Highway Systems and Transport Engineering

Stephane Werner

Makenzie Mccutcheon

First Edition, 2012

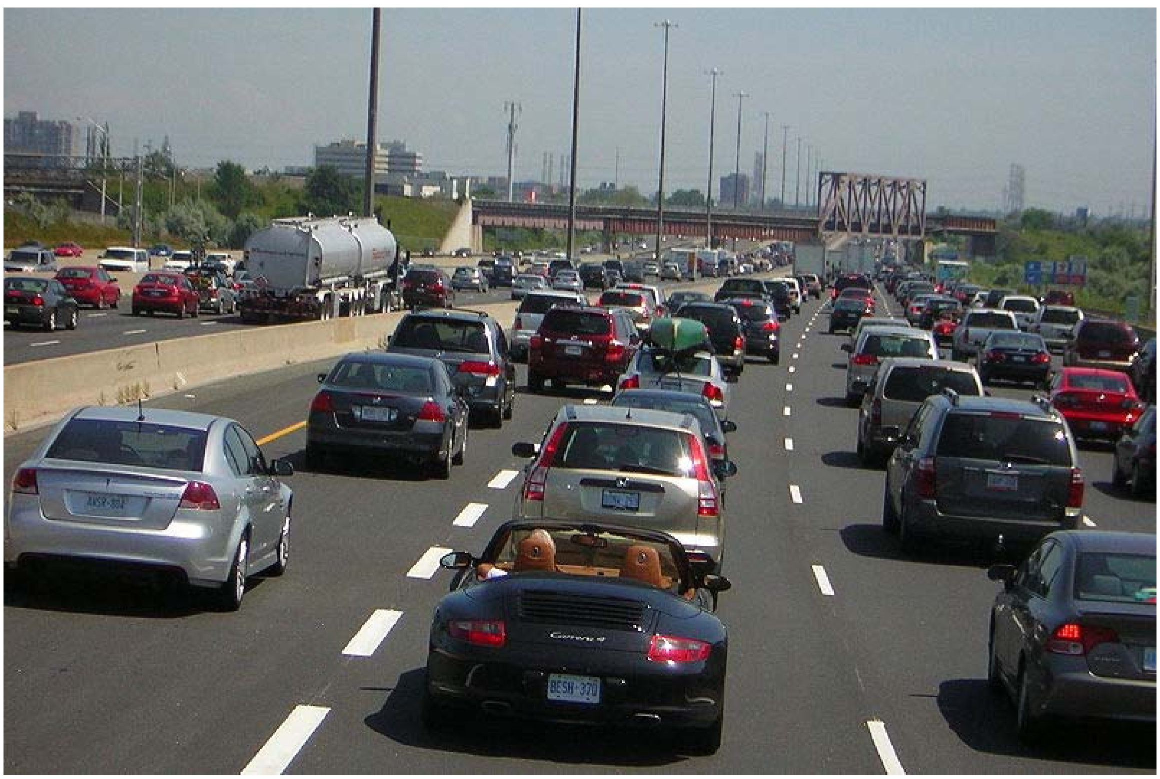
ISBN 978-81-323-0936-9

Published by:
Academic Studio
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

Chapter 1

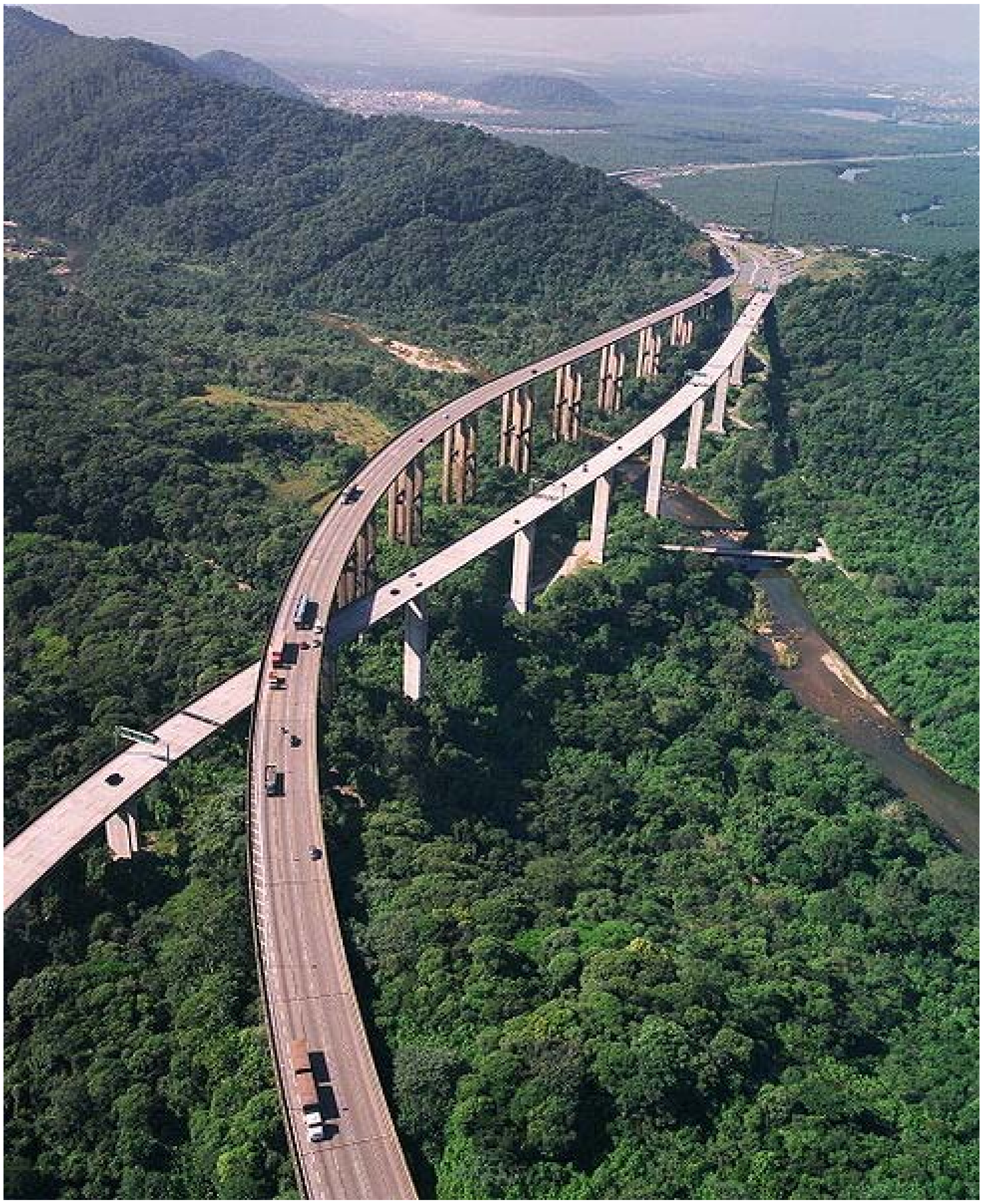
Highway



Highway 401, the busiest highway in North America.

A German Autobahn in Lehrte.

The Makran Coastal Highway was an ancient road within Pakistan. Now it's a major road leading to the city of Gwadar



The SP-160, known as Rodovia dos Imigrantes, in southeastern Brazil.

A **highway** is a public road, especially a major road connecting two or more destinations. Any interconnected set of highways can be variously referred to as a "highway system", a "highway network", or a "highway transportation system". Each country has its own national highway system.

Overview

Major highways are often named and numbered by the governments that typically develop and maintain them. Australia's Highway 1 is the longest national highway in the world at over 14,500 km (9,000 miles) and runs almost the entire way around the continent. The United States has the world's largest network of highways, including both the Interstate Highway System and the U.S. Highway System. At least one of these networks is present in every state and they interconnect most major cities. Some highways, like the Pan-American Highway or the European routes, span multiple countries. Some major highway routes include ferry services, such as U.S. Route 10, which crosses Lake Michigan.

Traditionally highways were used by people on foot or on horses. Later they also accommodated carriages, bicycles and eventually motor cars, facilitated by advancements in road construction. In the 1920s and 1930s many nations began investing heavily in progressively more modern highway systems to spur commerce and bolster national defense.

Major modern highways that connect cities in populous developed and developing countries usually incorporate features intended to enhance the road's capacity, efficiency, and safety to various degrees. Such features include a reduction in the number of locations for user access, the use of dual carriageways with two or more lanes on each carriageway, and grade-separated junctions with other roads and modes of transport. These features are typically present on highways built as *motorways* (*freeways*).

Terminology

In English law, parliament and more formal situations the term is used to denote *any* public road used which include streets and lanes as well as main roads, trunk roads and motorways. Acts of parliament have used the term throughout history from the Highways Act 1555 through to the Highways Act 1980. The rules of the road are outlined in the Highway Code.

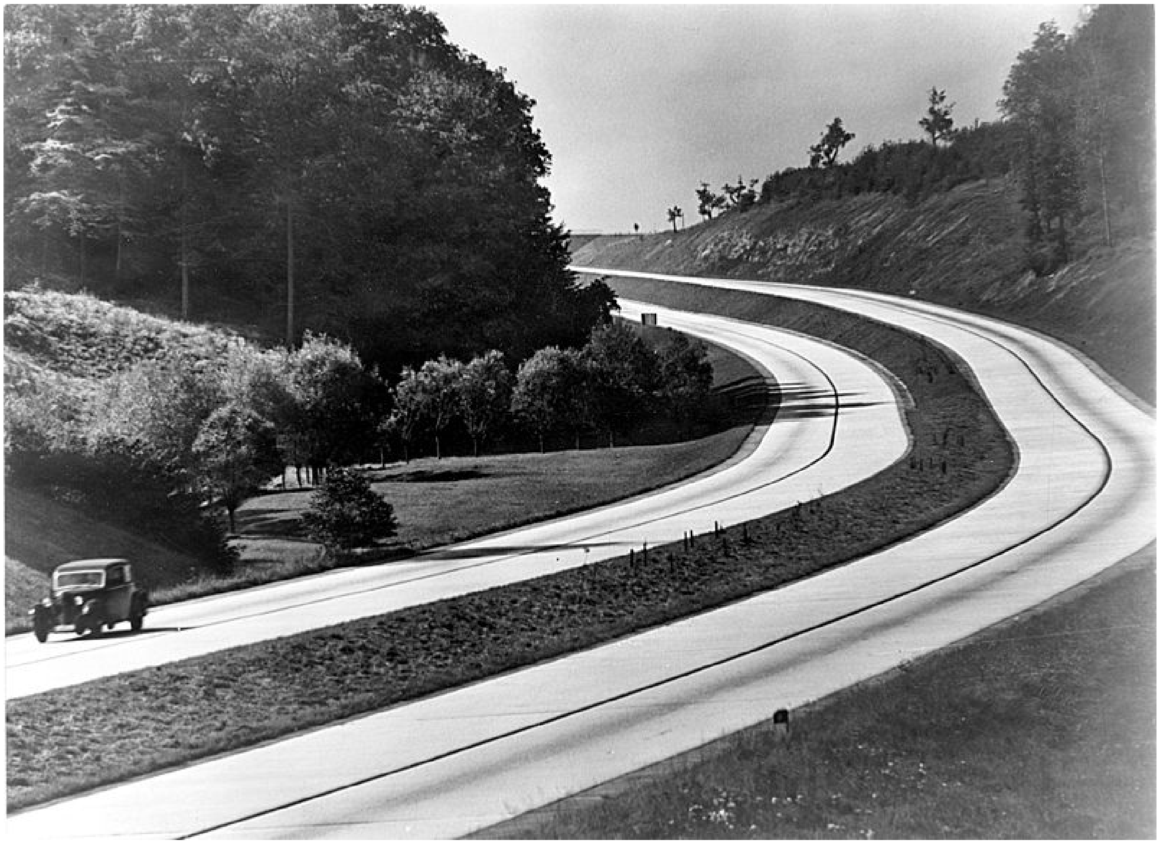
In England and Wales, a "Public Highway" is a road or footpath over which the public has the right of access, i.e. the opposite of a "private road".

In American law, the word "highway" is sometimes used to denote any public way used for travel, whether major highway, freeway, turnpike, street, lane, alley, pathway, dirt track, footpaths, and trails, and navigable waterways; however, in practical and useful meaning, a "highway" is a major and significant, well-constructed road that is capable of carrying reasonably-heavy to extremely-heavy traffic. Highways generally have a route number designated by the state and federal road comptroller offices.

California Vehicle Code, Sections 360, 590, define a "highway" as only a way open for use of motor vehicles, but the California Supreme Court has held that "the definition of 'highway' in the Vehicle Code is used for special purposes of that act," and that canals in the town of Venice, California, are "highways" that are entitled to be maintained with state highway funds.

Smaller roads may be termed byways.

History



A German autobahn in the 1930s

Modern highway systems developed in the 20th century as the automobile gained popularity. The world's first limited access road was constructed in Italy in 1922. Construction of the Bonn-Cologne autobahn began in 1929 and was opened in 1932 by the mayor of Cologne.

The Special Roads Act 1949 in the United Kingdom provided the legislative basis for roads for restricted classes of vehicles (later termed motorway). The first section of motorway in the UK opened in 1958 (part of the M6 motorway) and then in 1959 the first section of the M1 motorway.

The Federal Aid Highway Act of 1956 provided appropriating $25 billion for the construction of 41,000 miles (66,000 km) of Interstate Highways over a 20-year period in the United States.

Social effects

Reducing travel times relative to city or town streets, modern highways with limited access and grade separation create increased opportunities for people to travel for business, trade or pleasure and also provide trade routes for goods. Modern highways reduce commute and other travel time but additional road capacity can also create new induced traffic demand. If not accurately predicted at the planning stage, this extra traffic may lead to the new road becoming congested sooner than anticipated. More roads add on to car-dependence, which can mean that a new road brings only short-term mitigation of traffic congestion.

Where highways are created through existing communities, there can be reduced community cohesion and more difficult local access. Consequently property values have decreased in many cutoff neighborhoods, leading to decreased housing quality over time.

Economic effects

In transport, demand can be measured in numbers of journeys made or in total distance travelled across all journeys (e.g. passenger-kilometres for public transport or vehicle-kilometres of travel (VKT) for private transport). Supply is considered to be a measure of capacity. The price of the good (travel) is measured using the generalised cost of travel, which includes both money and time expenditure.

The effect of increases in supply (capacity) are of particular interest in transport economics, as the potential environmental consequences are significant.

In addition to providing benefits to their users, transport networks impose both positive and negative externalities on non-users. The consideration of these externalities - particularly the negative ones - is a part of transport economics. Positive externalities of transport networks may include the ability to provide emergency services, increases in land value and agglomeration benefits. Negative externalities are wide-ranging and may include local air pollution, noise pollution, light pollution, safety hazards, community severance and congestion. The contribution of transport systems to potentially hazardous climate change is a significant negative externality which is difficult to evaluate quantitatively, making it difficult (but not impossible) to include in transport economics-based research and analysis. Congestion is considered a negative externality by economists.

Environment effects

Highways are extended linear sources of pollution:

Roadway noise increases with operating speed so major highways generate more noise than arterial streets. Therefore, considerable noise health effects are expected from highway systems. Noise mitigation strategies exist to reduce sound levels at nearby sensitive receptors. The idea that highway design could be influenced by acoustical engineering considerations first arose about 1973.

Air quality issues: Highways may contribute fewer emissions than arterials carrying the same vehicle volumes. This is because high, constant-speed operation creates an emissions reduction compared to vehicular flows with stops and starts. However, concentrations of air pollutants near highways may be higher due to increased traffic volumes. Therefore, the risk of exposure to elevated levels of air pollutants from a highway may be considerable, and further magnified when highways have traffic congestion.

New highways can also cause habitat fragmentation, encourage urban sprawl and allow human intrusion into previously untouched areas, as well as (counterintuitively) increasing congestion, by increasing the number of intersections. They can also reduce the use of public transport, indirectly leading to greater pollution.

High-occupancy vehicle lanes are being added to some newer/reconstructed highways in North America and other countries around the world to encourage carpooling and mass-transit. These lanes help reduce the number of cars on the highway and thus reduces pollution and traffic congestion by promoting the use of carpooling in order to be able to use these lanes. However, they tend to require dedicated lanes on a highway, which makes them difficult to construct in dense urban areas where they are the most effective.

Road traffic safety

Road traffic safety aims to reduce the harm (deaths, injuries, and property damage) on the highway system from traffic collisions and includes the design, construction and regulation of the roads, the vehicles that use them and also the training of drivers and other road-users. Improvement of road safety needs to be balanced with the provision of an effective efficient transport system. A report published by the World Health Organization in 2004 estimated that some 1.2m people were killed and 50m injured on the roads around the world each year and was the leading cause of death among children 10 – 19 years of age. The report also noted that the problem was most severe in developing countries and that simple prevention measures could halve the number of deaths. For reasons of clear data collection, only harm involving a road vehicle is included. A person tripping with fatal consequences or dying for some unrelated reason on a public road is not included in the relevant statistics.

Statistics

International sign used widely in Europe denoting the start of special restrictions for a section of highway.

The United States has the world's largest network of highways, including both the Interstate Highway System and the U.S. Highway System. At least one of these networks is present in every state and they interconnect most major cities.

China's highway network is the second most extensive in the world, with a total length of about 3.573 million km. China's expressway network is also the second longest in the world, and it is quickly expanding, stretching some 60,300 km at the end of 2008, In 2008 alone, 6,433 km expressways were added to the network.

- **Longest international highway:** the Pan-American Highway, which connects many countries in the Americas, is nearly 25,000 kilometres (15,534 mi) long as of 2005. The Pan-American Highway is discontinuous because there is a significant gap in it in southeastern Panama, where the rainfall is immense and the terrain is entirely unsuitable for highway construction.

- **Longest national highway (point to point):** The Trans-Canada Highway is 7,821 km (4,857 mi) long as of 2006. The T.C.H. runs east-west across southern Canada, the populated portion of the country, and it connects many of the major urban centers along its route crossing almost all of the provinces, and reaching almost all of the capital cities. The T.C.H. begins on the east coast in Newfoundland, traverses that island, and crosses to the mainland by ferry. It reaches most of the Maritime Provinces of eastern Canada, and a side route using ferries traverses the province of Prince Edward Island. After crossing the two most populous provinces of Quebec and Ontario, the T.C.H. continues westward across Manitoba, Saskatchewan, Alberta, and British Columbia. After reaching Vancouver, B.C., on the Pacific Coast, there is a ferry route west to Vancouver Island and the provincial capital city of Victoria, B.C.

- **Longest national highway (circuit):** Australia's Highway 1 at over 20,000 km (12,427 mi). It runs almost the entire way around the continent's coastline. With the exception of the Federal Capital of Canberra, which is far inland, Highway 1 links all of Australia's capital cities, although Brisbane and Darwin are not directly connected, but rather are bypassed short distances away. Also, there is a ferry connection to the island state of Tasmania, and then a stretch of Highway 1 that links the major towns and cities of Tasmania, including Launceston and Hobart (this state's capital city).

- **Largest national highway system:** The United States of America has approximately 6,430,366 kilometres (3,995,644 mi) of highway within its borders as of 2008.

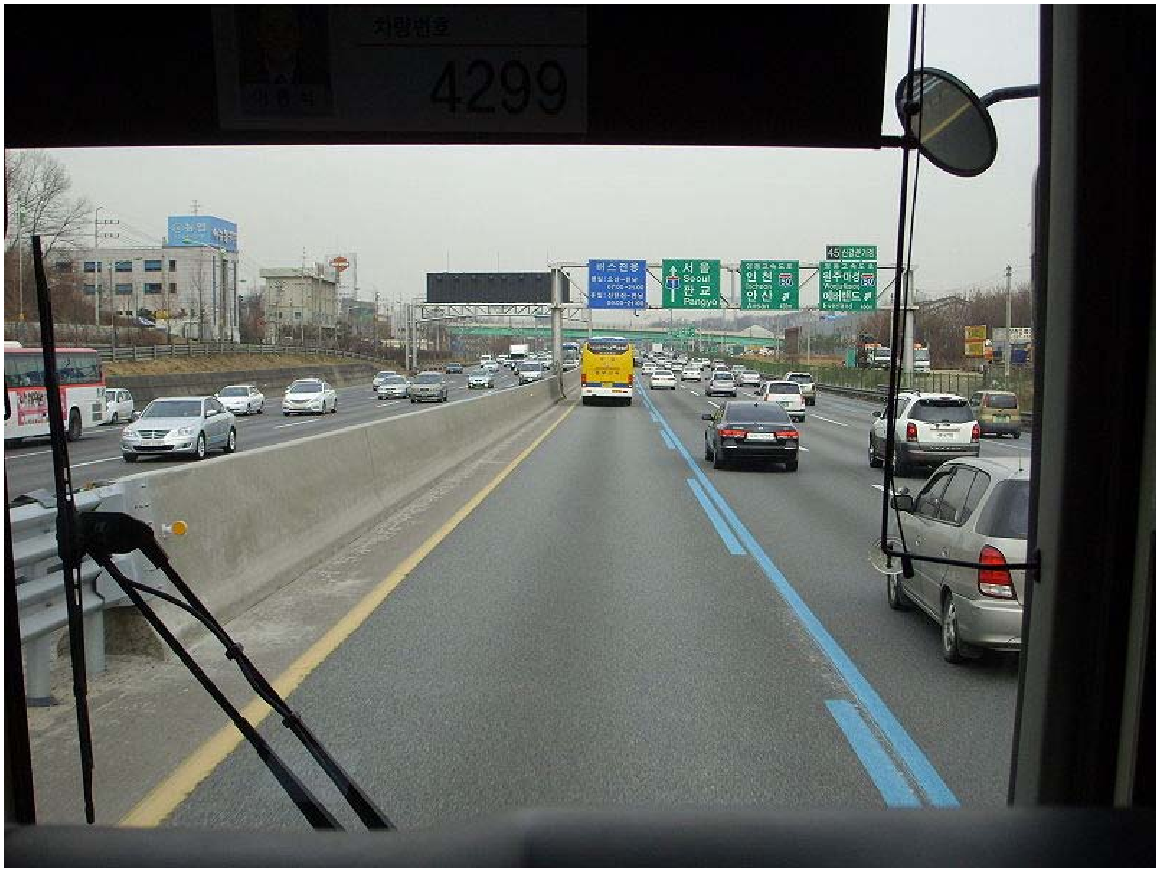
- **Busiest highway:** Highway 401 in Ontario, Canada, has volumes surpassing an average of 500,000 vehicles per day in some sections of Toronto as of 2006.

- **Widest highway (maximum number of lanes):** The Katy Freeway (part of Interstate 10) in Houston, Texas, has a total of 26 lanes in some sections as of 2007. However, they are divided up into general use/ frontage roads/ HOV lanes, restricting the traverse traffic flow.

- **Widest highway (maximum number of through lanes):** Interstate 5 along a 2-mile section between Interstate 805 and California State Route 56 in San Diego, California, which was completed in April 2007, is 22 lanes wide.

- **Highest international highway**: The Karakoram Highway, between Pakistan and China, is at an altitude of 4,693 m/15,397 ft.

Bus lane



Highway bus lane on Gyeongbu Expressway in South Korea

Some countries incorporate bus lanes onto highways.

Country	Highway	Bus lanes (km)	Section
Canada	Ontario Highway 417	7	Eagleson Road – Ontario Highway 417 (Ottawa)
Canada	Ontario Highway 403	6	Mavis Road – Winston Churchiil Blvd. (Mississauga)
South Korea	Gyeongbu Expressway	137.4	Hannam IC(Seoul) ~ Sintanjin IC(Daejeon)

Korea

In South Korea, in February 1995 - Bus lane (essentially an HOV-9) established between the northern terminus and Sintanjin for important holidays and in 1 July 2008 - Bus lane enforcement between Seoul and Osan (Sintanjin on weekends) becomes daily between 6

AM and 10 PM. On 1 October this is adjusted to 7 AM to 9 PM weekdays, 9 AM to 9 PM weekends.

Highway 401 in London, Ontario

A Polish expressway in Bielsko-Biała

The Pan-American Highway where it serves as the main street in Máncora, Peru

The Pan-American Highway in the Greater Buenos Aires (city of Florida), Argentina

Highway A1 near Bologna, Italy with 10 lanes

The Dr. Sun Yat-sen Memorial Freeway in Taipei, Taiwan

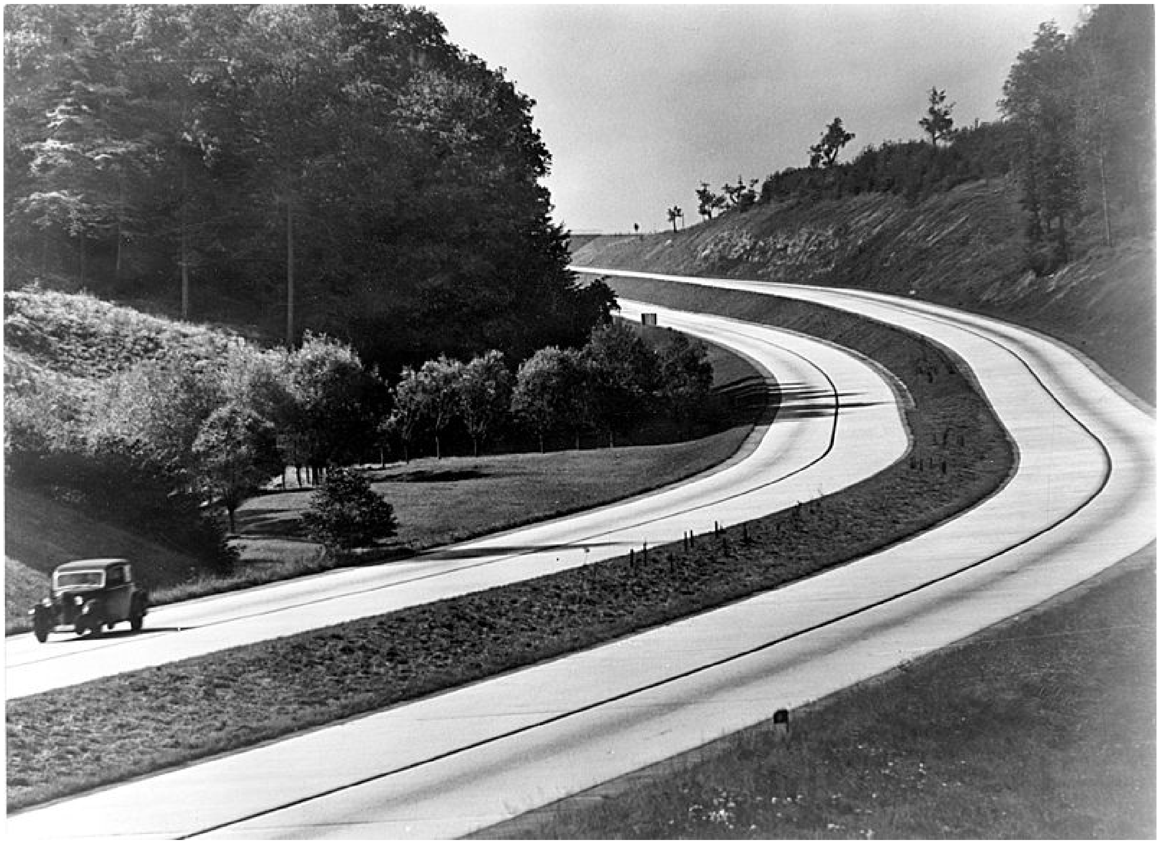
Chapter 2

Dual Carriageway

A typical British dual carriageway with central barrier on the A63(T) near Hull, England.

Ohio State Route 11 in Ohio

Clara Barton Parkway outside Washington, D.C.



A German dual carriageway in the 1930s

A **dual carriageway** (**divided highway**) is a highway in which the two directions of traffic are separated by a central barrier or strip of land, known as a central reservation (median). It may also have limited access and grade separated junctions. Where more than one lane is provided in each direction this type of road is usually able to carry a great deal more traffic than single carriageways (undivided highways). Each carriageway or roadway usually has at least two lanes for traffic. Dual carriageways generally have lower accident rates than single carriageways (undivided highways) due to the separation of traffic moving in opposing directions.

History

A very early example (perhaps the first) of a dual carriageway was the *Via Portuensis*, built in the 1st century by the Roman emperor Claudius between Rome and its port Ostia at the mouth of the Tiber.

In 1907 the Long Island Parkway opened and roughly 20% of it featured a semi-dual carriageway design. The New York City parkway system, which was built between 1907 and 1934, also pioneered the same design. However the majority of it featured concrete or brick railings as lane dividers as opposed to using grass medians.

In 1924 the first Italian autostrada was opened running 55 km (34 mi) from Milan to Varese. It featured a broad road bed and did not feature lane dividers except near cities and through the mountains.

The London end of the Great West Road became Britain's first dual carriageway when it was opened in 1925 by King George V.

In 1927 the Rome bypass was opened. It ran 92 km (57 mi) bypassing Rome to the east. Almost the entire length featured a dual carriageway design. In the early 1930s it was extended southward all the way to Naples and northward to Florence. Most of the original routing was destroyed by the Allies in the Second World War.

By 1930 several American and European cities had built dual carriageway highways mostly to control traffic jams and/or to provide bypass routes for traffic.

In 1932 the first German Autobahn opened between Cologne and Bonn. It ran 21 km (13 mi) and became a precedent for future highways. Although it, like the first Autostrada, did not feature a dual carriageway design, it inspired the mass construction of future high speed roadways.

During the 1930s, Germany, Italy, and the Soviet Union began construction of a network of dual carriageway expressways. By 1942, Germany had over 3,200 km (2,000 mi) of dual carriageway roads, Italy had nearly 1,300 km (800 mi), and the Soviet Union had 400 km (250 mi).

What may be the world's first long-distance intercity dual carriageway/freeway is the Queen Elizabeth Way in Southern Ontario in Canada, initially linking the large cities of Toronto and Hamilton together by 1939, with construction on this stretch of the present-day Queen Elizabeth Way beginning in 1936 as "Middle Road".

Opened to traffic in 1940, the 160 mile (257 km) long Pennsylvania Turnpike was the first rural dual carriageway built in the United States. By 1955 several states had built dual carriageway freeways and turnpikes and in 1957 the Interstate Highway System began. Completed in 1994, the major highway system links all the major cities of the United States.

European implementations

United Kingdom

In the UK, although the term *dual carriageway* applies to any road with physically separated lanes, it is frequently used as a descriptive term for major routes built in this style. Such major dual carriageways usually have two lanes of traffic in each direction, with the lane nearest the centre being reserved for overtaking. Occasionally dual carriageways have only one lane in each direction, or more than two lanes each way (usually to permit easier overtaking of slower uphill traffic). Different speed limits apply on dual carriageway sections from those that apply on single carriageway sections of the same class of road, except in cities and built-up areas where the dual carriageway is more of a safety measure, often intended to prevent pedestrians from crossing a busy road.

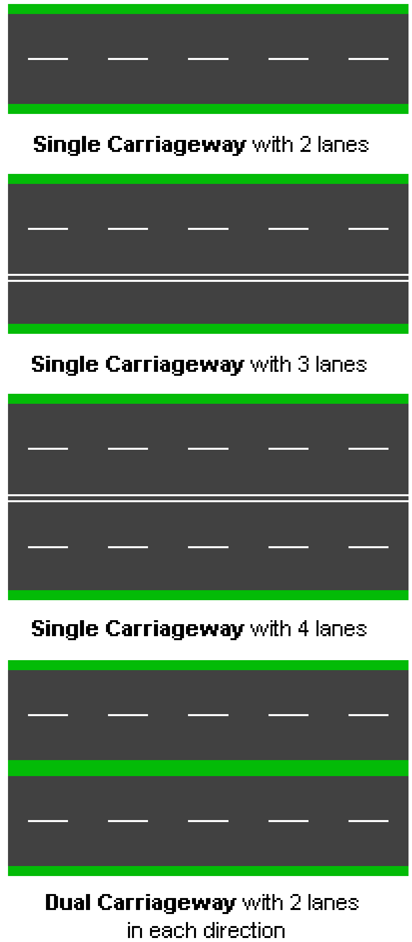


Diagram of types of road in the UK

A sign informing motorists of an oncoming Dual Carriageway in Seacroft, Leeds.

When first constructed, many dual carriageways - including the first motorways - had no crash or other barriers in the central reservation. Hence in the event of delays on the road, or if a driver missed his exit, there was a widespread problem of drivers making a U-turn onto the other carriageway; many accidents were caused as a result of their misjudging the speed of approaching traffic on the other carriageway when doing so. The majority of dual carriageway roads now have barriers. Some are heavy concrete obstructions which can have the effect of bouncing a vehicle back into the path of other traffic; others are made from steel ropes mounted on moderately weak posts, where the rope cuts into the vehicle body to slow the vehicle while keeping it against the barrier until it has stopped. Often on urban dual carriageways where the road has been converted from a four-lane single carriageway the central reservation will be nominal and often just a small steel divider to save space.

Turning right (that is, across the line of traffic heading in the opposite direction) is usually permitted only at specific locations. Often the driver will be required to turn left (away from the dual carriageway) in order to loop around to an access road that permits crossing the major road. Roundabouts on dual carriageways are relatively common, especially in cities or where the cost of a grade-separated junction would be prohibitive. Where space is even more limited, intersections may be controlled by traffic lights. Smaller residential roads adjoining urban dual carriageways may be blocked off at one end in an attempt to limit the number of junctions on a dual carriageway, often other roads may pass over or under the dual-carriageway without an intersection.

A dual carriageway with grade-separated junctions and which meets other requirements may be upgraded to motorway standard, denoted as an **(M)** added after the road number (e.g. "A1(M)" or "A38(M)"). Unlike in the Republic of Ireland there is no official terminology for 'high-quality dual carriageways', however many roads such as the A1, the A14, the A19 and the A42 are built to a high quality, in many places they are only intersected by grade-separated junctions, have full barriers at both the road side and the central reservations and in some cases three lanes of traffic, however for at least one reason they fall short of motorway standard (often this may be down to the height of overpasses or the quality of intersecting junctions.

Confusion

While most drivers are clear about what a motorway is, some are confused about the definition of a dual carriageway. For a road to be classed as a dual carriageway, the two directions of traffic flow must be physically separated by a central reservation. A road where the two directions of flow are separated only by lines painted on the road surface is a single carriageway, regardless of the number of traffic lanes that may be available to the traffic in each direction. So a road with three or four lanes is not a dual carriageway if there is no central reservation. Increasingly many such roads are being converted into dual carriageways with the erection of a central barrier.

Speed limits

The national speed limit applies on dual carriageways (unless it is in a 'built-up area', or a lower limit is posted), which is as follows:

National speed limits on dual carriageways in the UK

Type of vehicle	Speed limit
Car, motorcycle or a car-based van up to 2 metric tonnes	70 mph (110 km/h)
Car with caravan or trailer	60 mph (97 km/h)
Bus or coach up to 12 m long	60 mph (97 km/h)

Goods vehicle up to 7.5 t	60 mph (97 km/h)
Goods vehicle over 7.5 t	50 mph (80 km/h)

A dual carriageway in a built up area will have a statutory speed limit of 30 mph (48 km/h) unless otherwise sign-posted. It is common for such urban dual carriageways to have an increased speed limit of 40 mph (64 km/h). A built up road is indicated by the presence of street lights, on lit dual carriageways that are not considered to be in a built-up area, the speed limit will be clarified with intermittent signs.

Ireland

A typical modern Irish dual carriageway (opened 2004) along the N11, south of Newtownmountkennedy. On motorways, the yellow hard shoulder markings are unbroken.

An example of a 2+2 dual-carriageway in Ireland. This type is similar to many found in the UK.

Although in the Republic of Ireland the term dual carriageway technically applies to any road with physically separated lanes, it is usually used only to refer to those route sections that do not have a motorway designation. Most often it is *national roads* (roads with a route number prefix of **N**; e.g. N8) that are built as or upgraded to dual carriageway. A number of non-national roads (for example, *regional roads*) are dual carriageway, for example in urban areas near or in cities, or where the road was formerly part of a national route.

Dual carriageways of this class differ from motorways in a number of ways. The hard shoulder is demarcated with a dashed yellow line (as opposed to an unbroken yellow line on motorways). The standard speed limit of 100 km/h (62 mph) for national routes usually applies (by default the limit is 80 km/h (50 mph) for non-national roads, even if dual carriageway). Local authorities have the power to apply a limit of up to 120 km/h (75 mph) as used on most motorways (The High Quality Dual Carriageway section of the N1 between the end of the M1 and the border with Northern Ireland and the N25/N22 Ballincollig Bypass in Cork are the only route sections with such special limits). Traffic lights and junctions are permitted at grade on dual carriageways. For older sections of dual carriageway, this has resulted in fewer flyover junctions. Newer dual carriageway sections are usually near motorway standard, with grade-separated junctions, but may not be designated as motorways due to the need to preserve access to adjoining property or to the absence of a non-motorway alternative route. Also, dual carriageways that are not motorway classified do not need to be equipped with emergency phones.

Motorway restrictions only apply to motorway sections, rather than all dual carriageway sections of national roads (these are signposted with the **N** prefix on the route number, rather than **M**). Some *national secondary roads*, and *regional roads* in particular often have houses, schools and other developments fronting on to them. Less important *national primary roads*, and older sections not yet upgraded may also feature such developments built before the introduction the Irish Planning system in 1964. Today Irish planning policy prohibits such development on National Primary or National Secondary roads where the speed limit exceeds 60 km/h (37 mph). This policy results from concerns expressed by the National Roads Authority. However, a local authority is not obliged to implement this policy and can disregard this policy at it own discretion. This would usually only occur in exceptional circumstances or where planners are over ruled by elected councillors using section 140 of the Local Government Act 2001. Accordingly, hard shoulders are included wherever feasible to provide for the resulting pedestrian and cyclist traffic, and are present on much of the national route network. These hard shoulders may also be used as running lanes by motorised traffic under certain conditions.

Until 2004/2005, many motorways and dual carriageways in Ireland did not have crash barriers in the central reservation, the policy being to use a wider median instead. Crash barriers are now mandatory for such routes, and wire cabling or full crash barriers (depending on whether or not the route is a motorway, and median width) have been fitted to existing routes.

As of 2008 three major types of dual carriageway are being built on national road schemes in the Republic of Ireland:

- High Quality Dual Carriageways (HQDC) - these are being built mainly on the major inter-urban routes, to full motorway standard but without motorway regulations. The Roads Act 2007 allows for these roads to be redesignated as motorways by ministerial order. Many of the sections of HQDC on the major inter-urban routes have been redesignated as motorways and full motorway regulations will apply when the redesignations come into effect.
- Standard dual carriageway of the traditional type is mainly planned for schemes on the N11 road, the N18 road and the N25 road. Plans for this type of dual carriageway on the N20 road have been superseded by newer plans to build a motorway, the M20, to replace most of this route. Traditionally this type of dual-carriageway had a mixture of at grade junctions (including roundabouts), grade separated junctions, and median crossings. Nowadays they are similar to HQDCs, but minor at grade exits - generally left turn only - are allowed and the design speed (by Irish standards) is only 100 km/h. Median crossings and roundabouts are no longer generally found on these schemes. An example of a standard dual carriageway scheme, opened in 2006, is the Ennis bypass although this road has grade separated junctions and no median crossings. This route has now been upgraded to motorway status.
- 2+2 roads - officially these roads are designated as *Type 2* dual carriageways by the National Roads Authority (NRA). They will be created by widening existing

roads or building new roads, and will have two lanes in each direction with a steel cable barrier in the middle but no hard shoulder. Most junctions will be at grade. With the exception of the restricted median width and the lack of lay-bys, this type of dual carriageway is similar to many dual carriageways found in the UK. The first 2+2 scheme (and the only example as of 2008), the N4 Dromod Roosky bypass, opened on 7 December 2007.

