Precautionary Safety for Autonomous Driving Systems: Adapting Driving Policies to Satisfy Quantitative Risk Norms

Gabriel Rodrigues de Campos, Roozbeh Kianfar, and Mattias Brännström

Abstract— **Road safety has been studied for more than 80 years with the objective to prevent traffic accidents, injuries, and fatalities. Proper road design, standards, traffic rules, driver education and law enforcement are continuously improved to reduce the crash risk and enable mitigating actions. While significant advancements have been made, humans still frequently do mistakes, sometimes with severe consequences. In this paper we introduce the notion of Precautionary Safety and propose a methodology such that autonomous vehicles can adjust their trajectory planning to their capabilities, external conditions, and knowledge on human mistakes in order to satisfy overall requirements on accident-, injury- and fatality rates. More precisely, we describe how to make adjustments to existing driving policies so to satisfy real-world safety requirements, rather than only obeying to the law and traffic rules. As an illustrative example, the methods are applied to accident scenarios between vehicles and jaywalking pedestrians.**

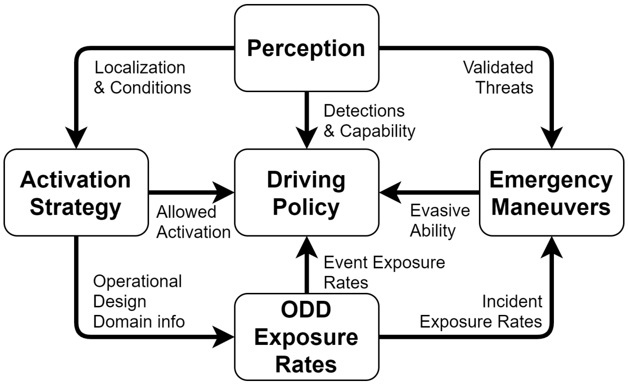


Fig. 1. Precautionary Safety Driving Policy for Autonomous Driving, adapting the trajectory planning to the ability to perform evasive maneuvers.

I. Introduction

Every year, over 1.3 million fatalities are caused by traffic accidents around the world. To tackle accidents related with human error or incorrect situation assessment, responsible for up to 99% of the traffic accidents in the US [1], several generations of Advanced Driver Assistance Systems (ADAS) have been developed and deployed in the last few decades, with remarkable effects on real-world traffic and traffic related injuries/fatalities. The reader can refer to [2] for a recent report of the European Commission on intelligent safety systems and their effects on the reduction of critical accidents and casualties.

During the last decade, and in particular in the last few years, the automotive industry is going through structural transformations powered by new breakthroughs on electrification, connectivity and automation. In particular, recent advances in perception and compute technologies, as well as on active safety and advanced cruising features, have led to high expectations on a rapid development of Autonomous Driving (AD) systems. But even if human-supervised cruising features and AD systems can be perceived as very similar, given that both control the longitudinal and lateral motion of the vehicle, they present clear differences: while AD systems are responsible for driving safely, cruising features only support the driver and rely in his/her supervision and responsibility. That is, cruising features are designed to do their best, whereas AD systems are to be designed to operate only if they can do it safely, without human supervision.

In this paper, we focus on safe trajectory planning and decision-making for AD, often referred to as a **safe driving policy**. Before introducing the proposed driving policy design methodology, it is important to first discuss the notion of safety. Some researchers define safe driving as *legal safety*, i.e., in the sense that AD systems are considered safe if they obey to a set of rules at all times [3], [4]. As a starting point, this is a indeed a fair statement and ambition. After all, it is easy to draw the conclusion that, if everyone just followed the rules, there would be no accidents. However, the underlying assumption that other road users always follow rules, is questionable. In fact, many people violate traffic rules, either on purpose or by mistake: driving faster than the speed limits, getting distracted, taking way when changing lanes or when driving through intersections. Fortunately, the infrastructure is built to be resilient to human errors, and other road users are fairly good at counteracting other's mistakes by using a combination of proactive and reactive actions. It is therefore important to acknowledge that people do make mistakes and to design AD systems that are resilient to human errors. Thus, instead of using the legal safety concept alone, we propose to define safe driving as a low accident rate with low severity, no matter whose fault is it.

The contribution of this paper is a novel approach for the design of safe driving policies, summarized in Fig. 1. The advantages with the proposed methodology are:

- any existing driving policy can be used as a base;
- any emergency maneuver algorithm can be utilized;
- perception capabilities, evasive ability and knowledge on exposure to risky situations are jointly assessed to identify how the driving policy needs be adjusted to stay safe, and/or when it is safe to activate the AD system.

* The authors are with Zenseact AB, Lindholmspiren 2, 417 56 Göteborg, Sweden. This work was also supported by the Swedish Strategic Vehicle Research and Innovation (FFI) programme, via SALIENCE4CAV (ref. 2020-02946). E-mail:{gabriel.campos, roozbeh.kianfar, mattias.brannstrom} @zenseact.com

The paper is organized as follows. First, Section II provides some background and context to this work. Section III focuses on quantitative safety requirements aspects, while Section IV describes the novel concept of precautionary safety (PCS), and how to enforce safety adaptations to existing driving policies accordingly. Section V discusses adaptations to jaywalkers, providing a deeper discussion on the derivation of PCS adaptations and some numerical examples. Finally, Section VI presents some concluding remarks and avenues for future research.

II. Background

Safety assurance and safe planning is one of the most challenging tasks in the development of AD systems. In [5], [6], [7], [8], [9], [10], various formal method techniques are used for synthesizing decision & control software, where the software is guaranteed by design to not violate traffic rules and other safety requirements. However, reduced scalability and the inability to deal with probabilistic uncertainties and quantitative requirements are limiting factors for practical usability of such techniques [11].

