length of vehicle j. In addition, the effect of the force is also related to the speeds and relative speed of units i and j. Such an effect can be incorporated into the formulation of driver i's desired spacing s_{ij}^*.

Therefore, a simple way to represent the force is to use an exponential function. The general idea of this model is to set the desired spacing s_{ij}^* as a baseline, beyond which the intrusion by unit i is translated exponentially to the repelling force acting on the unit. As such, a more specific but still quite generic form of the force is

$$F_i^j = f(\mathrm{e}^{s_{ij}^* - s_{ij}}), \tag{21.11}$$

where $s_{ij}^* - s_{ij}$ represents how far unit i intrudes into s_{ij}^*.

Combining the above, we can express the effort that is required by driver i to control his or her vehicle in the longitudinal direction as

$$m_i \ddot{x}_i = G_i - R_i - F_i^j = m_i g_i \left[1 - \left(\frac{\dot{x}_i}{v_i}\right)^{\delta}\right] - f(\mathrm{e}^{s_{ij}^* - s_{ij}}) \tag{21.12}$$

or

$$m_i \ddot{x}_i = m_i g_i \left[1 - \left(\frac{\dot{x}_i}{v_i}\right)^{\delta} - f(\mathrm{e}^{s_{ij}^* - s_{ij}})\right] \tag{21.13}$$

if the coefficient of F_i^j is chosen properly. Though Equation 21.13 can be instantiated in many possible ways, the following special case is of particular interest. Putting time t and response delay τ back in and eliminating vehicle mass m from both sides, we find Equation 21.13 can take the following special form:

$$\ddot{x}_i(t + \tau_i) = g_i \left[1 - \left(\frac{\dot{x}_i(t)}{v_i}\right)^{\delta} - \mathrm{e}^{\frac{s_{ij}(t)^* - s_{ij}(t)}{Z}}\right]. \tag{21.14}$$

If one chooses $\delta = 1$ and $Z = s_{ij}(t)^*$, the above equation reduces to a more specific form:

$$\ddot{x}_i(t + \tau_i) = g_i \left[1 - \left(\frac{\dot{x}_i(t)}{v_i}\right) - \mathrm{e}^{\frac{s_{ij}(t)^* - s_{ij}(t)}{s_{ij}(t)^*}}\right]. \tag{21.15}$$

This model will be further elaborated and analyzed in the next chapter.

21.4.2 Motion in a Lateral Direction

A driver-vehicle unit's motion in the lateral y direction involves a decision at two levels: lane change and gap acceptance. A lane-change decision concerns the driver's desire to change to an adjacent lane to better achieve his or her goals such as mobility and safety, a situation which typically happens when the driver is blocked by a slow leader in the current lane. A gap-acceptance decision addresses the execution of the lane change decision by physically moving the vehicle into the target lane when an opportunity comes up (e.g., a safe gap is available in the target lane). Figure 21.8 illustrates the scenario in Figure 21.7 from a different and broader perspective involving the subject unit i and its leader j in the right lane and a trailing neighbor p in the left lane. The top part of Figure 21.8 illustrates unit

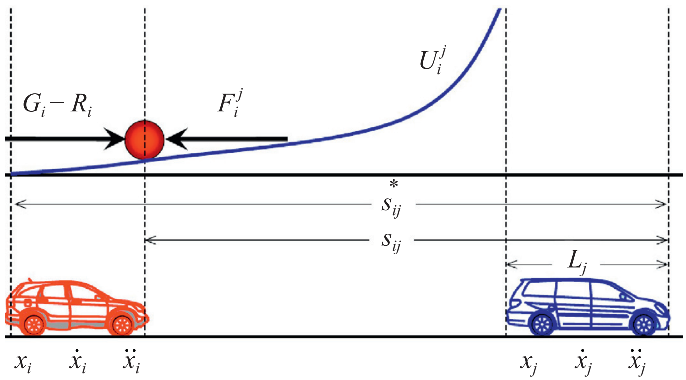

Figure 21.7 Action at a distance.