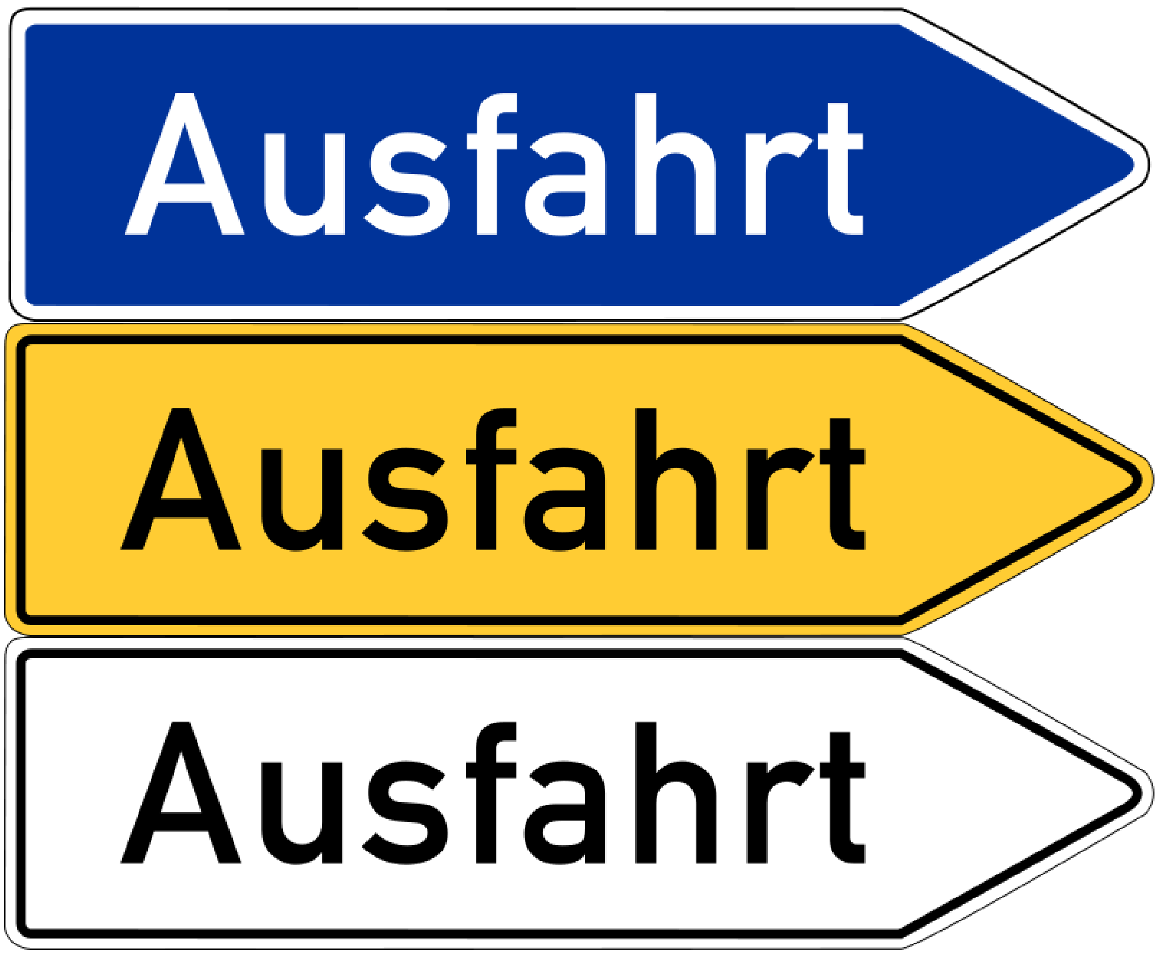
- 2+1 roads - officially these roads are designated as *Type 3* dual carriageways by the NRA. They have two lanes in one direction and one lane in the other, alternating every few kilometres, and usually separated with a steel cable barrier. Sections of 2+1 road have been built on the N20 and the N2. In July 2007, the NRA announced that it would no longer build 2+1 roads and 2+2 roads will be built instead.

Germany

B6(M) near Abbenrode

In Germany the term "Schnellstraße" (expressway) refers to streets that allow for higher speed traffic than is common on other streets which in turn requires them to have a dual carriageway in most cases. A special exception is the 2+1 road system in some rural areas that are also referred to as expressways (three lanes on a single carriageway where bypassing on the lane of the opposite direction is prohibited so that speed restriction is not required to increase safety - they are generally built to a 100 or 120 km/h traffic standard). This is based on the definition of expressways as having an increased amount of grade-separated sections. Technically a dual carriageway with central reservation is the legal foundation that no blank speed limit exists - comparable to the German Autobahn - but expressways are factually always given speed limit signs.

Expressways however can be upgraded to motorway standard which is colloquially referred to as "Gelbe Autobahn" (yellow motorway) because they have the same technical standard of the German autobahn but they use black on yellow signs instead of the white on blue signs on the Autobahn motorway network. These are generally high speed arterial roads of larger cities and important streets inside a federal state that do not connect major cities so that they do not fall under the federal budgeting program for the Autobahn network. The federal road Bundesstraße 27 is an example where about half of its length is upgraded to a high speed motorway standard. Some of these non-Autobahn motorways do not have an actual speed limit (design speed 130 km/h).



Exit signs.

The increasing importance of motorways outside of the Autobahn network is subject to upcoming legislation which does not exist so far in Germany where technically these are just rural roads that refer to the ministerial sign code for high speed dual carriageways. At the moment some (blue) motorways have been taken out of the Autobahn network program while still using the blue signs and on the other hand some former non-Autobahn (yellow) motorways have been added to Autobahn budgeting but the signs were not changed either. Motorways that are neither in the Autobahn network nor in the

Bundesstraße network are given black on white signs following the same sign code for high speed dual carriageways - this is mostly seen with urban trunk roads.

Italy

Italian national road *SS 1 Via Aurelia* near Livorno

In Italy, a dual carriageway is often called *superstrada* (meaning *expressway*), but this name is unofficial.

Italian Highway Code (*Codice della strada*) divides dual carriageways into three different classifications:

- *strada extraurbana principale* (meaning *main highway*) or *type-B road*: a road with separate carriageways, at least two lanes for each direction, paved shoulder on the right and no cross-traffic. This type of road is quite similar to an *autostrada* or *type-A road* (Italian official name for motorways or freeways), but its building standards are lower. Access limitations and drive behaviour on type-B roads are exactly the same of the motorways (no pedestrians, bicycles and other low vehicles), as well as the signage (except for the background color, that is blue instead of green). Speed limit on type-B roads is up to 110 kilometers per hour. Type-B roads are always toll-free.

- *strada extraurbana secondaria* (meaning *less importance road*) or *type-C road*. This category contains all the roads in non-urban context that are neither *autostrada* (type A) nor *strada extraurbana principale* (type B). This means that a dual carriageway that may not be classified as type-B road, since it does not meet such quality standards, belongs to this category. For type-C roads, there aren't neither special signage nor access restrictions, unless a specific sign is placed. Speed limit is 90 kilometers per hour, on both single and dual carriageways.
- *strada urbana di scorrimento* (meaning *urban expressway*) or *type-D road*: a road in urban context, with separate carriageways, and at least two lanes for each direction. At-level junctions with smaller roads, regulated by traffic lights, are allowed as well as roundabouts. Unless a prohibition sign is placed, there are not access restrictions. Speed limit on this type of road is up to 70 kilometers per hour.

Italian type-B and type-C roads do not follow a specific numbering criterion. They may be numbered as state roads (SS), regional roads (SR), provincial roads (SP) or municipal roads (SC).

Croatia

Dual carriageways or expressways in Croatia (Croatian: *brza cesta*) are non-tolled roads with 2 or more lanes in each direction, but without emergency lanes. The main highways/motorways in Croatia are also dual carriageways, but they have emergency lanes and tolls.

Many bypasses and beltways of smaller cities in Croatia have been recently constructed or planned as dual carriageways. All dual carriageways in Croatia house a central median, usually fitted with guardrails.

The most heavily used dual carriageway in Croatia is the B28 expressway, connecting capital Zagreb to a satellite town, Vrbovec. The D28 is currently finished up to the Gradec interchange. It is undergoing extensions which will increase the traffic traversing it.

Other regional implementations

United States

A divided highway (U.S. Route 52) in the state of Indiana.

Savery Avenue in Massachusetts
First Divided Highway in the U.S.

In the U.S., this sign tells motorists they are crossing a divided highway

In the United States, this type of road may be called a divided highway, boulevard, parkway, expressway, freeway, or interstate, and has a grassy median or Jersey barrier separating the traffic directions. With few exceptions, all roads in the federally funded Interstate Highway System are fully-controlled access divided highways known as freeways. A broader definition, expressways, includes both freeways and partial limited-access divided highways, and "expressway" is often used specifically to refer to the latter. United States Numbered Highways, state highways and other locally maintained highways may also be divided. Speed limits on rural divided highways range from 65–75 miles per hour (105–121 km/h), with some portions as high as 80 miles per hour (130 km/h). Urban divided highways which are at grade and typically have much lower speed limits are sometimes called boulevards.

In keeping with the U.S. Department of Transportation's *Manual of Uniform Traffic Controls and Devices* (MUTCD), since the early 1970s all divided highways are striped by color to show the direction of traffic flow. Two-way undivided roads have an amber center line, with a broken line indicating passing zones and a solid line indicating no passing zones and solid white baseline shoulder stripes. On undivided roads with more than one lane in each direction, the center is normally marked with a double solid line. The double solid stripe denotes that it is illegal to pass on the other side of the center line. Multilane one-way carriageways use broken white lines between lanes; the median-side baseline is solid amber, and the right sideline is solid white. Frequently in the U.S. the two carriageways are separated by some distance (wide medians with small forests or even hills in them), but drivers can always tell whether the roadway is two-way or one-way—and, if one-way, the direction in which the traffic flows—by looking at the striping coloration.

Canada

Highway 401 in Ontario, Canada, uses a divided highway, collector / express system to separate local traffic from longer distance travelers.

In Canada, "divided highway" is used for this type of road, and the segment between the roadways is referred to as a "median". More informally, a divided highway may be referred to as "twinned". This stems from the practice of "twinning" an existing two-lane highway (usually controlled-access) and converting it into a divided highway. On some portions of Ontario's 400-series highway network, the median may be either steel guardrail or an Ontario tall-wall barrier rather than an unpaved strip.

Like the US, there are two types of divided highways, fully-controlled access divided routes known as freeways, while expressways may include both freeways and partial limited-access divided highways. Canadians often use "highway" to refer to freeways. Partial limited-access divided highways such as the Hanlon Parkway and Black Creek Drive have at-grade intersections and private entrances but have sufficient right-of-way to convert them to full freeways if traffic warrants. There are also RIRO expressways, such as Highway 11 and a portion of Highway 35, which are not full freeways since they allow access to existing properties, but traffic speeds are faster than regular roads due to a median barrier preventing left turns (motorists have to use a "turnabout" overpass to access exits on the opposing direction).

Junctions may be at-grade or grade-separated, and there may be gaps in the median strip to allow turning and crossing. Divided highways are seldom equipped with traffic circles, roundabouts, or rotaries.

Australia

A typical dual carriageway in Melbourne, Australia.

Examples of dual carriageways on non-urban roads in Australia include the Hume Highway and the Pacific Highway (Australia); the Hume Highway by 2012, will be 100% dual carriageway and the Pacific Highway by 2016 will also be 100% dual carriageway. Today, 90% of the Hume Highway is dual carriageway and only 40% or 280 km (174 mi) of the Pacific Highway is dual carriageway, plus 10% of the Pacific Highway or 78 kilometres (48 miles) is under construction. The Federal Highway between the Hume Highway at Goulburn and Canberra is 100% dual carriageway, completed before the 2000 Summer Olympic Games. Some parts of the Princes Highway, Great Western Highway (A32) and the Barton Highway are also dual carriageway. Most non-urban dual carriageway highways/freeways are speed limited to 110 km/h (100 km/h for heavy vehicles), except for a short section in the Australian Capital Territory on the Federal Highway which is state limited to 100 km/h. The Warrego Highway between Brisbane and Toowoomba in Queensland is also dual carriageway with a maximum speed limit of 100km/h.

In Australia there is much confusion among the local population between a freeway and a dual carriageway, particularly those roads found in rural areas. While to be labelled a freeway, a road in Australia must have graded junctions, a physical barrier to separate the carriageways is not required. Speed limits do not change when road grades change from freeway to dual carriageway and vice versa. Stopping and U-turn prohibitions on Australian freeways are routinely ignored and rarely enforced. The Hume Freeway in Victoria, is actually a dual-carriageway grade road for much of its length with many ungraded junctions. Despite this, it has officially been called the "Hume Freeway" and given a route marking of M31. An **M** marking is traditionally used for Motorway or Freeway lengths of roads in international jurisdictions that adopt the same MABC route labelling system as Victoria.

China

The best examples of dual carriageways in mainland China can be seen on the China National Highways. On some routes, such as China National Highway 106, there is a central reservation.

Singapore

A high proportion of roads in Singapore are dual carriageways with central reservations; examples include Clementi Road, Commonwealth Avenue and Holland Road. Often there might be railings erected on the central reservation to prevent pedestrians from dashing across the road. These usually have traffic lights along the way but flyovers and road tunnels (or 'underpasses') can be built to minimise the use of traffic lights; for example, at the Holland Road-Farrer Road-Queensway junction there are three levels of roads. Before the 1980s roundabouts were popular but since then many have been changed to traffic-light controlled junctions.

These dual carriageways are to be distinguished from motorways, known in Singapore as expressways such as the Pan-Island Expressway (PIE) and Ayer Rajah Expressway (AYE) where no traffic lights are used.

Malaysia

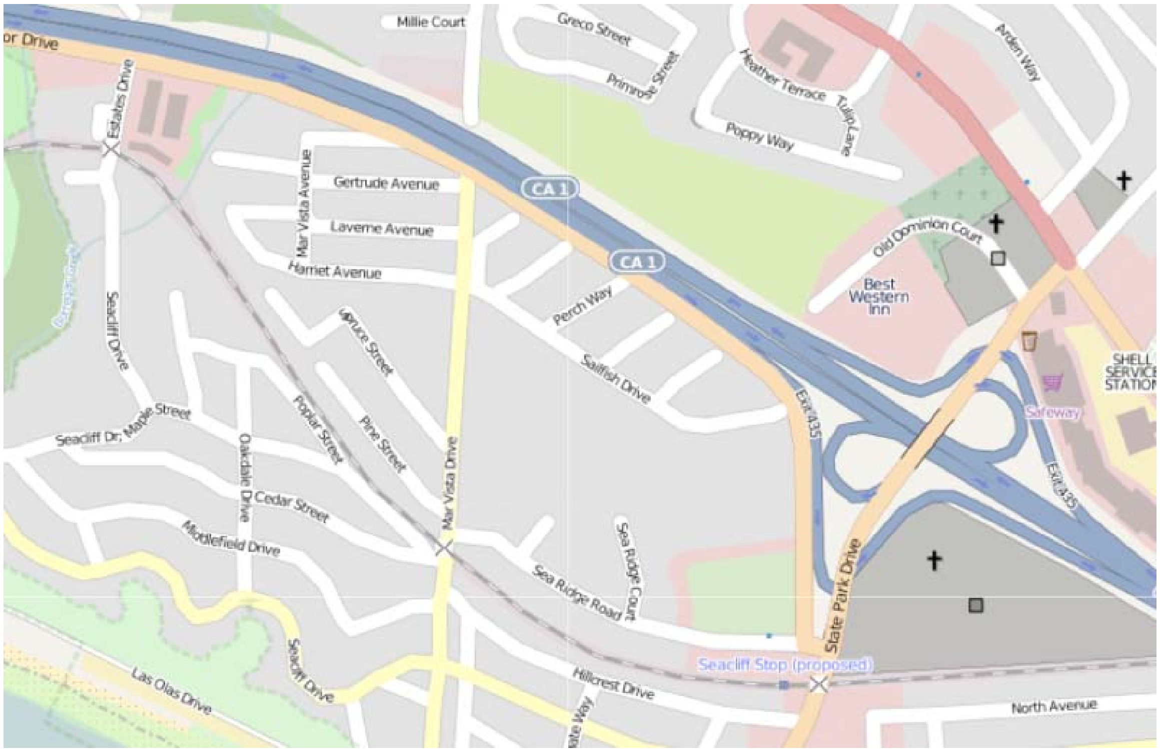
There are dual expressways, such as East Coast Expressway.

Chapter 3

Limited-Access Road & Toll Road

Limited-Access Road

Limited access to the A1214 main road in the Ipswich (UK). Map source: Ordnance Survey



California State Route 1 is showing parallel service roads for local traffic. Map source OpenStreetMap

A **limited-access road** (also **limited-access highway**, **controlled-access road**) is a highway to which access from adjacent properties is limited in some way; a limited access road (for example a freeway) may also have a divided highway, grade separated junctions and prohibit some modes of transport such as bicycles or horses; however others may be city streets where controls are placed on the number of driveway accesses by the provision of a parallel service road or other means. The precise definition of these terms varies by jurisdiction.

History

The first implementation of limited-access roadways in the United States was of the Bronx River Parkway in New York, in 1907. The New York State Parkway System was constructed as a network of high speed roads in and around New York City.

The first limited access highway built is thought to be the privately built Long Island Motor Parkway in Long Island, New York.

The concept evolved into uninterrupted arterial roads that are commonly known as expressways in some parts of the world and by other names including motorway and autobahn in others.

Regional implementations

In the United States, the national Manual on Uniform Traffic Control Devices (MUTCD) uses "full control of access" only for freeways. Expressways are defined as having "partial control of access" (or *semi-controlled access*), meaning that major roads typically use interchanges and commercial development is accessed via cross roads or frontage roads, while minor roads can cross at grade and farms can have direct access. This definition is also used by some states, some of which also restrict freeways only to motor vehicles capable of maintaining a certain speed. Some other states use "controlled access" to mean a higher standard than "limited access", while others reverse the two terms.

In the United Kingdom, ribbon development became a problem with the advent of the motorcar and was controlled by the Restriction of Ribbon Development Act 1935. Some roads from this period started to be built parallel service roads. The Special Roads Act 1949 provided for the construction of motorways including the M1 motorway which opened in 1958 which incorporated limited access, divided highways and grade separated junctions.

Toll Road

A **toll road** (or **tollway**, **turnpike**, **toll highway** or an **express toll route**) is a privately or publicly built road for which a driver pays a **toll** (a fee) for use. Structures for which tolls are charged include toll bridges and toll tunnels. Non-toll roads are financed using other sources of revenue, most typically fuel tax or general tax funds. The building or facility in which a toll is collected may be called a **toll booth**, **toll house**, **toll plaza**, **toll station**, **toll bar** or **toll gate**. This building is usually found on either side of a bridge and at exits.

Road tolling concepts

Road tolls were levied traditionally for a specific access (e.g. city) or for a specific infrastructure (e.g. roads, bridges). These concepts were widely used until the last century. However, the evolution in technology made it possible to implement road tolling policies based on different concepts. The different charging concepts are designed to suit different requirements regarding purpose of the charge, charging policy, the network to the charge, tariff class differentiation etc.:

Time Based Charges and Access Fees: In a time-based charging regime, a road user has to pay for a given period of time in which he may use the associated infrastructure. For the practically identical access fees, the user pays for the access to a restricted zone for a period or several days.

Motorway and other Infrastructure Tolling: The term tolling is used for charging a well-defined special and comparatively costly infrastructure, like a bridge, a tunnel, a mountain pass, a motorway concession or the whole motorway network of a country.

Classically a toll is due when a vehicle passes a tolling station, be it a manual barrier-controlled toll plaza or a free-flow multi-lane station.

Kilometre or Area Charging: In a kilometre or area charging system concept, vehicles are charged per total distance driven in a defined area.

Variations

Three systems of toll roads exist: open (with mainline barrier toll plazas); closed (with entry/exit tolls) and all-electronic toll collection (no toll booths, only electronic toll collection gantries at entrances and exits, or at strategic locations on the mainline of the road).

On an open toll system, all vehicles stop at various locations along the highway to pay a toll. While this may save money from the lack of need to construct tolls at every exit, it can cause traffic congestion, and drivers may be able to avoid tolls (shunpiking) by exiting and re-entering the highway.

With a closed system, vehicles collect a ticket when entering the highway. In some cases, the ticket displays the toll to be paid on exit. Upon exit, the driver must pay the amount listed for the given exit. Should the ticket be lost, a driver must typically pay the maximum amount possible for travel on that highway. Short toll roads with no intermediate entries or exits may have only one toll plaza at one end, with motorists traveling in either direction paying a flat fee either when they enter or when they exit the toll road. In a variant of the closed toll system, mainline barriers are present at the two endpoints of the toll road, and each interchange has a ramp toll that is paid upon exit or entry. In this case, a motorist pays a flat fee at the ramp toll and another flat fee at the end of the toll road; no ticket is necessary.

In an all-electronic system (such as that used on Highway 407 in the Canadian province of Ontario and the Fort Bend Westpark Tollway in the U.S. state of Texas), no cash toll collection takes place, tolls are usually collected with the use of a transponder mounted on the windshield of each vehicle, which is linked to a customer account which is debited for each use of the toll road. On some roads, such as Highway 407, automobiles and light trucks without transponders are permitted to use the road (though trucks with a gross vehicle weight over 5,000 kilograms must have a transponder) - a bill for the toll due is then sent to the registered owner of the vehicle by mail; by contrast, the Fort Bend Westpark Tollway requires all vehicles to be equipped with a transponder.

Modern toll roads often use a combination of the three, with various entry and exit tolls supplemented by occasional mainline tolls.

Some toll roads charge a toll in only one direction, such as where the M4 in Great Britain crosses the River Severn on either of the two Severn Bridges. On these bridges, it is free to travel from Wales into England, but a toll must be paid on the return journey. Crossings between Pennsylvania and New Jersey operated by Delaware River Port

Authority, and crossings between New Jersey and New York operated by Port Authority of New York and New Jersey, use this method (in coordination with the E-ZPass electronic transponder system) given the distance between the bridges along the river, commuter traffic between the two states, and similar tolls on each bridge. This is practical where the detour to avoid the toll is large or the toll differences are small.

Toll payments may be made in cash, by credit card, by pre-paid card, or by an electronic toll collection system. In some European countries, payment is made using stickers which are affixed to the windscreen. Some toll booths are automated. Tolls may vary according to the distance traveled, the building and maintenance costs of the motorway, and the type of vehicle.

Early toll roads

The history of tolls stretches back to Greek mythology where Charon the ferryman charged a toll to carry the dead across the rivers Acheron and Styx to Hades. If the soul paid a toll, Charon ferried it across the river. If not, it wandered between death and life for eternity.

Tolls have been placed on roads at various times in history, often to generate funds for repayment of toll revenue bonds used to finance constructions and/or operation.

Toll roads are at least 2700 years old, as tolls had to be paid by travellers using the Susa–Babylon highway under the regime of Ashurbanipal, who reigned in the seventh century BC. Aristotle and Pliny refer to tolls in Arabia and other parts of Asia. In India, before the 4th century BC, the Arthasastra notes the use of tolls. Germanic tribes charged tolls to travellers across mountain passes. Tolls were used in the Holy Roman Empire in the 14th century and 15th century.

A 14th century example (though not for a road) is Castle Loevestein in the Netherlands, which was built at a strategic point where 2 rivers meet, and charged tolls on boats sailing along the river.

Many modern European roads were originally constructed as toll roads in order to recoup the costs of construction. In 14th century England, some of the most heavily used roads were repaired with money raised from tolls by pavage grants. Turnpike trusts were established in England from 1706 onwards, and were ultimately responsible for the maintenance and improvement of most main roads in England and Wales, until they were gradually abolished from the 1870s. Most trusts improved existing roads, but some new ones, usually only short stretches of road, were also built. Thomas Telford's Holyhead road (now the A5 road) is exceptional as a particularly long new road, built in the early 19th century with many toll booths along its length.

19th century toll booth in Kings County, New York

Some cities in Canada had toll roads in the 19th Century. Roads radiating from Toronto required users to pay at toll gates along the street (Yonge Street, Bloor Street, Davenport Road, Kingston Road) and disappeared after 1895.

19th century plank roads were usually operated as toll roads. One of the first U.S. motor roads, the Long Island Motor Parkway (which opened on October 10, 1908) was built by William Kissam Vanderbilt II, the great-grandson of Cornelius Vanderbilt. The road was closed in 1938 when it was taken over by the state of New York in lieu of back taxes.

In the 20th century, road tolls have been introduced in Europe for financing the construction of motorway networks and specific road infrastructure such as bridges and tunnels. Italy has been the first European country to apply the use of motorway tolls on a 50 km motorway section near Milan in 1924. It was followed by Greece, which made users to pay for the network of motorways around and between its cities in 1927. Later in the 1950's and 1960's, also France, Spain and Portugal started to build motorways largely with the aid of concessions, allowing rapid development of this infrastructure without massive State debts. Since then, road tolls have been introduced in the majority of the EU Member States.

National toll-road differences

Toll roads are found in many countries. The way they are funded and operated may differ from country to country. Some of these toll roads are privately owned and operated. Others are owned by the government. Some of the government-owned toll roads are privately operated.

Some toll roads are managed under such systems as the Build-Operate-Transfer (BOT) system. Private companies build the roads and are given a limited franchise. Ownership is transferred to the government when the franchise expires. Throughout the world, this type of arrangement is prevalent in Australia, India, South Korea, Japan, Philippines, and Canada. The (BOT) system is a fairly new concept that is gaining ground in the United States, with Arkansas, California, Delaware, Florida, Illinois, Indiana, Mississippi, Texas, and Virginia already building and operating toll roads under this scheme. Pennsylvania, Massachusetts, New Jersey, and Tennessee are also considering the BOT methodology for future highway projects.

The more traditional means of managing toll roads in the United States is through semi-autonomous public authorities. New York, Massachusetts, New Hampshire, New Jersey, Maryland, Ohio, Pennsylvania, Kansas, Oklahoma, and West Virginia manage their toll roads in this manner. While most of the toll roads in California, Delaware, Florida, Texas, and Virginia are operating under the BOT arrangement, a few of the older toll roads in these states are still operated by public authorities.

In France, all toll roads are operated by private companies, and the government takes a part of their profit.

Critics of toll roads

Toll roads have been criticized as being inefficient in various ways:

1. They require vehicles to stop or slow down, manual toll collection wastes time and raises vehicle operating costs.
2. Collection costs can absorb up to one-third of revenues, and revenue theft is considered to be comparatively easy.
3. Where the tolled roads are less congested than the parallel "free" roads, the traffic diversion resulting from the tolls increases congestion on the road system and reduces its usefulness.
4. By tracking the vehicle locations, their drivers are subject to an effectual restriction of their freedom of movement and freedom from excessive surveillance.

Toll collection technology

An adaptation of military "identification friend or foe" or RFID technology, called electronic toll collection, is lessening the delay incurred in toll collection. The electronic system determines whether a passing car is enrolled in the program, alerts enforcers if it is not. The accounts of registered cars are debited automatically without stopping or even opening a window. Currently, DSRC is used as a wireless protocol. Other systems are based on GPRS/GSM and GPS technology. Such a system (for trucks only) in Germany launched successfully in January 2005 and by the end of its first year of operation will have charged tolls for around 22 billion driven kilometres. One of the advantages of GPS-based systems is their ability to adapt easily and quickly to changes in charge parameters (road classes, vehicle types, emission levels, time slots, etc.). Another advantage is the systems' ability to support other value-added services on the same technology platform. These services might include fleet and vehicle engine management systems, emergency response services, pay-as-you-drive insurance services and navigation capabilities.

The first major deployment of an RFID electronic toll collection system in the United States was on the Dallas North Tollway in 1989 by Amtech. The Amtech RFID technology used on the Dallas North Tollway was originally developed at Sandia Labs for use in tagging and tracking livestock. In the same year, the Telepass active transponder RFID system was introduced across Italy.

Highway 407 in the province of Ontario, Canada has no toll booths, and instead reads a transponder mounted on the windshields of each vehicle using the road (the rear license plates of vehicles lacking a transponder are photographed when they enter and exit the highway). This made the highway the first all-automated highway in the world. A bill is mailed monthly for usage of the 407. Lower charges are levied on frequent 407 users who carry electronic transponders in their vehicles. The approach has not been without controversy: In 2003 the 407 ETR settledPDF a class action with a refund to users. The same method is used on Highway 6 in Israel and the reversible lanes of the Lee Roy Selmon Crosstown Expressway in Hillsborough County, Florida (in the latter case, the system reads SunPass transponders).

Throughout most of the East Coast of the United States, E-ZPass (operated under the brands I-Pass in Illinois, i-Zoom in Indiana, and Fast Lane in Massachusetts) is accepted on almost all toll roads. Similar systems include SunPass in Florida, FasTrak in California, and ExpressToll in Colorado. The systems use a small radio transponder mounted in or on a customer's vehicle to deduct toll fares from a pre-paid account as the vehicle passes through the toll barrier. This reduces manpower at toll booths and increases traffic flow and fuel efficiency by reducing the need for complete stops to pay tolls at these locations.

By designing a tollgate specifically for electronic collection, it is possible to carry out open-road tolling, where the customer does not need to slow at all when passing through the tollgate. The U.S. state of Texas is testing a system on a stretch of Texas 121 that has no toll booths. Drivers without a TollTag have their license plate photographed automatically and the registered owner will receive a monthly bill, at a higher rate than those vehicles with TollTags.

Another feature of many electronic toll collection systems is interagency interoperability, where the same transponder is accepted at many toll agencies. For instance, the E-ZPass tag is accepted at most toll facilities in the Eastern United States, from Virginia to Maine, west to the Peace Bridge spanning the Niagara River, and in Ohio, Indiana, and Illinois. The TxTAG system allows interoperability throughout the state of Texas, but is not compatible with systems used outside of Texas.

Electronic toll collection systems also have drawbacks. A computer glitch can result in delays several miles long. Some U.S. state turnpike commissions have debated implementing E-ZPass but have found that such a system would be ineffective because most of the people who use the turnpike are not commuters, are from states that have no ETS on turnpikes, or are from states that don't have a turnpike at all. The toll plazas of some turnpikes are antiquated because they were originally built for traffic that stops to pay the toll or get a ticket.

The technology does have its limits. For instance, the Highway 407 automatic number plate recognition technology has a reputation for the occasional misread plate, leading to bills being sent to motorists in remote parts of Ontario who have never been near the tollway. The Ontario government responded to complaints by hiring an ombudsman to address 407 toll complaints.

One of the first all-electric toll roads, the Triangle Parkway, will open at the end of 2011 in North Carolina.

Closed system

For toll roads, a "closed system" refers to a road where a motorist obtains a ticket upon entering the toll road, then pays a toll upon exiting the expressway. The toll is calculated by the distance travelled on the toll road. In the United States, for instance, the Kansas Turnpike, Pennsylvania Turnpike, Ohio Turnpike, and portions of Florida's Turnpike currently implement closed systems. In contrast, a toll road using an 'open system' consists of mainline toll plazas (a.k.a., toll barriers) at set intervals; it is possible for motorists to get on an 'open toll road' after one toll barrier and exit before the next one, thus travelling on the toll road toll-free. Most open toll roads have ramp tolls or partial access junctions to prevent this. The Massachusetts Turnpike or "MassPike," the Pennsylvania Turnpike and the New York State Thruway implement both systems in different sections.

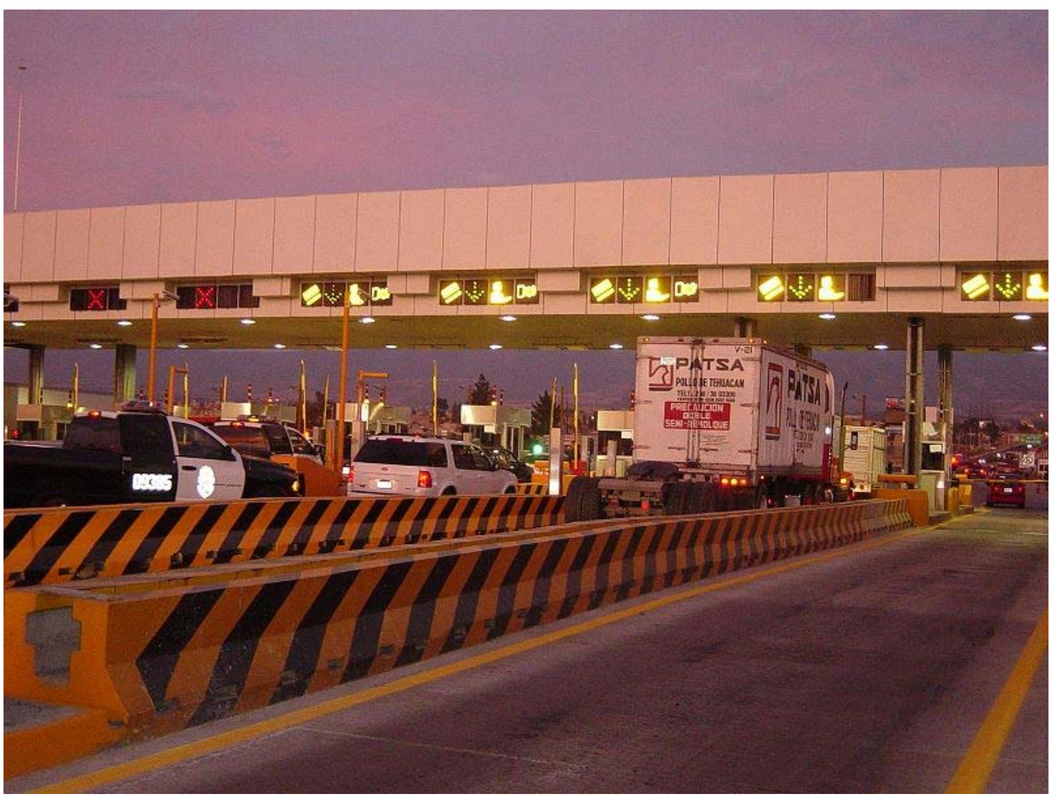
Toll road leading to the Kingsway Tunnel which connects Wallasey and Liverpool under the River Mersey in England, United Kingdom

A high-speed toll booth on SR 417 near Orlando, Florida, United States

On the Sayama bypass (Saitama prefectural road 397) in Japan

A New Jersey Turnpike Toll Gate for Exit 8A in Monroe Township, New Jersey, United States

The open road tolling lanes at the West 163rd Street toll plaza, on the Tri-State Tollway near Hazel Crest, Illinois, United States

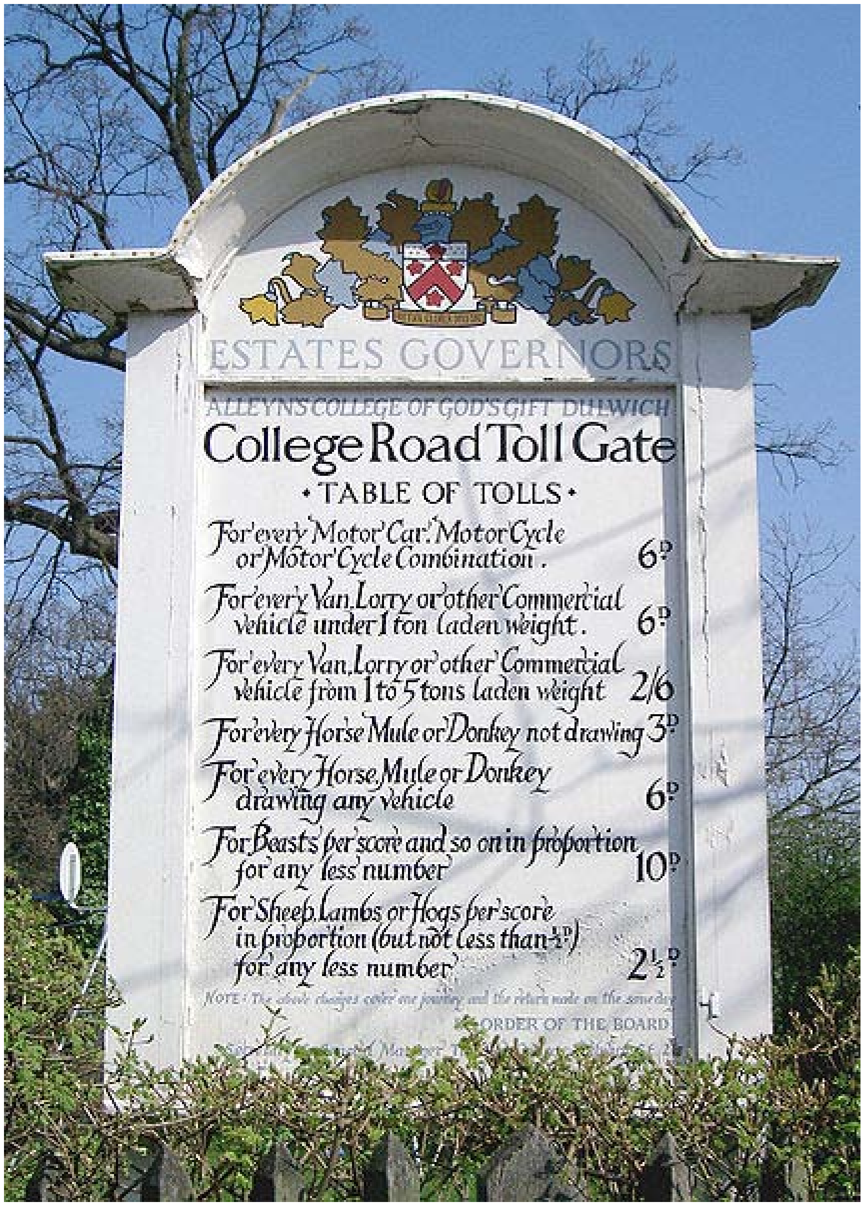


Toll gate *San Marcos*, at the México-Puebla autopista, Mexico

M6 Toll at Great Wyrley near Birmingham in England, United Kingdom

Venice Toll Gate in Autostrada A57 in Italy

ERP gantry at North Bridge Road. Singapore



A table of tolls in pre-decimal currency for the College Road, Dulwich, London SE21 tollgate.

On the Dom Pedro I Highway near the city of Itatiba, Brazil

The Balintawak Toll Barrier of the North Luzon Expressway located in Caloocan City, Philippines.

Service areas provide food, fuel, and travel information on the Ohio Turnpike, in the United States.

407ETR overhead cameras used to capture rear license plates in Ontario, Canada.

Northern Terminus of Cipularang Toll Road, West Java, Indonesia.

Chapter 4

Road Traffic Safety

Sidewalks, curbs and traffic signals in Maryland, United States

Speed limits in different areas, unusually with only a "recommended" limit (130km) for the Autobahn

The term **road traffic safety** is an indication of how safe individual users are on some particular road, or on the roads belonging to some region. The main danger to road users is the likelihood of a traffic collision. Such dangers can be reduced by individual road users operating cautiously and defensively, by building roads in alignment with competent traffic engineering practices, by the application of rational traffic control methods, and by designing road vehicles so they are more able to avoid and survive collisions.

Background

Guardrails save a vehicle from a long fall c. 1920.

Road traffic crashes are one of the world's largest public health and injury prevention problems. The problem is all the more acute because the victims are overwhelmingly healthy prior to their crashes. According to the World Health Organization (WHO), more than a million people are killed on the world's roads each year. A report published by the WHO in 2004 estimated that some 1.2m people were killed and 50m injured in traffic collisions on the roads around the world each year and was the leading cause of death among children 10 – 19 years of age. The report also noted that the problem was most severe in developing countries and that simple prevention measures could halve the number of deaths.

The standard measures used in assessing road safety interventions are fatalities and Killed or Seriously Injured (KSI) rates, usually per billion (10^9) passenger kilometres. In the United States, crashes per million vehicle miles is typically used for road safety.