In [12], NVIDIA describes its Safety Force Field concept, as a safety layer for obstacle avoidance which guarantees that the AD vehicle does not expose other road users to dangerous behaviors. Similarly, Mobileye has proposed a white-box, interpretable, mathematical model for safety assurance, denoted as Responsibility-Sensitive Safety [3]. Similar methods are also proposed in [13], [14], which can guarantee that, as long as other road users act according to certain assumed behaviors, the AD vehicle will not crash and will mitigate collisions for unforeseen behaviour of other road-users.

Formal methods are good starting points for establishing safe driving policies, especially in a world with only AD vehicles are present, or if all road users always obey to a set of rules. However, in practice, AD vehicles will always have to interact with other road users that frequently violate rules. To begin with, the first AD vehicles will have to interact with existing human-driven vehicles. A long, co-existence phase will follow until manually driven vehicles are phased out and, even if they are completely phased out, AD vehicles would still have to interact with rule breaking pedestrians/animals.

In addition to obeying traffic rules, AD systems therefore need to be resilient to the behaviour of human road users. People are not perfect, and one cannot expect them to avoid all mistakes, no matter how much we educate them. Instead, it is important to observe how people behave in the real world and then ensure that the AD systems adapt their behaviour to people, and not vice versa. In Fig. 1 the core principles of the proposed PCS concept are illustrated, showing how the Driving Policy can be adapted to account for exposure to external incidents, the perception ability of the AD system, and its ability to identify and execute emergency maneuvers.

III. Quantitative Risk Norm

Real-world safety is often described in quantitative numbers, e.g., low accident or fatality rates [15], [16]. In order achieve this, multiple methodologies have been proposed, including Waymo's combination of system-level testing (simulation, test track and public road) and component and subsystem testing [17], [18]. Similar proposals are also under development in industry standards, e.g. the Safety of the Intended Functionality [19], combined with ISO 26262 [20].

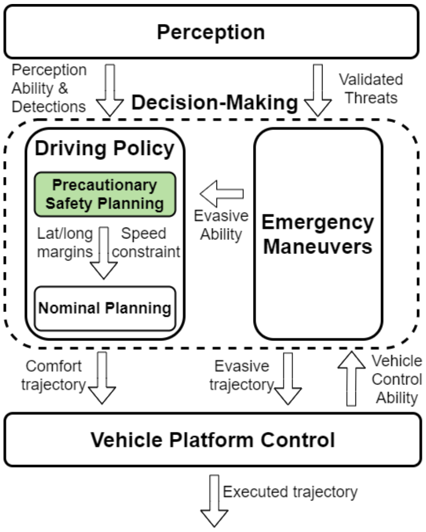


Fig. 2. Autonomous Driving (AD) system architecture. The Precautionary Safety Planning module is added to guide the Nominal Planning module with additional speed constraints and/or longitudinal/lateral margins in order to ensure that the Emergency Module is given the prerequisites it needs to satisfy a given quantitative risk norm. Depending on the design of the AD system and its original Driving Policy, the Emergency Maneuver module can optionally be incorporated directly in the Nominal Planning module.

In this paper, as a complement to existing work, we propose a structured way to adapt AD driving policies to satisfy quantitative safety requirements. Simply put, AD systems are proposed to adapt their driving policies to their abilities, such that they can satisfy any given quantitative safety requirement, denoted hereafter as Quantitative Risk Norm (QRN), as introduced in [15]. The QRN can advantageously be split into QRNs for different accident types and their severity, to ensure that real-world safety is achieved for all types of road users, in any given Operational Design Domain.

IV. Precautionary Safety driving policy

This section presents the proposed PCS concept and design methodology for safe driving policies. A high level illustration of the functional architecture is provided in Fig. 2. The main responsibility of decision-making is, apart from obeying traffic rules, to design and adapt planned trajectories to the system's capabilities, external conditions and knowledge on road user mistakes, in order to satisfy the QRN.

A. *Advisor/precautionary safety planning:*

Nowadays, most ADAS systems are reactive, in the sense that they only act when hazardous situations are detected. For example, Automatic Emergency Braking (AEB) systems can brake if the system judges that an accident is imminent. Obviously, the performance of such systems depends on how early they can detect a situation before it is too late. Although reactive measures can contribute to the safety of AD, it will, alone, never be sufficient to reach the desired QRN. To cope with such shortcomings, precautionary measures need be taken by the decision-making module. For anticipated human behaviors, precautionary measures are easy to introduce, but to handle jaywalkers and other unexpected situations there is a need for combined precautionary and reactive measures.

The core idea behind the PCS module, depicted in Fig. 2, is to drive with precaution to facilitate collision avoidance/mitigation by emergency maneuvers in case of unexpected events. Precautionary measures can be defined as a set of advisory inputs to the trajectory planner. In particular, a set of speed constraints and lateral/longitudinal margins can be provided to the planner to reduce the risk for accidents due to unexpected events. But the amount of measures taken by the PCS planner should depend on the ability of the AD system to detect and react to critical situations, and the prescribed QRN requirements. Specifically, the AD vehicle's ability to detect and react can be categorized into three main sources:

1) perception limitations;
2) planning and prediction limitations;
3) vehicle control limitations.

1) Adaptation to perception limitations: For an AD vehicle with low perception performance, the driving policy should take more precaution to be able to cope with unexpected events. Such limitations, e.g., a limited sensing range/unobserved areas, can impose speed restrictions. Moreover, whenever the AD vehicle has limited ability to detect and react to certain unexpected events, e.g. moose crossing the road, there is a need to drive with extra caution on roads with high exposure to such events, whereas it can drive faster on roads with low exposure. One key element of the PCS principles is therefore how to estimate the impact of perception limitations on the overall performance of emergency maneuvers, and account for it in the planning phase. The adaptations can either be made dynamically or statically, depending on whether the perception performance can be estimated in real-time or using offline validation.