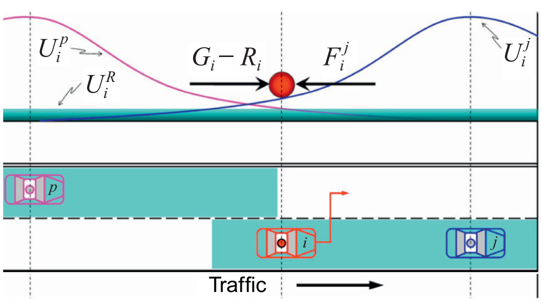

Figure 21.8 Lane change.

i (the ball) in its perceived potential fields: U_i^j due to unit j in the same lane, U_i^p due to unit p in the left lane, and U_i^{R} due to the road barrier (lane line). Driven by its desire for mobility, unit i climbs up onto U_i^j, during which time unit i has to adapt to unit j's speed while achieving a balance between $(G_i - R_i)$ and F_i^j. Under this circumstance, unit i reaches its decision on a lane change in order to satisfy its desire for mobility. With this decision, driver i begins to seek opportunities in adjacent lanes. In this particular example, the right side is obviously not an option since it is prohibitive to move off the road. Hopefully, an opportunity exists in the left lane because the elevation of unit i (where the ball rides) is higher than both the lane barrier U_i^{R} and the front of field U_i^p. Therefore, unit i initiates a smooth transition by laterally rolling off the tail of U_i^j, crossing over U_i^{R}, and landing on the front of U_i^p, the effect of which is shown in the middle part of Figure 21.8.

We can further simplify the above lateral control model by reducing a smooth field to a flat "personal space" into which intrusion by another unit is undesirable. For example, the bottom part of Figure 21.8 illustrates units j and p's personal spaces after elimination of the lane barrier. A lane-change decision is reached whenever a unit intrudes into another unit's personal space, which certainly applies to unit i. With such a decision, unit i begins to search for open spaces in adjacent lanes, and one happens to be available in the left lane. Hence, the result of the gap-acceptance decision is to abruptly switch unit i to the left lane.

21.5 DISCUSSION OF THE FIELD THEORY

Responding to the two questions posed at the beginning of this chapter, we state the field theory can serve as a unifying framework that is able to coherently relate existing models to each other. This will be the topic of Chapter 23. Meanwhile, the field theory is proposed with its roots in both physical science and social science and, therefore, establishes the foundation that allows transportation to be treated as a science. Such a role of the field theory is discussed further below.

21.5.1 Tentative Definition of Two Vague Terms

On the basis of the field theory, it is possible to quantify two vague terms—namely, mobility and congestion—which are frequently used in the transportation profession without rigorous definition.

Mobility

The dictionary definition of *mobility* is "the quality of moving freely." As such, the quality reaches 100% if an individual is able to move as he or she desires, while the quality drops to 0 if the person is stuck in a traffic jam. Therefore, the phenomenological interpretation of personal mobility $M_i(t)$ at an instant of time t perceived by driver i can be expressed as the portion of his or her desired speed which has been satisfied—that is,

$$M_i(t) = \frac{\dot{x}_i(t)}{v_i}, \tag{21.16}$$

where $\dot{x}_i(t)$ is driver i's actual speed and v_i is the desired speed. Since v_i is (typically) greater than $\dot{x}_i(t)$, personal instant mobility $M_i(t)$ ranges between 0 and 1. The mobility perceived by the same driver over the course of a journey can be represented as the average of $M_i(t)$ over trip time T_i:

$$M_i = \frac{1}{T_i} \int_0^{T_i} \frac{\dot{x}_i(t)}{v_i} \mathrm{d}t. \tag{21.17}$$

Again, personal mobility M_i falls between 0 and 1. Similarly, the mobility perceived by all drivers in a traffic system can be calculated as

$$M = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{T} \int_0^{T_i} \frac{\dot{x}_i(t)}{v_i} \mathrm{d}t \right), \tag{21.18}$$

where N is number of drivers in the traffic system. Note that system mobility $M \in [0, 1]$ can be used as an indication of the level of service provided by the traffic system and perceived by drivers.