Speed is a key goal of modern road design, but impact speed affects the severity of injury to both occupants and pedestrians. For occupants, Joksch (1993) found the probability of death for drivers in multi-vehicle accidents increased as the fourth power of impact speed (often referred to by the mathematical term δv ("delta V"), meaning change in velocity). Injuries are caused by sudden, severe acceleration (or deceleration), this is difficult to measure. However, crash reconstruction techniques can be used to estimate vehicle speeds before a crash. Therefore, the change in speed is used as a surrogate for acceleration.

Interventions take many forms. Contributing factors to highway crashes may be related to the driver (such as driver error, illness or fatigue), the vehicle (brake, steering, or throttle failures) or the road itself (lack of sight distance, poor roadside clear zones, etc.). Interventions may seek to reduce or compensate for these factors, or reduce the severity of crashes that do occur. A comprehensive outline of interventions areas can be seen in Management systems for road safety.

Infrastructure design

For road traffic safety purposes it can be helpful to classify roads into ones in built-up area, non built-up areas and then major highways (Motorways/Freeways etc.)

Most casualties generally occur in on roads in built-up areas and major highways are the safest. Reported Road Casualties Great Britain for 2008 show that most fatalities occur on non built-up roads but that the vast majority of serious injuries and injuries occur in built-up areas:

Road Type	Killed	Serious injury	Slight injury	total injury	ref	Note
Non Built-up (excludes motorways)	1,323	8,342	48,810	58,475		52% of the total killed, 32% of total seriously injured, 25% of total with slight injuries
Built-up	1,057	16,823	143,079	160,959		42% of the total killed, 65% of total seriously injured, 70% of total with slight injuries
Motorway	158	869	10,444	11,471		6% of the total killed, 3% of total seriously injured, 5% of total with slight injuries. Fatalities on motorways have decreased by 9 per cent since 1994-98 in a period with traffic levels increased by 28%"
All casualties	**2,538**	**26,034**	**202,333**	**230,905**		

Built-up areas

Pedestrian crossing, line markings and street furniture.

A curb extension at a mid-block crosswalk

Utrecht has specially-painted bicycle-only lanes.

On neighborhood roads where many vulnerable road users, such as pedestrians and bicyclists can be found, traffic calming can be a tool for road safety. Shared space schemes, which rely on human instincts and interactions, such as eye contact, for their effectiveness, and are characterised by the removal of traditional traffic signals and signs, and even by the removal of the distinction between carriageway (roadway) and footway (sidewalk), are also becoming increasingly popular. Both approaches can be shown to be effective.

Modern safety barriers are designed to absorb impact energy and minimize the risk to the occupants of cars, and bystanders. For example, most side rails are now anchored to the ground, so that they cannot skewer a passenger compartment, and most light poles are designed to break at the base rather than violently stop a car that hits them. Some road fixtures such as road signs and fire hydrants are designed to collapse on impact. Highway authorities have also removed trees in the vicinity of roads; while the idea of "dangerous trees" has attracted a certain amount of skepticism, unforgiving objects such as trees can cause severe damage and injury to any errant road users.

Most roads are cambered (crowned), that is, made so that they have rounded surfaces, to reduce standing water and ice, primarily to prevent frost damage but also increasing traction in poor weather. Some sections of road are now surfaced with porous bitumen to enhance drainage; this is particularly done on bends. These are just a few elements of highway engineering. As well as that, there are often grooves cut into the surface of cement highways to channel water away, and rumble strips at the edges of highways to rouse inattentive drivers with the loud noise they make when driven over. In some cases, there are raised markers between lanes to reinforce the lane boundaries; these are often reflective. In pedestrian areas, speed bumps are often placed to slow cars, preventing them from going too fast near pedestrians.

Poor road surfaces can lead to safety problems. If too much asphalt or bitumenous binder is used in asphalt concrete, the binder can 'bleed' or flush' to the surface, leaving a very smooth surface that provides little traction when wet. Certain kinds of stone aggregate become very smooth or polished under the constant wearing action of vehicle tyres, again leading to poor wet-weather traction. Either of these problems can increase wet-weather crashes by increasing braking distances or contributing to loss of control. If the pavement is insufficiently sloped or poorly drained, standing water on the surface can also lead to wet-weather crashes due to hydroplaning.

Lane markers in some countries and states are marked with Cat's eyes or Botts dots, bright reflectors that do not fade like paint. Botts dots are not used where it is icy in the winter, because frost and snowplows can break the glue that holds them to the road, although they can be embedded in short, shallow trenches carved in the roadway, as is done in the mountainous regions of California.

Road hazards and intersections in some areas are now usually marked several times, roughly five, twenty and sixty seconds in advance so that drivers are less likely to attempt violent manoeuvres.

Most road signs and pavement marking materials are retro-reflective, incorporating small glass spheres or prisms to more efficiently reflect light from vehicle headlights back to the driver's eyes.

Designing for pedestrians and cyclists

Pedestrians and Cyclists are among the most vulnerable road users, and in some countries constitute over half of all road deaths. Interventions aimed at improving safety of non-motorised users:

- Sidewalks of suitable width for the expected pedestrian traffic
- pedestrian crossings close to the desire line which allow pedestrians to cross roads safely
- segregated pedestrian routes and cycle lanes away from the main highway
- Overbridges (tend to be unpopular with pedestrians and cyclists due to additional distance and effort)
- Underpasses (these can pose heightened risk from crime is not designed well, can work for cyclists in some cases)
- traffic calming and speed humps

- low speed limits that are rigorous enforced, possibly by speed cameras
- shared space schemes giving ownership of the road space and equal priority to all road users, regardless of mode of use
- pedestrian barriers to prevent pedestrians crossing dangerous locations

Pedestrians' advocates question the equitability of schemes if they impose extra time and effort on the pedestrian to remain safe from vehicles, for example overbridges with long slopes or steps up and down, underpasses with steps and addition possible risk of crime and at-grade crossings off the desire line. The Make Roads Safe was criticised in 2007 for proposing such features. Successful pedestrian schemes tend to avoid over-bridges and underpasses and instead use at-grade crossings (such as pedestrian crossings) close the intended route. Successful cycling scheme by contrast avoid frequent stops even if some additional distance is involved given that the main effort required for cyclists is starting off.

In Costa Rica 57% of road deaths are pedestrians, however a partnership between AACR, Cosevi, MOPT and iRAP has proposed the construction of 190 km of pedestrian footpaths and 170 pedestrian crossings which could save over 9000 fatal or serious injuries over 20 years.

Shared space

A shared space in Brighton (UK).

By 1947 the Pedestrians' Association was suggesting that many of the safety features being introduced (speed limits, traffic calming, road signs and road markings, traffic lights, Belisha beacons, pedestrian crossings, cycle lanes etc.) were potentially self defeating because "every nonrestrictive safety measure, however admirable in itself, is treated by the drivers as an opportunity for more speeding, so that the net amount of danger is increased and the latter state is worse than the first."

During the 1990s a new approach, known as 'shared space' was developed which removed many of these features in some places has attracted the attention of authorities around the world. The approach was developed by Hans Monderman who believed that "if you treat drivers like idiots, they act as idiots" and proposed that trusting drivers to behave was more successful than forcing them to behave. Professor John Adams, an expert on risk compensation suggested that traditional traffic engineering measures assumed that motorists were "selfish, stupid, obedient automatons who had to be protected from their own stupidity" and non-motorists were treated as "vulnerable, stupid, obedient automatons who had to be protected from cars – and their own stupidity".

Reported results indicate that the *shared space* approach leads to significantly reduced traffic speeds, the virtual elimination of road casualties, and a reduction in congestion. Living Streets share some similarities with Shared Spaces. The woonerven also sought to reduce traffic speeds in community and housing zones by the use of lower speed limits enforced by the use of special signage and road markings, the introduction of traffic calming measures, and by giving pedestrians priority over motorists.

Major highways

Guard rail on road in Kaluga Oblast (Russia)

The Pan-American Highway with central median and no freestanding obstructions

Major highways including motorways, freeways, Autobahnen and Interstates are designed for safer high-speed operation and generally have lower levels of injury per vehicle km than other roads.

Safety features include:

- limited access from properties and local roads.
- Grade separated junctions
- Median dividers between opposite-direction traffic to reduce likelihood of head-on collisions
- Removing roadside obstacles.
- Prohibition of more vulnerable road users and slower vehicles.
- Placements of energy attenuation devices (e.g. guard rails, wide grassy areas, sand barrels).
- Eliminating road toll booths

The ends of some guard rails on high-speed highways in the United States are protected with impact attenuators, designed to gradually absorb the kinetic energy of a vehicle and slow it more gently before it can strike the end of the guard rail head on, which would be devastating at high speed. Several mechanisms are used to dissipate the kinetic energy.

Fitch Barriers, a system of sand-filled barrels, uses momentum transfer from the vehicle to the sand. Many other systems tear or deform steel members to absorb energy and gradually stop the vehicle.

In some countries major roads have "tone bands" impressed or cut into the edges of the legal roadway, so that drowsing drivers are awakened by a loud hum as they release the steering and drift off the edge of the road. Tone bands are also referred to as "rumble strips," owing to the sound they create. An alternative method is the use of "Raised Rib" markings, which consists of a continuous line marking with ribs across the line at regular intervals. They were first specially authorised for use on motorways as an edge line marking to separate the edge of the hard shoulder from the main carriageway. The objective of the marking is to achieve improved visual delineation of the carriageway edge in wet conditions at night. It also provides an audible/vibratory warning to vehicle drivers, should they stray from the carriageway, and run onto the marking.

Better motorways are banked on curves in order to reduce the need for tire-traction and increase stability for vehicles with high centers of gravity.

An example of the importance of roadside clear zones can be found on the Isle of Man TT motorcycle race course. It is much more dangerous than Silverstone because of the lack of runoff. When a rider falls off at Silverstone he slides along slowly losing energy, so minimal injuries. When he falls of in the Manx he impacts with trees and walls. Similarly, a clear zone alongside a freeway or other high speed road can prevent off-road excursions from becoming fixed-object crashes.

The U.S. has developed a prototype automated roadway, to reduce driver fatigue and increase the carrying capacity of the roadway. Roadside units participating in future Wireless vehicle safety communications networks have been studied.

Motorways are far more expensive and space-consumptive to build than ordinary roads, so are only used as principal arterial routes. In developed nations, motorways bear a significant portion of motorized travel; for example, the United Kingdom's 3533 km of motorways represented less than 1.5% of the United Kingdom's roadways in 2003, but carry 23% of road traffic.

The proportion of traffic borne by motorways is a significant safety factor. For example, even though the United Kingdom had a higher fatality rates on both motorways and non-motorways than Finland, both nations shared the same overall fatality rate in 2003. This result was due to the United Kingdom's higher proportion of motorway travel.

Similarly, the reduction of conflicts with other vehicles on motorways results in smoother traffic flow, reduced collision rates, and reduced fuel consumption compared with stop-and-go traffic on other roadways.

The improved safety and fuel economy of motorways are common justifications for building more motorways. However, the planned capacity of motorways is often exceeded in a shorter timeframe than initially planned, due to the under estimation of the extent of the suppressed demand for road travel. In developing nations, there is significant public debate on the desirability of continued investment in motorways.

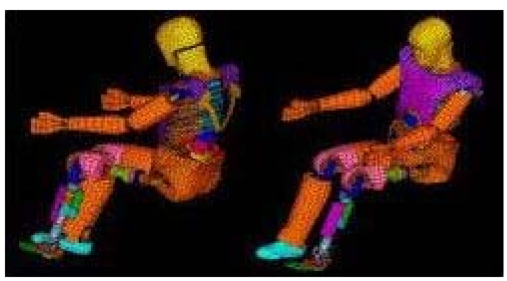
Motorways around the world are subject to a broad range of speed limits. Recent experiments with variable speed limits based on automatic measurements of traffic density have delivered both improvements in traffic flow and reduced collision rates, based on principles of turbulent flow analysis.

With effect from January 2005 and based primarily on safety grounds, the UK's Highways Agency's policy is that all new motorway schemes are to use high containment concrete step barriers in the central reserve. All existing motorways will introduce concrete barriers into the central reserve as part of ongoing upgrades and through replacement as and when these systems have reached the end of their useful life. This change of policy applies only to barriers in the central reserve of high speed roads and not to verge side barriers. Other routes will continue to use steel barriers.

30% of highway crashes that occur in the vicinity of toll collection booths in the countries that have them, these can be eliminated by switching to a system of government taxation for road use

Vehicle safety

Cars



Simulated crashes using crash test dummies can help improve automobile design

Safety can be improved by reducing the chances of a driver making an error, or by designing vehicles to reduce the severity of crashes that do occur. Most industrialized countries have comprehensive requirements and specifications for safety-related vehicle devices, systems, design, and construction. These may include:

- Passenger restraints such as seat belts — often in conjunction with laws requiring their use — and airbags
- Crash avoidance equipment such as lights and reflectors
- Driver assistance systems such as Electronic Stability Control
- Crash survivability design including fire-retardant interior materials, standards for fuel system integrity, and the use of safety glass

- *Sobriety detectors*: These interlocks prevent the ignition key from working if the driver breathes into one and it detects significant quantities of alcohol. They have been used by some commercial transport companies, or suggested for use with persistent drunk-driving offenders on a voluntary basis

Trucks

According to the European Commission Transportation Department "it has been estimated that up to 25% of accidents involving trucks can be attributable to inadequate cargo securing" . Improperly secured cargo can cause severe accidents and lead to loss of cargo, loss of lives, loss of vehicles and can be a hazard for the environment. One way to stabilize, secure and protect cargo during transportation on the road is by using Dunnage Bags which are placed in the void between the cargo and are designed to prevent the load from moving during transport.

Regulation of road users

Various types of road user regulations are in force or have been tried in most jurisdictions around the world, some these are discussed by road user type below.

Motor vehicle users

Dependent on jurisdiction, driver age, road type and vehicle type, motor vehicle drivers may be required to pass a driving test (public transport and goods vehicle drivers may need additional training and licensing), conform to restrictions on driving after consuming alcohol or various drugs, comply with restrictions on use of mobile phones, be covered by compulsory insurance, wear seat belts and comply with certain speed limits. Motorcycle riders may additionally be compelled to wear a motorcycle helmet. Drivers of certain vehicle types may be subject to maximum driving hour regulations.

Some jurisdictions such as Virginia, U.S. and in Maryland, U.S. are targeting specific regulations such as the prohibiting mobile phone use and limiting passenger numbers at young and inexperienced drivers. It has been noticed that more of these types of serious collision occur at night, when the car has multiple occupants and when seat belt use is less.

Insurance companies have proposed that the following restrictions should be imposed on new drivers: a "curfew" imposed on young drivers to prevent them driving at night, an experienced supervisor to chaperone the less experienced driver, forbidding the carrying of passengers, zero alcohol tolerance, raising the standards required for driving instructors and improving the driving test, vehicle restrictions (e.g. restricting access to 'high performance' vehicles), a sign placed on the back of the vehicle (an N- or P-Plate) to notify other drivers of a novice driver and encouraging good behaviour in the post-test period.

Some countries or states have already implemented some of these ideas. Pay-as-you-drive adjusts insurance costs according to when and where the person drives.

Pedal bicycle users

Dependent on jurisdiction, road type and age, pedal cyclists may be required conform to restrictions on driving after consuming alcohol or various drugs, comply with restrictions on use of mobile phones, be covered by compulsory insurance, wear a bicycle helmet and comply with certain speed limits.

Pedestrians

Dependent on jurisdiction, jaywalking may be prohibited.

Statistics

Rating roads for safety

Since 1999 the EuroRAP initiative has been assessing major roads in Europe with a road protection score. This results in a star rating for roads based on how well its design would protect car occupants from being severely injured or killed if a head-on, run-off, or intersection accident occurs, with 4 stars representing a road with the best survivability features. The scheme states it has highlighted thousands of road sections across Europe where road-users are routinely maimed and killed for want of safety features, sometimes for little more than the cost of safety fencing or the paint required to improve road markings.

There are plans to extend the measurements to rate the probability of an accident for the road. These ratings are being used to inform planning and authorities' targets. For example, in Britain two-thirds of all road deaths in Britain happen on rural roads, which score badly when compared to the high quality motorway network; single carriageways claim 80% of rural deaths and serious injuries, while 40% of rural car occupant casualties are in cars that hit roadside objects, such as trees. Improvements in driver training and safety features for rural roads are hoped to reduce this statistic.

The number of designated traffic officers in the UK fell from 15–20% of Police force strength in 1966 to seven per cent of force strength in 1998, and between 1999 and 2004 by 21%. It is an item of debate whether the reduction in traffic accidents per 100 million miles driven over this time has been due to robotic enforcement.

KSI by country

Country	Killed per 1 Billion Veh·km (Motorways in 2003)	Killed per 1 Billion Veh·km (Non-Motorways in 2003)	Motorway AADT	Road Travel by Motorway	km/h (mph) Motorway 2003 Speed Limit
Austria	5.9	13.4	30,077	23%	130 (80)
Czech Republic	9.9	34.3	25,714	11%	130 (80)
Denmark	3.0	11.9	29,454	25%	110 (70)
Finland	1.4	8.3	22,780	10%	120 (75)
France	4.0	12.8	31,979	21%	130 (80)
Germany	3.8	12.4	48,710	31%	none (130 (80) advisory)
Ireland	7.4	11.0	26,730	4%	120 (75)
Italy	13.0	Unknown	Unknown	Unknown	130 (80)
Japan	4.0	11.9	26,152	9%	100 (60)
Netherlands	2.1	11.7	66,734	41%	120 (75)
Slovenia	8.1	18.7	15,643	19%	130 (80)
Spain	62.3	Unknown	Unknown	Unknown	130 (80)
Sweden	2.5	9.9	24,183	21%	110 (70)
Switzerland	2.8	11.8	43,641	33%	120 (75)
United Kingdom	2.0	9.3	85,536	23%	110 (70)
United States	5.2	10.7	39,634	24%	120 (75)

definition: AADT - average annual daily traffic. The bi-direction traffic count representing an average 24-hour day in a year. Sometimes called "traffic density" although it ignores or assumes a constant number of travel lanes.

Advocacy groups

The Automobile Association was established in 1905 in the United Kingdom to help motorists avoid police speed traps. They became involved in other safety issues and also erected thousands of roadside warning signs.

The Pedestrians Association in the United Kingdom was formed in 1929 to press for better road safety. Other groups have been active in other countries.

Motoring advocacy groups including the Association of British Drivers (UK), Speed cameras.org (UK), National Motorists Association (USA/Canada) argue that the strict enforcement of speed limits does not necessarily result in safer driving, and may even have |negative effect on road safety in general. Safe Speed is a UK group set up specifically to campaign against the use of Speed cameras. The Association of British Drivers also argues that speed humps result in increased air pollution, increased noise pollution, and even unnecessary vehicle damage.

In 1965, Ralph Nader put pressure on car manufactures in his book Unsafe at Any Speed detailing resistance by car manufacturers to the introduction of safety features, like seat belts, and their general reluctance to spend money on improving safety. The GM President James Roche was later forced to appear before a United States Senate subcommittee, and to apologize to Nader for the company's campaign of harassment and intimidation. Nader later successfully sued GM for excessive invasion of privacy.

RoadPeace was formed in 1991 in the United Kingdom to advocate for better road safety and founded World Day of Remembrance for Road Traffic Victims in 1993 which received support from the United Nations General Assembly in 2005.

There is some controversy over the way that the motor advocacy groups has been seen to dominate the road safety agenda. Some road safety activists use the term "road safety" (in quotes) to describe measures such as removal of "dangerous" trees and forced segregation of the vulnerable to the advantage of motorized traffic. Orthodox "road safety" opinion fails to address what Adams describes as the top half of the risk thermostat, the perceptions and attitudes of the road user community.

Criticisms

Some road-safety groups argue that the problem of road safety is largely being stated in the wrong terms because most road safety measures are designed to increase the safety of drivers, but many road traffic casualties are *not* drivers (in the UK only 40% of casualties are drivers), and those measures which increase driver safety may, perversely, increase the risk to these others, through risk compensation.

The core elements of the thesis are:

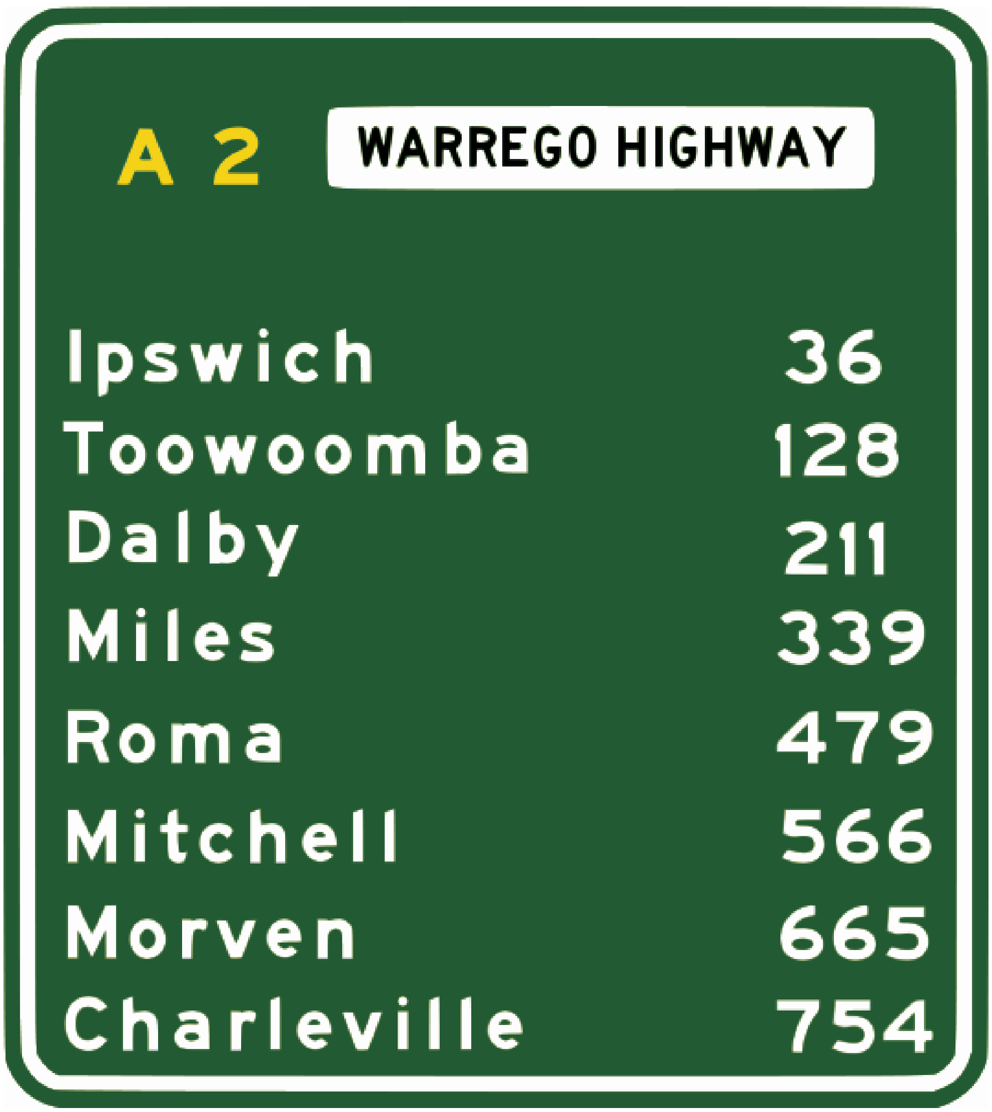
- that vulnerable road users are marginalised by the "road safety" establishment
- that "road safety" interventions are often centred around reducing the severity of results from dangerous behaviours, rather than reducing the dangerous behaviours themselves
- that improved "road safety" has often been achieved by making the roads so hostile that those most likely to be injured cannot use them at all
- that the increasing "safety" of cars and roads is often counteracted wholly or in part by driver responses (risk compensation).

RoadPeace and other groups have been strongly critical of what they see as moves to solve the problem of danger posed to vulnerable road users by motor traffic through increasing restrictions on vulnerable road users, an approach which they believe both blames the victim and fails to address the problem at source. This is discussed in detail by Dr Robert Davis in the book *Death on the Streets: Cars and the mythology of road safety*, and the core problem is also addressed in books by Professor John Adams, Mayer Hillman and others.

For example; the UK publishes Road Casualties Great Britain each year detailing reported road fatalities and injuries and claims to have among the best pedestrian safety in Europe with falling injury rates, as measured in pedestrian KSI per head of population. A study published by the British Medical Journal in 2006 suggested that the reduction in injury levels was due to lower levels of reporting not reducing levels of injury as such. Considerable under-reporting was confirmed by a second report prepared for the UK Department for Transport. and the UK government now acknowledges the issue of under-reporting but is not convinced that the reductions in reported injury levels do not reflect an actual decline. Another independent report investigated if the roads were actually sufficiently dangerous as to deter pedestrians from using them at all.

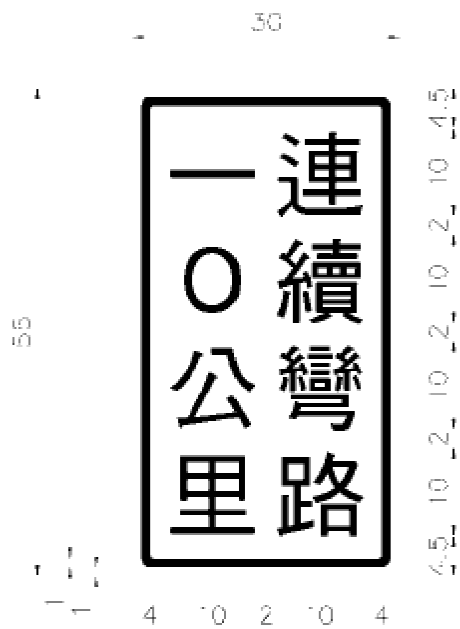
Chapter 5

Traffic Sign



A "route confirmation" sign on the Warrego Highway in Queensland, Australia, informing motorists of their distance (in kilometres) from the places listed

A sign in the United States of America indicating a local speed limit in miles per hour



The design specifications for a Taiwanese sign warning of double bends ahead

Two or more signs may be displayed on one post. Here a Canadian end-of-road marker appears together with a rural airport sign.

Traffic signs or **road signs** are signs erected at the side of roads to provide information to road users. With traffic volumes increasing over the last eight decades, many countries have adopted pictorial signs or otherwise simplified and standardized their signs to facilitate international travel where language differences would create barriers, and in general to help enhance traffic safety. Such pictorial signs use symbols (often silhouettes) in place of words and are usually based on international protocols. Such signs were first developed in Europe, and have been adopted by most countries to varying degrees.

Categories

A group of green-colored directional signs on the National Highway No. 1 in Kaohsiung, Taiwan.

Traffic signs can be grouped into several types. For example, Annexe 1 of the Vienna Convention on Road Signs and Signals (1968), which at 30 June 2004 had 52 signatory countries, defines eight categories of signs: ***A.** Danger warning signs

- B. Priority signs
- C. Prohibitory or restrictive signs
- D. Mandatory signs
- E. Special regulation signsÁ
- F. Information, facilities, or service signs
- G. Direction, position, or indication signs
- H. Additional panels

In the United States, Canada and Australia signs are categorised as follows:

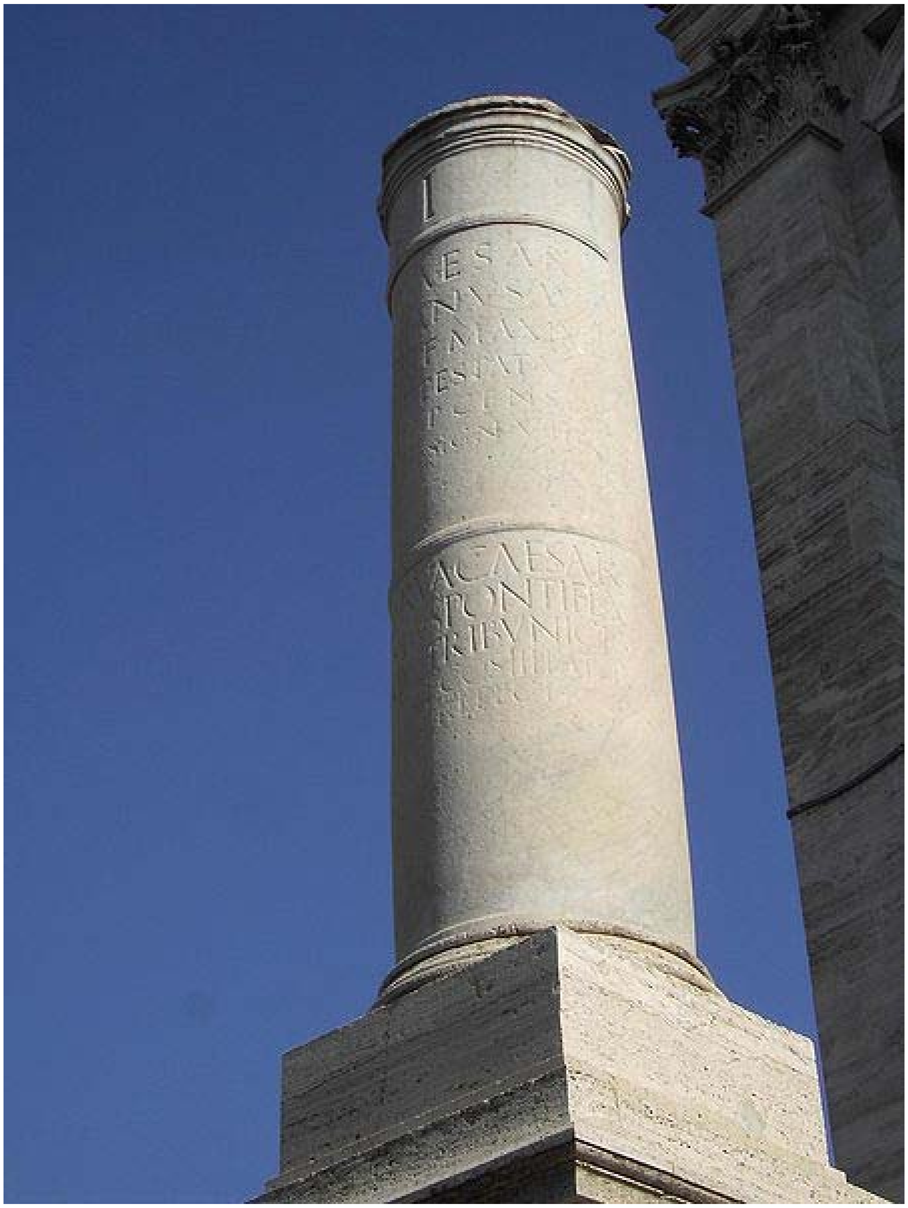
- Regulatory signs
- Warning signs
- Guide signs
    - Street signs
    - Route marker signs
    - Expressway signs

    - Freeway signs
    - Welcome Signs
    - Informational signs
    - Recreation and cultural interest signs
- Emergency management (civil defense) signs
- Temporary traffic control (construction or work zone) signs
- School signs
- Railroad and light rail signs
- Bicycle signs

In the United States, the categories, placement, and graphic standards for traffic signs and pavement markings are legally defined in the Federal Highway Administration's *Manual on Uniform Traffic Control Devices* as the standard.

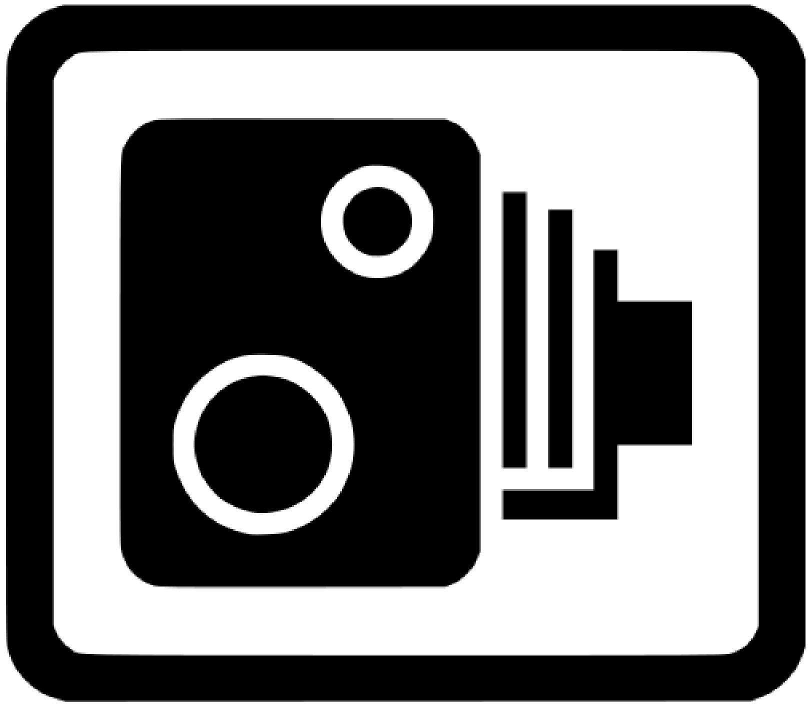
A rather informal distinction among the directional signs is the one between advance directional signs, interchange directional signs, and reassurance signs. Advance directional signs appear at a certain distance from the interchange, giving information for each direction. A number of countries do not give information for the road ahead (so-called "pull-through" signs), and only for the directions left and right. Advance directional signs enable drivers to take precautions for the exit (e.g., switch lanes, double check whether this is the correct exit, slow down). They often do not appear on lesser roads, but are normally posted on expressways and motorways, as drivers would be missing exits without them. While each nation has its own system, the first approach sign for a motorway exit is mostly placed at least 1000 m from the actual interchange. After that sign, one or two additional advance directional signs typically follow before the actual interchange itself.

History



Roman milestone

Rural municipality weight limit sign on a dirt road.



Speed camera sign used in Canada, Cyprus, the Czech Republic, Estonia, Finland, Georgia, Hong Kong, Iceland, Ireland, Latvia, Malta, Ukraine, and the United Kingdom

The earliest road signs were milestones, giving distance or direction; for example, the Romans erected stone columns throughout their empire giving the distance to Rome. In the Middle Ages, multidirectional signs at intersections became common, giving directions to cities and towns.

The first modern roadsigns erected on a wide scale were designed for riders of high or 'ordinary' bicycles in the late 1870s and early 1880s. These machines were fast, silent and their nature made them difficult to control, moreover their riders travelled considerable distances and often preferred to tour on unfamiliar roads. For such riders, cycling organisations began to erect signs that warned of potential hazards ahead (particularly steep hills), rather than merely giving distance or directions to places, thereby contributing the sign type that defines 'modern' traffic signs.

The development of automobiles encouraged more complex signage systems using more than just text based notices. One of the first modern-day road sign systems was devised by the Italian Touring Club in 1895. By 1900, a Congress of the International League of

Touring Organizations in Paris was considering proposals for standardization of road signage. In 1903 the British government introduced four 'national' signs based on shape, but the basic patterns of most traffic signs were set at the 1908 International Road Congress in Rome. In 1909, nine European governments agreed on the use of four pictorial symbols, indicating "bump", "curve", "intersection", and "grade-level railroad crossing". The intensive work on international road signs that took place between 1926 and 1949 eventually led to the development of the European road sign system. Both Britain and the United States developed their own road signage systems, both of which were adopted or modified by many other nations in their respective spheres of influence. The UK adopted a version of the European road signs in 1964 and, over past decades, North American signage began using some symbols and graphics mixed in with English.

Over the years, change was gradual. Pre-industrial signs were stone or wood, but with the development of Darby's method of smelting iron using coke, painted cast iron became favoured in the late eighteenth and nineteenth centuries. Cast iron continued to be used until the mid twentieth century, but it was gradually displaced by aluminium or other materials and processes, such as vitreous enamelled and/or pressed malleable iron, or (later) steel. Since 1945 most signs have been made from sheet aluminium with adhesive plastic coatings, these are normally retroreflective for nighttime and low-light visibility. Before the development of reflective plastics, reflectivity was provided by glass reflectors set into the lettering and symbols.

New generations of traffic signs based on electronic displays can also change their text (or, in some countries, symbols) to provide for "intelligent control" linked to automated traffic sensors or remote manual input. In over 20 countries, real-time Traffic Message Channel incident warnings are conveyed directly to vehicle navigation systems using inaudible signals carried via FM radio, 3G cellular data and satellite broadcasts. Finally, cars can pay tolls and trucks pass safety screening checks using video numberplate scanning, or RFID transponders in windshields linked to antennae over the road, in support on-board signalling, toll collection and travel time monitoring.

Yet another "medium" for transferring information ordinarily associated with visible signs is RIAS (Remote Infrared Audible Signage), e.g., "talking signs" for print-handicapped (including blind/low-vision/illiterate) people. These are infra-red transmitters serving the same purpose as the usual graphic signs when received by an appropriate device such as a hand-held receiver or one built into a cell phone.

North America, Australia and New Zealand

handicap sign

One of Catskill Park's distinctive brown town signs with yellow text, showing the hamlet of Pine Hill

Yellow and black warning signs for kangaroos are common in Australia.

Color schemes

The North American, Australian and New Zealand colours normally have these meanings:

- red with white for stop signs, yield, and forbidden actions (such as No Parking)
- green with white letters for informational signs, such as directions, distances, and places
- brown with white for signs to parks, historic sites, ski areas, forests, and campgrounds
- blue with white for rest areas, food, gasoline, hospitals, lodging, and other services
- white with black (or red) letters for regulatory signs, such as speed limits (or parking)
- yellow with black letters and symbols for warning signs, such as curves and school zones
- orange with black letters for temporary traffic control zones and detours associated with road construction
- purple for "lanes restricted to use only by vehicles with registered electronic toll collection (ETC) accounts", such as EZPass.

The U.S. *Manual on Uniform Traffic Control Devices* prescribes four other colors:

- fluorescent yellow-green for school zone, school bus stop, pedestrian, playground, and bicycle warning signs
- fluorescent pink for incident management signs
- coral and light blue, which are unassigned.

Regulatory signs are also sometimes seen with white letters on red or black signs. In Quebec, blue is often used for tourist attractions and brown public services such as rest areas; many black-on-yellow signs are red-on-white instead.

Many U.S. states and Canadian provinces now use fluorescent orange for construction signs.