2) Adaptation to planning and prediction limitations: Predicting the intent of other road users is not easy, especially when people make mistakes. Moreover, it would be impossible to drive if one always assumes that all other road users will always make the worst mistake possible. Thus, AD vehicles need to plan trajectories despite their inherent prediction and planning limitations. However, if one can estimate how frequently a given driving policy gets exposed to conflicts due to prediction limitations, and how well the reactive part of the AD system, i.e. the emergency maneuvering, would be able to solve such conflicts, then one can derive additional precautionary measures that need to be enforced on the driving policy so to satisfy the QRN.

3) Adaptation to platform limitations: Vehicle platform limitations is another factor that the decision-making module should take into account when establishing a driving policy. Similar as for the perception limitations, vehicle platform limitations can arise from fundamental limitations, e.g. limited steering torque or slow response in the brake systems, or from dynamical limitations, e.g. low friction on a slippery road. Such limitations are also dependent on the planned maneuver by the decision-making module, i.e. an evasive maneuver requires higher road-tyre friction when compared to a slow lane-change maneuver, for instance. Hence, the planning module should know which performance to expect from the vehicle platform for a given maneuver, at least in probabilistic terms, in order to ensure that platform limitations do not lead to a violation of the QRN.

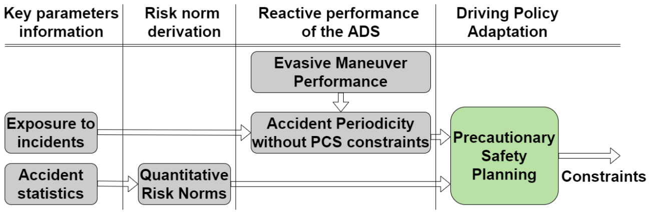

Key parameters information
Risk norm derivation
Reactive performance of the ADS
Driving Policy Adaptation
Evasive Maneuver Performance
Exposure to incidents
Accident Periodicity without PCS constraints
Precautionary Safety Planning
Accident statistics
Quantitative Risk Norms
Constraints

Fig. 3. Illustration of the proposed methodology for adapting existing driving policies to reach low accident rates in Operational Design Domains where autonomous vehicles get exposed to human mistakes (incidents).

B. Nominal/precautionary planning:

The nominal planner utilizes the surrounding information from perception and advised inputs from PCS planner to determine a smooth, comfortable and legal trajectory. For instance, in a vehicle following scenario in a highway, the nominal planner should receive information about the road and surrounding objects, together with advised inputs, such as the advised time gap, from the PCS planner. It should then output a trajectory that maintains a sufficiently large distance to the lead vehicle and does not violate the posted speed, and ensures that the QRN for rear-end accidents is satisfied. In other words, an adequate driving policy should keep sufficient distance to both ensure that it doesn't collide with the lead vehicle, and only rarely needs to use hard braking, in order to minimize the risk of being rear-ended.

C. Emergency/reactive planning:

The emergency or reactive planning module is responsible to exploit the full capability of vehicle platform to deal with conflicts and to contribute to the fulfilment of the QRN. This is similar to how traditional collision avoidance/mitigation systems, such as AEB, operate, even if AD systems need to handle many more scenarios than traditional AEB systems. It is important to highlight that the emergency planning module is designed to only act in conflict scenarios, whereas the nominal planning module is responsible for interacting with other road users in the first place, and is designed to avoid as many conflicts as possible. For the design of a PCS driving policy, we therefore propose the ability of the emergency module to detect and react to unexpected events to be analysed using simulations and directed testing at test tracks, and the outcome used to put precautionary constraints on the nominal planning to ensure that the QRN is satisfied.

D. Deriving PCS constraints - methodology

Leveraging the notion of PCS detailed before, an illustration of the proposed methodology for adapting driving policies to PCS constraints is given in Fig. 3. Accident statistics observed from human drivers, and exposure to incidents using an existing driving policy/manual drive, serve as a basis to derive the QRN and accident periodicity with and without PCS adaptations, considering the AD system's evasive maneuver performance. In the sequel, the proposed precautionary safety concept and driving policy derivation methodology will be applied to an illustrative example considering jaywalking pedestrians. Note that both lateral and/or longitudinal adaptations can be considered to increase evasive performance. For the sake of simplicity, though, we only discuss speed adaptations in the remainder of this paper.

V. Speed adaptation to jaywalkers

This section provides an example case to which the proposed Precautionary Safety (PCS) concepts will be applied. PCS will be used to enforce speed constraints on a driving policy, considering the capabilities of the reactive part of the system (subject to perception, planning and actuation limitations), as well as a numerical Quantitative Risk Norm (QRN) for pedestrian jaywalkers for different road segments. It is worth noting, though, that the proposed approach and methodology can be used for alternative traffic scenarios, other sources of data, or to enforce lateral motion constraints.

In terms of traffic interactions, we consider here the EUNCAP Car-to-Pedestrian Nearside Adult (CPNA) certification case to illustrate the novel concepts and methodology. The scenario is represented in Fig. 4a, and corresponds to a collision situation where an adult pedestrian crosses the ego vehicle's path from the nearside, e.g., jaywalking. The collision point, here identified as α, corresponds to a percentage value of the ego vehicle's width at which the ego vehicle hits the pedestrian, if no collision avoidance maneuver is attempted. It is defined between 0 and 1, where 0 means that the collision point is the right corner of the front bumper and 1 the left corner (of the front bumper).