Congestion

The dictionary definition of *congestion* is "a state that is so crowded as to hinder or prevent freedom of movement." One candidate quantification of congestion, C, can be the opposite of mobility—that is, $C = 1 - M$, which is expressed relatively as a percentage. Another, perhaps more meaningful, way to quantify congestion is to recognize the "stress" experienced by a driver when he or she moves in a traffic system. Therefore, the phenomenological interpretation of personal congestion $C_i(t)$ at an instant of time t can be expressed in absolute terms as the stress (or equivalently hazard or potential) $U_i(t)$ perceived by driver i—that is,

$$C_i(t) = U_i(t). \tag{21.19}$$

As defined in Equation 21.5, the overall potential consists of potentials due to moving units U_i^{B}, roadways U_i^{R}, and traffic control devices U_i^{C}. Therefore, a driver would experience no congestion if he or she moves on a roadway that is free of impedance from other moving units, the need for lane changes, and traffic control devices. Therefore, any increase of these would add to the driver's perception of congestion. Consequently, personal congestion perceived by the same driver over the course of a journey can be represented as the sum of $C_i(t)$ over trip time T_i:

$$C_i = \int_0^{T_i} C_i(t)\mathrm{d}t. \tag{21.20}$$

Further, the congestion experienced by all drivers in a traffic system can be calculated as the sum of personal congestion over the driver population:

$$C = \sum_{i=1}^{N} C_i. \tag{21.21}$$

21.5.2 Connection to the Existing Knowledge Base

The purpose of this discussion is to place the field theory in a broader context and show how the field theory relates to the existing knowledge base and, in return, how successful experience of related fields can be transferred to solve our problems at hand.

Connection to Other Traffic Flow Theories

Existing microscopic traffic flow models emphasize the application of social rules or human factors in the modeling of car-following behavior, whereas the work presented in this chapter attempts to integrate both social rules and physical principles in the modeling of traffic flow. Readers are referred to Chapter 23, where such a connection is elaborated and presented from a unified perspective.

Connection to Other Engineering Disciplines

The Lennard-Jones potential plays an important role in engineering, particularly in granular flow and molecular dynamics. In molecular dynamics, computer simulation is employed to trace the time evolution of a set of interacting particles (e.g., atoms or molecules) by integrating their equations of motion. The Lennard-Jones potential is the underlying model

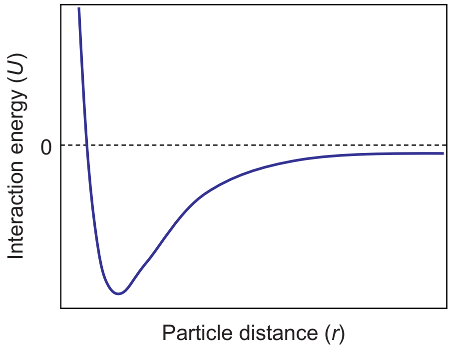

Figure 21.9 Lennard-Jones potential.

to determine the motion and interaction of these particles. In materials engineering and granular flow, the Lennard-Jones potential is typically used as the constitutive law to determine the interaction of two particles. With a clear understanding of the constitutive law of two particles, systems consisting of a large quantity of these particles (e.g., many-body systems) can be simulated and analyzed. Illustrated in Figure 21.9, the Lennard-Jones potential takes the following form:

$$U(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - (\frac{\sigma}{r})^{6}\right], \tag{21.22}$$

where U is the Lennard-Jones potential due to particle interaction, r is the distance between two particles, ϵ is the depth of the potential well, and σ is the distance at which the interparticle potential is zero. The equation is actually a superposition of two terms: a long-range attraction term $(\frac{\sigma}{r})^6$ and a short-range repulsion term $(\frac{\sigma}{r})^{12}$.

The phenomenology, in particular Equation 21.12, takes a similar form. For example, the long-range attraction $(m_i\ddot{x}_i = m_i g_i[1 - (\frac{\dot{x}_i}{v_i})])$ is due to a driver's desire for mobility and the short-range repulsion $f(e^{s^*_{ij}-s_{ij}})$ is due to safety rules. In addition, the interaction between two vehicles is a function of the spacing s_{ij} (equivalent to r) between them, and there is an equilibrium spacing s^*_{ij} (equivalent to σ) around which the attraction equals the repulsion. Therefore, the Lennard-Jones potential in a transportation system can be derived from the phenomenology. With such a bridge, transportation and related engineering disciplines are able to not only learn from but also shed light on each other.