Highway symbols and markers

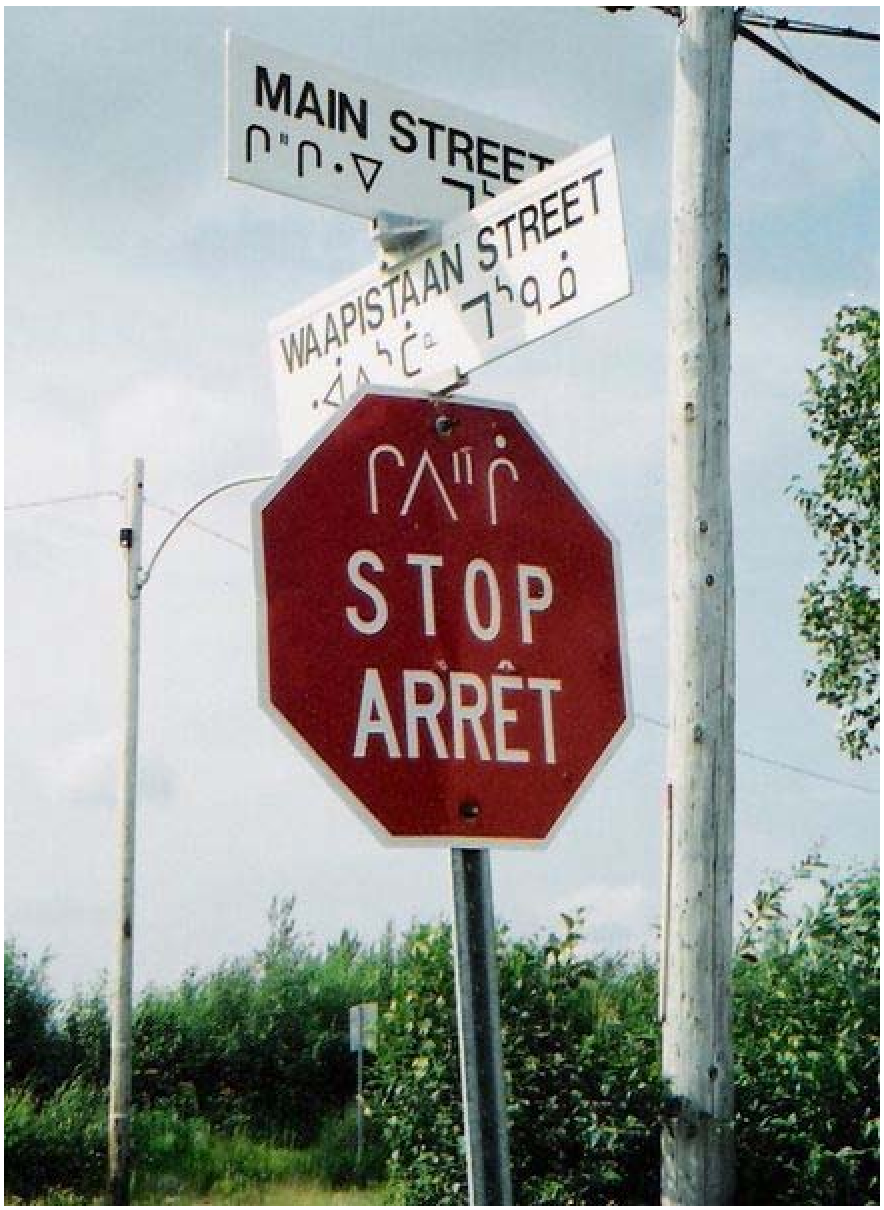
Rural highway sign, Saskatchewan.

Every state and province has different markers for its own highways, but use standard ones for all federal highways. Many special highways– such as the Queen Elizabeth Way, Trans-Canada Highway, and various auto trails in the U.S. – have used unique signs. Counties in the U.S. sometimes use a pentagonal blue sign with yellow letters for numbered county roads, though the use is inconsistent even within states.

Units

Distances on traffic signs generally follow the measurement system in use by the country. Most U.S. road signs use the "U.S. Customary Units" system of miles (or a fraction) or yards (not to be confused with the British "Imperial Measure System", adopted throughout all British possessions following the American Revolution), although the federal Department of Transportation has developed metric standards for all signs. United Kingdom signs also display distances in miles. Elsewhere, metric distances are in very wide use, though not universal.

Languages



Multilingual road signs in Mistissini, Quebec in Cree, English and French.

Where signs use a language, the recognized language/s of the area is normally used. Signs in most of the U.S., Canada, Australia, and New Zealand are in English. Quebec uses French, while New Brunswick and the Jacques-Cartier and Champlain bridges, in Montreal (as well as some parts in the West Island), use both English and French, and a number of other provinces and states, such as Ontario, Manitoba, and Vermont use bilingual French–English signs in certain localities. Mexico and Spain use Spanish. Signs in Belgium are in French, Dutch and German depending on region. In the Brussels Capital Region, road signs are in French and Dutch. Signs in Finland are in Finnish and Swedish. Signs in Germany and Austria are in German. Signs in Luxembourg are in French and German. Signs in Switzerland are in French, German, and Italian. Within a few miles of the U.S.-Mexico border, road signs are often in English and Spanish in places like San Diego, Yuma, and El Paso. Indigenous languages, mainly Nahuatl as well as some Mayan languages, have been used as well.

Typefaces

The typefaces predominantly used on signs in the U.S. and Canada are the FHWA alphabet series (Series B through Series F and Series E Modified). Details of letter shape and spacing for these alphabet series are given in "Standard Alphabets for Traffic Control Devices," first published by the Bureau of Public Roads (BPR) in 1945 and subsequently updated by the Federal Highway Administration (FHWA). It is now part of Standard Highway Signs (SHS), the companion volume to the MUTCD which gives full design details for signfaces.

Initially, all of the alphabet series consisted of uppercase letters and digits only, although lowercase extensions were provided for each alphabet series in a 2002 revision of SHS. Series B through Series F evolved from identically named alphabet series which were introduced in 1927.

Straight-stroke letters in the 1927 series were substantially similar to their modern equivalents, but unrounded glyphs were used for letters such as B, C, D, etc., to permit more uniform fabrication of signs by illiterate painters. Various state highway departments and the federal BPR experimented with rounded versions of these letters in the following two decades.

The modern, rounded alphabet series were finally standardized in 1945 after rounded versions of some letters (with widths loosely appropriate for Series C or D) were specified as an option in the 1935 MUTCD and draft versions of the new typefaces had been used in 1942 for guide signs on the newly constructed Pentagon road network.

The mixed-case alphabet now called Series E Modified, which is the standard for destination legend on freeway guide signs, originally existed in two parts: an all-uppercase Series E Modified, which was essentially similar to Series E, except for a larger stroke width, and a lowercase-only alphabet. Both parts were developed by the California Division of Highways (now Caltrans) for use on freeways in 1948–1950.

Initially, the Division used all-uppercase Series E Modified for button-reflectorized letters on ground-mounted signs and mixed-case legend (lowercase letters with Series D capitals) for externally illuminated overhead guide signs. Several Eastern turnpike authorities blended all-uppercase Series E Modified with the lowercase alphabet for destination legends on their guide signs.

Eventually, this combination was accepted for destination legend in the first manual for signing Interstate highways, which was published in 1958 by the American Association of State Highway Officials and adopted as the national standard by the BPR.

Uses of non-FHWA typefaces

Some traffic signs, such as the left-turn prohibition sign hanging from this gantry, are lit for better visibility, particularly at night or in inclement weather.

The U.S. National Park Service uses NPS Rawlinson Roadway, a serif typeface, for guide signage; it typically appears on a brown background. Rawlinson has replaced Clarendon as the official NPS typeface, but some states still use Clarendon for recreational signage.

Georgia, in the past, used uppercase Series D with a custom lowercase alphabet on its freeway guide signs; the most distinctive feature of this typeface is the lack of a dot on lowercase *i* and *j*. More recent installations appear to include the dots.

The Clearview typeface, developed by U.S. researchers to provide improved legibility, is permitted for light legend on dark backgrounds under FHWA interim approval. Clearview has seen widespread use by state departments of transportation in Arkansas, Illinois, Maryland, Michigan, Ohio, Pennsylvania, Texas, and Virginia. In Canada, the Ministry of Transportation for the Province of British Columbia specifies Clearview for use on its highway guide signs, and its usage has shown up in Ontario on the Don Valley Parkway and Gardiner Expressway in Toronto and on new 400-series highway installations in Hamilton, Halton and Niagara, as well as street signs in various parts of the province. The font is also being used on newer signs in Alberta, Manitoba, and Quebec.

A new Clearview typeface sign beside an old FHWA typeface, Quebec

It is common for local governments, airport authorities, and contractors to fabricate traffic signs using typefaces other than the FHWA series; Helvetica and Arial are common choices.

Moose crossing warning with kill-counter, Alaska

New Zealand

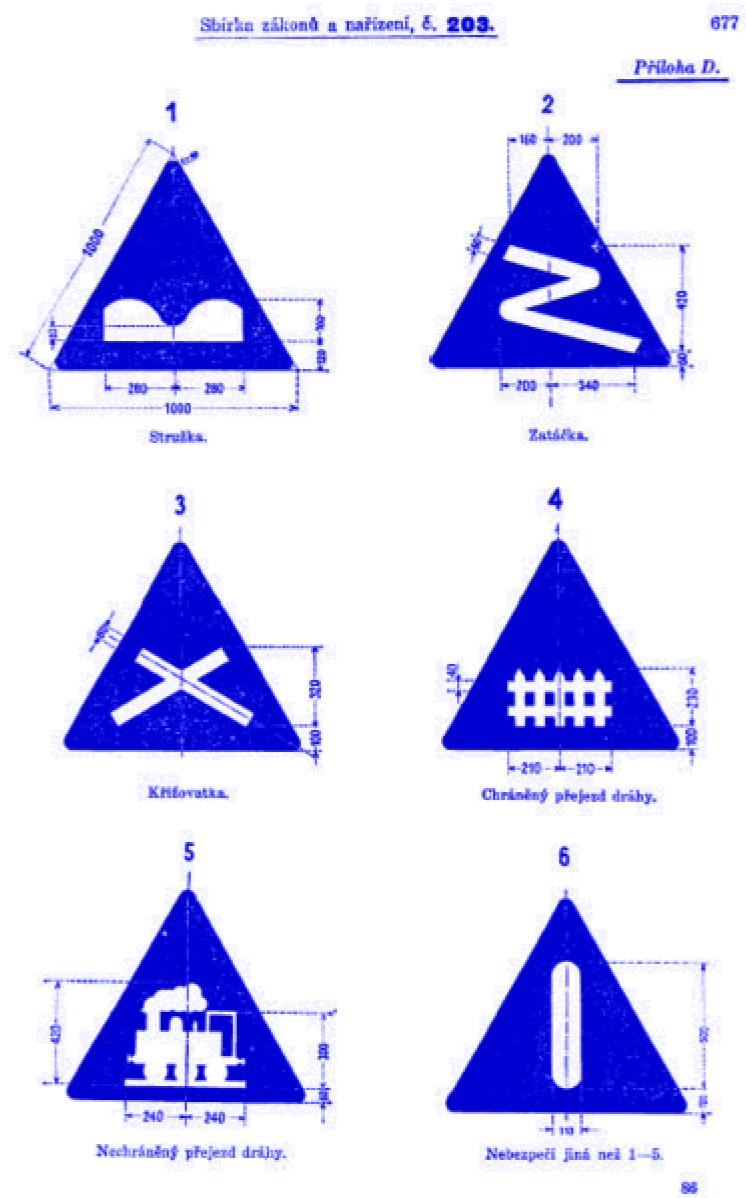
New Zealand road signs are influenced both by American and European practices.

Warning signs are diamond shaped with a yellow background for permanent warnings, and an orange background for temporary warnings. They are somewhat more pictorial than their American counterparts. This is also true for Canadian signage.

Regulatory signs follow European practice, with a white circle with a red border indicating prohibitive actions, and a blue circle indicating mandatory actions. White rectangular signs with a red border indicate lane usage directions. Information and direction signs are rectangular, with a green background indicating a state highway, a blue background for all other roads and all services (except in some, where directional signage is white), and a brown background for tourist attractions.

Before 1987, most road signs had black backgrounds - diamonds indicated warnings, and rectangles indicated regulatory actions (with the exception of the Give Way sign (an inverted trapezium), and Stop sign and speed limit signs (which were the same as today)). Information signs were yellow, and direction signage was green on motorways and black everywhere else.

Europe



The first road signs established in Czechoslovakia on November 1, 1935: six blue-white danger warning signs. They were later supplanted with red-white-black signs.

Road sign in Beussent, France.

Keep right, Portugal.

Romantic road sign in southern Germany.

In 1968, the European countries signed the Vienna Convention on Road Traffic treaty, with the aim of standardizing traffic regulations in participating countries in order to facilitate international road traffic and to increase road safety. Part of the treaty was the Vienna Convention on Road Signs and Signals, which defined the traffic signs and signals. As a result, in Western Europe the traffic signs are well standardised, although there are still some country-specific exceptions, mostly dating from the pre-1968 era.

The principle of the European traffic sign standard is that shapes and colours are to be used for indicating same purposes. Triangular shapes (white or yellow background) are used in warning signs. Additionally, the Vienna convention allows an alternative shape for warning signs, a diamond shape, which is rarely used in Europe. The prohibition signs in Europe are round with a red border. Informative and various other secondary signs are of rectangular shape. Animals shown on warning signs include moose, frogs, deer, ducks, cows, sheep, horses, polar bears (on Svalbard), and monkeys (in Gibraltar). The Convention allows any animal image to be used.

Directional signs have not been harmonised under the Convention, at least not on ordinary roads. As a result, there are substantial differences in directional signage throughout Europe. Differences apply in typeface, type of arrows and, most notably, colour scheme. The convention however specifies a difference between motorways and ordinary roads, and that motorways use white-on-green (e.g., Italy, Switzerland, Denmark, Sweden, Finland, Slovenia, Croatia, Czech Republic, Greece, Cyprus, Bulgaria, Romania, Slovakia, Serbia, Republic of Macedonia, Albania) or white-on-blue (e.g., Norway, Germany, the Republic of Ireland, France, United Kingdom, Spain, Netherlands, Belgium, Austria, Luxembourg, Poland, Portugal, Latvia). Hungary switched from white-on-green to white-on-blue in the early 2000s during the reconstruction of existing and construction of new motorways.

Differences are greater for non-motorways: white-on-blue in Italy, Switzerland, Sweden, Czech Republic, Greece, Cyprus, Slovakia, Bulgaria, Romania, Latvia, Estonia, Finland and Netherlands (in this case the same as motorways), white-on-green in France, United Kingdom, Republic of Ireland, Poland and Portugal, black-on-yellow in Germany, Luxembourg, Norway, Slovenia, Serbia and Croatia, red-on-white in Denmark (though white-on-blue on motorway exits and all overhead gantries), and black-on-white in Spain.

Secondary roads are different from primary roads in France, United Kingdom, Finland, Republic of Ireland, Switzerland and Portugal, always signposted in black-on-white. In Germany, Italy, Romania and Sweden, black-on-white indicates only urban roads or urban destinations.

Signposting road numbers differs greatly as well. Only the European route number, if signposted, will always be placed in white letters on a green rectangle. European route numbers are not signed at all in the United Kingdom.

Some signs like "STOP", "ZONE" etc. are recommended to be in English, but the local language is also permitted. If the language uses non-Latin characters, the names of cities and places should also be in Latin transcription. Road signs in the Republic of Ireland are bilingual, using Irish and English. Wales is also the same, with bilingual Welsh-English signs; some parts of Scotland also have bilingual Scottish Gaelic-English signs. Finland also uses bilingual signs, in Finnish and Swedish.

European countries use the metric system on road signs (distances in kilometres or metres, heights/widths in metres) with the notable exception of the UK, where distances are indicated in miles, and on remaining finger post signs in the Republic of Ireland erected before 1977, where distances are also indicated in miles (which were formally used for all directional signage in the Republic of Ireland prior to 1977 and on speed limits prior to 2005). For countries driving on the left, the convention stipulates that the traffic signs should be mirror images of those used in countries driving on the right. This practice, however, is not systematically followed in the four European countries driving on the left, Cyprus, the Republic of Ireland, Malta and the United Kingdom. The convention permits the use of two background colours for danger and prohibit signs, white or yellow. Most countries use white with a few exceptions like Sweden, Finland, Iceland and Poland, as yellow tends to be more visible in areas in which snow is prevalent.

The European traffic signs have been designed with the principles of heraldry on mind; i.e., the sign must be clear and able to be resolved with one single glance. Most traffic signs conform to heraldic tincture rules, and rather use symbols than written texts for better semiotic clarity.

United Kingdom

One of the more unusual UK road signs, at the Magic roundabout in Swindon

Traffic signing in the UK conforms broadly to European norms, though a number of signs are unique to Britain and direction signs omit European route numbers. The current sign system, introduced on 1 January 1965, was developed in the late 1950s and early 1960s by the Anderson Committee, which established the motorway signing system, and by the Worboys Committee, which reformed signing for existing all-purpose roads.

Britain remains the only European Union member nation and the only major Commonwealth country to use non-metric (Imperial) measurements for distance and speed, although metric "authorised-weight" signs were prescribed in 1981 and there is now a dual-unit (imperial first) option for restriction signage, used on safety grounds where foreign drivers may use the routes so that they may better understand the restriction and/or advice about a hazard ahead.

Three colour schemes exist for direction signs. A road may be a motorway (white on blue), a primary route (white on dark green with yellow route numbers), or a non-primary route (black on white). A fourth colour scheme, black on yellow, is seen on temporary signs, for example marking a diversionary route avoiding a road closure.

Two typefaces are specified for British road signs. Transport *Medium* or Transport *Heavy* are used for all text on fixed permanent signs and most temporary signage, depending on the colour of the sign and associated text color; dark text on a white background is normally set in *Heavy* so that it stands out better. This is except for route numbers on motorway signs, for which a taller limited character set typeface called Motorway is used.

Signs are generally in English although bilingual signs are used in Wales (English/Welsh) and are beginning to be seen in parts of the Scottish Highlands (English/Scottish Gaelic).

The Netherlands

Road signs in The Netherlands follow the Vienna Convention. Directional signs (which have not been harmonised under the Convention) always use blue as the background colour. The destinations on the sign are printed in white. If the destination is not a town (but an area within town or some other kind of attraction), that destination will be printed in black on a separate white background within the otherwise blue sign.

The Netherlands always signpost European road numbers where applicable (i.e., on the advance directional signs, the interchange direction signs and on the reassurance signs). Dutch national road numbers are placed on a rectangle, with motorways being signposted in white on a red rectangle (as an Axx) and primary roads in black on a yellow rectangle (as Nxx). When a motorway changes to a primary road, its number remains the same, but the A is replaced by the N. So at a certain point the A2 becomes N2, and when it changes to a motorway again, it becomes A2 again.

Signs intended for bike-riders always go on white signs with red or green letters.

The Dutch typeface, known as ANWB-Ee, is based on the US typeface. A new font, named ANWB-Uu (also known as Redesign), has been developed in 1997 and appears on many recent Dutch signs. On the motorways however the typeface remains the ANWB-Ee or a similar typeface. The language of the signs is typically Dutch, even though bilingual signs may be used, when the information is relevant for tourists.

Sweden

Swedish elk warning sign

The road signs in Sweden mostly follow the Vienna Convention with a few adaptations, however, allowed within the convention:

- the background of warning signs is yellow
- warning signs for elk and reindeer
- the background of direction signs is blue with white text
- the background of motorway direction signs is green with white text
- when applicable, the language is Swedish in Sweden.

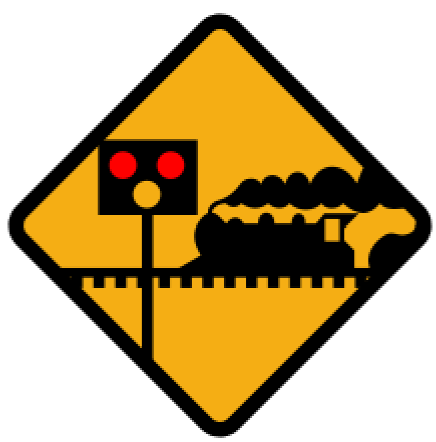
The signage typeface Tratex is used exclusively in Sweden and is available as freeware.

Croatia

Traffic signs in Croatia are the same as traffic signs in the rest of the former Yugoslavia. Croatian road signs follow the Vienna convention (SFR Yugoslavia was the original signatory for Croatia, which is now a contracting party itself). The most common signs are yellow and black signs *for direction*, blue and white signs *for information* and white-on-green signs are used on the highways.

Ireland

Advance directional sign in for a roundabout in Ireland. The green background indicates that this sign is on a national road, with the blue patches left and right indicating a motorway (with symbol) and the white patches indicating a regional road or local road.



Irish warning signs

Until the partition of Ireland in 1922 and the independence of Southern Ireland (now the Republic of Ireland) British standards applied across the island. In 1926 road sign standards similar to those used in the UK at the time were adopted. Law requires that the signs be written in both Irish and English.

In 1956, road signs in the Republic were changed to markedly differ from the UK standard, with the adoption of U.S.-style "diamond" signs for many road hazard warnings (junctions, bends, railway crossings, traffic lights). Some domestic signs were also invented, such as the keep-left sign (a black curved arrow pointing to the upper-left, although some are similar to the European "white arrow on blue disk" signs), while some other signs are not widely adopted outside Ireland, such as the no-entry sign (a black arrow pointing ahead in a white circle with a red slashed circumference).

Directional signage is still firmly based on the United Kingdom standard, however, with the basic design of directional signs remaining the same as the UK in most cases. The same colours are used for directional signs in Ireland as in the UK, and the UK Transport and Motorway fonts are used. However, signage in the Republic of Ireland is bilingual, with the Irish text in mixed case italics, while the English text is in all upper-case.

The 'wild animals' warning sign, used in Italy, Spain, Germany, Latvia, and other countries and in the other former republics of the Soviet Union

In January 2005 Ireland adopted metric speed limits. Around 35,000 existing signs were replaced and a further 23,000 new signs erected bearing the speed limit in kilometres per hour. To avoid confusion with the old signs, each speed limit sign now has "km/h" beneath the numerals. Also, since the adoption of signs based on the *Warboys Committee* standard in 1977, Irish directional signs have used the metric system, however, unlike with the later speed limit change over, there was no effort made to change the existing signage, and as of 2007 many finger posts still remain on rural roads with distances in miles, although the numbers continue to decline as roads are improved.

In late 2007 Ireland started to radically replace signs and posts. Good examples are the M1 (Dublin - Dundalk) and the M50 (Dublin). While being mostly the same as the old signs, it is welcome as a lot of the signs were damaged / stained. About 1/2 of the new posts are now two medium posts with crosshatched metal posts in-between instead of one large pole to minimize the damage in case of a crash.

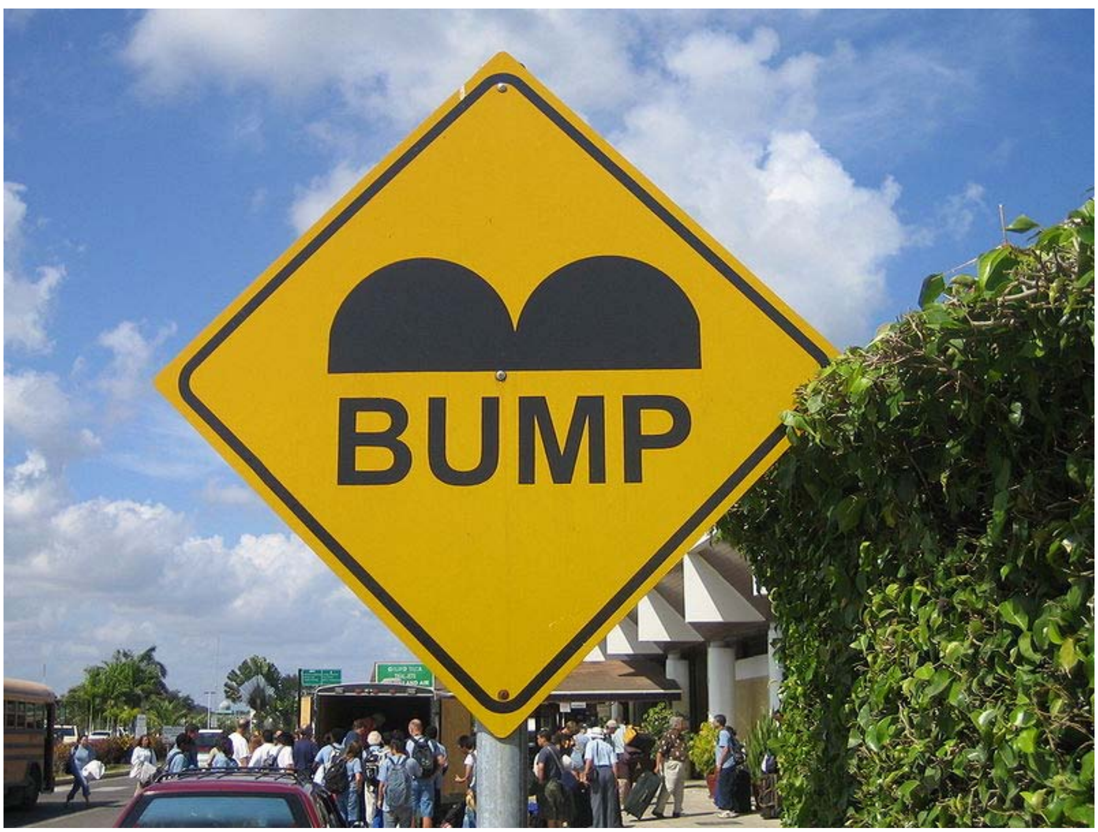
Iceland

Road signs in Iceland mainly follow the Vienna Convention, but use a variant of the colour scheme and minor design changes similar to the signs in Sweden.

Latvia

Road signs in Latvia mainly of Vienna Convention regulations, only the design is different from many other European countries. The signs have many design issues common with Russian road signs since the times of USSR.

Mexico, South and Central America



Speed bump sign in Belize.

Road signs in Mexico, Central America, and South America vary from country to country. For the most part, conventions in signage tend to resemble United States signage conventions more so than European and Asian conventions. For example, warning signs are typically diamond shaped and yellow rather than triangular and white. Some variations include the "No Parking" sign, which uses a letter *E* instead of *P* (the Spanish word for "parking" is *estacionamiento* and *estacionamento* in Brazilian Portuguese), as well as the Stop sign, which usually reads "Pare" or "Alto". Notable exceptions include speed limit signs, which follow the European conventions, and the "No Entry" sign, often replaced with a crossed upwards arrow.

Colombia

Traffic signs in Colombia are classified into three categories. These are Warning signs, Mandatory signs and Information signs.

Warning signs are very similar to warning signs in United States. They are yellow diamond shaped with a black symbol (the yellow color is changed to an orange color in areas under construction). In certain cases, the yellow color is shifted to fluorescent yellow (in the School area sign and Chevron sign).

Mandatory signs are similar to European signs. They are circular with a red border, a white background and a black symbol. Stop sign and Yield sign are as European, except the word “Stop” is changed for “Pare” and the Yield sign has no letters, it is a red triangle with white center.

Information signs have many shapes and colors. Principally they are blue with white symbols and in many cases these signs have an information letter below the symbol.

Asia

Philippines

By law road signs in the Philippines follow the Vienna Convention, however, in reality most road signs are of various character and can occur in different styles and shapes. For example MMDA's (Metro Manila Development Authority) uses pink colored traffic signs within Metro Manila.

China

Warning signs in China are triangular with a black border, yellow background and black symbol. Mandatory signs generally follow European conventions (circular with red border/blue circle)with some local variations. Direction signs are green for expressways, brown for tourist attractions and blue for other roads. Occasionally black on white is used for directions to local facilities.

Japanese stop sign with the word Tomare (止まれ), meaning Stop

Japan

Road signs in Japan are either controlled by local police authorities under Road Traffic Law (道路交通法) or by other road-controlling entities including Ministry of Land, Infrastructure and Transport, local municipalities, NEXCO (companies controlling expressways), under Road Law (道路法). Most of the design of the road signs in Japan are similar to the signs on the Vienna Convention, except for some significant variances, such as stop sign with a red downward triangle. The main signs are categorized into four meaning types: guidance (white characters on blue in general - on green in expressways), warning (black characters and symbols on yellow diamond), regulation (red or blue circle, depending on prohibition or regulation), and instruction (mostly white characters or symbols on blue square).

Automatic traffic sign recognition

Cars are beginning to feature cameras with automatic traffic sign recognition, beginning with the Opel Insignia. It mainly recognises speed limits and no-overtaking areas.

Street sign theft

Street sign theft occurs when street signs are stolen, often to be used as decorations, but also sometimes to avoid obeying the law by claiming later the sign was not there. Although the theft often seems arbitrary, signs that are unusual or amusing tend to be stolen more frequently. Sometimes considered to be a prank by the perpetrators, the theft is often costly and inconveniencing for the municipality or agency that owns the sign. In the United States, each street sign generally costs between $100 and $500 to replace.

Popular culture can act as a catalyst to street sign theft. Popular bands The Beatles and Lynyrd Skynyrd have inadvertently perpetuated street sign theft as their songs and albums include real place names including Penny Lane, Blue Jay Way, Abbey Road, and Brickyard Road.

Another commonly stolen sign marks the entrances to the village of Fucking, Austria. The sign simply says the name of the village. Because of the vulgar connotation of the word, "fucking," this sign has been repeatedly stolen by pranksters.

Chapter 6

Freeway

Interstate 75 and Interstate 85 (Downtown Connector) in Downtown Atlanta, Georgia, United States: a typical American freeway (MUTCD definition)

The High Five Interchange in Dallas, Texas, United States, an extreme example of stack interchange design.

A **freeway** is a limited access divided highway with grade separated junctions and without traffic lights or stop signs. This term is used in the United States, Canada, Australia, and South Africa.

Etymology

The word *freeway* was coined by the "Father of American Zoning," Edward M. Bassett, in an influential article published in February 1930. Bassett argued that roads should be classified into three basic types: highways, parkways, and freeways. In Bassett's zoning and property law-based system (he was a Columbia-trained lawyer), abutting property owners have the rights of light, air, and access to highways, but not parkways and freeways; the latter two are distinguished in that the purpose of a parkway is recreation, while the purpose of a freeway is movement. Thus, as originally conceived, a freeway is simply a strip of public land devoted to movement to which abutting property owners do not have rights of light, air, or access.

In the United States, the term *freeway* is used nationwide. In some regions of the U.S., other terms are also used, including *expressway*, *Interstate Highway*, *thruway*, *highway*, and *turnpike*. While some people use these terms interchangeably, turnpikes and thruways usually have associations with toll roads, such as the New Jersey Turnpike, Ohio Turnpike, Pennsylvania Turnpike, West Virginia Turnpike, Florida's Turnpike, and the New York State Thruway.

Consequently, the term *freeway* is often used to refer to a toll-free highway as opposed to its original meaning – in which the component "free" implies freedom from traffic interference rather than "at no cost" – still used in other countries and in parts of the U.S.

General characteristics

High-capacity freeway interchange in Los Angeles, California, United States

Freeways, by definition, have no at-grade intersections with other roads, railroads or multi-use trails. Movable bridges, such as the Interstate Bridge on Interstate 5 between Oregon and Washington, do require drivers to stop for ship traffic.

State Route 11 in Ohio.

Confounding the term, not all highways bearing the name of freeway are in fact "freeways" by definition; for example, the William L. Wilson Freeway (U.S. Route 340) by Harpers Ferry, West Virginia, is a two-lane undivided roadway featuring at-grade intersections.

The crossing of freeways by other routes is typically achieved with grade separation either in the form of underpasses or overpasses. In addition to sidewalks (footpaths) attached to roads that cross a freeway, specialized pedestrian footbridges or tunnels may also be provided. These structures enable pedestrians and cyclists to cross the freeway at that point without a detour to the nearest road crossing.

Access to freeways is typically provided only at grade-separated interchanges, though lower-standard right-in/right-out access can be used for direct connections to side roads. In many cases, sophisticated interchanges allow for smooth, uninterrupted transitions between intersecting freeways and busy arterial roads. However, sometimes it is necessary to exit onto a surface road to transfer from one freeway to another. An example of this would be Interstate 70 in the town of Breezewood, Pennsylvania.

The M25 motorway near Reigate in the United Kingdom. In most Commonwealth nations, freeway-like roads are referred to as motorways.

Speed limits are generally higher on freeways and are occasionally nonexistent (as on much of Germany's Autobahn network). Because higher speeds reduce decision time, freeways are usually equipped with a larger number of guide signs than other roads, and the signs themselves are physically larger. Guide signs are often mounted on overpasses or overhead gantries so that drivers can see where each lane goes. Exit numbers are commonly derived from the exit's distance in miles or kilometers from the start of the freeway. In some areas, there are public rest areas or service areas on freeways, as well as emergency phones on the shoulder at regular intervals.

In the United States, mileposts start at the southern or westernmost point on the freeway (either its terminus or the state line). California, Ohio, and Nevada use milepost systems in which the markers indicate mileage through the state's individual counties. However, in Nevada and Ohio, and freeways that pass through Kern County, California, also use the standard milepost system concurrently with their respective postmile systems.

Cross sections

An 18 lane, quad roadway highway in southern Ontario, an extreme example of multilane freeway design.

Two-lane freeways, often undivided, are sometimes built when traffic volumes are low or right-of-way is limited; they may be designed for easy conversion to one side of a four-lane freeway. Otherwise, freeways typically have at least two lanes in each direction; some busy ones can have as many as 16 or more lanes in total.

In Mississauga, Ontario, Highway 401 uses collector-express lanes for a total of 18 lanes through its intersection with 403/410 and 427. In San Diego, California, Interstate 5 has a similar system of express and local lanes for a maximum width of 21 lanes on a two-mile segment between Interstate 805 and California State Route 56.

These wide freeways may use separate collector and express lanes to separate through traffic from local traffic, or special high-occupancy vehicle lanes, either as a special restriction on the innermost lane or a separate roadway, to encourage carpooling. These HOV lanes, or roadways open to all traffic, can be reversible lanes, providing more capacity in the direction of heavy traffic, and reversing direction before traffic switches. Sometimes a collector/distributor road, a shorter version of a local lane, shifts weaving between closely-spaced interchanges to a separate roadway or altogether eliminates it.

In some parts of the world, notably parts of the U.S., frontage roads form an integral part of the freeway system. These parallel surface roads provide a transition between high-speed "through" traffic and local traffic. Frequent slip-ramps provide access between the freeway and the frontage road, which in turn provides direct access to local roads and businesses.

Except on some two-lane freeways (and very rarely on wider freeways), a median separates the opposite directions of traffic. This strip may be as simple as a grassy area, or may include a crash barrier such as a "Jersey barrier" or an "Ontario Tall Wall" to prevent head-on collisions. On some freeways, the two carriageways are built on different alignments; this may be done to make use of available corridors in a mountainous area or to provide narrower corridors through dense urban areas.

Some roads in Ohio that conform to freeway criteria use at-grade intersections in lieu of over/under-passes, with occasional interchanges to avoid signalized traffic interruption (i.e., traffic lights are omitted). Examples include US 23 between OH-15's eastern terminus and Delaware, Ohio, along with Highway 15 between its eastern terminus and I-75, US-30, OH-29/US-33, and US-35 in western and central Ohio. These highways are fundamentally expressways, but expressways tend to have lower design speeds, and signalized at-grade intersections.

Access restrictions

A bicyclist using a freeway legally

To reduce the probability that high-speed freeway traffic will have to slow down for slower same-direction traffic, access to freeways is usually limited to drivers of motor vehicles that are powerful enough to maintain a certain minimum speed. Some East Asian countries partially restrict the use of motorcycles or ban them completely from freeways (or expressways in countries where that term is used).

Travelers in a low-powered transportation class (such as pedestrians, bicyclists, equestrians, and moped drivers) are banned at all times from the freeways in many areas by default. In some jurisdictions, these classes are allowed on the shoulders of certain freeways (usually where the freeway has completely replaced an existing road) or on sidepaths.

Legal definitions

United States

Santa Clara County Route G4 (Montague Expressway), an American expressway under the MUTCD definition

In the United States, a *freeway* is defined by the federal government's Manual on Uniform Traffic Control Devices as a divided highway with full control of access. This means two things. First, adjoining property owners do not have a legal right of access, meaning that they cannot connect their lands to the highway by constructing driveways, although frontage roads provide access to properties adjacent to a freeway in many places. When an existing road is converted into a freeway, all existing driveways must be removed and access to adjacent private lands must be blocked with fences or walls.

Second, traffic on a freeway is "free-flowing". All cross-traffic (and left-turning traffic) is relegated to overpasses or underpasses, so that there are no traffic conflicts on the main line of the highway which must be regulated by traffic lights, stop signs, or other traffic control devices. Achieving such free flow requires the construction of many overpasses, underpasses, and ramp systems. The advantage of grade-separated interchanges is that freeway drivers can almost always maintain their speed at junctions since they do not need to yield to vehicles crossing perpendicular to mainline traffic.

In contrast, an expressway is defined as a divided highway with partial control of access. Expressways may have driveways and at-grade intersections, though these are usually less numerous than on ordinary arterial roads.

This distinction was apparently first developed in 1949 by the Special Committee on Nomenclature of what is now the American Association of State Highway and Transportation Officials. In turn, the definitions were incorporated into AASHTO's official standards book, the Manual on Uniform Traffic Control Devices, which would become the national standards book of the U.S. Department of Transportation under a 1966 federal statute. The same distinction has also been codified into the statutory law of eight states: California, Minnesota, Mississippi, Missouri, Nebraska, North Dakota, Ohio, and Wisconsin.

However, each state codified the federal distinction slightly differently. California expressways do not necessarily have to be divided, though they must have at least partial access control. For both terms to apply, in Wisconsin, a divided highway must be at least four lanes wide; and in Missouri, both terms apply only to divided highways at least 10 miles (16 km) long that are not part of the Interstate Highway System. In North Dakota and Mississippi, expressways may have "full or partial" access control and "generally" have grade separations at intersections; a freeway is then defined as an expressway with full access control. Ohio's statute is similar, but instead of the vague word *generally*, it imposes a requirement that 50% of an expressway's intersections must be grade-separated for the term to apply. Only Minnesota enacted the exact MUTCD definitions, in May 2008.

The term *expressway* is also used for what the federal government calls "freeways". Where the terms are distinguished, freeways can be characterized as expressways upgraded to full access control, while not all expressways are freeways.

In contrast, there are at least four highways in the United States which are technically expressways (under the federal definition) but contain the word "freeway" in their names: State Fair Freeway in Kansas, Chino Valley Freeway, Rockaway Freeway in New York, and Shenango Valley Freeway (a portion of U.S. Route 62) in Pennsylvania.

South Africa

In South Africa, the term *freeway* differs from most parts of the world. A freeway is a road where certain restrictions apply. These are not allowed on a freeway:

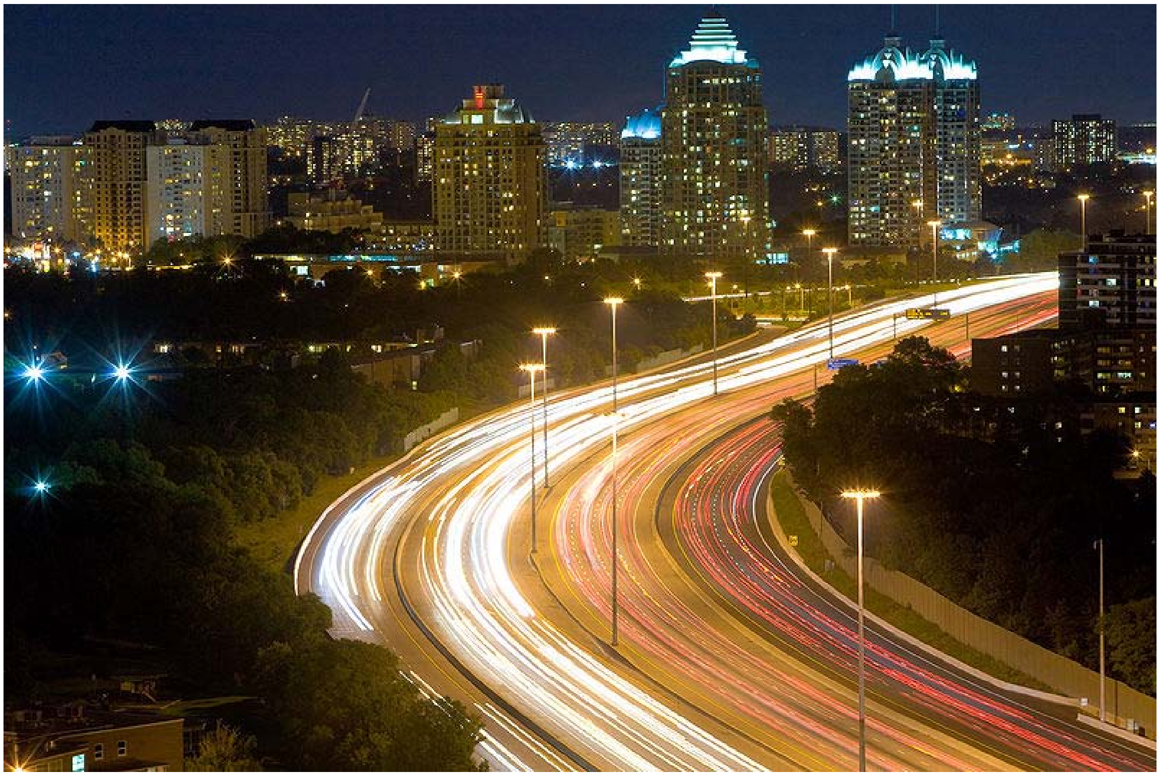
- a vehicle drawn by an animal;
- a pedal cycle (such as a bicycle);
- a motor cycle having an engine with a cylinder capacity not exceeding 50 cm^3 or that is propelled by electrical power;
- a motor tricycle or motor quadrucycle;
- pedestrians

Drivers may not use hand signals on a freeway (except in emergencies) and the minimum speed on a freeway is 60 km/h (37 mph). Drivers in the rightmost lane of multi-carriageway freeways must move to the left if a faster vehicle approaches from behind to overtake.