Fig. 4b also illustrates the considered traffic domains, also called later as Operational Design Domain (ODD): urban/suburban driving and freeway/highway driving. Different road segment speeds, defined by posted speed signs or the typical driving speed in dense traffic conditions, are also considered ($[30-100]$ km/h). As a representative example of a complex driving routine, we assume here a routine based on the two different ODDs corresponding to, for example, a daily/recurrent office-commuting-residence routine. Note, though, that more complex routines or different traffic elements/situations can be used, and the numeric elements proposed in this paper extended to such cases.

A. *Quantitative Risk Norms (QRN)*

We consider here Quantitative Risk Norms (QRN) as a safety requirement on AD systems, also referred to in literature as acceptable safety risk or positive risk balance [21]. One particularly important aspect for defining risk norms concerns the severity of any potential accident, which for jaywalkers is shown to be related to the impact speed. Indeed, for any given road, higher severity of injuries are correlated with higher impact speeds, due to clear physical relationships: when the impact speed increases, the amount of energy that is released also increases. While part of the shock energy will be absorbed by the human body, the body can only tolerate a limited amount of external forces. Hence, higher speeds result in more severe injuries whenever unprotected road users, such as pedestrians and cyclists, collide with motorized vehicles. According to [22], pedestrians have a 90% chance of surviving car crashes at 30 km/h or below, but less than a 50% chance of surviving impacts at 45 km/h or above. Based on such insights, we have classified in Table I the different type of collisions with respect to different impact speeds values.

TABLE I

Quantitative Risk Norm (QRN) for Jaywalkers

Impact speed $[km/h]$	Severity [pedestrian accident]	Risk Norm [hours between accidents]
≤ 10	Minor accident	100.000
$[10-20]$	Light injury risk	1.000.000
$[20-30]$	Severe injury risk	10.000.000
$[30-40]$	Low fatality risk	100.000.000
> 40	High fatality risk	1.000.000.000

In order to derive representative and meaningful values for QRNs, we consider here accident statistics from the Strada[1] (Swedish Traffic Accident Data Acquisition) database, managed by the Swedish Transport Authority. The Strada database is continuously updated and is based on information from two sources: i) the police, reporting road traffic accidents with personal injuries; ii) the healthcare system, providing information on people who have sought care for an injury in the road traffic system. According to the Strada database 2006-2019, every year are reported:

- 1.250 light VRU injuries;
- 500 severe VRU injuries;
- 50 VRU fatalities.

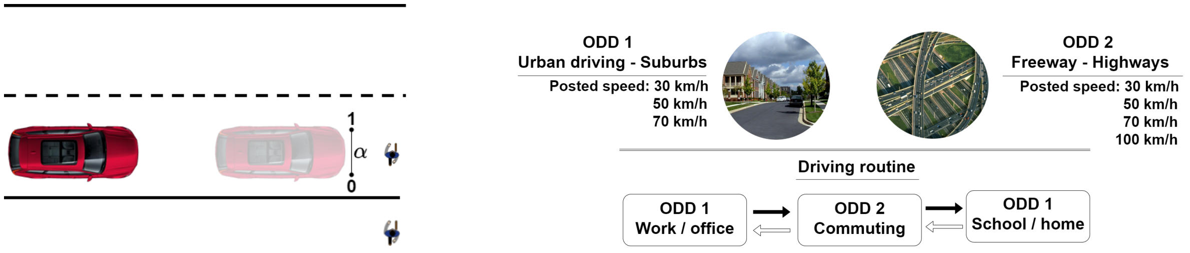
Moreover, and still according to Strada, there has been, on average, 5 millions registered cars in Sweden. Assuming that each car operates 100 hours in urban environment per year and 100 hours in highway environment per year, this leads to a total of 500 million hours of operation in total per each of the ODDs. By dividing the number of injuries/fatalities over the total number of driven hours we obtain:

- 1 light VRU injury per 0.8 million of driven hours;
- 1 severe VRU injury per 2 million of driven hours;
- 1 VRU fatality per 20 millions of driven hours.

For the sake of simplicity, the above metrics assume that the occurrence of injuries and fatalities are evenly distributed between urban and non-urban environments, which is obviously not the case in reality. Hence, more precise datasets can be used to refine these computations such as, e.g., [23].

Considering that the above detailed metrics correspond to the average human driving capabilities/failure rate, we consider in this paper Quantitative Risk Norms (QRN), prescribed for an AD system, that are considerably superior to the average human driver capabilities derived before. Such QRNs, established in terms of hours between accidents, are

[1] More details on the Strada database: www.transportstyrelsen.se/sv/vagtrafik/statistik/olycksstatistik/om-strada/



(a) (b)

Fig. 4. Considered accident scenario (a) and operational design domains (b).

given in Table I for different impact speeds. One can see that the risk norm for low speed collisions, leading to minor or light injuries, is considerably lower than for high impact speeds, for which there is a high risk of fatalities.

It is important to highlight that the derivation of pertinent quantitative risk norms is a vast research field on its own, and different proposals have been presented in the last few years, such as, e.g., in [16]. Being the scope and contribution of this paper broader that the derivation of risk norms topic itself, the above described method should therefore be taken only as a (realistic) example, that will be used in the scope of this paper for the derivation of a driving policy satisfying the prescribed safety standards expressed in the form of QRN. The authors envision, however, to detail other derivation methods in future research.

B. Decision & control reactive performance

To exemplify the proposed methodology described in Fig. 3, we consider in this paper a reference Autonomous Braking System (AEB) based on the algorithms presented in [24]. Such an emergency braking system can be seen as the emergency planning module of an autonomous system, as illustrated in Fig. 2.