Connection to Physical Science

The phenomenology proposed herein represents a body of knowledge that originates from empirical observations and, in return, that is able to explain real-world phenomena. Rather than being derived directly from first principles, the phenomenology is formulated in a way that is consistent with fundamental theory. For example, the essence of the field theory in the phenomenology is the explanation of an individual driver's action by recourse to his or her position in relation to others. The driver's position in the field in turn gives rise to a force acting on the person, but such a force is motivated from within as opposed to being applied from without. As another example, Equation 21.9 is a special form of Newton's second law of motion if one ignores the driver's perception-reaction time τ and directional response γ. In addition, action at a distance as a means of interaction between drivers becomes a hard collision if the driver's need for safety disappears (i.e., the potential field as a function of spacing $U(s_{ij})$ becomes a spike). Moreover, Newton's third law of motion holds if drivers respond to their surroundings isotropically. Furthermore, isotropic response, together with a hard collision, gives rise to the laws of momentum and energy conservation. Therefore, the phenomenology represents a special form of Newton's laws in a social setting (i.e., a transportation system involving human drivers). With its interpretation of the mean free path (i.e., desired car-following distance s^*_{ij}) and a molecular collision (i.e., action at a distance between vehicles), the phenomenology allows the application of other physical principles (such as kinetic theory) to further understand transportation systems as an ensemble.

21.6 SUMMARY

Involving both physical objects (e.g., vehicles) and living entities (e.g., drivers), a transportation system shares many commonalities with social and physical systems. The social side of the transportation system has long been recognized, as evidenced by applications of social rules and human factors in microscopic traffic flow modeling such as car following, lane changing, and route choice. In contrast, the physical side of the system has yet to receive proper attention. The transportation system does, however, exhibit many physical properties, which provides part of motivation for the proposed field theory of traffic flow.

To pave the foundation for the field theory, some physical phenomena in traffic flow were analyzed in relation to its social properties, in particular motivations for drivers' decisions observed from their driving experiences. These phenomena, including those of mechanics, electromagnetics, waves, and statistical mechanics, strongly suggest that it is meaningful to integrate both physical and social principles into the modeling of traffic flow.

With the above understanding, the field theory was progressively formulated on the basis of a series of postulates, two of which are physical and two are social. The first postulate (physical) assumes that a roadway is a physical field in which a vehicle is subject to a roadway gravity and also a resistance due to the driver's willingness to observe traffic rules (e.g., speed limits). The second postulate (social) accounts for the driver's directional responsiveness to his or her surroundings. The third postulate (physical) imposes an action at a distance between two neighboring vehicles, and such an interaction is mediated by a potential field which is perceived as the danger of a collision. The distinction between a field perceived by a driver and a physical field is that the former impinges from the inside of the driver through motivation as opposed to through external compulsion. The fourth postulate (social) interprets driving strategy as a social rule—that is, a driver always tries to achieve gains (e.g., mobility and safety) and avoid losses (e.g., collisions and violation of traffic rules). From combination of the above postulates, the field theory was generically formulated as a system of equations governing the motion of a vehicle on a roadway in relation to other vehicles.

PROBLEMS

1. Use the field theory to explain the following phenomena:
   - **a.** A vehicle begins to accelerate after an emergency stop on the hard shoulder.
   - **b.** The vehicle gradually settles at its desired speed,
   - **c.** After a momentary speeding, the driver beings to speed up to the desired speed.
   - **d.** The driver applies the brakes when approaching a slow vehicle.
   - **e.** The driver adopts the leading vehicle's speed by following the leader.
   - **f.** The driver brakes again because a third vehicle cuts in between two vehicles in car-following mode.
   - **g.** The vehicle being cut off from the leading vehicle changes lane to seek speed gains.

2. Elaborate your strategies on how to capture the effect of the following traffic control devices with use of the field theory:
   a. A signal
   b. A STOP sign
   c. A speed limit sign
3. Use the field theory to explain a rear-end accident.
4. Comment on how realistic the field theory is. Is it feasible to find evidence to prove or disprove the existence of a field?