Despite popular opinion that "freeway" means a road with at least two lanes, single carriageway freeways exist, as is evidenced by the statement that "the roads include 1,400 km of dual carriageway freeway, 440 km of single carriageway freeway and 5 300 km of single carriage main road with unlimited access." The Afrikaans translation of *freeway* is *snelweg* (literally *fast road* or *expressway*).

Effects and controversy

Rush hour on I-45, downtown Houston, Texas.



Highway lighting can have a negative influence on those living close to the freeway. High mast lighting is an alternative as it concentrates the light on the road, but the tall structures can also have a NIMBY effect.

Freeways have been constructed both between urban centers and within them, leading to the sprawling suburban development found near most modern cities. Freeways reduced travel times and accident rates, though the higher speeds have increased the severity and death rates of the collisions that do occur.

Freeways have been heavily criticized by environmentalists, urbanists, and preservationists for the noise, pollution, and economic shifts they bring. Additionally, they have also been criticized by the driving public for the inefficiency with which they handle peak hour traffic.

Often, rural freeways open up vast areas to economic development, generally raising property values. In contrast to this, above ground freeways in urban areas are often a source of lowered property values, contributing to urban decay. Even with overpasses and underpasses, above ground freeways divide neighborhoods — especially impoverished ones where residents are less likely to own a car, or to have the political and economic influence to resist construction efforts. Beginning in the early 1970s, the U.S. Congress identified freeways and other urban highways as responsible for most of the noise exposure of the U.S. population. Subsequently, computer models were developed to analyze freeway noise and aid in their design to help minimize noise exposure.

Some cities have implemented freeway removal policies in which freeways have even been demolished and reclaimed as boulevards or parks, notably in Portland (Harbor Drive), New York City (West Side Highway), Boston (Central Artery), San Francisco (Embarcadero Freeway) and Milwaukee (Park East Freeway).

An alternative to surface or above ground freeway construction has been the construction of underground urban freeways using tunnelling technologies. This has been extremely successful in the Australian cities of Sydney (which has five such freeways) and Melbourne (which has two such freeways). This has had the benefit of removing traffic from surface roads and in the case of Melbourne's Eastlink Motorway, has helped preserve an ecologically sensitive area from destruction.

Other Australian cities face similar problems (lack of available land, cost of home acquisition, aesthetic problems, and community opposition). Brisbane, which also has to contend with physical boundaries (the river) and heavy population increases, has embraced underground tunnel freeways. There are currently three under active development, one of which (the North-South Bypass Tunnel) is currently under construction. All of the planned tunnels include provisions for public transport, whether underground or in reclaimed space on the surface.

Similarly, a below-grade freeway, Sweden's National Highway 75, the Stockholm Southern Link was constructed under an area of dense development south of downtown Stockholm and opened to traffic in 2004.

Freeway opponents have found that freeway expansion is often self-defeating: expansion simply generates more traffic. That is, even if traffic congestion is initially shifted from local streets to a new or widened freeway, people will begin to run errands and commute to more remote locations. Over time, the freeway and its environs become congested again as both the average number and distance of trips increases. This idea is known as induced demand.

Interstate H-1 eastbound into Honolulu, Hawaii.

Urban planning experts such as Drusilla Van Hengel, Joseph DiMento, and Sherry Ryan argue that although properly designed and maintained freeways may be convenient and safe, at least in comparison to uncontrolled roads, they may not expand recreation, employment and education opportunities equally for different ethnic groups, or for people located in certain neighborhoods of a given city. Still, they may open new markets to some small businesses.

Construction of urban freeways for the U.S. Interstate Highway System, which began in the late 1950s, led to the demolition of thousands of city blocks, and the dislocation of many more thousands of people. The citizens of many inner city areas responded with the freeway and expressway revolts. Through the study of Washington's response, it can be shown that the most effective changes came not from executive or legislative action, but instead from policy implementation. One of the foremost rationales for the creation of the U.S. Department of Transportation (DOT) was that an agency was needed to mediate between the conflicting interests of interstates and cities. Initially, these policies came as regulation of the state highway departments. Over time, DOT officials re-focused highway building from a national level to the local scale. With this shift of perspective came an encouragement for alternative transportation, and locally based planning agencies.

At present, freeway expansion has largely stalled in the United States, due to a multitude of factors that converged in the 1970s: higher due process requirements prior to taking of private property, increasing land values, increasing costs for construction materials, local opposition to new freeways in urban cores, the passage of the National Environmental Policy Act (which imposed the requirement that each new federally funded project must have an environmental impact statement or report), and falling gas tax revenues as a result of the nature of the flat-cent tax (it is not automatically adjusted for inflation), the tax revolt movement, and growing popular support for high-speed mass transit in lieu of new freeways.

History

The concept of limited-access automobile highways dates back to the New York City area Parkway system, whose construction began in 1907–1908; but parkways are traditionally distinguished from freeways by lower design speeds and a ban on commercial traffic. Some parkways, notably the Taconic Parkway and Saw Mill Parkway have at-grade intersections, although direct access to property adjacent to the parkways is prohibited. Designers elsewhere also researched similar ideas, especially in Germany, where the Autobahn would become the first national freeway system.

However, in 1925, Italy was technically the first country to build a freeway-like road, which linked Milan to Lake Como. It is known in Italy as the Autostrada dei Laghi.

Meanwhile, in Great Britain, the related concept of the motorway was first proposed by Sidney Webb in a 1910 book, *The King's Highway*, but was not formally embraced by the government until the passage of the Special Roads Act 1949. In 1926, the English intellectual Hilaire Belloc recognized the necessity of grade-separated roads for "rapid and heavy traffic", but thought they would be the exception rather than the rule:

> The creation of a great network of local highways suitable for rapid and heavy traffic is impossible. Even if the wealth of the community increases, the thing would be impossible, because it would mean the destruction of such a proportion of buildings as would dislocate all social life.

Connecticut's Merritt Parkway was one of the first long-distance freeways in America when it opened on June 29, 1938, but its design standards are well below modern best practices. Two years later, the Pennsylvania Turnpike opened on October 1, 1940, connecting the outskirts of Harrisburg and Pittsburgh. It was designed so that straightaways could handle maximum speeds of 102 miles per hour, and curves could be taken as fast as 90.

In stages from 1938 to 1940, California opened its first significant freeway, the Arroyo Seco Parkway (now called the Pasadena Freeway) which connected Pasadena with Los Angeles. And in 1942, Detroit, Michigan opened the world's first urban depressed freeway, the Davison Freeway. Portions of the first freeway in Texas and the Southern United States, the Gulf Freeway in Houston, opened in 1948. Meanwhile, traffic in Los Angeles continued to deteriorate and local officials began planning the huge freeway network for which the city is now famous.

Today, many freeways in the United States belong to the extensive Interstate Highway system (most of which was completed between 1960 and 1990). Starting in the 1970s freeways began to consider environmental factors, particularly noise and air quality in their location and design. Nearly all Interstate Highways are freeways. The earlier United States highway system and the highway systems of U.S. states also have many sections that are built to controlled-access standards (though these systems are mostly composed of uncontrolled roads). Only a handful of sections of the Interstate system are not freeways, such as I-81 as it crosses the American span of the 2-lane Thousand Islands Bridge and a segment of Interstate 93 through Franconia Notch, New Hampshire that is a 2-lane road with partial access control.

Recent developments

An intersection with the Southern Expressway in Adelaide, South Australia, is the world's longest reversible freeway.

Australia has been innovative in using the newest tunneling technologies to bring freeways into its high-density central business districts (Sydney, Perth and Melbourne). Brisbane currently has three major freeway tunnels under development. The city of Adelaide pioneered the concept of a dedicated reversible freeway. The M2 Southern Expressway operates toward the city centre in the morning and away from the city centre in the afternoon and evening. Its ramps are designed so that they can double as on- or off-ramps, depending upon the time of day. Gates and electronic signage prevent motorists from driving in the wrong direction.

Around the world, major progress has been made in making existing freeways and expressways more efficient. Innovations include the addition of high-occupancy vehicle lanes (HOV lanes) to discourage driving solo, and building new roads, or retro-fitting existing roads with train tracks down the median (or overhead). In the U.S., California's Caltrans has been very innovative in squeezing HOVs into limited right-of-way (by elevating them), and in building special HOV-only ramps so that HOVs can switch freeways or exit the freeway without having to merge across regular traffic. Many states have added truck-only ramps or lanes on heavily congested routes, so that cars need not weave around slow-moving big rigs.

Intelligent transportation systems are also increasingly used, with cameras to monitor and direct traffic, so that police, fire, ambulance, tow, or other assistance vehicles can be dispatched as soon as there is a problem, and to warn drivers via variable message signs, radio, television, and the Web to avoid problem areas. Research has been underway for many years on how to partly automate cars by making smart roads with such things as buried magnets to guide sensor-equipped vehicles, with on-board GPS to determine location, direction, and destination. While these systems may eventually be used on surface streets as well, they are most practical in a freeway setting.

Public-private partnerships in the United States

Until the late 1990s, funding of construction and maintenance of the Interstate Highway System was by the national gasoline tax. Originally, revenues generated by the national gasoline tax were intended solely for the maintenance and expansion of the country's highway system. During the Clinton Administration, federal legislation was passed allowing the use of gasoline tax revenues to fund other government programs and projects not related to highways or transportation. Since this reduced the amount of money available for the intended purpose of maintaining America's road network, many projects were either delayed, canceled, or scaled back.

Additionally, the original Highway Act of 1956 prohibited states from collecting tolls on Interstate-funded freeways. As more miles of freeways were completed, the cost of maintaining the infrastructure increased dramatically. A major issue that has slowed new freeway construction in America has been the application of highway funds to maintaining and repairing existing infrastructure. Most of the freeways in America are near or have exceeded their designed life span, which necessitates replacing of bridges and overpasses and reconstruction of the driving surfaces on many freeways nationwide.

To address the issue of lack of funding for new freeways and maintenance of existing roads, legislation enacted in 1998 gives states greater flexibility in funding major highway projects. Specifically the legislation, known as TEA-21 in official documents, authorizes states to add tolls to Interstate-funded freeways. Additionally, it gave states the latitude to enter into public-private partnership P3 arrangements to facilitate expansion and maintenance of the freeway network.

Texas, Florida, Virginia, and California quickly took advantage of the TEA-21 legislation and began on massive projects to expand their respective states' freeway networks, complementing existing Interstate freeways with privately funded and operated tollways. In 2004, Illinois sealed a $1.8 billion deal with Macquarie Infrastructure Group and Cintras to operate the Chicago Skyway for 99 years. In a similar P3 arrangement in Indiana, the Cintras-Macquarie joint venture assumed responsibility for the Indiana East-West Toll Road for 75 years on June 30, 2006 in a very controversial $3.8 billion deal, which for political purposes was dubbed Major Moves. As of late 2006, Pennsylvania is actively pursuing the P3 toll road concept, but still has to clear challenges in the state legislature before such an arrangement can be implemented on the Pennsylvania Turnpike.

Also in late 2006 Delaware has plans to enter into an agreement with a private firm to design, build, and operate a planned 17 mile (27 km) bypass of U.S. Route 301 between Delaware Route 1 and the Maryland state line. Meanwhile in New York and Massachusetts, the respective state public authorities that operate the New York State Thruway and Massachusetts Turnpike have generated enough revenue to assume maintenance of other freeways beyond the roads on which tolls are collected. The Massachusetts Turnpike Authority provided more than 50 percent of the funding to complete the Big Dig project in Boston, and later assumed responsibility for operating the Central Artery, the Sumner Tunnel, and the Callahan Tunnel following the project's completion in 2005.

As federal funding dries up for expanding and maintaining America's freeway network, states are looking to innovative solutions using a combination of state and federal funding, toll collection through public authorities, and private sector investment.

Construction techniques

Express to collector transfer on Highway 401, a freeway with a collector-express system.

The most frequent way freeways are laid out is usually by building them from the ground up after things such as forestry or buildings are cleared away. Sometimes they deplete farmland, but other methods have been developed for economic, social, and even environmental reasons.

Full freeways are sometimes made by converting at-grade expressways or by replacing at-grade intersections with overpasses; however, any at-grade intersection that ends a freeway remains. Often, when there is a two-lane undivided freeway or expressway, it is converted by constructing a twin corridor on the side by leaving a median between the two travel directions. The opposing side for the old two-way corridor becomes a passing lane.

Other techniques involve building a new carriageway on the side of a divided highway that has a lot of private access on one side and sometimes has long driveways on the other side since an easement for widening comes into place, especially in rural areas.

When a "third" carriageway is added, sometimes it can shift a directional carriageway by 50–200 ft (or maybe more depending on land availability) as a way to retain privrate access on one side that favors over the other. Other instances involve constructing a service drive that shortens the long driveways typically by less than 100 m.

Chapter 7

Bicycle Transportation Engineering

Engineering for bicycles - From no provision in hostile environments.

Engineering for bicycles - to dedicated cycle facilities.

Engineering for bicycles - to providing environments in which cyclists and others can share the road safely.

Bicycle transportation engineering is the study of transportation engineering as it affects bicycles and cycling. It includes the design of dedicated transport facilities for cyclists, but also the study of how mixed-mode environments (i.e. where cyclists are not segregated from other traffic) can be made to work safely.

Roads

Various methods of altering or reallocating the roadway right-of-way to facilitate bicycling and create bikeways have been added to many of the manuals used by transport planners and engineers.

Sidepath or Shared use path

A shared use path is a bikeway that is physically separated from motorized vehicle traffic by an open space (generally landscaping) or some form of barrier. While often perceived as safe and pleasant by cyclists, this type of facility can create additional hazards and problems of it's own, especially at driveways, side streets, and due to the need for extra land required compared to less extensive facilities. Depending on design and jurisdictional rules regarding crossing side streets, these paths may also be much slower for the rider, making them less attractive for sports or commutter cyclists.

Cycle track

Cycle tracks, also called *separated cycle lanes*, are bicycle exclusive facilities that provide physical separation from motorized vehicle traffic within the road right-of-way, generally by means of kerbs or other barriers. Cycle tracks can either incorporate bicycle-only signal phases at intersections (for 100% separation) or utilize "mixing zones" to merge bicycle and motor vehicle traffic. A cycle track combines the user experience of a separated path with the on-road infrastructure of a bike lane.

Bicycle lane

A designated bicycle lane, according to the 1998 United States Manual on Uniform Traffic Control Devices, is:

- a portion of a roadway or shoulder which has been designated for use by bicyclists
- marked by a white (usually solid) stripe painted on the pavement
- significantly narrower than traffic lanes
- found at the side of the traffic lanes

Cycle lanes require less space compared with the option of providing side paths or separate tracks, and generally do not have the legal / priority issues that facilities separated from the general traffic may have (depending on the locally applicable road laws). However, if traffic is heavy or fast, the lack of physical separation from motor vehicles may make cycle lanes unattractive to less confident cyclists. The lack of separation also means they are more apt to be blocked by (illegally) parked vehicles.

Non-segregated facilities

Wide outside lane (WOL)

Wide outside through lanes help motorists pass slower cyclists without having to decrease speed or change lanes. This method is held by some to be particularly important on roads with a high proportion of wide vehicles, such as buses or heavy trucks. These lanes also provide more room for cyclists to filter past queues of other vehicles in congested conditions. The use of such lanes is specifically endorsed by *Cycling: the way ahead for towns and cities*, the European Commission policy document on cycle promotion.

Cycle friendly infrastructure argues for a marked lane width of 4.25 m (14 ft). It is argued that, on undivided roads, this width provides cyclists with adequate clearance from passing wide vehicles while being sufficiently narrow to deter motorists from attempting to “double up” and form two lanes. This “doubling up” effect may be related to junctions. At non-junction locations, greater width might be preferable if this effect can be avoided.

Shared bus lane

A bus and cycle lane in Mannheim, Germany

A shared bus lane is a bus lane that allows cyclists to use it. Depending on the width of the lane, the speeds and number of buses, and other local factors, the safety and popularity of this arrangement vary.

Guidance produced for Cycling England endorses bus lanes as providing cyclists with a *direct and barrier free route into town centres* and as avoiding the difficulties associated with other provisions such as shared-use footways. According to a French survey 42% of cyclists described themselves as "enthusiasts" for shared bus bike lanes versus 33% who were of mixed opinion and 27% who were opposed. Many cycling activists view these as being more attractive than cycle paths, while others object to being in close proximity to bus exhausts. The sharing of bus lanes has been described as "generally very popular" with cyclists in London.

As of 2003, mixed bus/cycle lanes accounted for 118 km of the 260 km of cycling facilities in Paris. The French city of Bordeaux has 40 km of shared bus cycle lanes. It is reported that that in the city of Bristol, a showcase bus priority corridor, where road space was re-allocated along a 14 km stretch also resulted in more space for cyclists and had the effect of increasing cycling. The reverse effect has also been suggested, a review carried out in London reports that cycling levels fell across Kew bridge following the removal of a bus lane - this was despite a general increase in cycling level in the city generally. In addition, it is arguably easier, politically speaking, to argue for funding of joint facilities rather than the additional expense of both segregated cycling facilities and bus-only lanes.

In some instances. bus lane proposals have run into vehement opposition from cycling representatives - a typical theme is the perceived generation of conflict due to the narrowing of other lanes already shared by cars/cyclists so as to create space for the bus lanes The TRL reports that cyclists and bus drivers tend to have low opinions of each other There have been reports in Dublin of conflict as cyclists choose to cycle in the bus lanes and a bus driver apparently expected them to use adjacent cycle tracks instead. In some other cities the arrangements seem to work successfully, with bus companies and cyclists' groups taking active steps to ensure that understanding is improved between the two groups of road users.

Streets

Shared space

Shared space schemes, which are characterized by the removal of road markings, signs and signals, give all street users equal priority and equal responsibility for each others safety. Experiences where these schemes are in use, show that street users, particularly motorists, undirected by signs, curbs or road markings, reduce their speed and establish eye contact with other users. Results from the thousands of such implementations worldwide all show casualty reductions and most also show reduced journey times. Following the partial conversion of London's Kensington High Street to *shared space*, accidents were reduced there by 44% (the London average was 17%).

Bicycle boulevards

A bicycle boulevard is a low speed street which has been designed to discourage cut-through motor vehicle traffic and to give priority to to cyclists as through-going traffic.

Bike paths

Bike paths are bicycle routes that follow an independent right-of-way. There are two distinct types of bike paths: those used exclusively by bicycles and those shared with pedestrians. In the United States almost all bike paths are shared with pedestrians.

Bike paths are often found on old railroad right-of-ways, called rail trails.

Traffic lights

How traffic signals are designed and implemented directly impacts cyclists. For instance, poorly adjusted vehicle detector systems, used to trigger signal changes, may not correctly detect cyclists. Traffic managers in Copenhagen link cyclist-specific traffic signals on a major arterial bike lane to provide green waves for rush hour cycle-traffic. Cycling-specific measures that can be applied at traffic signals include the use of advanced stop lines and/or bypasses.

The frequency with which lights change is important to cyclists who may conserve energy by anticipating green lights ahead i.e. the shorter the interval the better for cyclists. Intersection clearance times and green wave timing may also be relevant.

Road surface

Bicycle tires being narrow, road surface is more important than for other transport, for both comfort and safety. The type and placement of storm drains, manholes, surface markings, and the general road surface quality should all be taken into account by a bicycle transportation engineer. Drain grates, for example, must not catch wheels.

Parking

Efficient two tiered Dutch Bicycle Rack

Bicycle parking is another important part of Bicycle Transportation Engineering. In most of the United States, bicycle parking facilities are scarce, or are so inadequate that nearby trees or parking meters are used. The hitching post type of bicycle stand is an improvement over the old type that had a slot for the front wheel but only allow for two bicycles per post. The Netherlands, where bicycles are much in use, has two-tiered bicycle racks giving high density (the handlebars overlap) and security (the bicycle is held well and is easy to lock).

Improved parking meter

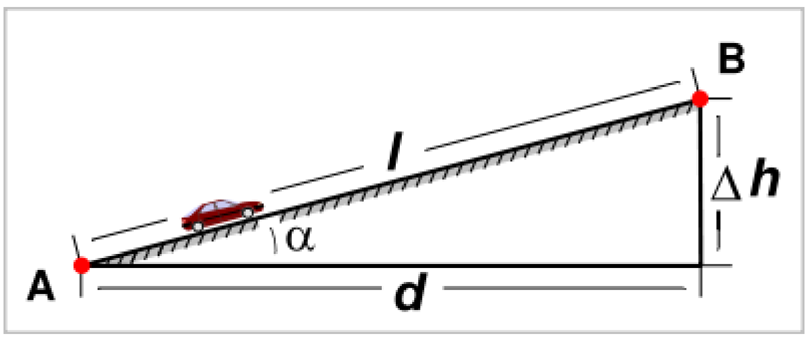
Secure bicycle parking is argued to be a key factor influencing the decision to cycle. To be considered secure, the parking must be of a suitable design: allowing the bicycle to be locked via the frame. A readily observable location can also permit so-called passive security from passers-by. Weather protection is also desirable. As a rule, where cycling is encouraged as an alternative to motoring, efforts are made to make bicycle parking more convenient and attractive to use than nearby car parking arrangements. This usually means providing a wide distribution of visible, clearly designated parking spots, close to the entrances of destinations being served.

Storage rooms or bicycle lockers may also be provided. In some cases large concentrations of bike parking may be more appropriate, sometimes being supervised and sometimes charging a fee. Examples include large bike parking facilities at public transport interchanges such as railway, subway, tram, bus stations, or at bike stations where they may be useful in Mixed-mode commuting.

Conversely, where cycling is seen as an unwelcome or inappropriate activity, or due to lack of knowledge about best practices, bicycle parking may simply not be provided or else placed at awkward, distant, and out-of-sight locations. Cyclists may be expressly forbidden from parking their bicycles at the most convenient locations. In April 2007, the authorities at the University of California's Santa Barbara campus started confiscating bicycles not parked at the allegedly inconvenient official bike racks. Often, property owners display signage on fencing to discourage bicyclists from locking their bicycles.

Chapter 8

Grade (Slope)



d = run
Δh = rise
l = slope length
α = angle of inclination

The **grade** (also called **slope**, **incline**, **gradient**, **pitch** or **rise**) of a physical feature, topographic landform or constructed element, refers to the amount of inclination of that surface to the horizontal. It is a special case of the gradient in calculus where zero indicates *gravitational level*. A larger number indicates higher or steeper degree of "tilt". Often slope is calculated as a ratio of "rise" to "run", or as a fraction ("rise over run") in which *run* is the horizontal distance and *rise* is the vertical distance.

Grade or slope is applied to measuring existing physical features (such as canyon and hillsides, stream and river banks and beds), or in designing and engineering new elements for construction (such as roads, landscape and garden grading, roof pitches, railroads, aqueducts, and pedestrian-handicapped-bicycle circulation routes).

Expression nomenclature

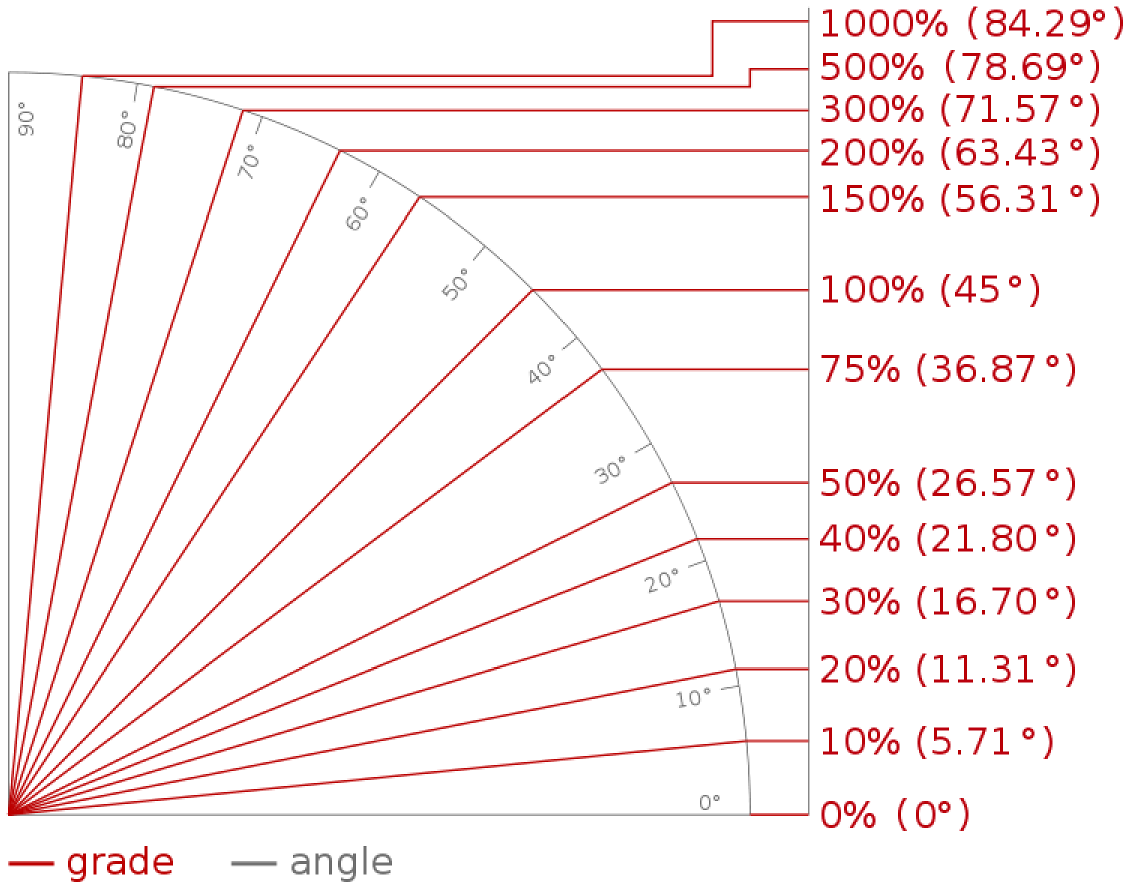


Illustration of grades in percent and angles in degrees

There are several systems for expressing slope:

1. as an *angle* of inclination to the horizontal. (This is the angle α opposite the "rise" side of a triangle with a right angle between vertical rise and horizontal run.)
2. as a *percentage*, the formula for which is $100\frac{\text{rise}}{\text{run}}$ which could also be expressed as the tangent of the angle of inclination times 100. In the U.S., this percentage "grade" is the most commonly used unit for communicating slopes in transportation, surveying, construction, and civil engineering.
3. as a *per mille* figure, the formula for which is $1000\frac{\text{rise}}{\text{run}}$ which could also be expressed as the tangent of the angle of inclination times 1000. This is commonly used in Europe to denote the incline of a railway.
4. as a *ratio* of one part rise to so many parts run. For example, a slope that has a rise of 5 feet for every 100 feet of run would have a slope ratio of 1 in 20. (The word "in" is normally used rather than the mathematical ratio notation of "1:20")

Any one of these expressions may be used interchangeably to express the characteristics of a slope. Grade is usually expressed as a percentage, but this may easily be converted to the angle α from horizontal since that carries the same information.

There is a method in which slope may be expressed when the horizontal run is not known: rise divided by the hypotenuse (the slope length). This is *not* a usual way to measure slope. This follows the sine function rather than the tangent function and this method diverges from the "rise over run" method as angles start getting larger.

Many of the mathematical principles of slope that follow from the definition are applicable in topographic practice. In the UK, for road signs, maps and construction work, the gradient was traditionally expressed as a ratio such as 1 in 12, but signs showing gradient expressed as a percentage are becoming more common.

In civil engineering applications and physical geography, the slope is calculated along a particular direction of interest which is normally the route of a highway or railway road bed.

Mathematical equations

Grades can be related using the following equations with symbols from the figure at top.

Tangent as a ratio

$$\tan\alpha = \frac{\Delta h}{d}$$

This ratio can also be expressed as a percentage by multiplying by 100.

Angle from a tangent gradient

$$\alpha = \arctan\frac{\Delta h}{d}$$

If the tangent is expressed as a percentage, the angle can be determined as:

$$\alpha = \arctan\frac{\%\,\text{slope}}{100}$$

If the angle is expressed as a ratio *(1 in n)* then:

$$\alpha = \arctan\frac{1}{n}$$

Roads

In vehicular engineering, various land-based designs (cars, SUVs, trucks, trains, etc.) are rated for their ability to ascend terrain. (Trains typically rate much lower than cars.) The highest grade a vehicle can ascend while maintaining a particular speed is sometimes termed that vehicle's "gradeability" (or, less often, "grade ability"). The lateral slopes of a highway geometry are sometimes called fills or cuts where these techniques have been used to create them.

slope warning sign, Netherlands

25% slope sign, Wales

A 1371-metre long stretch of railroad with a 20‰ (2%) slope, Czech Republic

Slope warning sign, 30 % over 1500 m. La route des Crêtes, Cassis, France

Environmental design

Grade, pitch, and slope are important components in landscape design, garden design, landscape architecture, and architecture; for engineering and aesthetic design factors. Drainage, slope stability, circulation of people and vehicles, complying with building codes, and design integration are aspects of slope considerations in environmental design.

Railways

Grade indicator near Bellville, Western Cape, South Africa, showing 1:150 and 1:88 grades.

Steep gradients limit the size of load that a locomotive can haul, including the weight of the locomotive itself. A 1% gradient (1 in 100) halves the load. Early railways in the United Kingdom were laid out with very gentle gradients, such as 0.05% (1 in 2000), because the early locomotives (and their brakes) were so feeble. Steep gradients were concentrated in short sections of lines where it was convenient to employ assistant engines or cable haulage, such as from Euston to Camden Town, about 1.2 km. Extremely steep gradients need the help of cables, or some kind of rack railway.

The steepest non-rack railway lines include:

- 13.5 % - Lisbon tram, Portugal
- 11.6 % - Pöstlingbergbahn, Linz, Austria
- 11 % Cass Scenic Railway USA (former logging line)
- 9.0 % - Ligne de Saint Gervais - Vallorcine, France
- 7% - Bernina Railway, Switzerland
- 5.6% (1 in 18) - Flåm, Norway.
- 5.1% - Saluda Grade, North Carolina, United States
- 4.0% - Cologne-Frankfurt high-speed rail line
- 4.0% (1 in 25) - Tarana - Oberon branch, New South Wales, Australia.

- 4.0% (1 in 25) - Matheran Light Railway, India
- 3.7% (1 in 27) - Ecclesbourne Valley Railway, Heritage Line, Wirksworth, Derbyshire, UK

It is customary for civil engineers to refer to the steepest grade on a section of rail line as the ruling grade for that section. Civil engineering works such as cuttings, embankments and tunnels are employed to reduce this maximum grade.

Compensation for curvature

Gradients on sharp curves are effectively a bit steeper than the same gradient on straight track, so to "compensate" for this and make the ruling grade uniform throughout, the gradient on those sharp curves should be reduced slightly.

Effects of grade

The greater a grade, the more power an animal or a machine requires to climb it; therefore routes with lower grades are preferred, so long as they do not have other disadvantages, such as causing the overall travel distance to increase significantly.

Vehicles proceeding up a steep grade demand more fuel consumption with typically increased air pollution generation. Sound level increases are also produced by motor vehicles travelling upgrade.

Chapter 9

Geometric Design of Roads

The **geometric design of roadways** deals with the portioning of the physical elements of the roadway according to standards and constraints. The basic objective in geometric design is to provide a smooth-flowing, crash-free facility. The American Association of State Highway and Transportation Officials (AASHTO) has established guidelines relating to the design of roadways. Geometric roadway design can be broken into two main parts: vertical curves and horizontal curves. Combined, vertical and horizontal curves provide a three-dimensional layout for a roadway. Using guidelines provided by AASHTO, an engineer can design a roadway that is comfortable, safe, and appealing to the eye. The AASHTO guidelines take into account speed, vehicle type, road grade (slope), view obstructions, and stopping distance.

Vertical Curves

Vertical curves are used to provide a gradual change from one road slope to another, so that vehicles may smoothly navigate grade changes as they travel.

Sag vertical curves are those that have a tangent slope at the end of the curve that is higher than that of the beginning of the curve. When driving on a road, a sag curve would appear as a valley, with the vehicle first going downhill before reaching the bottom of the curve and continuing uphill or level.

Crest vertical curves are those that have a tangent slope at the end of the curve that is lower than that of the beginning of the curve. When driving on a crest curve, the road appears as a hill, with the vehicle first going uphill before reaching the top of the curve and continuing downhill.

Terminology

G1 = initial roadway (tangent)slope

G2 = final roadway (tangent)slope

A = absolute value of the difference in grades (initial minus final, expressed in percent)

L = curve length (along the x-axis)

BVC = begin of vertical curve

PVI = point of vertical interception (intersection of initial and final grades)

EVT = end of vertical tangent

x = distance from PTC/BVC

tangent elevation = elevation of a point along the initial tangent

Y (offset) = vertical distance from the initial tangent to the curve

Y' = curve elevation = tangent elevation - offset

Sag Curves

Sag vertical curves are curves which have a negative A-value - that is, the tangent slope at the end of the curve is greater than the tangent slope at the beginning of the curve. The most important design criteria for these curves is headlight sight distance. When a driver is driving on a sag curve at night, the sight distance is limited by the higher grade in front of the vehicle. This distance must be long enough that the driver can see an obstruction in the road and stop the vehicle within the headlight sight distance. The headlight sight distance (S) is determined by the angle of the headlight and angle of the tangent slope at the end of the curve. By first finding the headlight sight distance (S) and then solving for the curve length (L) in each of the equations below, the correct curve length can be determined. If the S<L curve length is greater than the headlight sight distance, then this number can be used. If it is smaller, this value cannot be used. Similarly, if the S>L curve length is smaller than the headlight sight distance, then this number can be used. If it is larger, this value cannot be used.

Sight Distance < Curve Length (S<L)

$L = AS^2 / (400 + 3.5S)$

Sight Distance > Curve Length (S>L)

$L = 2S - (400 + 3.5S) / A$

Crest Curves

Crest vertical curves are curves which have a positive A-value - that is, the tangent slope at the end of the curve is less than the tangent slope at the beginning of the curve. The most important design criteria for these curves is stopping sight distance. As a vehicle comes over the top of a crest vertical curve, the driver must be able to see any obstructions that lie ahead within a safe stopping distance. If a car on the other side of the hill is stalled or there is an animal in the vehicle's lane, the driver must see the obstruction and be able to stop the car within the stopping sight distance. The stopping sight distance (S) is determined by the speed limit on a road. By first finding the stopping sight distance (S) and then solving for the curve length (L) in each of the equations below, the correct curve length can be determined. If the S<L curve length is greater than the stopping sight distance, then this number can be used. If it is smaller, this value cannot be used. Similarly, if the S>L curve length is smaller than the stopping sight distance, then this number can be used. If it is larger, this value cannot be used.

Sight Distance > Curve Length (S>L)

$L = 2S - (2158 / A)$

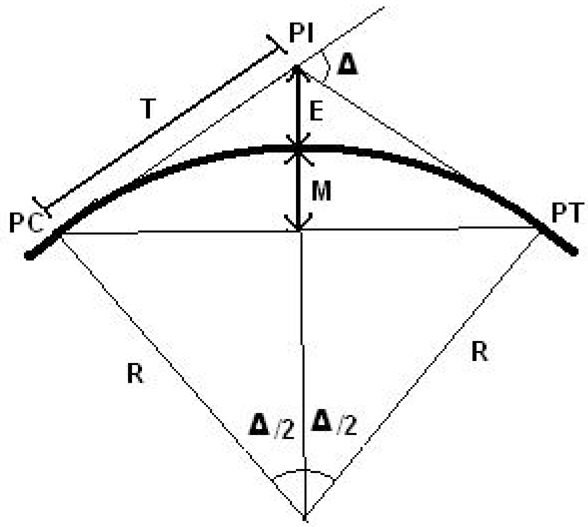
Sight Distance < Curve Length (S<L)

$L = AS^2 / 2158$

Horizontal Curves

Horizontal alignment in road design consists of straight sections of road, known as tangents, connected by horizontal curves. The design of a horizontal curve entails the determination of a minimum radius (based on speed limit), curve length, and objects obstructing the view of the driver. Using AASHTO standards, an engineer works to design a road that is safe and comfortable. If a horizontal curve has a high speed and a small radius of curvature, an increased superelevation (bank) is needed in order to assure safety. If there is an object obstructing the view around a corner or curve, the engineer must work to ensure that drivers can see far enough to stop to avoid an accident or accelerate to join traffic.

Terminology



R = Radius

PC = Point of Curvature (point at which the curve begins)

PT = Point of Tangent (point at which the curve ends)

PI = Point of Intersection (point at which the two tangents intersect)

T = Tangent Length

C = Long Chord Length (straight line between PC and PT)

L = Curve Length

M = Middle Ordinate, now known as HSO - Horizontal Sightline Offset (distance from sight-obstructing object to the middle of the outside lane)

e = Rate of Superelevation

fs= Coefficient of Side Friction

u = Vehicle Speed

Δ = Deflection Angle

Geometry

$T = R\tan(\Delta / 2)$

$C = 2R\sin(\Delta / 2)$

$L = R\Delta\pi / 180$

Sight Distance

$M = R(1 - \cos((28.65S) / R))$

Superelevation

This concept deals with the horizontal or lateral(cross) slope of a roadway. Superelevation of a roadway is used in order to accomplish two objectives: to provide adequate drainage flow for water on the roadway surface and to aid in steering of motor vehicles on curved roadways.

In a tangent section, a common superelevation or cross slope of 1-2% is applied in order to achieve drainage flow of surface water off of the subject roadway. Cross slopes of this magnitude especially when applied in both directions of travel with a crown point along the centerline of a roadway are commonly referred to as "normal crown" and are generally imperceptible to traveling motorists. In a curved section, calculated superelevations generally reaching a maximum of 6 or 8% (depending on selected design criteria) are applied in order to aid motorists in safely traversing these sections while maintaining entry speed of the vehicle along the length of the curve.

The equation for the desired radius of a curve is found below taking into account factors for speed and superelevation rates (e) of a given roadway section. This equation can be rearranged algebraically to provide instead for desired rates of superelevation based on the design speed of a roadway and the radius to be used in a curved section.