In order to derive meaningful and comprehensive results, we have performed a large set of simulations leveraging in-house developed simulation environments that include realistic threat-assessment and decision making algorithms and simplified sensor models. To derive a large set of simulation results, different values of the parameters, defining the car-to-pedestrian scenario illustrated in Fig. 4a, were used such as, for example, the ego vehicle speed, pedestrian speed as well as different collision points. Please note that the scope of this paper is not to propose and discuss new decision & control methods, but rather to use representative metrics of emergency systems' response to design an adaptive driving policy. Such reactive performance metrics can therefore be replaced by any other choice of emergency system, and the remaining of the proposed approach applied to it.

Table II presents the probability of accidents for the reference AEB system, for Ego Vehicle (EV) speeds up to 100 km/h and for the different levels of impact speed, correlated to the severity of the collision as detailed in Table I. The reactive performance presented here is driven by two main aspects: i) sensing and perception limitations, highlighted in blue cells, incorporating non-detections or late object classifications, for instance; ii) decision & control limitations and shortcomings, in the white cells, representing late interventions, inaccurate predictions, inadequate decision making, or insufficient actuation power, for example. Note, however, that additional or more refined sources of failure or uncertainty could also influence the numerical values of each cell. From Table II one can see that, up to 40 km/h, almost all collisions are avoided, which seems consistent with the known range of high efficiency for AEB systems. For ego vehicle speeds above 50 km/h, the performance of the reference collision avoidance system deteriorates leading to occasional low-speed collisions. For ego vehicle's speed of 100 km/h, around 50% of the cases can lead to collisions, some of them at high speed, which can potentially lead to high fatality risk cases in about 10% of the cases.

C. Exposure to incidents

For safety argumentation purposes, an important aspect to be taken into consideration is the exposure to risky situations and incidents. Table III presents the exposure rate for the considered ODDs and for different road segment speeds. The values in Table III, expressed in terms of hours between incidents, correspond to the exposure to situations where the AD vehicle would need to perform evasive braking and/or steering action to avoid collisions with jaywalkers on a given road segment. It can be seen that one is more exposed to car-to-pedestrian incidents whenever driving in urban environments, and particularly in low-speed segments (30 km/h). The exposure rate to incidents is however assumed to be considerably lower (i.e., events are more seldom) when driving in highway segments, considering that many high speed infrastructures are reserved for vehicles only and prevent, as much as possible, the presence of pedestrians.

It is worth mentioning that the exposure metrics in Table III represent an engineering judgement of the authors,

TABLE II

IMPACT SPEED PROBABILITY WITH JAYWALKERS, GIVEN PERCEPTION LIMITATIONS (BLUE) AND EVASIVE ABILITY LIMITATIONS (WHITE)

EV speed	Impact speed at jaywalker incidents [km/h]					
[km/h]	≤ 10	[10-20]	[20-30]	[30-40]	> 40	Avoidance
30	0.01%	0.001%	0.0001%	0.00001%	N.A.	$\sim$ 100%
40	0.1%	0.01%	0.001%	0.001%	N.A.	$\sim$ 99.9%
50	5%	0.1%	0.01%	0.001%	0.0001%	$\sim$ 94.9%
60	5%	5%	0.1%	0.01%	0.001%	$\sim$ 89.9%
70	10%	5%	5%	0.1%	0.01%	$\sim$ 79.9%
80	10%	10%	5%	5%	0.1%	$\sim$ 69.9%
90	10%	10%	10%	5%	5%	60%
100	10%	10%	10%	10%	10%	50%

TABLE III

INCIDENT EXPOSURE: DEFINED AS EVENTS WHEN EVASIVE BRAKING/STEERING IS NEEDED TO AVOID A COLLISION WITH A JAYWALKER

Avg road segment speed	**Urban driving**		**Highway driving**	
[km/h]	Segment type	[hours between incidents]	Segment type	[hours between incidents]
30	Low speed road segment	100	Traffic jam	100.000
50	Medium speed road segment	1.000	Dense traffic	1.000.000
70	High speed road segment	10.000	Low density traffic	10.000.000
100	-	-	Free flow traffic	10.000.000

established for the purpose of this work, but believed to be realistic. More precise values and information could be estimated using, e.g., accident databases in combination with observations from ADAS/AD vehicle fleets. One could also improve such metrics by considering, for instance, exposure metrics with respect to weather conditions, geographical locations or the time of the day. In such cases, the driving policy derivation detailed later in SectionV-E, can be established and adjusted depending on the different parameters.

D. Accident periodicity

The combination of information on the AD system's reactive performance and the exposure to incidents yields a crucial understanding on the accident rates of a given AD system with and without driving policy adaption, see Fig 3. The accident periodicity metrics, established in terms of hours between accidents, are presented in Table V. The numerical values are computed as:

$$\text{accident periodicity} = \frac{\text{hours between incidents}}{\text{accident probability}} \, ,$$

where the exposure rates values (i.e., hours between incidents) are retrieved from Table III and the accident probability (i.e., collision avoidance/reactive performance) from Table II. Table V is organized as follows. In the vertical direction, the table is separated in two panels for each of the considered ODD's, i.e., urban driving on the left-hand side and highway driving on the right-hand side, respectively. In the horizontal direction, the table is separated in four different panels, each one corresponding to a different road segment speed, from 30 km/h on the top to 100 km/h on the bottom. In each sub-panel, the accident periodicity, established in hours between accidents, is presented for different values of ego vehicle speed and for different values of impact speed. The color map is defined according to the risk norms given in Table I: i) for the green cells, the accident periodicity is lower that the established risk norm (for a given impact speed); ii) for the yellow cells, the accident periodicity is equal to the established risk norm (for a given impact speed); iii) for the red cells, the accident periodicity is higher than the established risk norm, i.e., the risk norm is violated. It is worth highlighting that, even if some of the periodicity values seem very strict (i.e., accidents are seldom) whenever applied to a single vehicle, such values take a completely different importance for OEMs and fleet operators that sell or operate hundreds of thousands of vehicles per year.