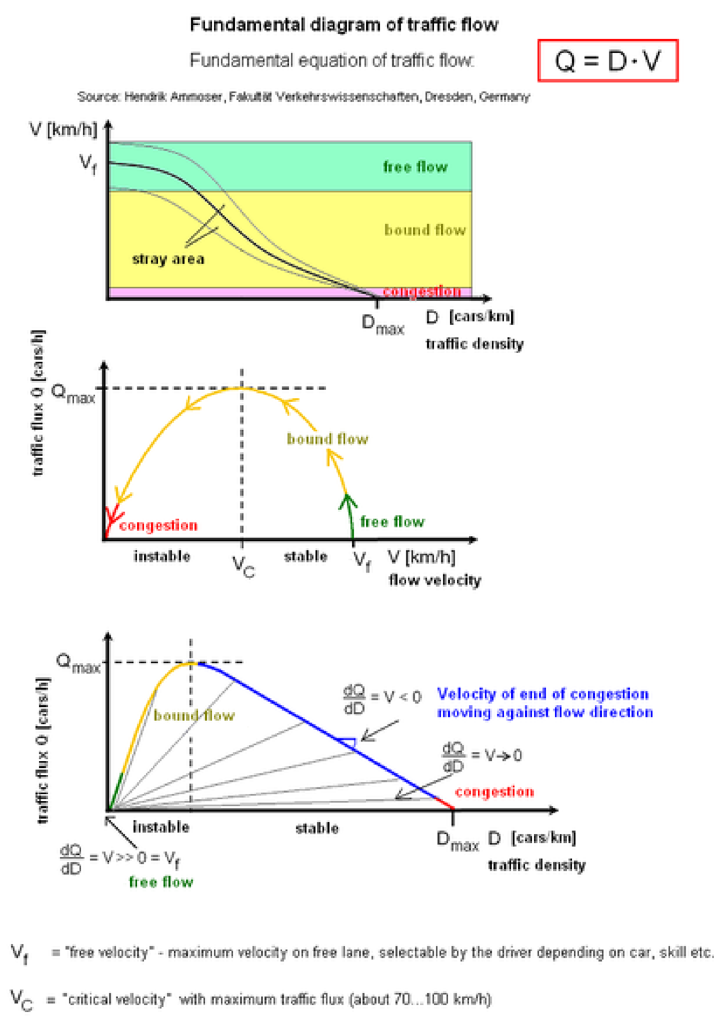
$R = u^2 / (15(e + fs))$

The American Association of State Highway and Transportation officials (AASHTO) provides a table from which desired superelevation rates can be easily interpolated based on the prescribed magnitudes for both design speed and radius of a curved section of roadway. This table can also be seen reprinted in many state roadway design guides and manuals in the U.S.

Chapter 10

Fundamental Diagram of Traffic Flow

The **fundamental diagram of traffic flow** is a diagram that gives a relation between the traffic flux (vehicles/hour) and the traffic density (vehicles/km). A macroscopic traffic model involving traffic flux, traffic density and velocity forms the basis of the fundamental diagram. It can be used to predict the capability of a road system, or its behaviour when applying inflow regulation or speed limits.



Fundamental Diagram of traffic flow

Basic statements

- There is a connection between traffic density and vehicle velocity: The more vehicles are on a road, the slower their velocity will be.
- To prevent congestion and to keep traffic flow stable, the number of vehicles entering the control zone has to be smaller or equal to the number of vehicles leaving the zone in the same time.
- At a critical traffic density and a corresponding critical velocity the state of flow will change from stable to unstable.
- If one of the vehicles brakes in unstable flow regime the flow will collapse.

The primary tool for graphically displaying information in the study traffic flow is the fundamental diagram. Fundamental diagrams consist of 3 different graphs: flow-density, speed-flow, and speed-density. The graphs are two dimensional graphs. All the graphs are related by the equation "flow = speed * density"; this equation is the essential equation in traffic flow. The fundamental diagrams were derived by the plotting of field data points and giving these data points a best fit curve. With the fundamental diagrams researchers can explore the relationship between speed, flow , and density of traffic.

Speed-Density

The speed-density relationship is linear with a negative slope; therefore, as the density increases the speed of the roadway decreases. The line crosses the speed axis, y, at the freeflow speed, and the line crosses the density axis, x, at the jam density. Here the speed approaches freeflow speed as the density approaches zero. As the density increases, the speed of the vehicles on the roadway decreases. The speed reaches zero when the density equals the jam density.

Flow -Density

In the study of traffic flow theory, the flow-density diagram is used to determine the traffic state of a roadway. Currently, there are two types of flow density graphs. The first is the parabolic shaped flow-density curve, and the second is the triangular shaped flow density curve. Academia views the triangular shaped flow-density curve as more the accurate representation of real world events. The triangular shaped curve consists of two vectors. The first vector is the freeflow side of the curve. This vector is created by placing the freeflow velocity vector of a roadway at the origin of the flow-density graph. The second vector is the congested branch, which is created by placing the vector of the shock wave speed at zero flow and jam density. The congested branch has a negative slope, which implies that the higher the density on the congested branch the lower the flow; therefore, even though there are more cars on the road, the number of cars passing a single point is less than if there were fewer cars on the road. The intersection of freeflow and congested vectors is the apex of the curve and is considered the capacity of the roadway, which is the traffic condition at which the maximum number of vehicles can pass by a point in a given time period. The flow and capacity at which this point occurs is the optimum flow and optimum density, respectively. The flow density diagram is used to give the traffic condition of a roadway. With the traffic conditions, time-space diagrams can be created to give travel time, delay, and queue lengths of a road segment.

Speed-Flow

Speed flow diagrams are used to determine the speed at which the optimum flow occurs. There are currently two shapes of the speed-flow curve. The speed-flow curve also consists of two branches, the freeflow and congested branches. The diagram is not a function, allowing the flow variable to exist at two different speeds. The flow variable existing at two different speeds occurs when the speed is higher and the density is lower or when the speed is lower and the density is higher, which allows for the same flow rate. In the first speed-flow diagram, the freeflow branch is a horizontal line, which shows that the roadway is at freeflow speed until the optimum flow is reached. Once the optimum flow is reached, the diagram switches to the congested branch, which is a parabolic shape. The second speed flow diagram is a parabola. The parabola suggests that the only time there is freeflow speed is when the density approaches zero; it also suggests that as the flow increases the speed decreases. This parabolic graph also contains an optimum flow. The optimum also flow divides the freeflow and congested branches on the parabolic graph.

Chapter 11

Banked Turn

A **banked turn** is a turn or change of direction in which the vehicle banks or inclines, usually towards the inside of the turn. The bank angle is the angle at which the vehicle is inclined about its longitudinal axis with respect to its path.

Turn on flat surfaces

If the bank angle is zero the surface is flat, simplifying the calculations, and the vehicle is just driving in a circle. The normal force is vertically upwards, so the only force keeping the vehicle turning on its path is friction, or traction. This must be large enough to provide the centripetal acceleration, a relationship which can be expressed as an inequality,

$$\mu mg > \frac{mv^2}{r}.$$

The expression on the right hand side is the centripetal acceleration multiplied by mass, so the force required to turn the vehicle. The left hand side is the maximum frictional force, which equals the coefficient of friction μ multiplied by the normal force. Rearranging the maximum cornering speed is

$$v < \sqrt{r\mu g}.$$

Note that μ can be the coefficient for static or dynamic friction. In the latter case, where the vehicle is skidding around a bend, the friction is at its limit and the inequalities becomes equations. This also ignores effects such as downforce which can increase the normal force and cornering speed.

Normal reaction in a banked turn

In the case of a car being parked on a banked turn, the Normal force would simply be:

$mg\cos\theta$

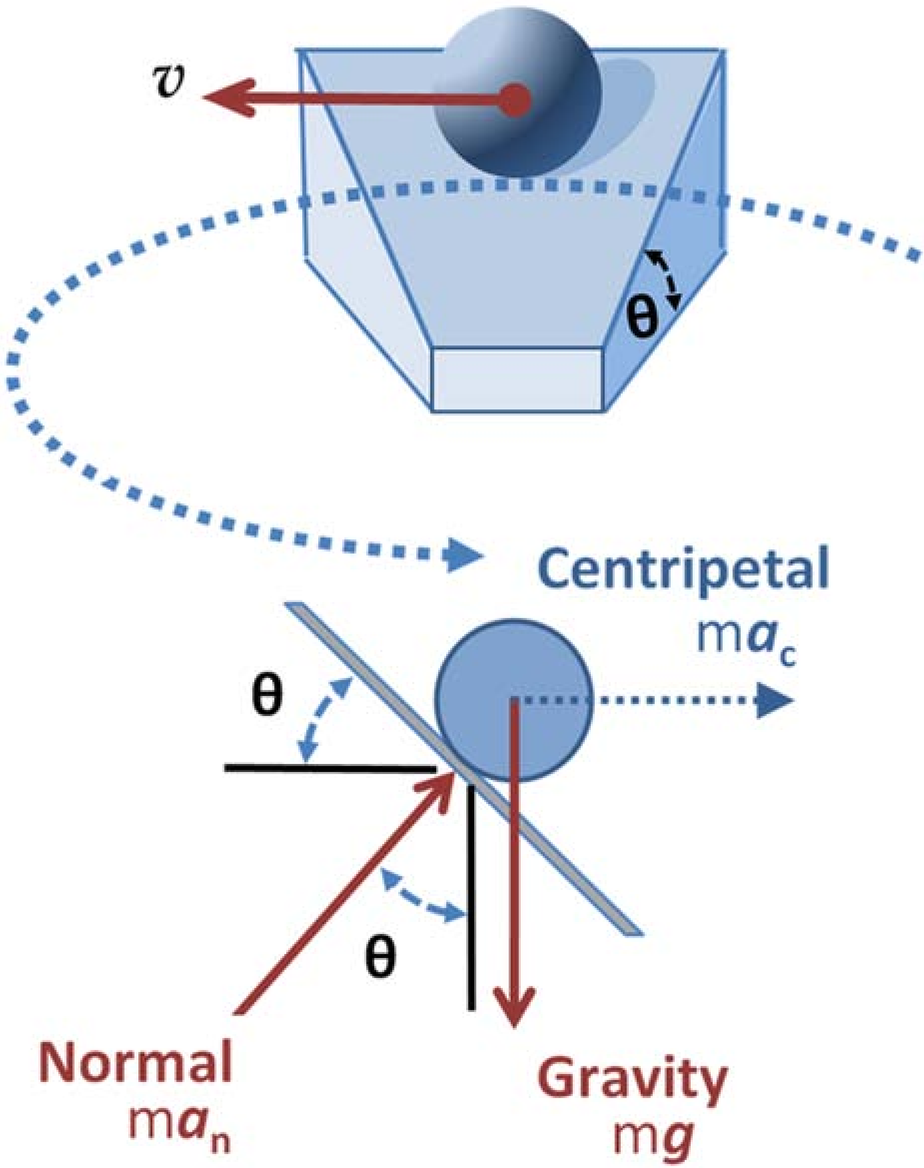
But once the car starts to move on a banked turn, it 'collides' with the turn itself, the turn feels this force, and returns it in the normal, causing the car to move circularly.

The Normal is thus greater than simply its gravitational component.

$$N = mg\cos\theta + \frac{mv^2}{r}\sin\theta$$

If the Normal were simply the component to gravity, one could not say mg = $N\cos\theta$ = $mg\cos^2\theta$ when the vertical acceleration is 0.

Frictionless banked turn



Upper panel: Ball on a banked circular track moving with constant speed *v*; Lower panel: Forces on the ball. The resultant or net force on the ball found by vector addition of the normal force exerted by the road and vertical force due to gravity must equal the required force for centripetal acceleration dictated by the need to travel a circular path.

As opposed to a car riding along a flat circle, inclined edges add an additional force that keeps the car in its path and prevents it from being "dragged into" or "pushed out of" the circle. This force is the horizontal component of the car's normal force. In the absence of friction, the normal force is the only one acting on the car in the direction of the center of the circle. Therefore, as per Newton's second law, we can set the horizontal component of the normal force equal to mass multiplied by centripetal acceleration:

$$N \sin\theta = \frac{mv^2}{r}$$

Because there is no motion in the vertical direction, the sum of all vertical forces acting on the system must be zero. Therefore we can set the vertical component of the car's normal force equal to its weight:

$$N\cos\theta = mg$$

Solving the above equation for the normal force and substituting this value into our previous equation, we get:

$$\frac{mv^2}{r} = mg\tan\theta$$

Solving for velocity we have:

$$v = \sqrt{rg\tan\theta}$$

This provides the velocity that in the absence of friction and with a given angle of incline and radius of curvature, will ensure that the car will remain in its designated path. The magnitude of this velocity is also known as the "rated speed" of a turn or curve. Notice that the rated speed of the curve is the same for all massive objects, and a curve that is not inclined will have a rated speed of 0.

Banked turn with friction

When considering the effects of friction on the system, once again we need to note which way the friction force is pointing. When calculating a maximum velocity for our automobile, friction will point down the incline and towards the center of the circle. Therefore we must add the horizontal component of friction to that of the normal force. The sum of these two forces is our new net force in the centripetal direction:

$$\frac{mv^2}{r} = \mu_s N\cos\theta + N\sin\theta$$

Once again, there is no motion in the vertical direction, allowing us to set all opposing vertical forces equal to one another. These forces include the vertical component of the normal force pointing upwards and both the car's weight and vertical component of friction pointing downwards:

$$N\cos\theta = \mu_s N\sin\theta + mg$$

By solving the above equation for mass and substituting this value into our previous equation we get:

$$\frac{v^2\left(N\cos\theta-\mu_s N\sin\theta\right)}{rg}=\mu_s N\cos\theta+N\sin\theta$$

Solving for v we get:

$$v=\sqrt{\frac{rg\left(\sin\theta+\mu_s\cos\theta\right)}{\cos\theta-\mu_s\sin\theta}}$$

This equation provides the maximum velocity for the automobile with the given angle of incline, coefficient of static friction and radius of curvature. By a similar analysis of minimum velocity, the following equation is rendered:

$$v=\sqrt{\frac{rg\left(\sin\theta-\mu_s\cos\theta\right)}{\cos\theta+\mu_s\sin\theta}}$$

The difference in the latter analysis comes when considering the direction of friction for the minimum velocity of the automobile (towards the outside of the circle). Consequently opposite operations are performed when inserting friction into equations for forces in the centripetal and vertical directions.

Improperly banked road curves increase the risk of run-off-road and head-on crashes. A 2% deficiency in superelevation (say, 4% superelevation on a curve that should have 6%) can be expected to increase crash frequency by 6%, and a 5% deficiency will increase it by 15%. Up until now, highway engineers have been without efficient tools to identify improperly banked curves and to design relevant mitigating road actions. A modern profilograph can provide data of both road curvature and cross slope (angle of incline).

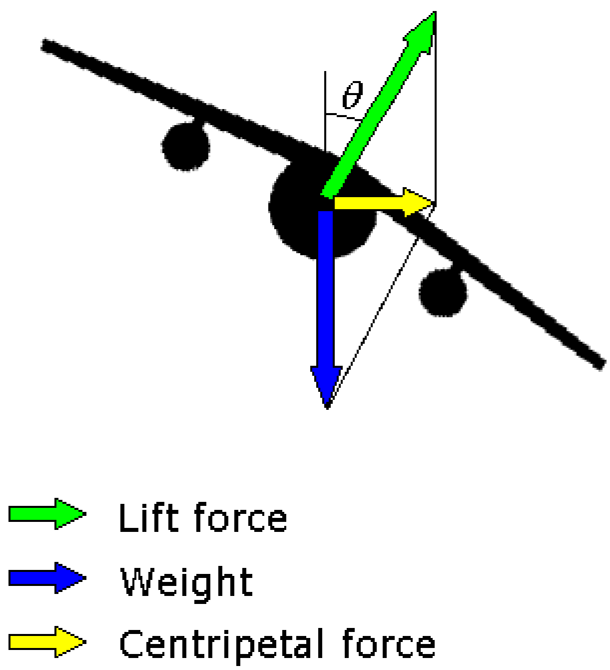
Banked turn in aeronautics

Douglas DC-3 banking to make a left turn

When a fixed-wing aircraft is making a turn (changing its direction) the aircraft must roll to a banked position so that its wings are angled towards the desired direction of the turn. When the turn has been completed the aircraft must roll back to the wings-level position in order to resume straight flight.

When any moving vehicle is making a turn, it is necessary for the forces acting on the vehicle to add up to a net inward force, to cause centripetal acceleration. In the case of an aircraft making a turn, the force causing centripetal acceleration is the horizontal component of the lift acting on the aircraft.

In straight, level flight, the lift acting on the aircraft acts vertically upwards to counteract the weight of the aircraft which acts downwards. During a balanced turn where the angle of bank is θ the lift acts at an angle θ away from the vertical. It is useful to resolve the lift into a vertical component and a horizontal component. If the aircraft is to continue in level flight (i.e. at constant altitude), the vertical component must continue to equal the weight of the aircraft. The horizontal component is unbalanced, and is thus the net force causing the aircraft to accelerate inward and execute the turn.



Vector diagram showing lift, weight and centripetal force acting on a fixed-wing aircraft during a banked turn.

During a banked turn in level flight the lift on the aircraft must support the weight of the aircraft, as well as provide the necessary component of horizontal force to cause centripetal acceleration. Consequently, the lift required in a banked turn is greater than that one required in straight, level flight and can be achieved either by increasing the angle of attack of the wing (typically by pulling on the elevator control) or by deploying flaps. The maneuver is usually complemented by an increase in power, in order to maintain airspeed.

Because centripetal acceleration is:

$$a = \frac{v^2}{r}$$

Newton's second law in the horizontal direction can be expressed mathematically as:

$$L \sin\theta = \frac{mv^2}{r}$$

where:

L is the lift acting on the aircraft
θ is the angle of bank of the aircraft
m is the mass of the aircraft
v is the true airspeed of the aircraft
r is the radius of the turn

In straight flight, lift is approximately equal to the aircraft weight. In turning flight the lift exceeds the aircraft weight, and is equal to the weight of the aircraft (*mg*) divided by the cosine of the angle of bank:

$$L = \frac{mg}{\cos\theta}$$

where *g* is the gravitational field strength.

The radius of the turn can now be calculated:

$$r = \frac{v^2}{g\tan\theta}$$

This formula shows that the radius of turn is proportional to the square of the aircraft's true airspeed. With a higher airspeed the radius of turn is larger, and with a lower airspeed the radius is smaller.

This formula also shows that the radius of turn is inversely proportional to the angle of bank. With a higher angle of bank the radius of turn is smaller, and with a lower angle of bank the radius is greater.

The angle of bank is the sole determinant of the aircraft's load factor during the turn.

Chapter 12

Red Light Camera

A red light camera in Chicago.

A red-light camera in use in Beaverton, Oregon, USA

A **red light camera** is a traffic camera which captures an image of a vehicle moving through a signalized intersection traffic moving in the direction of that vehicle has a red signal. Red light cameras enforce traffic laws and deter violation to traffic laws by photographing vehicles passing through signalized intersections illegally. Red light cameras automatically photograph vehicles whose drivers run red lights. The cameras are connected to the traffic signal and to sensors that monitor traffic flow just before the crosswalk or stop line. The system continuously monitors the traffic signal, and the camera captures any vehicle that doesn't stop during the red phase. Many red light camera programs provide motorists with grace periods of up to half a second after the light switches to red.

Red light cameras make great engineering tools too. The red light running data they collect expose engineering failures such as short yellows, dilemma zone problems and over-sensitive right-turn-on-red actuators. As the red light camera data shows, most red light running is caused by engineering defects, including the defect of the dilemma zone which departments of transportation build into every intersection in the world. This defect guarantees a steady stream of cars running red lights. Today's usage of red light cameras is as an enforcement device. The usage is based on the premise that red light running is only caused by bad driving, as opposed to bad engineering.

Depending on the particular technology, a series of photographs and/or a video clip show the red light prior to a car entering the intersection on a red signal, as well as the vehicle's progression through the intersection. Cameras record the date, time of day, time elapsed since the beginning of the red signal, vehicle speed, and license plate. Tickets typically are mailed to owners of violating vehicles, based on a review of photographic evidence.

Fines carried out against drivers for red light camera-caught violations in the U.S. are regulated at the State and City level and range from $50 in North Carolina to over $500 in California.

Studies

In the United States of America, a number of studies have addressed the question of whether, on balance, red-light cameras produce a safety benefit.

Most recently, the Insurance Institute for Highway Safety found a reduction in the rate of fatal crashes at signalized intersections both with and without red light cameras in cities where red light cameras are in use. After controlling for population density and land area, the rate of fatal red light running crashes during 2004-08 for cities with camera programs was an estimated 24 percent lower than what would have been expected without cameras. However Raleigh, North Carolina and Bakersfield, California showed increases. The rate of all fatal crashes at signalized intersections during 2004-08 for cities with camera programs was an estimated 17 percent lower than what would have been expected without cameras.

A July 2004 study of 17,271 crashes by Mark Burkey, PhD., and Kofi Obeng, PhD of North Carolina A & T University showed that the presence of red light cameras only increased the overall number of crashes by 40%. "The results do not support the view that red light cameras reduce crashes. Instead, we find that RLCs are associated with higher levels of many types of severity categories of crashes." However, this research received no scientific peer review and exhibited major research flaws.

A U.S. study , for example, found that red light cameras led to a decrease in right-angle crashes and an increase in the number of rear-end collisions. The total number of collisions remained essentially unchanged. The study applied estimates from a 1997 study of the cost of accidents based on severity to conclude the cameras yielded a positive overall cost benefit from a reduction in more expensive right-angle injury collisions.

This FHWA study was criticized in two follow-up studies. In addition to a study co-director having performed research for the Insurance Institute for Highway Safety, which represents an industry that profits significantly from red light camera surcharges, the follow up studies found critical methodological flaws and that the FHWA study completely missed an increase in fatalities associated with red light camera use. Additionally, "the authors spotlight the statistical difficulties of including the cost of fatalities, while ignoring the practical implications of such events", assuming that each angle injury crash had a societal cost of $64,468, when in fact the cost was $82,816 before camera use and $100,176 after camera use.

A 2005 Virginia Department of Transportation study of the long-term effect of camera enforcement in the state found a decrease in the number of right-angle crashes, but an increase in rear-end crashes and an overall increase in the number of accidents causing injuries .

A 2004 Texas Transportation Institute study found, "crashes decrease with an increase in yellow interval duration and a reduction in speed limit." After 1.0 second was added to the yellow signal timing at test intersections, accidents dropped by 35 to 40%. This compares with a 6.4% reduction for "area-wide officer enforcement of intersection traffic control devices... during the time of the enforcement activity" .

A 2005 study of the Raleigh, North Carolina red light camera program conducted by the Institute for Transportation Research and Education at North Carolina State University compared "before" and "after" red-light camera intersection data and found right-angle crashes dropped by 42 percent, rear-end crashes dropped by 25 percent and total accidents dropped by 17 percent . However, this contradicts a later report done by the Institute of Highway Safety (February 2011) which showed that crashes and fatalities increased in Raleigh. Also, the study contradicts the experience of several other cities such as:

- Albuquerque, NM, which found that crashes increased at camera-monitored intersections while declining at unmonitored intersections.
- Aurora, CO, which after starting camera enforcement at 4 intersections, saw 100% crash increases at two, 175% crash increase at one, and a 60% crash decrease at one intersection.
- Chicago, IL, where "there is no evidence that the red light camera have had a significant safety benefit".
- Los Angeles, CA, where cameras "failed to adequately demonstrate an improvement in safety".
- Minneapolis, MN, which saw a 11% reduction in crashes after disabling its red light camera system.
- Winnipeg, Canada, where crashes significantly increased in the years following red light camera introduction.

A 2005 meta analysis compared the results of 10 controlled before-after studies of red light cameras in the United States, Australia and Singapore. "Five studies found that use of red-light cameras cut the number of crashes in which there were injuries. In the best conducted of these studies, the reduction was nearly 30%... The evidence is less conclusive on total collisions, specific casualty collision types and violations, where reductions achieved could be explained by the play of chance. Most evaluations did not adjust for RTM or spillover, affecting their accuracy. Larger and better controlled studies are needed."

Usage

As of May 2009, red light cameras are used in more than 400 communities in the United States. Major cities that use red light cameras include Albuquerque, Atlanta, Austin, Baltimore, Baton Rouge, Chicago, Dallas, Denver, Houston, Los Angeles, New Orleans, New York City, Philadelphia, Phoenix, Raleigh, San Diego, San Francisco, Seattle, Tucson and Washington, DC.

On November 2, 2010, four years after the original installation, Houston, Texas voted 53% to 47% against the continued use of red cameras.

Some states have chosen to prohibit the use of red light cameras. These include Connecticut, Nevada, New Hampshire, West Virginia, Wisconsin, and Mississippi with other states, such as Louisiana considering a ban. In some states such as Wisconsin, the ban comes from decisions by the state supreme court declaring that type of device unconstitutional.

Some countries in other parts of the world that use this technology extensively include Australia, Austria, Belgium, Canada, Germany, Israel, Malta, the Netherlands, Singapore, South Africa, India, Switzerland, Taiwan and the United Kingdom.

Issues

While some public officials support the use of red-light cameras, various groups and individuals oppose them. They believe that the use of these devices raises legal issues and violates the privacy of citizens. They also question the effectiveness of red light cameras and if they really help traffic safety. In a few U.S. states (including California), the cameras *are* set up to get a "face photo" of the driver;. This has been done because in those states, red light camera tickets are criminal violations, and criminal charges must always name the actual violator. In California, that need to identify the actual violator has led to the creation of a unique investigatory tool.

In some areas, red light enforcement cameras are installed and maintained by private firms such as Affiliated Computer Services, American Traffic Solutions, Inc., and Redflex Traffic Systems. Many people disagree with this privatization of a police function. In Texas, red light violators caught by a red light camera are served with a civil citation rather than a criminal citation. The civil infraction (civil fine of $75, no traffic points) conflicts with the same criminal infraction (fines of $1 to $200, and traffic points). A December 2008 lawsuit against Dallas County's program, challenging a private camera operating company's right to hand out citations, was dismissed in March 2009.

Throughout the USA, red light cameras were used in 380 individual communities as of October 2008 and speed cameras were used in 45 communities plus Arizona on state roads and Illinois in work zones as of September 2008. Arkansas, Nebraska, Nevada, New Jersey, Utah, West Virginia and Wisconsin have also enacted various prohibitions on photo enforcement.

On February 19, 2010, Arizona completed a study of their statewide 76 photo enforcement cameras and decided that they would not renew their contract with Redflex in 2011. Less than expected revenue, mixed public acceptance and mixed accident data were cited, among other findings.

In New York State, red light cameras are allowed in New York City, and since 2009, Nassau County, Yonkers, Buffalo, Rochester, and Suffolk County,

On January 31, 2008, then New York State Assembly member Ivan Lafayette introduced New York State Assembly Bill A09877 with intent to prohibit the sale or use of a product that alerts of the presence of a red light camera under a civil penalty up to 500 USD. The memo of the bill claimed New York City's Red-Light Camera Program to be a revenue-neutral success to reduce the running of red lights and improve safety. It was referred to consumer affairs and protection on the same day.

Chapter 13

Three-Phase Traffic Theory

Three-phase traffic theory is an alternative theory of traffic flow developed by Boris Kerner between 1996 and 2002. It focuses mainly on the explanation of the physics of traffic breakdown and resulting congested traffic on highways. Kerner describes three phases of traffic, while the classical theories based on the fundamental diagram of traffic flow have two phases: *free flow* and *congested traffic*. Kerner's theory divides congested traffic into two distinct phases, *synchronized flow* and *wide moving jam*, bringing the total number of phases to three:

Synchronized flow

Jam

1. Free flow (*F*)
2. Synchronized flow (*S*)
3. Wide moving jam (*J*)

A **phase** is defined as a *state in space and time.*

Free flow (F)

In free traffic flow, empirical data show a positive correlation between the flow rate *q* (in vehicles per time interval) and vehicle density *k* (in vehicles per kilometres). This relation has an upper boundary at the maximal point of free flow at the flow rate at the corresponding critical density k_{crit} (Figure 1).

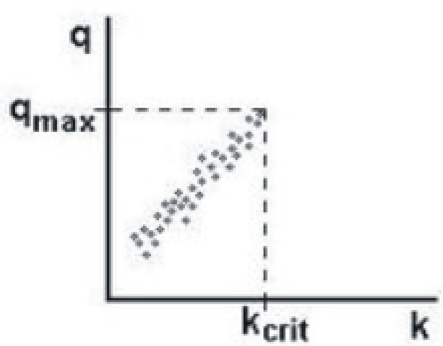


Figure 1:Measured data of flow rate related to vehicle density in free flow

Congested traffic

In congested traffic, the vehicle speed is lower than the minimal possible vehicle speed v_{free}^{min} in free flow, i.e., the line with the slope of the minimal speed $v_{\text{free}}^{\text{min}} = \frac{q_{\text{max}}}{k_{\text{crit}}}$ in free flow (dotted line in Figure 2) divides all empirical data on the flow-density plane into two regions: on the left side the data of free flow and on the right side the data of the congested traffic.

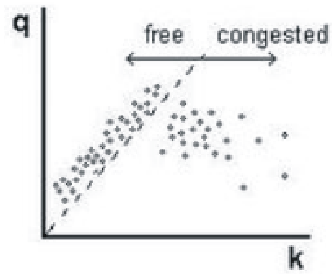


Figure 2: Measured data of flow rate related to vehicle density in free flow and congested traffic

Definitions of the phases J *and* S *in congested traffic*

Kerner's phase definitions [*J*] and [*S*] in congested traffic are the result of common spatial-temporal features of real traffic data. These phase definitions are as follows:

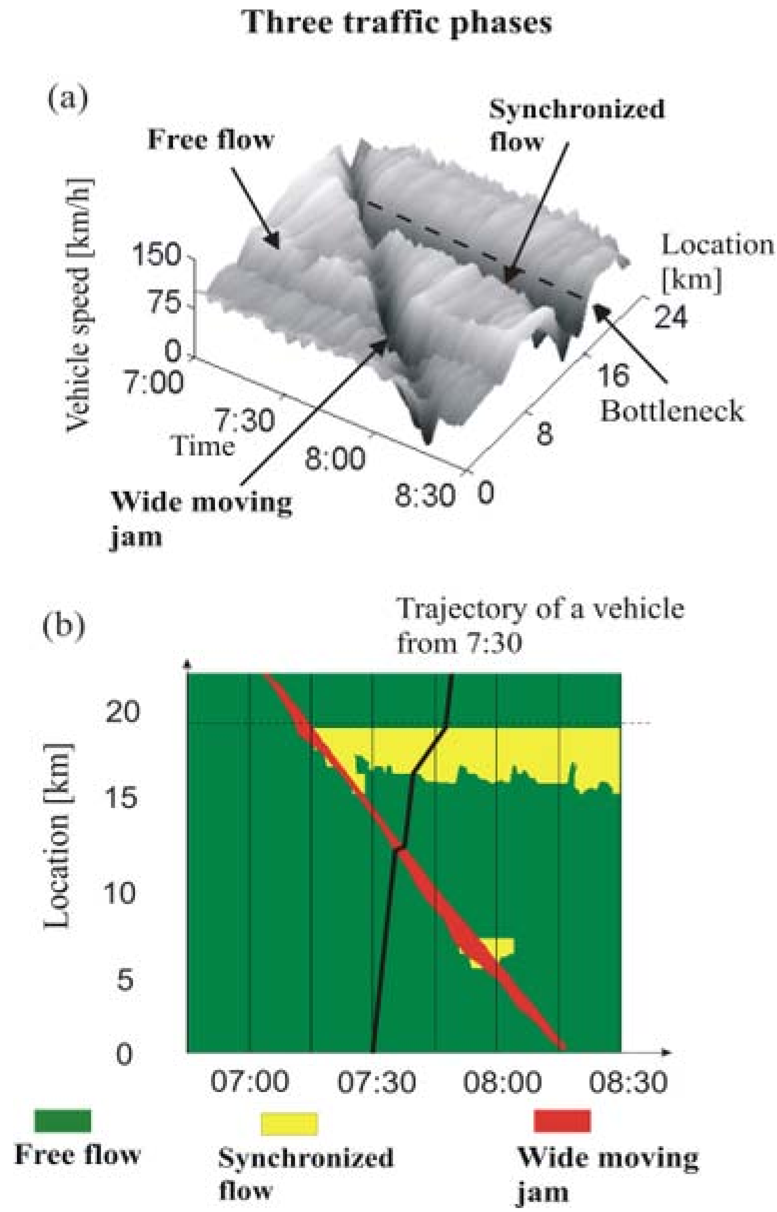


Figure 3: Measured data of speed in time and space (a) and its representation on the time-space plane (b)

The traffic phase "wide moving jam" (*J*)

A wide moving jam moves through a highway bottleneck while maintaining the mean velocity v_g of the jam downstream front at which vehicles escaping the jam accelerate either to free flow or synchronized flow. This is the characteristic feature of the wide moving jam.

The traffic phase "synchronized flow" (S)

The downstream front of synchronized flow, where the vehicles accelerate to free flow, does not show this characteristic feature of the wide moving jam. Specifically, the downstream front of synchronized flow is mostly fixed at the bottleneck.

Explanation of the traffic phase definitions based on measured traffic data

Measured data of the averaged vehicle speeds (Figure 3 (a)) illustrate the definitions [J] and [S]. There are two spatial-temporal patterns of congested traffic with low vehicle speeds in Figure 3 (a). One pattern of congested traffic propagates upstream with almost constant velocity of the downstream pattern front through the freeway bottleneck. According to the definition [J] this pattern of congested traffic belongs to the "wide moving jam" traffic phase. In contrast, the downstream front of the other pattern of the congested traffic is fixed at the bottleneck. According to the definition [S] this pattern of congested traffic belongs to the "synchronized flow" traffic phase (Figure 3 (a) and (b)).

The fundamental hypothesis of Kerner's three-phase traffic theory

The fundamental hypothesis of Kerner's three-phase traffic theory is associated with steady states of synchronized flow. A steady state of synchronized flow is a *hypothetical* state of synchronized flow of identical vehicles and drivers in which all vehicles move with the same time-independent speed and have the same space gaps (a space gap is a net distance between two following each other vehicles), i.e., this synchronized flow is homogeneous in time and space. The fundamental hypothesis is as follows: Steady states of synchronized flow cover a two-dimensional (2D) region in the flow-density plane (2D-region S in Figure 4(a)). The multitudes of free flow states (F) overlap steady states of synchronized flow in the vehicle density. The free flow states on a multi-lane road and steady states of synchronized flow are separated by a gap in the flow rate and, therefore, by a gap in the speed at a given density: at each given density the synchronized flow speed is lower than the free flow speed. In accordance with the fundamental hypothesis of Kerner's three-phase theory, at a given speed in synchronized flow, the driver can make an *arbitrary choice* in the space gap to the preceding vehicle within a finite range of space gaps associated with the 2D-region of steady states of synchronized flow (Figure 4(b)): the driver accepts different space gaps at different times and does not control a fixed space gap to the preceding vehicle.

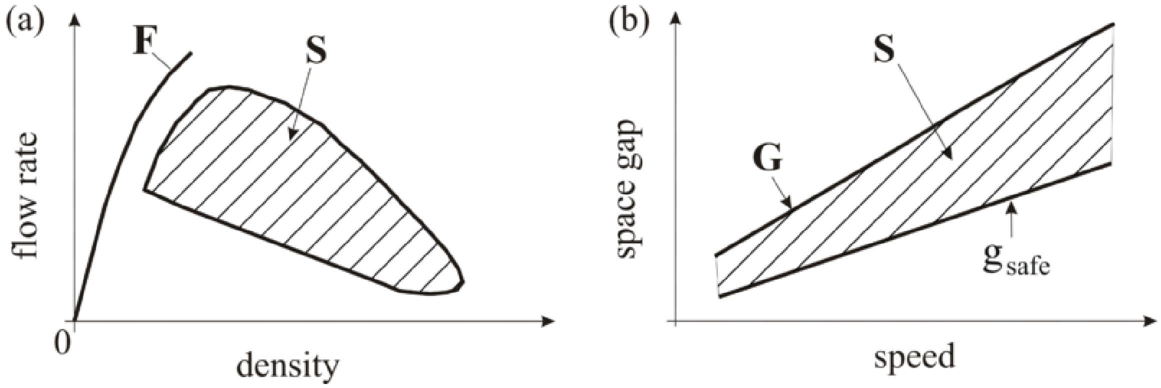


Figure 4: Fundamental hypothesis of Kerner' three-phase traffic theory: (a) Qualitative representation of free flow states (F) and 2D-region of steady states of synchronized flow (dashed region S) on a multi-lane road in the flow-density plane. (b) A part of the 2D-region of steady states of synchronized flow shown in (a) in the space-gap-speed plane (dashed region S). In (b), *G* and g_{safe}, are respectively a synchronization space gap and safe space gap between two vehicles following each other.

The fundamental hypothesis of Kerner's three-phase traffic theory contradicts the hypothesis of earlier traffic flow theories about the fundamental diagram of traffic flow that is a 1D-relationship between the vehicle density and flow rate.

Car following in three-phase traffic theory

In Kerner's three-phase theory, a vehicle accelerates when the space gap *g* to the preceding vehicle is greater than a synchronization space gap *G*, i.e., at $g > G$ (labelled by *acceleration* in Figure 5); the vehicle decelerates when the gap g is smaller than a safe space gap g_{safe}, i.e., at $g < g_{\text{safe}}$ (labelled by *deceleration* in Figure 5).

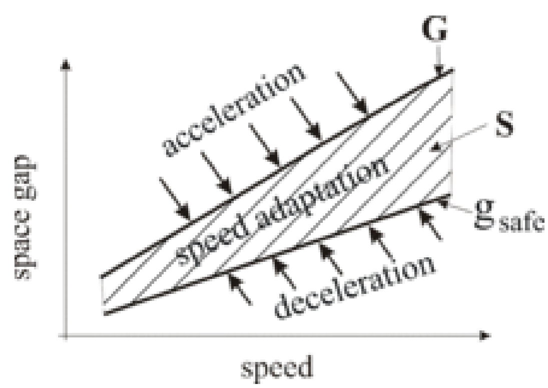


Figure 5: Qualitative explanation of car-following in Kerner's three-phase traffic theory: A vehicle accelerates at a space gap $g > G$ and decelerates at space gaps $g < g_{\text{safe}}$, whereas under condition $g_{\mathbf{safe}} \leq g \leq G$ the vehicle adapts its speed to the speed of the preceding vehicle without caring what the precise space gap is. A dashed region of synchronized flow is taken from Figure 4(b).

The synchronization space gap G is a space gap g between the vehicle and the preceding vehicle within which the vehicle tends to adapt the speed to the speed of the preceding vehicle without caring, what the precise space gap is, as long as this space gap is not smaller than the safe space gap g_{safe} (labelled by *speed adaptation* in Figure 5). Thus the space gap g in car following in the framework of Kerner's three-phase theory can be any space gap within the space gap range $g_{\mathrm{safe}} \leq g \leq G$.

Traffic breakdown - a F → S *phase transition*

In measured data, congested traffic most often occurs at highway bottlenecks, e.g., at on-ramps, off-ramps or at roadwork. Such a transition of free flow to congested traffic is known as traffic breakdown. In Kerner's three-phase traffic theory such a traffic breakdown is explained by a F → S phase transition. This explanation is supported by available measurements, because in measured traffic data after a traffic breakdown at a bottleneck the downstream front of the congested traffic is fixed at the bottleneck. Therefore, the emerging congested traffic after a traffic breakdown fulfils the definition [S] of the "synchronized flow" traffic phase.

Spontaneous and induced *F* → *S* transitions

Kerner states by using empirical measured data that synchronized flow can emerge in free flow spontaneously (spontaneous F → S phase transition) or externally induced (induced F → S phase transition). A spontaneous F → S phase transition means that the traffic breakdown occurs in the case that there has been free flow before at the bottleneck as well as both up- and downstream of the bottleneck. This implies that a spontaneous F → S phase transition occurs through the growth of an internal disturbance in free flow in a neighbourhood of a bottleneck. Alternately, an induced F → S phase transition occurs through disturbance of traffic flow that has initially emerged at a different road location than the bottleneck location. Normally, this correlates with an upstream propagation of a synchronized flow region or a wide moving jam. An empirical example of an induced breakdown at a bottleneck leading to synchronized flow can be seen in Figure 3: synchronized flow emerges through the upstream propagation of a wide moving jam.

Physical explanation of traffic breakdown

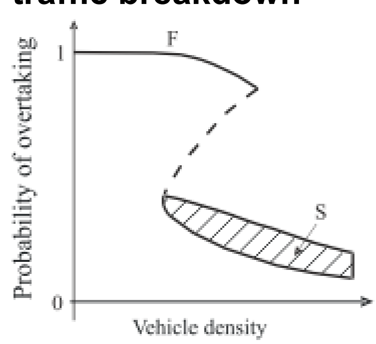


Figure 6: Explanation of traffic breakdown by a Z-like non-linear interrupted function of the probability of overtaking in Kerner's three-phase traffic theory. The dotted curve illustrates the critical probability of overtaking as function of traffic density.

Kerner explains the nature of the F → S phase transitions by a spatial-temporal competition of the vehicle acceleration through the overtaking of a slower vehicle ahead and the vehicle deceleration to the speed of the slower moving vehicle ahead ("speed adaptation"). The overtaking supports the further existence of free flow. In contrast, "speed adaptation" leads to synchronized flow. Such a speed adaptation will occur, if an overtaking is not possible. Kerner states that the probability of overtaking is an *interrupted function of the vehicle density* (Figure 6): at a given vehicle density, the probability of overtaking in free flow is much higher than in synchronized flow.

Infinite number of highway capacities

Maximum and minimum highway capacities

The spontaneous traffic breakdown, i.e., a spontaneous F → S phase transition, might occur in a wide range of flow rates in free flow. Kerner states based on empirical traffic data that because of the possibility of spontaneous or induced traffic breakdowns at the same freeway bottleneck there are an infinite number of highway capacities of free flow at the bottleneck. This infinite number of freeway capacities is between a minimum capacity q_{th} and a maximum capacity q_{max} of free flow (Figure 7).

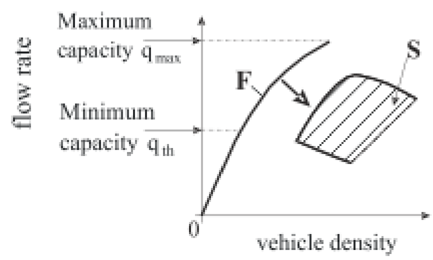


Figure 7: Maximum and minimum highway capacities in Kerner's three-phase traffic theory

Highway capacities and metastability of free flow

Already small disturbances in free flow at the bottleneck will lead to a spontaneous F → S phase transition, if the flow rate in free flow is close to the maximum capacity q_{max}. On the other hand, only very large disturbances in free flow at the bottleneck will lead to a spontaneous F → S phase transition, if the flow rate is close to the minimum capacity q_{th}. The probability of a smaller disturbance in free flow is much higher than of a larger disturbance. Therefore, the higher the flow rate in free flow at a bottleneck, the higher the probability of the spontaneous F → S phase transition. If the flow rate in free flow is lower than the minimum capacity q_{th}, there will be no traffic breakdown (no F → S phase transition) at the bottleneck.

The infinite number of highway capacities at a bottleneck can be illustrated by the meta-stability of free flow at the flow rates q with

$$q_{th} \leq q < q_{\max}.$$

Meta-stability of free flow means, that for small disturbances, free flow can remain stable (free flow remains), but at larger disturbances, the free flow becomes unstable and a F → S phase transition to synchronized flow occurs.

Discussion of capacity definitions

The infinite number of highway capacities at a bottleneck of Kerner's three-phase traffic theory contradicts fundamentally the classical traffic theories and methods for traffic management and traffic control, which at any time assume the existence of a *particular* highway capacity. In contrast, in Kerner's three-phase traffic theory *at any time* there are an infinite number of highway capacities, which are within the abovementioned flow rate range between the minimum capacity q_{th} and maximum capacity q_{max}. The values q_{th} and q_{max} can depend considerably on traffic parameters (the percentage of long vehicles in traffic flow, weather, bottleneck characteristics, etc).

Wide moving jams (J)

A wide moving jam will be called "wide", if the width of the wide moving jam (in direction of the flow) clearly exceeds the width of the jam fronts. The average vehicle speed within the wide moving jams is much lower than the average speed in free flow. At the downstream front the vehicles might accelerate to the free flow speed. At the upstream jam front the vehicles come from free flow or synchronized flow and must reduce their speed. According to the definition [J] the wide moving jam keeps its mean velocity of the downstream front v_g, even if the jam propagates through other traffic phases or bottlenecks. The flow rate is sharply reduced within a wide moving jam.

Characteristic parameters of wide moving jams

Kerner's empirical results show, that some characteristic features of wide moving jams are independent of the traffic volume and bottleneck features (e.g. where and when the jam has emerged). However, these characteristic features are dependent on weather conditions, road conditions, vehicle technology, percentage of long vehicles, etc.. The velocity of the downstream front of a wide moving jam v_g (in direction upstream) is a characteristic parameter, as well as the flow rate downstream of the downstream jam front q_{out} (with free flow at this location, see Figure 8). This means, that several wide moving jams have similar features under similar conditions. These parameters are relatively predictable due to theses reasons. The movement of the downstream jam front can be illustrated in the flow-density plane by a line, which is called "Line J" (Line J in Figure 8). The slope of the Line J is the velocity of the downstream jam front v_g, because the co-ordinate of the Line J at the flow rate zero is the vehicle density within the jam, k_{max}.

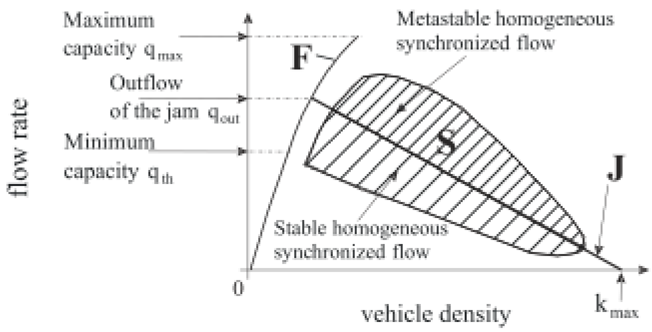


Figure 8: Three traffic phases on the flow-density plane in Kerner's three-phase traffic theory

Minimum highway capacity and outflow from wide moving jam

Kerner emphasizes, that the minimum capacity q_{th} and the outflow of a wide moving jam q_{out} describe two *qualitatively different features* of free flow: the minimum capacity q_{th} characterizes an F → S phase transition at a bottleneck, i.e., a traffic breakdown. The outflow of a wide moving jam q_{out} characterizes conditions of the existence of the jam, i.e., the traffic phase J. Depending on traffic parameters like weather, percentage of long vehicles, etc., and of bottleneck characteristics, where the F → S phase transition can occur, the minimum capacity q_{th} might be smaller (as in Figure 8), or greater than the jam's outflow q_{out}.

*Synchronized flow phase (*S*)*

In contrast to wide moving jams, both the flow rate and vehicle speed might vary significantly in the synchronized flow phase. The downstream front of synchronized flow is often spatially fixed, normally at the bottleneck at a certain road location. The flow rate in this phase could remain similar to the one in free flow, even if the vehicle speeds are sharply reduced.

Because the synchronized flow phase does not have the characteristic features of the wide moving jam phase J, Kerner's three-phase traffic theory assumes that hypothetical homogeneous states of synchronized flow cover a two-dimensional region in the flow-density plane (dashed regions in Figure 8).

S → J *phase transition*

Wide moving jams do not emerge spontaneously in free flow, but they can emerge in regions of synchronized flow. This phase transition is called a S → J phase transition.

"Jam without obvious reason" - F→S→J phase transitions

Therefore, the emergence of wide moving jam in free flow is observed as a cascade of F → S → J phase transitions: first, a region of synchronized flow emerges in a region of free flow. As explained above, such an F → S phase transition occurs mostly at a bottleneck. Within the synchronized flow phase a further "self-compression" occurs and vehicle density increases while vehicle speed decreases. This self-compression is called "pinch effect". In "pinch" regions of synchronized flow, narrow moving jams emerge. If these narrow moving jams grow, wide moving jams will emerge. Kerner notes that the frequency of the emergence of wide moving jams increases, if the density in synchronized flow increases. The wide moving jams propagate further upstream, even if they propagate through regions of synchronized flow or bottlenecks.

The physics of *S* → *J* transition

To further illustrate S → J phase transitions, it should be noted that in Kerner's three-phase traffic theory the Line J divides all homogeneous states of synchronized flow (Figure 8). States of homogeneous synchronized flow above Line J are meta-stable. States of homogeneous synchronized flow below Line J are stable in which no S → J phase transition can occur. Meta-stable homogeneous synchronized flow means that for small disturbances, traffic state remains stable. However, when larger disturbances occur, synchronized flow becomes unstable, and a S → J phase transition occurs.

Traffic patterns of S *and* J

Caused by F → S and S → J phase transitions, very complex congested patterns can be observed.

Classification of synchronized flow traffic patterns (SP)

A congestion pattern of synchronized flow (Synchronized Flow Pattern (SP)) with a fixed downstream and a not continuously propagating upstream front is called Localised Synchronized Flow Pattern (LSP).

Frequently the upstream front of a SP propagates upstream. If only the upstream front propagates upstream, the related SP is called Widening Synchronised Flow Pattern (WSP). The downstream front remains at the bottleneck location and the width of the SP increases.

It is possible that both upstream and downstream front propagate upstream. The downstream front is no longer located at the bottleneck. This pattern has been called Moving Synchronised Flow Pattern (MSP).

Catch effect of synchronized flow at highway bottleneck

The difference between the SP and the wide moving jam becomes visible when the WSP or the MSP reaches an upstream bottleneck: the so-called "catch-effect" occurs. The SP will be caught at the bottleneck and as a result a new congested pattern emerges. A wide moving jam will not be caught at a bottleneck and moves further upstream. In contrast to wide moving jams, the synchronized flow, even as an MSP has no characteristic parameters. As an example, the velocity of the downstream front of the MSP might vary significantly and can be different for different MSPs. These features of SP and wide moving jams are consequences of the phase definitions [S] and [J].

General congested traffic pattern (GP)

A very typical congested pattern contains both congested phases [S] and [J]. Such a pattern with [S] and [J] is called General Pattern (GP).

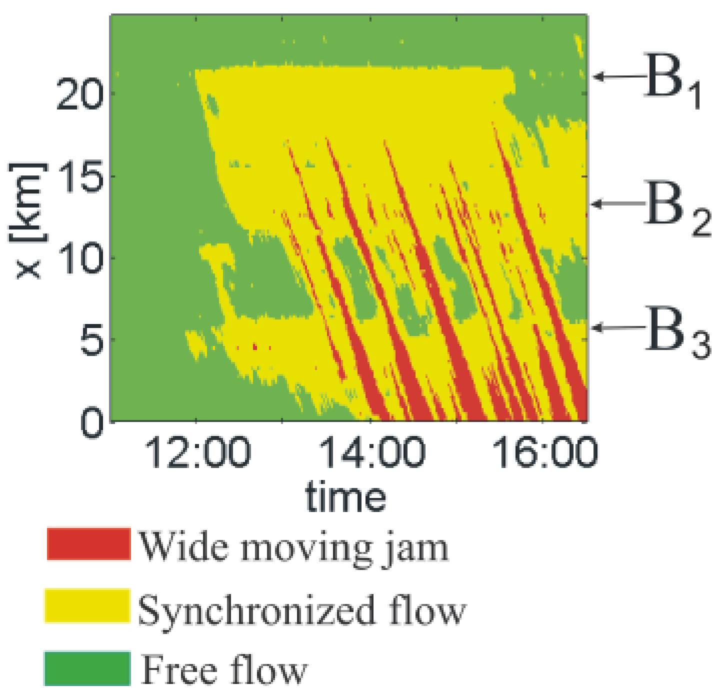


Figure 9: Measured EGP at three bottlenecks B_1, B_2 and B_3

In many freeway infrastructures bottlenecks are very close to each other. A congested pattern whose synchronized flow covering two or more bottlenecks is called Expanded Pattern (EP). An EP could contain synchronized flow only (called ESP: Expanded Synchronized Flow Pattern)), but normally wide moving jams are emerging in the synchronized flow. In those cases the EP is called EGP (Expanded General Pattern) (see Figure 9).

Applications of three-phase traffic theory in transportation engineering

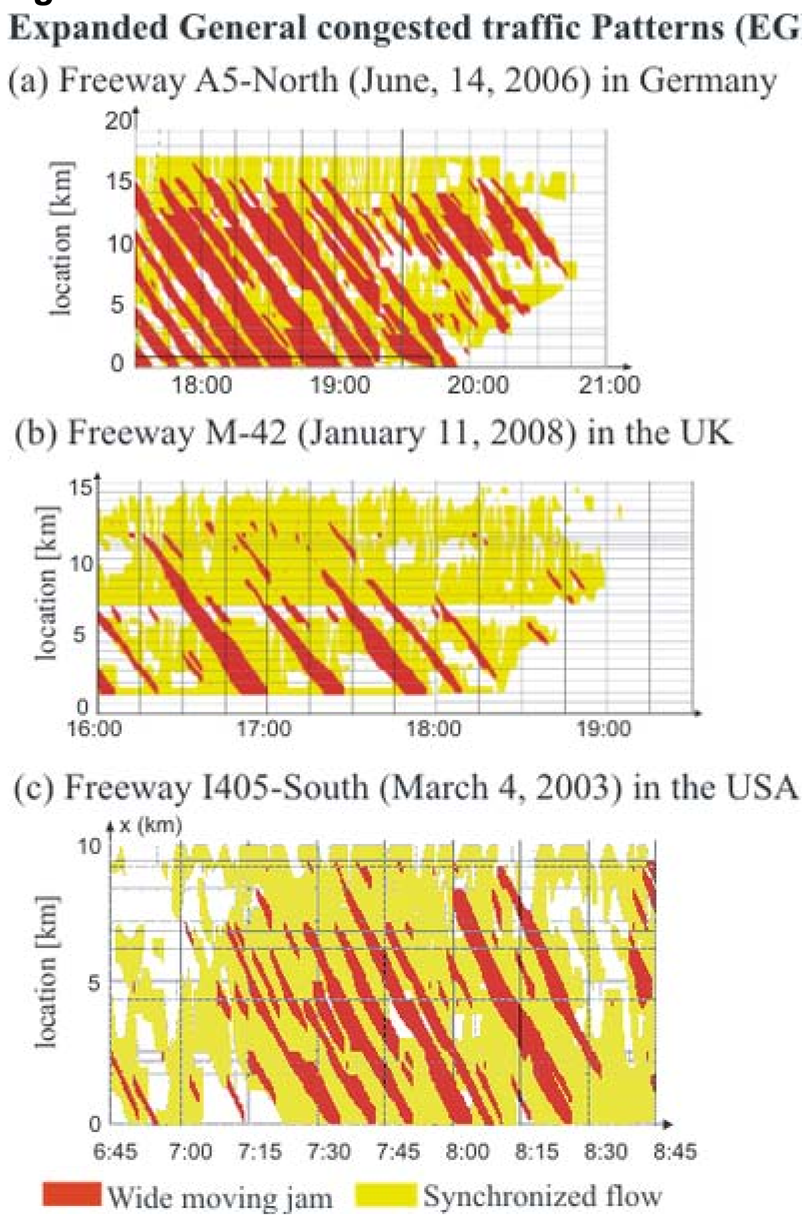


Figure 10: Traffic patterns in the ASDA/FOTO application in three countries

One of the applications of Kerner's three-phase traffic theory is the methods called ASDA/FOTO (**A**utomatische **S**tau**D**ynamik**A**nalyse (Automatic tracking of wide moving jams) and **F**orecasting **O**f **T**raffic **O**bjects). ASDA/FOTO is a software tool able to process large traffic data volumes quickly and efficiently on freeway networks. ASDA/FOTO works in an online traffic management system based on measured traffic data. Recognition, tracking and prediction of [S] and [J] are performed using the features of Kerner's three-phase traffic theory.

Further applications of the theory are seen in the development of traffic simulation models, ramp metering system (ANCONA), collective traffic control, traffic assistance and traffic state detection, as described in the books by Kerner.

Chapter 14

Traffic Barrier

Traffic barriers keep vehicles within their roadway and prevent vehicles from colliding with dangerous obstacles. Traffic barriers installed at the road side also prevent errant vehicles from traversing steep (non-recoverable) slopes. Traffic barriers installed at the medians of divided highways are also referred to as median barriers. The latter also prevent errant vehicles from entering the opposing carriageway of traffic and help to prevent head-on collisions.

Need and placement

Traffic barriers constitute hazards themselves and should only be used when the obstacle poses a greater threat than the barrier itself. In all cases, roadside hazards must be assessed for the danger they pose to traveling motorists based on size, rigidity and distance from the edge of travelway. For instance, small roadside signs and some large signs (ground-mounted breakaway post) often do not merit roadside protection as the barrier itself may pose a greater threat to general health and well-being of the public than the obstacle it intends to protect. In many regions of the world, the concept of clearzone is taken into account when examining the distance of an obstacle or hazard from the edge of travelway.

Clearzone also known as clear recovery area or horizontal clearance is defined (through study) as a lateral distance in which a motorist on a recoverable slope may travel outside of the travelway and return their vehicle safely to the roadway. This distance is commonly determined as the 85th percentile in a study comparable to the method of determining speed limits on roadways through speed studies and varies based on the classification of a roadway. In order to provide for adequate safety in roadside conditions, hazardous elements, whether they be obstacles or steep slopes can be placed outside of the clearzone in order to reduce or eliminate the need for roadside protection.

Common sites for installation of traffic barrier:

- Bridge ends
- Near steep slopes from roadway limits

- At drainage crossings or culverts where steep or vertical drops are present
- Near large signs/illumination poles or other roadside elements which may pose hazards

When barrier is needed, careful calculations are completed to determine length of need which take into account the aforementioned factors. Specifically, the traffic volumes and therefore, the classification of the roadway in addition to the distance of the hazard from the edge of travelway and the distance or offset of the barrier to be placed or installed from the edge of travelway. It is the case in current times, that barrier or rail that is to be used in construction and maintenance operations has undergone extensive testing in both government and private research facilities in order to determine proper 'crash-worthiness' and effectiveness in conditions which are prescribed for its use. In particular, most roadside protection, whether it be a concrete barrier or rail, or a metal beam fence will perform properly only when placed in adequate proximity to the travelway so as to prevent vehicle impacts at large (obtuse) angles. The method in which a barrier protects motorists from roadside hazards is in how it dissipates the energy of an impact.

Barrier types and performance

Traffic barriers are categorized in two ways: by the function they serve, and by how much they deflect when a vehicle crashes into them.

Barrier functions

Roadside barriers are used to protect traffic from roadside obstacles or hazards, such as slopes steep enough to cause rollover crashes, fixed objects like bridge piers, and bodies of water. Roadside barriers can also be used wide medians, to prevent vehicles from colliding with hazards within the median.

Median barriers are used to prevent vehicles from crossing over a median and striking an oncoming vehicle in a head-on crash. Unlike roadside barriers, they must be designed to be struck from either side.

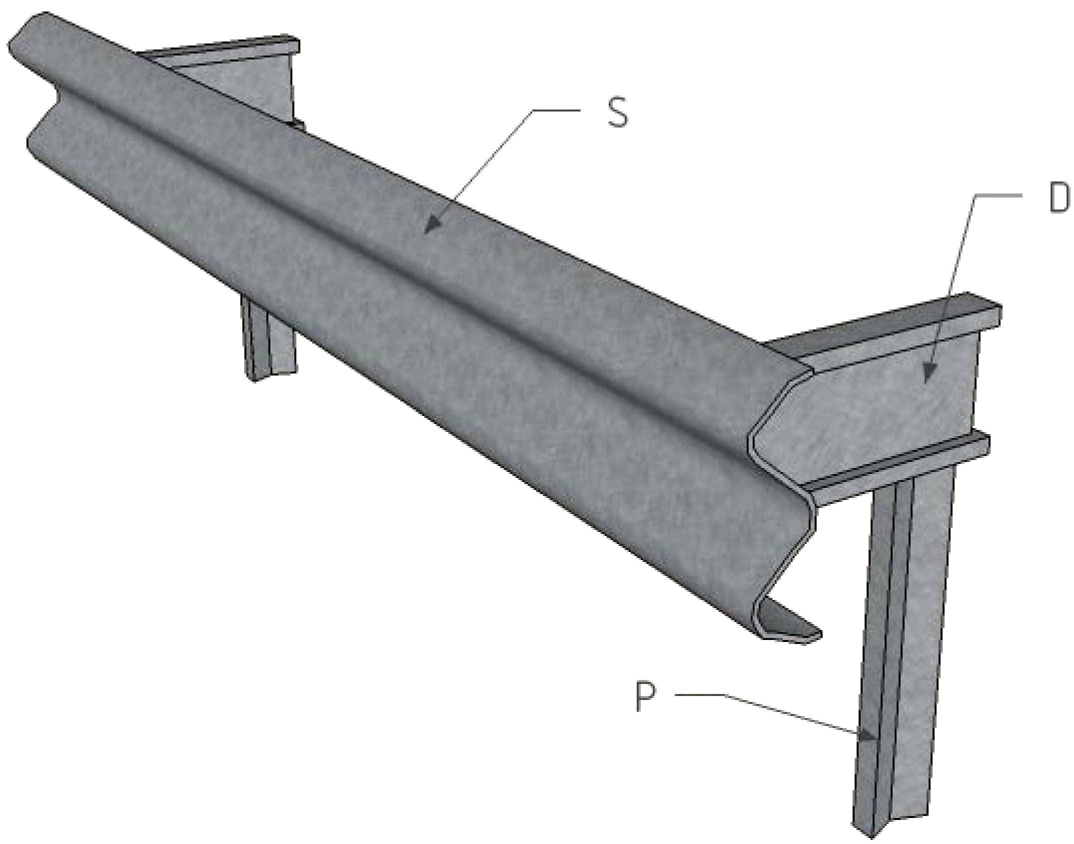
Bridge barrier is designed to restrain vehicles from crashing off the side of a bridge and falling onto the roadway, river or railroad below. It is usually higher than roadside barrier, to prevent trucks, buses, pedestrians and cyclists from vaulting or rolling over the barrier and falling over the side of the structure. Bridge rails are usually multi-rail tubular steel barriers or reinforced concrete parapets and barriers

Work zone barriers are used to protect traffic form hazards in work zones. Their distinguishing feature is they can be relocated as conditions change in the road works. Two common types are used: temporary concrete barrier and water-filled barrier. The latter is composed of steel-reinforced plasic boxes that are put in place where needed, linked together to form a longitudinal barrier, then ballasted with water. These have an advantage in that they can be assembled without heavy lifting equipment, but they cannot be used in freezing weather.

Barrier stiffness

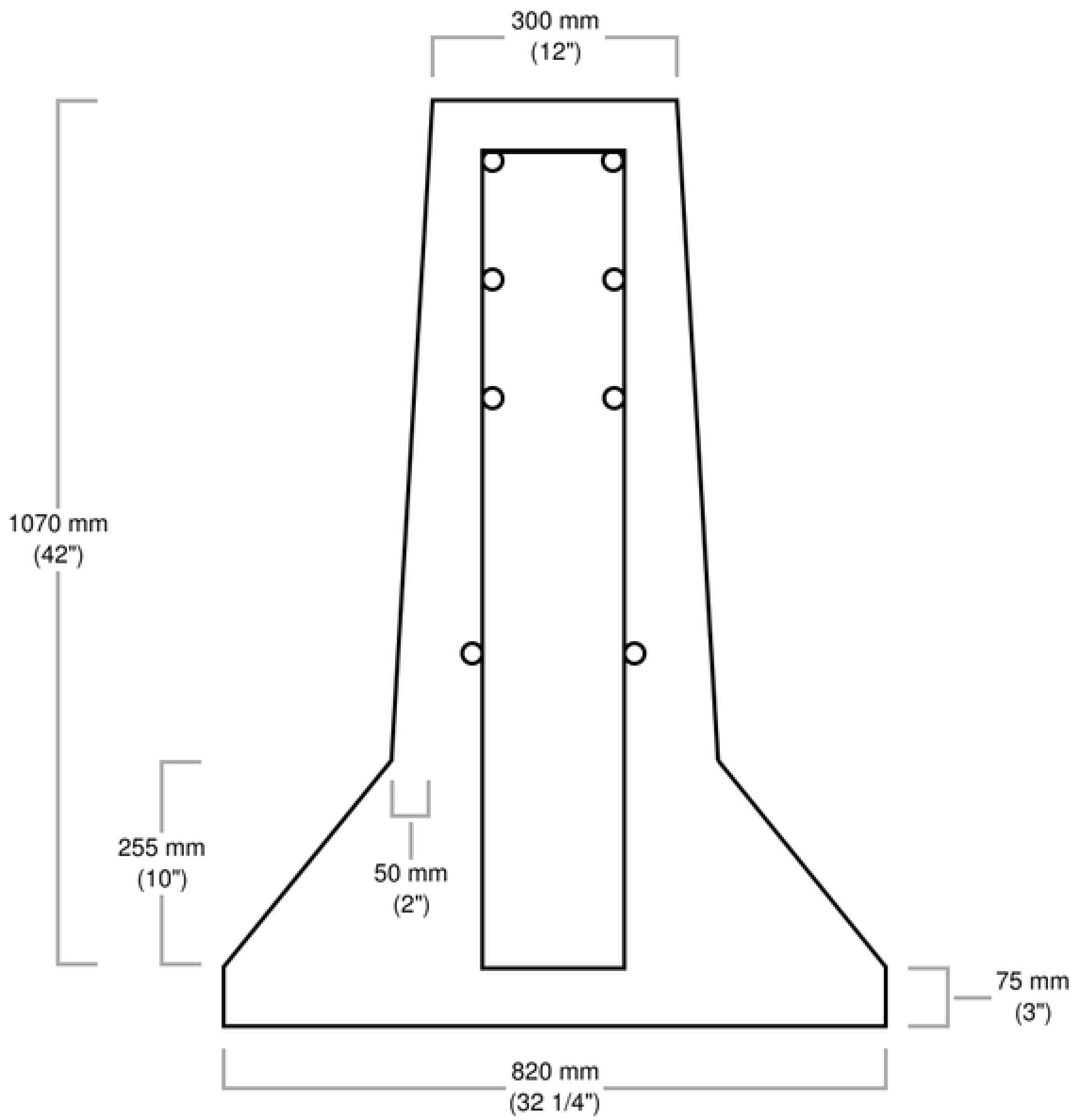
Barriers are divided into three groups, based on the amount they deflect when struck by a vehicle and the mechanism the barrier uses to resist the impact forces. In the United States, traffic barriers are tested and classified according to the AASHTO Manual for Assessing Safety Hardware (MASH) standards, which recently superseded Federal Highway Administration NCHRP Report 350. Barrier deflections listed below are results from crash tests with a 2000 kg pickup truck traveling 100 km/h, colliding with the rail at a 25 degree angle.

Flexible barriers include cable barriers and weak post corrugated guide rail systems. These are referred to as flexible barriers because they will deflect 1.6 m to 2.6 m when struck by a typical passenger car or light truck. Impact energy is dissipated through tension in the rail elements, deformation of the rail elements, posts, soil and vehicle bodywork, and friction between the rail and vehicle.



Components of a standard guardrail (A-profile): S – guardrail, D – distance piece/spacer, P – sigma post

Semi-rigid barriers include box beam guide rail, heavy post blocked out corrugated guide rail and thrie-beam guide rail. They deflect three to six feet: more than rigid barriers, but less than flexible barriers. Impact energy is dissipated through deformation of the rail elements, posts, soil and vehicle bodywork, and friction between the rail and vehicle. Box beam systems also spread the impact force over a number of posts due to the stiffness of the steel tube. These barrier systems deflect up to 1.5 m.



42 inches (110 cm)-high version of the Jersey barrier for deflecting automobiles and semi-trailer trucks.

Rigid barriers are usually constructed of reinforced concrete. A permanent concrete barrier will only deflect a negligible amount when struck by a vehicle. Instead, the shape of a concrete barrier is designed to redirect a vehicle into a path parallel to the barrier. This means they can be used to protect traffic from hazards very close behind the barrier, and generally require very little maintenance. Impact energy is dissipated through redirection and deformation of the vehicle itself. Jersey barriers and F-shape barriers also lift the vehicle as the tires ride up on the angled lower section. For low-speed or low-angle impacts on these barriers, that may be sufficient to redirect the vehicle without damaging the bodywork. The disadvantage is there is a higher likelihood of rollover with a small car than the single slope or step barriers. Impact forces are resisted by a combination of the rigidity and mass of the barrier. Deflection is usually negligible.

An early concrete barrier design was developed by the New Jersey State Highway Department. This led to the term Jersey barrier being used as a generic term, although technically it applies to a specific shape of concrete barrier. Other types include constant slope barriers, concrete step barriers and F-shape barriers.

Concrete barriers usually have smooth finishes. At some impact angles, coarse finishes allow the drive wheel of front wheel drive vehicles to climb the barrier, potentially causing the vehicle to roll over. However, along parkways and other areas where aesthetics are considered important, reinforced concrete walls with stone veneers or faux stone finishes are sometimes used. These barrier walls usually have vertical faces to prevent vehicles from climbing the barrier.

Barrier end treatments

Early traffic barrier designs often paid little attention to the ends of the barriers. Vehicles that struck blunt ends often stop abruptly or have steel rail sections penetrate into the passenger compartment. This often results in a severe injuries or fatalities.

As a result, barrier terminals were developed that brought the end of the terminal down to ground level. This prevented the rail from penetrating the vehicle, but could vault a vehicle into the air or cause it to roll over, since the barrier end formed a ramp.

To address the vaulting and rollover crashes, energy absorbing terminals were developed. These have a large steel impact head that engages the frame or bumper of the vehicle. The impact head is driven back along the guide rail, dissipating the vehicle's kinetic energy by bending or tearing the steel in the guide rail sections.

The final way to terminate a guide rail is bend it back to the point that the terminal is unlikely to be hit end-on, and, if possible, embed the end in a hillside or cut slope.

Chapter 15

Traffic Enforcement Camera

Gatso speed camera

A **traffic enforcement camera** (also **road safety camera**, **road rule camera**, **photo radar**, **speed camera**, **Gatso**) is an automated ticketing machine. It may include a camera which may be mounted beside on over a highway or installed in an enforcement vehicle to detect traffic regulation violations, including speeding, vehicles going through a red traffic light, unauthorized use of a bus lane, for recording vehicles inside a congestion charge area and others.

The latest automatic number plate recognition systems can be used for the detection of average speeds and raise concerns over loss of privacy and the potentially for governments to establish mass surveillance of vehicle movements and therefore by association also the movement of the vehicle's owner. Vehicles owners are often required by law to identify the driver of the vehicle and a case was taken to the European Court of Human Rights who found that the Human Rights Act 1998 was not being breached. Some groups, such as the National Motorists Association in the USA, claim that systems "encourage ... revenue-driven enforcement" rather than the declared objectives.

Types

Automatic speed enforcement gantry or "*Lombada Eletrônica*" with ground sensors at Brasilia, D.F.

Gatso Mobile Speed Camera, used in Victoria, Australia. The camera is mounted on the passenger side dash, whilst the black box on the front is the radar unit.

Bus lane enforcement

Some bus lane enforcement cameras use a sensor in the road which triggers a number plate recognition camera which compares the vehicle registation plate with a list of approved vehicles and records images of other vehicles. Other systems use a camera mounted on the bus, for example in London where they monitor Red routes on which stopping is not allowed for any purpose (other than taxis and disabled parking permit holders).

On Monday, February 23, 2009, New York City announced testing camera enforcement of bus lanes on 34th Street in Midtown Manhattan where a New York City taxi illegally using the bus lanes would face a fine of $150 adjudicated by the New York City Taxi and Limousine Commission.

Red light enforcement

Red light camera in Springfield, Ohio, USA.

A red light camera is a traffic camera that takes an image of a vehicle that goes through an intersection where the light is red. The system continuously monitors the traffic signal and the camera is triggered by any vehicle entering the intersection above a preset minimum speed and following a specified time after the signal has turned red.

Speed limit enforcement

Speed enforcement cameras are used to monitor compliance with speed limits which may use Doppler, LIDAR or Automatic number plate recognition. Other speed enforcement systems are also used which are not camera based.

Fixed or mobile speed camera systems that measure the time taken by a vehicle to travel between two or more fairly distant sites (from several hundred metres to several hundred kilometres apart) are called automatic number plate recognition (ANPR) cameras. These cameras time vehicles over a known fixed distance, then calculate the vehicle's average speed for the journey. The name derives from the fact that the technology uses infrared cameras linked to a computer to "read" a vehicle's registration number and identify it in real-time.

Number plate recognition systems

Automatic number plate recognition can be used for purposes unrelated to enforcement of traffic rules. In principle any agency or person with access to data either from traffic cameras or cameras installed for other purposes can track the movement of vehicles for any purpose.

In Australia's SAFE-T-CAM system, ANPR technology is used to monitor long distance truck drivers to detect avoidance of legally prescribed driver rest periods.

The United Kingdom's police ANPR system logs all the vehicles passing particular points in the national road network, allowing authorities to track the movement of vehicles and individuals across the country.

In the UK an 80-year-old pensioner John Catt and his daughter Linda (with no criminal record between them) were stopped by City of London Police while driving in London, UK in 2005, had their vehicle searched under section 44 of the Terrorism Act 2000 and were threatened with arrest if they refused to answer questions. After they complained formally, it was discovered they were stopped when their car was picked up by roadside ANPR CCTV cameras; it had been flagged in the Police National Computer database when they were seen near EDO MBM demonstrations in Brighton. Critics point out that the Catts had been suspected of no crime, however the UK's mass surveillance infrastructure allowed them to be targeted due to their association.

Other

- Congestion charge cameras to detect vehicles inside the chargeable area which have not paid the appropriate fee
- High-occupancy vehicle lane cameras to identify vehicles violating occupancy requirements.
- Level crossing cameras to identifying vehicles crossing railways at grade
- Noise pollution cameras that record evidence of heavy vehicles that break noise regulations by using engine braking
- Parking cameras which issue citations to vehicles which are illegally parked or which were not moved from a street at posted times.
- Toll-booth cameras to identify vehicles proceeding through a toll booth without paying the toll
- Turn cameras at intersections where specific turns are prohibited on red. This type of camera is mostly used in cities or heavy populated areas.
- Automatic number plate recognition systems can be used for multiple purposes, including identifying untaxed and uninsured vehicles, stolen cars and potentially mass surveillance of motorists.

Fixed camera systems can mounted in boxes or on poles beside the road or attached to gantries over the road, or to overpasses or bridges. Cameras can be concealed, for example in garbage bins.

Mobile speed cameras may be hand-held, tripod mounted, or vehicle-mounted. In vehicle-mounted systems, detection equipment and cameras can be mounted to the vehicle itself, or simply tripod mounted inside the vehicle and deployed out a window or door. If the camera is fixed to the vehicle, the enforcement vehicle does not necessarily have to be stationary, and can be moved either with or against the flow of traffic. In the latter case, depending on the direction of travel, the target vehicle's relative speed is either added or subtracted from the enforcement vehicle's own speed to obtain its actual speed. The speedometer of the camera vehicle needs to be accurately calibrated.

Some number plate recognition systems can be used from vehicles.

Controversy

Legal issues

There are a number of legal issues which arise as a result depending on local laws and the procedures used by the enforcing bodies. Various legal issues arise from such cameras and the laws involved in how cameras can be placed and what evidence is necessary to prosecute a driver varies considerably in different legal systems.

One issue is the potential conflict of interest when private contractors are paid a commission based on the number of tickets they are able to issue. Pictures from the San Diego red light camera systems were ruled inadmissible as court evidence in September 2001. The judge said that the "total lack of oversight" and "method of compensation" made evidence from the cameras "so untrustworthy and unreliable that it should not be admitted".

Some U.S. states and provinces of Canada such as Alberta operate "owner liability" where it is the registered owner of the vehicle who must pay all such fines regardless of whether he was driving at the time of the offense, although they do release the owner from liability if he signs a form identifying the actual driver and that individual pays the fine. These states do not issue demerit points for camera infractions which has has been criticized by some as giving a "license to speed" to those who can more easily afford speeding fines.

In Albuquerque, New Mexico, the city government attempted to bypass the legal issue of a defendant's right to cross-examine his accuser, as well as the issue of verifying the driver's identity. Automated red-light and speeding offenses are classed as public nuisances and fined to the vehicle's registered owner as civil violations, not as criminal offenses.

In April 2000 two motorists who were caught speeding in the United Kingdom challenged the Road Traffic Act 1988 which required the keeper of a driver to identify the driver at a particular time as being in contradiction to the Human Rights Act 1998 on the grounds that it amounted to a 'compulsory confession', also that since the camera partnerships included the police, local authorities, Magistrates Courts Service (MCS) and Crown Prosecution Service (CPS) which had a financial interest in the fine revenue that they would not get a fair trial. Their plea was initially granted by a judge then overturned but was the heard by the European Court of Human Rights (ECtHR), and the European Court of Justice (ECJ). In 2007 the European Court of Human Rights found there was no breach of article 6 in requiring the keepers of cars caught speeding on camera to provide the name of the driver.

Surveillance

- Police and government have been accused of "Big Brother tactics" in over-monitoring of public roads, and of "revenue raising" in applying cameras in deceptive ways to increase government revenue rather than improve road safety.

Revenue not safety

- In 2010 a campaign was set up against a speed camera on a dual carriageway in Poole, Dorset in a 30 mph area in the United Kingdom. which had generated £1.3m of fines every year since 1999. The initial Freedom of information request was refused and the information was only released after an appeal to the Information Commissioner.
- In May 2010 the new Coalition government said that the 'Labour's 13-year war on the motorist is over' and that the new government 'pledged to scrap public funding for speed cameras' In July Mike Penning, the Road safety minister reduced the Road Safety Grant for the current year to Local Authorities from £95 million to £57 million saying that local authorities had relied too heavily on safety cameras for far too long and that he was pleased that some councils were now focusing on other road safety measures. It is estimated that the as a result the Treasury is now distributing £40 million less in Road Safety Grant than is raised from fines in the year. Dorset and Essex announced plans to review camera provision with a view to possibly ending the scheme in their counties, however Dorset strongly affirmed its support for the scheme, albeit reducing financial contributions in line with the reduction in government grant. Seven counties also announced plans to turn off some or all of their cameras, amidst warnings from the country's most senior traffic policeman that this would result in an increase in deaths and injuries. Gloucestershire cancelled plans to update cameras and has reduced or cancelled maintenance contracts.

Unpopularity

Use of cameras is opposed by some motorists and motoring organisations. They have also be rejected in some places by referendum.

- The first speed camera systems in the USA was in Friendswood, Texas in 1986 and La Marque, Texas in 1987. Neither program lasted more than a few months before public pressure forced them to be dropped.
- In 1991 cameras have been rejected by voters in referenda in Peoria, Arizona voters were the first to reject cameras by a 2-1 margin. Speed cameras have since been installed on the highways in the Phoenix area since 2007.
- In 1992 cameras have been rejected by voters in referenda in Batavia, Illinois.
- Anchorage, Alaska rejected cameras in a 1997 referendum
- In 2002 the state of Hawaii experimented with speed limit enforcement vans but they were withdrawn months later due to public outcry.
- In 2005, the Virginia legislature declined to reauthorize its red light camera enforcement law after a study questioned their effectiveness, only to reverse itself in 2007 and allow cameras to return to any city with a population greater than 10,000.
- Steubenville, Ohio rejected cameras in a 2006 referendum.
- In 2009, a petition was started in the town of College Station, Texas which requested that all red light cameras be dismantled and removed from all of the town's intersections. Enough signatures were captured to put the measure on the November 2009 general election ballot. After an extensive battle between the College Station city council and the opposing sides, both for and against red light cameras, the voters voted to eliminate the red light cameras throughout the entire city. By the end of November the red light cameras were taken down. However, all citations issued are still valid and must be paid by the offenders.
- On May 4, 2010 an ordinance authorizing the use of speed cameras in the town of Sykesville, Maryland was put to a referendum, in which 321 out of 529 voters (60.4%) voted against the cameras. The turnout for this vote was greater than the number of voters in the previous local Sykesville election for mayor where 523 residents voted.
- Arizona decided to not renew their contract with Redflex in 2011 following a study of their statewide 76 photo enforcement cameras. Reasons given included less than expected revenue due to improved compliance, mixed public acceptance and mixed accident data.