E. Numerical example - adapting the driving policy

This section illustrates and exemplifies how to derive a safe driving policy by leveraging the proposed precautionary safety concepts. Ideally, an AD system is expected to be able to safely operate within a given environment and up to the specified road segment speed. Hence, we assume in the sequel that the vehicle's default speed policy is to travel at the road segment speed, and will establish additional adaptations in the form of maximum PCS driving speeds, subject to the specified QRNs. We will also discuss in which of the road segments AD is considered allowed, provided that it satisfies the prescribed QRNs. Recall that the road segment speed is considered to be defined by posted speed limits (e.g., speed sign) or environmental constraints (e.g., traffic jams).

Consider again Table V. For each of the sub-panels, the maximum safe driving speed (i.e, that satisfies the risk norms in Table I), is illustrated by the red lines and light gray cells on the driving speed column.

For urban driving segments, whenever driving in a road segment where the road segment speed is 30 km/h (left upper sub-panel), the maximum operating speed is 40 km/h, after which the QRNs are violated. Indeed, one can see that if the AD vehicle drive at 50 km/h, for example, the accident periodicity at all impact speeds is higher that the prescribed risk norm, leading for example to two minor collisions every 2000 driving hours and to high speed collisions (with a high risk of fatality) every 10.000.000 driving hours. In an analog way, whenever driving in a urban environment where the road segment speed is 50 km/h and 70 km/h (left middle- and lower sub-panels), the maximum operating speed, for a AD system with the reactive performance capabilities detailed in Table II, is 40 km/h and 50 km/h, respectively. Such an operating speed limit, largely below the road segment speed in the respective segments, is naturally an undesirable outcome, which can lead to decreased traffic flow and ultimately to traffic jams.

For highway segments (right-hand panels), the maximum operating is 60 km/h, 70 km/h and 80 km/h, all corresponding to higher values than the road segment speed for the respective road segments. However, the maximum speed is only 80 km/h for road segments where the road segment speed is 100 km/h, see the right lower sub-panel.

In order for the prescribed risk norms to be satisfied, a suitable driving policy/behavior planning needs to be defined and enforced by the decision & control system. In complex driving routines such as the one illustrated in Fig. 4b, covering both urban, suburb, and highway driving, the AD system should continuously adapt the driving policy for a given road segment. Considering the operational objective to operate within a given ODD at the road segment speed, an overview of the maximum safe driving speed for the different segments is given in Table IV. Here are also highlighted the

TABLE IV

PCS MAXIMUM SAFE SPEED VS ROAD SEGMENT SPEED AND QRN SATISFACTION TO DECIDE WHERE AD IS ALLOWED

Road segment	Urban driving			Highway driving		
speed (km/h)	Segment type	AD allowed	PCS driving policy	Segment type	AD allowed	PCS driving policy
30	Low speed segment	Yes, QRN satisfied	Max. speed = 40 km/h	Traffic jam	Yes, QRN satisfied	Max. speed = 60 km/h
50	Medium speed segment	No, blocks traffic	Max. speed = 40 km/h	Dense traffic	Yes, QRN satisfied	Max. speed = 70 km/h
70	High speed segment	No, blocks traffic	Max. speed = 50 km/h	Low density traffic	Yes, QRN satisfied	Max. speed = 80 km/h
100	-	-		Free flow traffic	No, blocks traffic	Max. speed = 80 km/h

segments for which AD is allowed given the prescribed PCS driving policy: for road segments where the speed has to be adjusted below the road segment speed, AD is not allowed in order to prevent traffic flow blockage by the AD vehicle.

Note that in this example only one of many accident types was considered, and only the nominal speed of the AD vehicle was adjusted when operating in different ODDs. Naturally, the driving policy can in addition be adapted to dynamically slow down and/or increase lateral margins to suspected jaywalkers/road users, to further reduce the incident exposure rate, thus enabling the AD vehicle to operate in a larger ODD. Nevertheless, performing dynamic adjustments with respect to suspected jaywalkers and/or when passing close to obscured areas from where they may appear is out of scope of the present paper, and will instead be considered for future work. In addition, performing similar dynamic adaptations for a dynamic perception- and vehicle control ability are also worth investigating when driving in various weather-, road- and traffic conditions.

VI. CONCLUSIONS

This paper proposes a new methodology to adjust existing driving policies for autonomous driving systems in order to ensure that challenging requirements, on low accident rates and low severity, are met. Specifically, we propose that driving policies shall be adapted with respect to their ability to detect and react to mistakes by human road users/external incidents. For that purpose, we have proposed the novel concept of Precautionary Safety and presented numerical examples supporting our methodology. By taking exposure rates to challenging situations into account, rather than hard limits based on worst case assumptions, the driving policy can be adapted to satisfy quantitative real-world safety requirements without becoming overly conservative.

The proposed approach presents two important and strong points. First, it is a add-on feature that supports the development and deployment of safe AD systems, and that can be used to monitor and improve, over time, the AD driving policies subject to the evolution of reactive performance of the emergency functionalities. Second, it incorporates perception, decision & control reactive performance, and knowledge on human mistakes, hence making it very modular and adaptive, and therefore suitable to be deployed and tuned for different operation domains, traffic interactions and deployment markets.

Future research should consider more complete driving policies, establishing adaptive longitudinal and lateral safety margins in order to increase the availability of unsupervised autonomous driving while satisfying quantitative risk norms.

REFERENCES

[1] D. L. Hendricks, M. Freedman, J. C. Fell *et al.*, “The relative frequency of unsafe driving acts in serious traffic crashes,” US National Highway Traffic Safety Administration, Tech. Rep., 2001.