Effectiveness

- The town of Swindon abandoned the use of fixed cameras in 2009, questioning their cost effectiveness with the cameras being replaced by vehicle activated warning signs and enforcement by police using mobile speed cameras: in the nine months following the switch-off there was a small reduction in accident rates which had changed slightly in similar periods before and after the switch off (Before: 1 fatal, 1 serious and 13 slight accidents. Afterwards: no fatalities, 2 serious and 12 slight accidents). The journalist George Monbiot claimed that the results were not statistically significant highlighting earlier findings across the whole of Wiltshire that there had been a 33% reduction in the number of people killed and seriously injured generally and a 68% reduction at camera sites during the previous 3 years.

- In January 2011 Edmonton, Alberta cancelled all 100,000 "Speed On Green" tickets issued in the previous 14 months due to concerns about camera reliability.

Avoidance/evasion

A GPS map showing speed camera POI information overlaid onto it

To avoid detection or prosecution drivers may:

- Brake just before a camera in order to travel past its sensor below the speed limit. This is however a cause of collisions.
- Use GPS navigation devices which contain databases of known camera locations to alert them in advance. These databases may in some cases be update in near-realtime. The use of GPS devices to locate speed cameras is illegal in some jurisdictions.
- Install passive laser detectors or radar detectors that detect when the vehicle's speed is being monitored and warn the driver. Use of these devices may be illegal in some jurisdictions.
- Install active laser jammer or radar jammer devices which actively transmit signals that interfere with the measuring device. These devices are illegal in many jurisdictions.
- Remove, falsify, obscure or modify vehicle license plate. Tampering with number plates is illegal in many jurisdictions.

In August 2010 a fast driving Swedish driver reportedly avoided several older model speed cameras, but was detected by a new model, as traveling at 186 mph (300 km/h), resulting in the world's largest speeding fine to date.

- In the past it was possible to avoid detection by changing lanes when SPECS average speed cameras were in use as they measured a vehicle's speed over distance in one lane only. As of 2011 the cameras are type approved to cover multiple lanes.

History

Older traffic enforcement camera in Ludwigsburg, Germany

The concept of the speed camera can be dated back to at least 1905; Popular Mechanics reports on a patent for a "Time Recording Camera for Trapping Motorists" that enabled the operator to take time-stamped images of a vehicle moving across the start and endpoints of a measured section of road. The timestamps enabled the speed to be calculated, and the photo enabled identification of the driver.

The Dutch company *Gatsometer BV*, which was founded in 1958 by rally driver Maurice Gatsonides, produced the 'Gatsometer'. Gatsonides wished to better monitor his average speed on a race track and invented the device in order to improve his lap times. The company later started supplying these devices as police speed enforcement tools. The first systems introduced in the late 1960s used film cameras to take their pictures. Gatsometer introduced the first red light camera in 1965, the first radar for use with road traffic in 1971 and the first mobile speed traffic camera in 1982;

From the late 1990s, digital cameras began to be introduced. Digital cameras can be fitted with a network connection to transfer images to a central processing location automatically, so they have advantages over film cameras in speed of issuing fines, maintenance and operational monitoring. However, film-based systems may provide superior image quality in the variety of lighting conditions encountered on roads, and are required by courts in some jurisdictions. New film-based systems are still being sold, but digital pictures are providing greater versatility and lower maintenance and are now more popular with law enforcement agencies.

A red-light and speed camera in Darwin, Northern Territory, Australia

Dazzle camouflaged speed camera in Loipersdorf, Austria

A red-light camera in use in Beaverton, Oregon, USA

Chapter 16

Intelligent Transportation System

The term ***intelligent transport system*** (ITS) refers to efforts to add information and communications technology to transport infrastructure and vehicles in an effort to manage factors that typically are at odds with each other, such as vehicles, loads, and routes to improve safety and reduce vehicle wear, transportation times, and fuel consumption.

Background

Interest in ITS comes from the problems caused by traffic congestion and a synergy of new information technology for simulation, real-time control, and communications networks. Traffic congestion has been increasing worldwide as a result of increased motorization, urbanization, population growth, and changes in population density. Congestion reduces efficiency of transportation infrastructure and increases travel time, air pollution, and fuel consumption.

The United States, for example, saw large increases in both motorization and urbanization starting in the 1920s that led to migration of the population from the sparsely populated rural areas and the densely packed urban areas into suburbs. The industrial economy replaced the agricultural economy, leading the population to move from rural locations into urban centers. At the same time, motorization was causing cities to expand because motorized transportation could not support the population density that the existing mass transit systems could. Suburbs provided a reasonable compromise between population density and access to a wide variety of employment, goods, and services that were available in the more densely populated urban centers. Further, suburban infrastructure could be built quickly, supporting a rapid transition from a rural/agricultural economy to an industrial/urban economy.

Recent governmental activity in the area of ITS – specifically in the United States – is further motivated by the perceived need for homeland security. Many of the proposed ITS systems also involve surveillance of the roadways, which is a priority of homeland security. Funding of many systems comes either directly through homeland security organizations or with their approval. Further, ITS can play a role in the rapid mass evacuation of people in urban centers after large casualty events such as a result of a natural disaster or threat. Much of the infrastructure and planning involved with ITS parallels the need for homeland security systems.

In the developing world, the migration of people from rural to urbanized habitats has progressed differently. Many areas of the developing world have urbanized without significant motorization and the formation of suburbs. In areas like Santiago, Chile, a high population density is supported by a multimodal system of walking, bicycle transportation, motorcycles, buses, and trains. A small portion of the population can afford automobiles, but the automobiles greatly increase the congestion in these multimodal transportation systems. They also produce a considerable amount of air pollution, pose a significant safety risk, and exacerbate feelings of inequities in the society.

Other parts of the developing world, such as China, remain largely rural but are rapidly urbanizing and industrializing. In these areas a motorized infrastructure is being developed alongside motorization of the population. Great disparity of wealth means that only a fraction of the population can motorize, and therefore the highly dense multimodal transportation system for the poor is cross-cut by the highly motorized transportation system for the rich. The urban infrastructure is being rapidly developed, providing an opportunity to build new systems that incorporate ITS at early stages.

Intelligent transport technologies

Intelligent transport systems vary in technologies applied, from basic management systems such as car navigation; traffic signal control systems; container management systems; variable message signs; automatic number plate recognition or speed cameras to monitor applications, such as security CCTV systems; and to more advanced applications that integrate live data and feedback from a number of other sources, such as parking guidance and information systems; weather information; bridge deicing systems; and the like. Additionally, predictive techniques are being developed to allow advanced modeling and comparison with historical baseline data. Some of the constituent technologies typically implemented in ITS are described in the following sections.

Wireless communications

Various forms of wireless communications technologies have been proposed for intelligent transportation systems.

Radio modem communication on UHF and VHF frequencies are widely used for short and long range communication within ITS.

Short-range communications (less than 500 yards) can be accomplished using IEEE 802.11 protocols, specifically WAVE or the Dedicated Short Range Communications standard being promoted by the Intelligent Transportation Society of America and the

United States Department of Transportation. Theoretically, the range of these protocols can be extended using Mobile ad-hoc networks or Mesh networking.

Longer range communications have been proposed using infrastructure networks such as WiMAX (IEEE 802.16), Global System for Mobile Communications (GSM), or 3G. Long-range communications using these methods are well established, but, unlike the short-range protocols, these methods require extensive and very expensive infrastructure deployment. There is lack of consensus as to what business model should support this infrastructure.

Computational technologies

Recent advances in vehicle electronics have led to a move toward fewer, more capable computer processors on a vehicle. A typical vehicle in the early 2000s would have between 20 and 100 individual networked microcontroller/Programmable logic controller modules with non-real-time operating systems. The current trend is toward fewer, more costly microprocessor modules with hardware memory management and Real-Time Operating Systems. The new embedded system platforms allow for more sophisticated software applications to be implemented, including model-based process control, artificial intelligence, and ubiquitous computing. Perhaps the most important of these for Intelligent Transportation Systems is artificial intelligence.

Floating car data/floating cellular data

"Floating car" or "probe" data collection is a set of relatively low-cost methods for obtaining travel time and speed data for vehicles traveling along streets, highways, freeways, and other transportation routes. Broadly speaking, three methods have been used to obtain the raw data:

- **Triangulation Method.** In developed countries a high proportion of cars contain one or more mobile phones. The phones periodically transmit their presence information to the mobile phone network, even when no voice connection is established. In the mid 2000s, attempts were made to use mobile phones as anonymous traffic probes. As a car moves, so does the signal of any mobile phones that are inside the vehicle. By measuring and analyzing network data using triangulation, pattern matching or cell-sector statistics (in an anonymous format), the data was converted into traffic flow information. With more congestion, there are more cars, more phones, and thus, more probes. In metropolitan areas, the distance between antennas is shorter and in theory accuracy increases. An advantage of this method is that no infrastructure needs to be built along the road; only the mobile phone network is leveraged. But in practice the triangulation method can be complicated, especially in areas where the same mobile phone towers serve two or more parallel routes (such as a freeway with a frontage road, a freeway and a commuter rail line, two or more parallel streets, or a street that is also a bus line). By the early 2010s, the popularity of the triangulation method was declining.

- **Vehicle Re-Identification.** Vehicle re-identification methods require sets of detectors mounted along the road. In this technique, a unique serial number for a device in the vehicle is detected at one location and then detected again (re-identified) further down the road. Travel times and speed are calculated by comparing the time at which a specific device is detected by pairs of sensors. This can be done using the MAC (Machine Access Control) addresses from Bluetooth devices, or using the RFID serial numbers from Electronic Toll Collection (ETC) transponders (also called "toll tags").
- **GPS Based Methods.** An increasing number of vehicles are equipped with in-vehicle GPS (sattelite navigation) systems that have two-way communication with a traffic data provider. Position readings from these vehicles are used to compute vehicle speeds.

Floating car data technology provides advantages over other methods of traffic measurement:

- Less expensive than sensors or cameras
- More coverage (potentially including all locations and streets)
- Faster to set up and less maintenance
- Works in all weather conditions, including heavy rain

Sensing technologies

Technological advances in telecommunications and information technology, coupled with state-of-the-art microchip, RFID (Radio Frequency Identification), and inexpensive intelligent beacon sensing technologies, have enhanced the technical capabilities that will facilitate motorist safety benefits for intelligent transportation systems globally. Sensing systems for ITS are vehicle- and infrastructure-based networked systems, i.e., Intelligent vehicle technologies. Infrastructure sensors are indestructible (such as in-road reflectors) devices that are installed or embedded in the road or surrounding the road (e.g., on buildings, posts, and signs), as required, and may be manually disseminated during preventive road construction maintenance or by sensor injection machinery for rapid deployment. Vehicle-sensing systems include deployment of infrastructure-to-vehicle and vehicle-to-infrastructure electronic beacons for identification communications and may also employ video automatic number plate recognition or vehicle magnetic signature detection technologies at desired intervals to increase sustained monitoring of vehicles operating in critical zones.

Inductive loop detection

Inductive loops can be placed in a roadbed to detect vehicles as they pass through the loop's magnetic field. The simplest detectors simply count the number of vehicles during a unit of time (typically 60 seconds in the United States) that pass over the loop, while more sophisticated sensors estimate the speed, length, and weight of vehicles and the distance between them. Loops can be placed in a single lane or across multiple lanes, and they work with very slow or stopped vehicles as well as vehicles moving at high-speed.

Video vehicle detection

Traffic flow measurement and automatic incident detection using video cameras is another form of vehicle detection. Since video detection systems such as those used in automatic number plate recognition do not involve installing any components directly into the road surface or roadbed, this type of system is known as a "non-intrusive" method of traffic detection. Video from black-and-white or color cameras is fed into processors that analyze the changing characteristics of the video image as vehicles pass. The cameras are typically mounted on poles or structures above or adjacent to the roadway. Most video detection systems require some initial configuration to "teach" the processor the baseline background image. This usually involves inputting known measurements such as the distance between lane lines or the height of the camera above the roadway. A single video detection processor can detect traffic simultaneously from one to eight cameras, depending on the brand and model. The typical output from a video detection system is lane-by-lane vehicle speeds, counts, and lane occupancy readings. Some systems provide additional outputs including gap, headway, stopped-vehicle detection, and wrong-way vehicle alarms.

Intelligent transport applications

Emergency vehicle notification systems

The in-vehicle eCall is an emergency call generated either manually by the vehicle occupants or automatically via activation of in-vehicle sensors after an accident. When activated, the in-vehicle eCall device will establish an emergency call carrying both voice and data directly to the nearest emergency point (normally the nearest E1-1-2 Public-safety answering point, PSAP). The voice call enables the vehicle occupant to communicate with the trained eCall operator. At the same time, a minimum set of data will be sent to the eCall operator receiving the voice call.

The minimum set of data contains information about the incident, including time, precise location, the direction the vehicle was traveling, and vehicle identification. The pan-European eCall aims to be operative for all new type-approved vehicles as a standard option. Depending on the manufacturer of the eCall system, it could be mobile phone based (Bluetooth connection to an in-vehicle interface), an integrated eCall device, or a functionality of a broader system like navigation, Telematics device, or tolling device. eCall is expected to be offered, at earliest, by the end of 2010, pending standardization by the European Telecommunications Standards Institute and commitment from large EU member states such as France and the United Kingdom.

Congestion pricing gantry at North Bridge Road, Singapore.

Automatic road enforcement

Automatic speed enforcement gantry or "*Lombada Eletrônica*" with ground sensors at Brasilia, D.F.

A traffic enforcement camera system, consisting of a camera and a vehicle-monitoring device, is used to detect and identify vehicles disobeying a speed limit or some other road legal requirement and automatically ticket offenders based on the license plate number. Traffic tickets are sent by mail. Applications include:

- Speed cameras that identify vehicles traveling over the legal speed limit. Many such devices use radar to detect a vehicle's speed or electromagnetic loops buried in each lane of the road.
- Red light cameras that detect vehicles that cross a stop line or designated stopping place while a red traffic light is showing.
- Bus lane cameras that identify vehicles traveling in lanes reserved for buses. In some jurisdictions, bus lanes can also be used by taxis or vehicles engaged in car pooling.
- Level crossing cameras that identify vehicles crossing railways at grade illegally.
- Double white line cameras that identify vehicles crossing these lines.
- High-occupancy vehicle lane cameras for that identify vehicles violating HOV requirements.
- Turn cameras at intersections where specific turns are prohibited on red. This type of camera is mostly used in cities or heavy populated areas.

Variable speed limits

Example variable speed limit sign in the United States.

Recently some jurisdictions have begun experimenting with variable speed limits that change with road congestion and other factors. Typically such speed limits only change to decline during poor conditions, rather than being improved in good ones. One example is on Britain's M25 motorway, which circumnavigates London. On the most heavily-traveled 14-mile (23 km) section (junction 10 to 16) of the M25 variable speed limits combined with automated enforcement have been in force since 1995. Initial results indicated savings in journey times, smoother-flowing traffic, and a fall in the number of accidents, so the implementation was made permanent in 1997. Further trials on the M25 have been thus far proved inconclusive.

Collision avoidance systems

Japan has installed sensors on its highways to notify motorists that a car is stalled ahead.

Dynamic Traffic Light Sequence

Intelligent RFID traffic control has been developed for dynamic traffic light sequence. It circumvents or avoids problems that usually arise with systems that use image processing and beam interruption techniques. RFID technology with appropriate algorithm and database were applied to a multi vehicle, multi lane and multi road junction area to provide an efficient time management scheme. A dynamic time schedule was worked out for the passage of each column. The simulation has shown that, the dynamic sequence algorithm has the ability to intelligently adjust itself even with the presence of some extreme cases. The real time operation of the system able to emulate the judgment of a traffic policeman on duty, by considering the number of vehicles in each column and the routing proprieties.

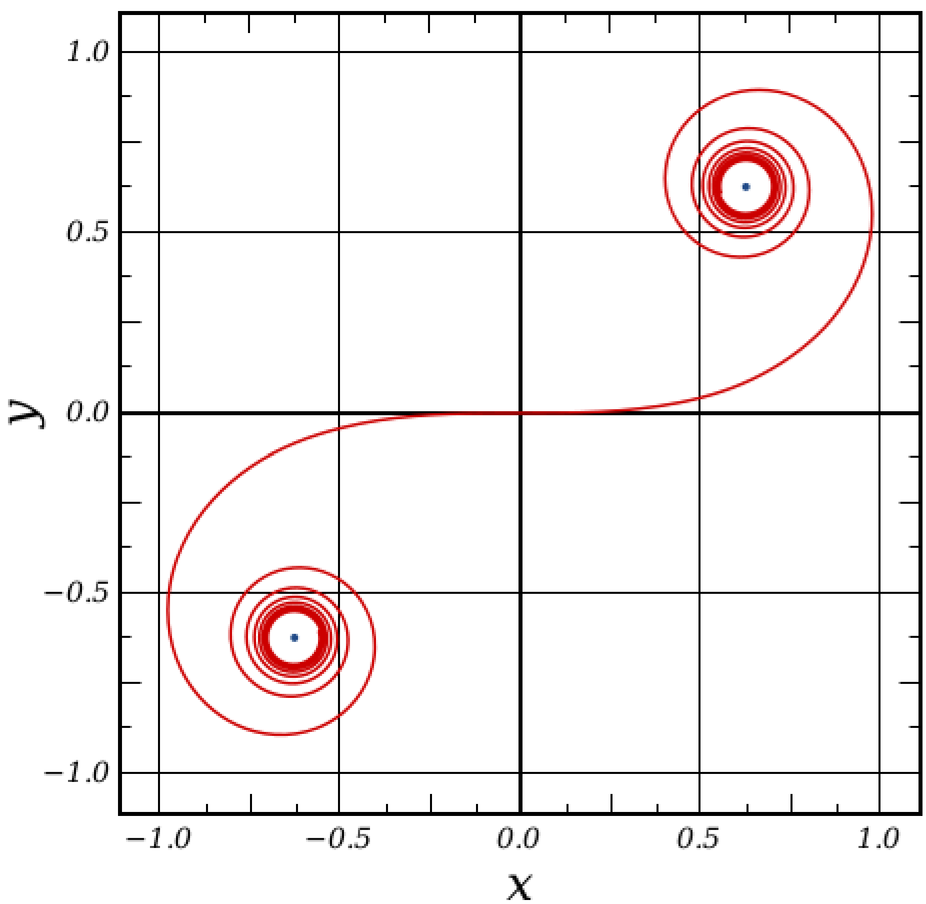
Cooperative systems on the road

Communication cooperation on the road includes car-to-car, car-to-infrastructure, and vice versa. Data available from vehicles is acquired and transmitted to a server for central fusion and processing. This data can be used to detect events such as rain (wiper activity) and congestion (frequent braking activities). The server processes a driving recommendation dedicated to a single or a specific group of drivers and transmits it wirelessly to vehicles. The goal of cooperative systems is to use and plan communication and sensor infrastructure to increase road safety. The definition of cooperative systems in road traffic is according to the European Commission:

> "Road operators, infrastructure, vehicles, their drivers and other road users will cooperate to deliver the most efficient, safe, secure and comfortable journey. The vehicle-vehicle and vehicle-infrastructure co-operative systems will contribute to these objectives beyond the improvements achievable with stand-alone systems."

Chapter 17

Euler Spiral



A double-end Euler spiral.

An **Euler spiral** is a curve whose curvature changes linearly with its curve length (the curvature of a circular curve is equal to the reciprocal of the radius). Euler spirals are also commonly referred to as **spiros**, **clothoids** or **Cornu spirals**.

Euler spirals have applications to diffraction computations. They are also widely used as transition curve in railroad/highway engineering for connecting and transiting the geometry between a tangent and a circular curve. The principle of linear variation of the curvature of the transition curve between a tangent and a circular curve defines the geometry of the Euler spiral:

- Its curvature begins with zero at the straight section (the tangent) and increases linearly with its curve length.
- Where the Euler spiral meets the circular curve, its curvature becomes equal to that of the latter.

Applications

Track transition curve

An object traveling on a circular path experiences a centripetal acceleration. When a vehicle traveling on a straight path approaches a circular path, it experiences a sudden centripetal acceleration starting at the tangent point; and thus centripetal force acts instantly causing much discomfort.

On early railroads this instant application of lateral force was not an issue since low speeds and wide-radius curves were employed (lateral forces on the passengers and the lateral sway was small and tolerable). As speeds of rail vehicles increased over the years, it became obvious that an easement is necessary so that the centripetal acceleration increases linearly with the traveled distance. Given the expression of centripetal acceleration V^2 / R, the obvious solution is to provide an easement curve whose curvature, $1 / R$, increases linearly with the traveled distance. This geometry is an Euler spiral.

Unaware of the solution of the geometry by Leonhard Euler, Rankine cited the cubic curve (a polynomial curve of degree 3), which is an approximation of the Euler spiral for small angular changes in the same way that a parabola is an approximation to a circular curve.

Marie Alfred Cornu (and later some civil engineers) also solved the calculus of Euler spiral independently. Euler spirals are now widely used in rail and highway engineering for providing a transition or an easement between a tangent and a horizontal circular curve.

Optics

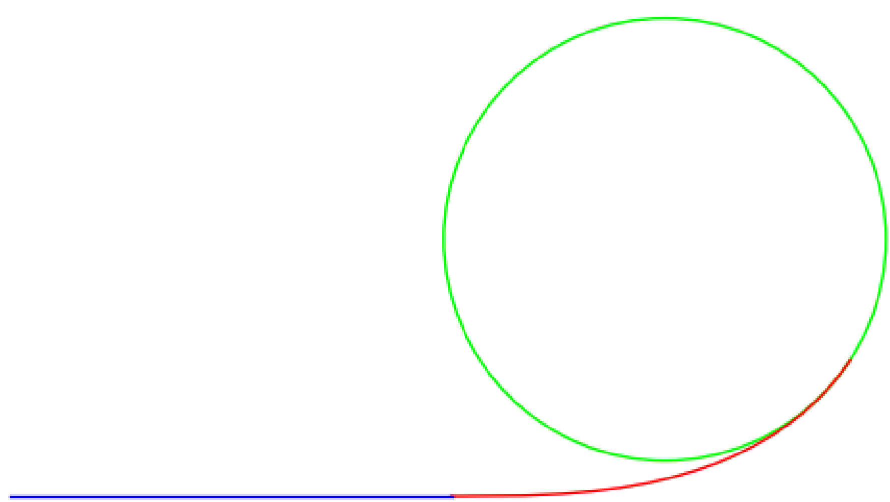
The Cornu spiral can be used to describe a diffraction pattern .

Formulation

Symbols

R	Radius of curvature
R_c	Radius of Circular curve at the end of the spiral
θ	Angle of curve from beginning of spiral (infinite *R*) to a particular point on the spiral.
	This can also be measured as the angle between the initial tangent and the tangent at the concerned point.
θ_s	Angle of full spiral curve
L, s	Length measured along the spiral curve from its initial position
L_s, s_o	Length of spiral curve

Derivation



The graph on the right illustrates an Euler spiral used as an easement (transition) curve between two given curves, in this case a straight line (the negative *x* axis) and a circle. The spiral starts at the origin in the positive *x* direction and gradually turns anticlockwise to osculate the circle.

The spiral is a small segment of the above double-end Euler spiral in the first quadrant.

From the definition of the curvature,

$$\frac{1}{R} = \frac{d\theta}{dL} \propto L$$

i.e.

$$RL = \text{constant} = R_c L_s$$

$$\frac{d\theta}{dL} = \frac{L}{R_c L_s}$$

We write in the format,

$$\frac{d\theta}{dL} = 2a^2 L$$

Where

$$2a^2 = \frac{1}{R_c L_s}$$

Or

$$a = \frac{1}{\sqrt{2R_c L_s}}$$

Thus

$$\theta = (aL)^2$$

Now

$$x = \int_0^L \cos\theta \, ds$$

$$= \int_0^L \cos\left[(as)^2\right] ds$$

If

$$L' = aL$$

Then

$$dL = \frac{dL'}{a}$$

Thus

$$x = \frac{1}{a}\int_0^{L'} \cos s^2 ds$$

$$y = \int_0^L \sin\theta \, ds$$

$$= \int_0^L \sin\left[(as)^2\right] ds$$

$$= \frac{1}{a}\int_0^{L'} \sin s^2 \, ds$$

Expansion of Fresnel integral

If a = 1, which is the case for normalized Euler curve, then the Cartesian coordinates are given by Fresnel integrals (or Euler integrals):

$$C(L) = \int_0^L \cos s^2 \, ds$$
$$S(L) = \int_0^L \sin s^2 \, ds$$

Expand C(L) according to power series expansion of cosine:

$$\cos\theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \cdots$$
$$C(L) = \int_0^L \cos s^2 \, ds$$
$$= \int_0^L (1 - \frac{s^4}{2!} + \frac{s^8}{4!} - \frac{s^{12}}{6!} + \cdots) \, ds$$
$$= L - \frac{L^5}{5 \times 2!} + \frac{L^9}{9 \times 4!} - \frac{L^{13}}{13 \times 6!} + \cdots$$

Expand S(L) according to power series expansion of sine:

$$\sin\theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \cdots$$
$$S(L) = \int_0^L \sin s^2 \, ds$$
$$= \int_0^L (s^2 - \frac{s^6}{3!} + \frac{s^{10}}{5!} - \frac{s^{14}}{7!} + \cdots) \, ds$$
$$= \frac{L^3}{3} - \frac{L^7}{7 \times 3!} + \frac{L^{11}}{11 \times 5!} - \frac{L^{15}}{15 \times 7!} + \cdots$$

Normalization and conclusion

For a given Euler curve with:

$$2RL = 2R_c L_s = \frac{1}{a^2}$$

or

$$\frac{1}{R} = \frac{L}{R_c L_s} = 2a^2 L$$

then

$$
\begin{aligned}
x &= \frac{1}{a}\int_0^{L'} \cos s^2 \, ds \\
y &= \frac{1}{a}\int_0^{L'} \sin s^2 \, ds
\end{aligned}
$$

where $L' = aL$ and $a = \dfrac{1}{\sqrt{2R_cL_s}}$.

The process of obtaining solution of (x, y) of an Euler spiral can thus be described as:

- Map L of the original Euler spiral by multiplying with factor a to L' of the normalized Euler spiral;
- Find (x', y') from the Fresnel integrals; and
- Map (x', y') to (x, y) by scaling up (denormalize) with factor $1 / a$. Note that $1 / a > 1$.

In the normalization process,

$$
\begin{aligned}
R'_c &= \frac{R_c}{\sqrt{2R_cL_s}} \\
&= \sqrt{\frac{R_c}{2L_s}} \\
L'_s &= \frac{L_s}{\sqrt{2R_cL_s}} \\
&= \sqrt{\frac{L_s}{2R_c}}
\end{aligned}
$$

Then

$$
\begin{aligned}
2R'_cL'_s &= 2\sqrt{\frac{R_c}{2L_s}}\sqrt{\frac{L_s}{2R_c}} \\
&= \tfrac{2}{2} \\
&= 1
\end{aligned}
$$

Generally the normalization reduces L' to a small value (<1) and results in good converging characteristics of the Fresnel integral manageable with only a few terms.

Illustration

Given:

$$R_c = 300\text{m}$$
$$L_s = 100\text{m}$$

Then

$$\begin{aligned}\theta_s &= \frac{L_s}{2R_c}\\ &= \frac{100}{2 \times 300}\\ &= 0.1667 \text{ radian}\end{aligned}$$

And

$$2R_cL_s = 60,000$$

We scale down the Euler spiral by √60,000, i.e.100√6 to normalized Euler spiral that has:

$$\begin{aligned}R'_c &= \tfrac{3}{\sqrt{6}}\text{m}\\ L'_s &= \tfrac{1}{\sqrt{6}}\text{m}\\ 2R'_cL'_s &= 2 \times \tfrac{3}{\sqrt{6}} \times \tfrac{1}{\sqrt{6}}\\ &= 1\end{aligned}$$

And

$$\begin{aligned}\theta_s &= \frac{L'_s}{2R'_c}\\ &= \frac{\frac{1}{\sqrt{6}}}{2 \times \frac{3}{\sqrt{6}}}\\ &= 0.1667 \text{ radian}\end{aligned}$$

The two angles θ_s are the same. This thus confirm that the original and normalized Euler spirals are having geometric similarity. The locus of the normalized curve can be determined from Fresnel Integral, while the locus of the original Euler spiral can be obtained by scaling back / up or denormalizing.

Other properties of normalized Euler spiral

Normalized Euler spiral can be expressed as:

$$x = \int_0^L \cos s^2 ds$$
$$y = \int_0^L \sin s^2 ds$$

Normalized Euler spiral has the following properties:

$$2R_c L_s = 1$$
$$\theta_s = \frac{L_s}{2R_c} = L_s^2$$

And

$$\theta = \theta_s \cdot \frac{L^2}{L_s^2} = L^2$$
$$\frac{1}{R} = \frac{d\theta}{dL} = 2L.$$

Note that $2R_cL_s = 1$ also means that $1 / R_c = 2L_s$, in agreement with the above.

Mathematical codes producing an Euler spiral

The following Sage code produces the second graph above. The first four lines express the Euler spiral component. Fresnel functions could not be found. Instead, the integrals of two expanded Taylor series are adopted. The remaining code expresses respectively the tangent and the circle, including the computation for the center coordinates.

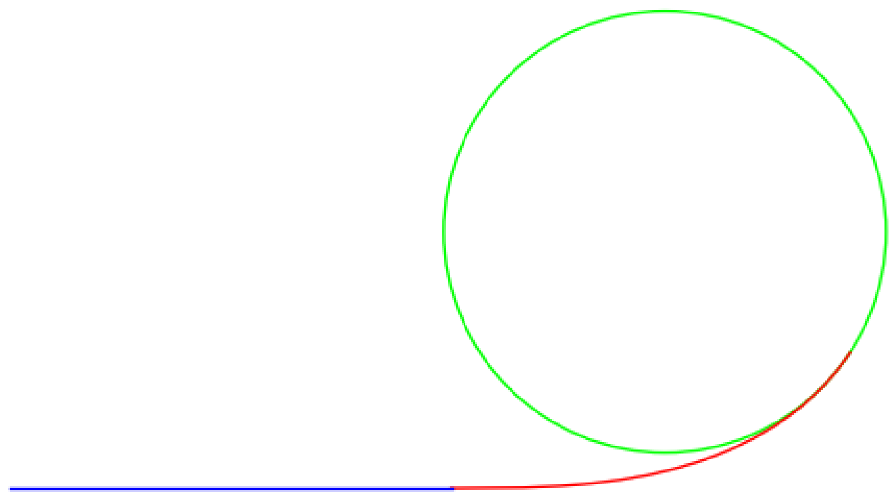
```
var('L')
p = integral(taylor(cos(L^2), L, 0, 12), L)
q = integral(taylor(sin(L^2), L, 0, 12), L)
r1 = parametric_plot([p, q], (L, 0, 1), color = 'red')

r2 = line([(-1.0, 0), (0,0)], rgbcolor = 'blue')

x1 = p.subs(L = 1)
y1 = q.subs(L = 1)
R = 0.5
x2 = x1 - R*sin(1.0)
y2 = y1 + R*cos(1.0)
r3 = circle((x2, y2), R, rgbcolor = 'green')
show(r1 + r2 + r3, aspect_ratio = 1, axes=false)
```

The following is Mathematica code for the Euler spiral component:

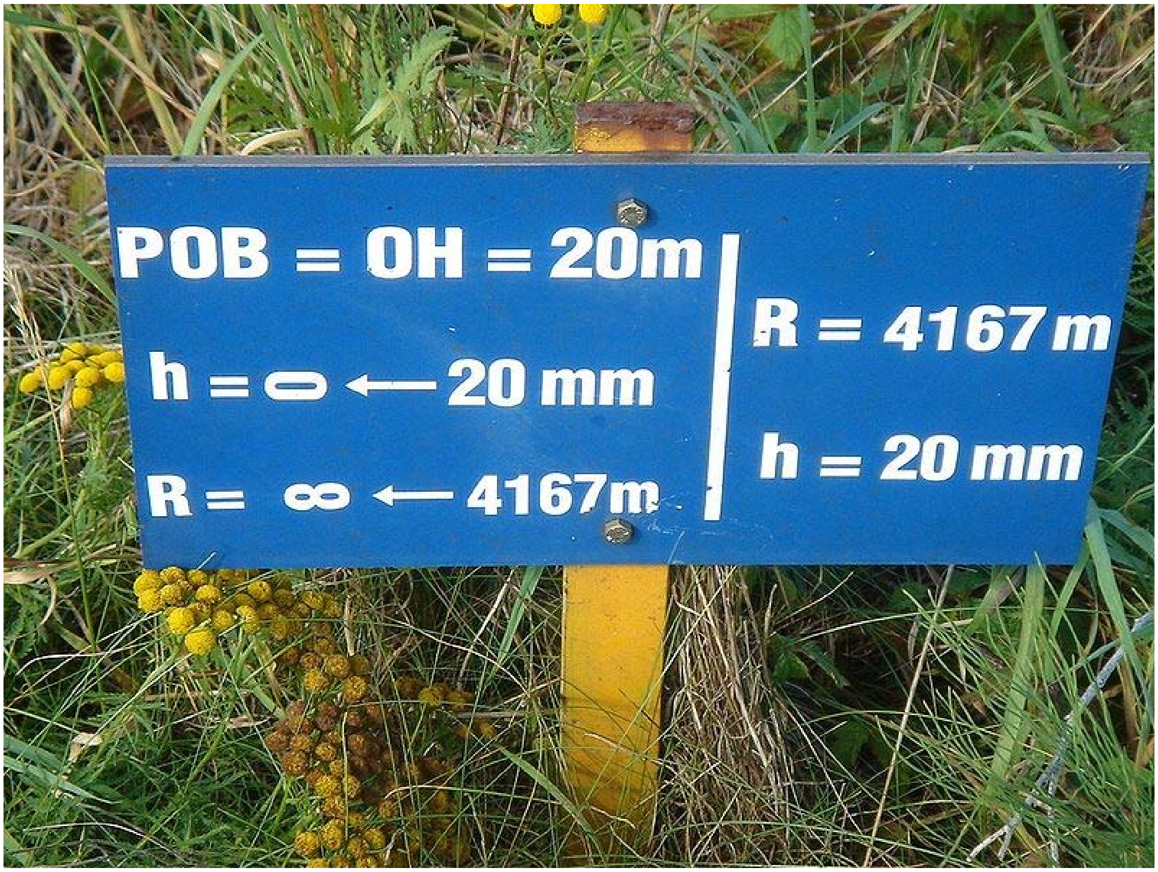
```
ParametricPlot[
 {FresnelC[Sqrt[2/\[Pi]] t]/Sqrt[2/\[Pi]],
  FresnelS[Sqrt[2/\[Pi]] t]/Sqrt[2/\[Pi]]},
 {t, -10, 10}]
```

Chapter 18

Track Transition Curve



The red **Euler spiral** is an example of an easement curve between a blue straight line and a circular arc, which shall be a segment of the green circle.



This sign aside a railroad (between Ghent and Bruges) indicates the start of the transition curve. A parabolic curve (*POB*) is used.

A **Track transition curve**, or **spiral easement**, is a mathematically calculated curve on a section of highway, or railroad track, where a straight section changes into a curve. It is designed to reduce the effects of centrifugal force experienced by users. In plan (i.e., the horizontal curve) the start of the transition is at infinite radius and at the end of the transition it has the same radius as the curve itself, thus forming a very broad spiral. At the same time, in the vertical plane, the outside of the curve is gradually raised until the correct degree of bank is reached.

If such easement were not applied, the lateral acceleration of a rail vehicle would change abruptly at one point – the tangent point where the straight track meets the curve – with undesirable results. With a road vehicle the driver naturally applies the steering alteration in a gradual manner and the curve is designed to permit this, using the same principle.

History

On early railroads, because of the low speeds and wide-radius curves employed, the surveyors were able to ignore any form of easement, but during the 19th century, as speeds increased, the need for a track curve with gradually increasing curvature became apparent. Rankine's 1862 "Civil Engineering" cites several such curves, including an

1828 or 1829 proposal based on the "curve of sines" by William Gravatt, and the *curve of adjustment* by William Froude around 1842 approximating the elastic curve. The actual equation given in Rankine is that of a cubic curve, which is a polynomial curve of degree 3. This was also known as cubic parabola at that time.

In the UK, only from 1845 when legislation and land costs began to constrain the laying out of rail routes and tighter curves were necessary, did the principles start to be applied in practice.

The "true spiral", where the curvature is exactly linear in arclength, requires more sophisticated mathematics (in particular, the ability to integrate its intrinsic equation) to compute than the proposals cited by Rankine. Several late-19th century civil engineers seem to have derived the equation for this curve independently (all unaware of the original characterization of this curve by Leonhard Euler in 1744). Charles Crandall gives credit to one Ellis Holbrook, in the Railroad Gazette, Dec. 3, 1880, for the first accurate description of the curve. Another early publication was *The Railway Transition Spiral* by Arthur N. Talbot, originally published in 1890. Some early 20th century authors call the curve "Glover's spiral", attributing it to James Glover's 1900 publication.

The equivalence of the railroad transition spiral and the clothoid seems to have been first published in 1922 by Arthur Lovat Higgins. Since then, "clothoid" is the most common name given the curve, even though the correct name (following standards of academic attribution) is "the Euler spiral".

Geometry

While railroad track geometry is intrinsically three-dimensional, for practical purposes the vertical and horizontal components of track geometry are usually treated separately.

The overall design pattern for the vertical geometry is typically a sequence of constant grade segments connected by vertical transition curves in which the local grade varies linearly with distance and in which the elevation therefore varies quadratically with distance. Here grade refers to the tangent of the angle of rise of the track. The design pattern for horizontal geometry is typically a sequence of straight line (i.e., a tangent) and curve (i.e. a circular arc) segments connected by transition curves.

In a tangent segment the track bed roll angle is typically zero. In the case of railroad track the track roll angle (*cant* or *camber*) is typically expressed as the difference in elevation of the two rails, a quantity referred to as the superelevation. A track segment with constant non-zero curvature will typically be superelevated in order to have the component of gravity in the plane of the track provide a majority of the centripetal acceleration inherent in the motion of a vehicle along the curved path so that only a small part of that acceleration needs to be accomplished by lateral force applied to vehicles and passengers or lading. The change of superelevation from zero in a tangent segment to the value selected for the body of a following curve occurs over the length of a transition curve that connects the tangent and the curve proper. Over the length of the transition the curvature of the track will also vary from zero at the end abutting the tangent segment to the value of curvature of the curve body, which is numerically equal to one over the radius of the curve body.

The simplest and most commonly used form of transition curve is that in which the superelevation and horizontal curvature both vary linearly with distance along the track. Cartesian coordinates of points along this spiral are given by the Fresnel integrals. The resulting shape matches a portion of an Euler spiral, which is also commonly referred to as a clothoid, and sometimes Cornu spiral.

A transition curve can connect a track segment of constant non-zero curvature to another segment with constant curvature that is zero or non-zero of either sign. Successive curves in the same direction are sometimes called progressive curves and successive curves in opposite directions are called reverse curves.

The Euler spiral has two advantages. One is that it is easy for surveyors because the coordinates can be looked up in Fresnel integral tables. The other is that it provides the shortest transition subject to a given limit on the rate of change of the track superelevation (i.e. the twist of the track). However, as has been recognized for a long time, it has undesirable dynamic characteristics due to the large (conceptually infinite) roll acceleration and rate of change of centripetal acceleration at each end. Because of the capabilities of personal computers it is now practical to employ spirals that have dynamics better than those of the Euler spiral.