[2] “Advanced driver assistance systems,” European Commission, Tech. Rep., 2018.

[3] S. Shalev-Shwartz, S. Shammah, and A. Shashua, “On a formal model of safe and scalable self-driving cars,” *arXiv:1708.06374*, 2018.

[4] C. Pek, S. Manzinger, M. Koschi, and M. Althoff, “Using online verification to prevent autonomous vehicles from causing accidents,” *Nature Machine Intelligence*, vol. 2, no. 9, 2020.

[5] Y. Shoukry, P. Nuzzo, I. Saha, A. L. Sangiovanni-Vincentelli, S. A. Seshia, G. J. Pappas, and P. Tabuada, “Scalable lazy smt-based motion planning,” in *IEEE Conference on Decision and Control*, 2016.

[6] T. Wongpiromsarn, S. Karaman, and E. Frazzoli, “Synthesis of provably correct controllers for autonomous vehicles in urban environments,” in *IEEE International Conference on Intelligent Transportation Systems*, 2011.

[7] J. Krook, A. Zita, R. Kianfar, S. Mohajerani, and M. Fabian, “Modeling and synthesis of the lane change function of an autonomous vehicle,” *IFAC Workshop on Discrete Event Systems WODES*, 2018.

[8] J. Tumova, G. Hall, S. Karaman, E. Frazzoli, and D. Rus, “Least-violating control strategy synthesis with safety rules,” in *HSCC*, 2013.

[9] A. Rizaldi and M. Althoff, “Formalising traffic rules for accountability of autonomous vehicles,” in *IEEE International Conference on Intelligent Transportation Systems*, 2015.

[10] M. Naumann, H. Konigshof, M. Lauer, and C. Stiller, “Safe but not overcautious motion planning under occlusions and limited sensor range,” in *IEEE Intelligent Vehicles Symposium*, 2019.

[11] P. Koopman and M. Wagner, “Autonomous vehicle safety: An interdisciplinary challenge,” *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 1, pp. 90–96, 2017.

[12] D. Nistér, H.-L. Lee, J. Ng, and Y. Wang, “The safety force field,” NVIDIA, Tech. Rep., 2018.

[13] M. Koschi and M. Althoff, “Set-based prediction of traffic participants considering occlusions and traffic rules,” *IEEE Transactions on Intelligent Vehicles*, 2020.

[14] S. Vaskov, S. Kousik, H. Larson, F. Bu, J. Ward, S. Worrall, M. Johnson-Roberson, and R. Vasudevan, “Towards provably not-at-fault control of autonomous robots in arbitrary dynamic environments,” *arXiv preprint:1902.02851*, 2019.

[15] F. Warg, M. Skoglund, A. Thorsén, R. Johansson, M. Brännström, M. Gyllenhammar, and M. Sanfridson, “The quantitative risk norm - a proposed tailoring of hara for ads,” in *IEEE/IFIP International Conference on Dependable Systems and Networks Workshops*, 2020.

[16] M. Lindman, I. Isaksson-Hellman, and J. Strandroth, “Basic numbers needed to understand the traffic safety effect of automated cars,” in *IRCOBI Conference*, 2017.

[17] N. Webb, D. Smith, C. Ludwick, T. Victor, Q. Hommes, F. Favaro, G. Ivanov, and T. Daniel, “Waymo’s safety methodologies and safety readiness determinations,” 2020.

[18] “Waymo safety report,” Waymo, Tech. Rep., 2021.

[19] “Road vehicles — safety of the intended functionality,” *ISO/PAS 21448:2019*.

[20] “Road vehicles – functional safety,” *ISO 26262-1:2018*.

[21] “Road vehicles – safety and cybersecurity for automated driving – design, verification and validation,” *ISO/NP TS 5083*.

[22] M. Peden and et al., “World report on road traffic injury prevention,” World Health Organization, Tech. Rep., 2004.

[23] “Annual accident report,” European Commission, Directorate General for Transport, Tech. Rep., 2018.

[24] M. Brännström, E. Coelingh, and J. Sjöberg, “Model-based threat assessment for avoiding arbitrary vehicle collisions,” *IEEE Trans. on Intelligent Transportation Systems*, vol. 1, no. 3, pp. 658–669, 2010.

TABLE V

Accident periodicity for the AD system defined according to the incident exposure rates established in Table III and the Quantitative Risk Norms defined in Table I. AD can be allowed on road segments where the maximum safe speed (red line) exceeds the road segment speed (red circle), otherwise the AD vehicle will block the traffic flow to drive safely.

Accident periodicity

[hours between accidents with jaywalkers]

	Urban driving						**Highway driving**					
Road segment speed: 30 km/h	**100 h between incidents with jaywalkers**						**100.000 h between incidents with jaywalkers**					
	EV speed (km/h)	Impact speed [km/h], periodicity [h]					EV speed (km/h)	Impact speed [km/h], periodicity [h]				
		≤ 10	[10-20]	[20-30]	[30-40]	> 40		≤ 10	[10-20]	[20-30]	[30-40]	> 40
	30	10^{6}	10^{7}	10^{8}	10^{9}	∞	30	10^{9}	10^{10}	10^{11}	10^{12}	∞
	40	10^{5}	10^{6}	10^{7}	10^{8}	∞	40	10^{8}	10^{9}	10^{10}	10^{11}	∞
	50	$2x10^{3}$	10^{5}	10^{6}	10^{7}	10^{8}	50	$2x10^{6}$	10^{8}	10^{9}	10^{10}	10^{11}
	60	$2x10^{3}$	$2x10^{3}$	10^{5}	10^{6}	10^{7}	60	$2x10^{6}$	$2x10^{6}$	10^{8}	10^{9}	10^{10}
	70	10^{3}	$2x10^{3}$	$2x10^{3}$	10^{5}	10^{6}	70	10^{6}	$2x10^{6}$	$2x10^{6}$	10^{8}	10^{9}
	80	10^{3}	10^{3}	$2x10^{3}$	$2x10^{3}$	10^{5}	80	10^{6}	10^{6}	$2x10^{6}$	$2x10^{6}$	10^{8}
	90	10^{3}	10^{3}	10^{3}	$2x10^{3}$	$2x10^{3}$	90	10^{6}	10^{6}	10^{6}	$2x10^{6}$	$2x10^{6}$
	100	10^{3}	10^{3}	10^{3}	10^{3}	10^{3}	100	10^{6}	10^{6}	10^{6}	10^{6}	10^{6}
Road segment speed: 50 km/h	**1.000 h between incidents with jaywalkers**						**1.000.000 h between incidents with jaywalkers**					
	EV speed (km/h)	Impact speed [km/h], periodicity [h]					EV speed (km/h)	Impact speed [km/h], periodicity [h]				
		≤ 10	[10-20]	[20-30]	[30-40]	> 40		≤ 10	[10-20]	[20-30]	[30-40]	> 40
	30	10^{7}	10^{8}	10^{9}	10^{10}	∞	30	10^{10}	10^{11}	10^{12}	10^{13}	∞
	40	10^{6}	10^{7}	10^{8}	10^{9}	∞	40	10^{9}	10^{10}	10^{11}	10^{12}	∞
	50	$2x10^{4}$	10^{6}	10^{7}	10^{8}	10^{9}	50	$2x10^{7}$	10^{9}	10^{10}	10^{11}	10^{12}
	60	$2x10^{4}$	$2x10^{4}$	10^{6}	10^{7}	10^{8}	60	$2x10^{7}$	$2x10^{7}$	10^{9}	10^{10}	10^{11}
	70	10^{4}	$2x10^{4}$	$2x10^{4}$	10^{6}	10^{7}	70	10^{7}	$2x10^{7}$	$2x10^{7}$	10^{9}	10^{10}
	80	10^{4}	10^{4}	$2x10^{4}$	$2x10^{4}$	10^{6}	80	10^{7}	10^{7}	$2x10^{7}$	$2x10^{7}$	10^{9}
	90	10^{4}	10^{4}	10^{4}	$2x10^{4}$	$2x10^{4}$	90	10^{7}	10^{7}	10^{7}	$2x10^{7}$	$2x10^{7}$
	100	10^{4}	10^{4}	10^{4}	10^{4}	10^{4}	100	10^{7}	10^{7}	10^{7}	10^{7}	10^{7}
Road segment speed: 70 km/h	**10.000 h between incidents with jaywalkers**						**10.000.000 h between incidents with jaywalkers**					
	EV speed (km/h)	Impact speed [km/h], periodicity [h]					EV speed (km/h)	Impact speed [km/h], periodicity [h]				
		≤ 10	[10-20]	[20-30]	[30-40]	> 40		≤ 10	[10-20]	[20-30]	[30-40]	> 40
	30	10^{8}	10^{9}	10^{10}	10^{11}	∞	30	10^{11}	10^{12}	10^{13}	10^{14}	∞
	40	10^{7}	10^{8}	10^{9}	10^{10}	∞	40	10^{10}	10^{11}	10^{12}	10^{13}	∞
	50	$2x10^{5}$	10^{7}	10^{8}	10^{9}	10^{10}	50	$2x10^{8}$	10^{10}	10^{11}	10^{12}	10^{13}
	60	$2x10^{5}$	$2x10^{5}$	10^{7}	10^{8}	10^{9}	60	$2x10^{8}$	$2x10^{8}$	10^{10}	10^{11}	10^{12}
	70	10^{5}	$2x10^{5}$	$2x10^{5}$	10^{7}	10^{8}	70	10^{8}	$2x10^{8}$	$2x10^{8}$	10^{10}	10^{11}
	80	10^{5}	10^{5}	$2x10^{5}$	$2x10^{5}$	10^{7}	80	10^{8}	10^{8}	$2x10^{8}$	$2x10^{8}$	10^{10}
	90	10^{5}	10^{5}	10^{5}	$2x10^{5}$	$2x10^{5}$	90	10^{8}	10^{8}	10^{8}	$2x10^{8}$	$2x10^{8}$
	100	10^{5}	10^{5}	10^{5}	10^{5}	10^{5}	100	10^{8}	10^{8}	10^{8}	10^{8}	10^{8}
Road segment speed: 100 km/h							**10.000.000 h between incidents with jaywalkers**					
							EV speed (km/h)	Impact speed [km/h], periodicity [h]				
								≤ 10	[10-20]	[20-30]	[30-40]	> 40
							30	10^{11}	10^{12}	10^{13}	10^{14}	∞
							40	10^{10}	10^{11}	10^{12}	10^{13}	∞
							50	$2x10^{8}$	10^{10}	10^{11}	10^{12}	10^{13}
							60	$2x10^{8}$	$2x10^{8}$	10^{10}	10^{11}	10^{12}
							70	10^{8}	$2x10^{8}$	$2x10^{8}$	10^{10}	10^{11}
							80	10^{8}	10^{8}	$2x10^{8}$	$2x10^{8}$	10^{10}
							90	10^{8}	10^{8}	10^{8}	$2x10^{8}$	$2x10^{8}$
							100	10^{8}	10^{8}	10^{8}	10^{8}	10^{8}

🟩 Accident periodicity **lower** than QRN, see Table I

🟨 Accident periodicity **equal** to QRN, see Table I

🟥 Accident periodicity **higher** than QRN, see Table I

⬜ Safe driving speed satisfying QRN