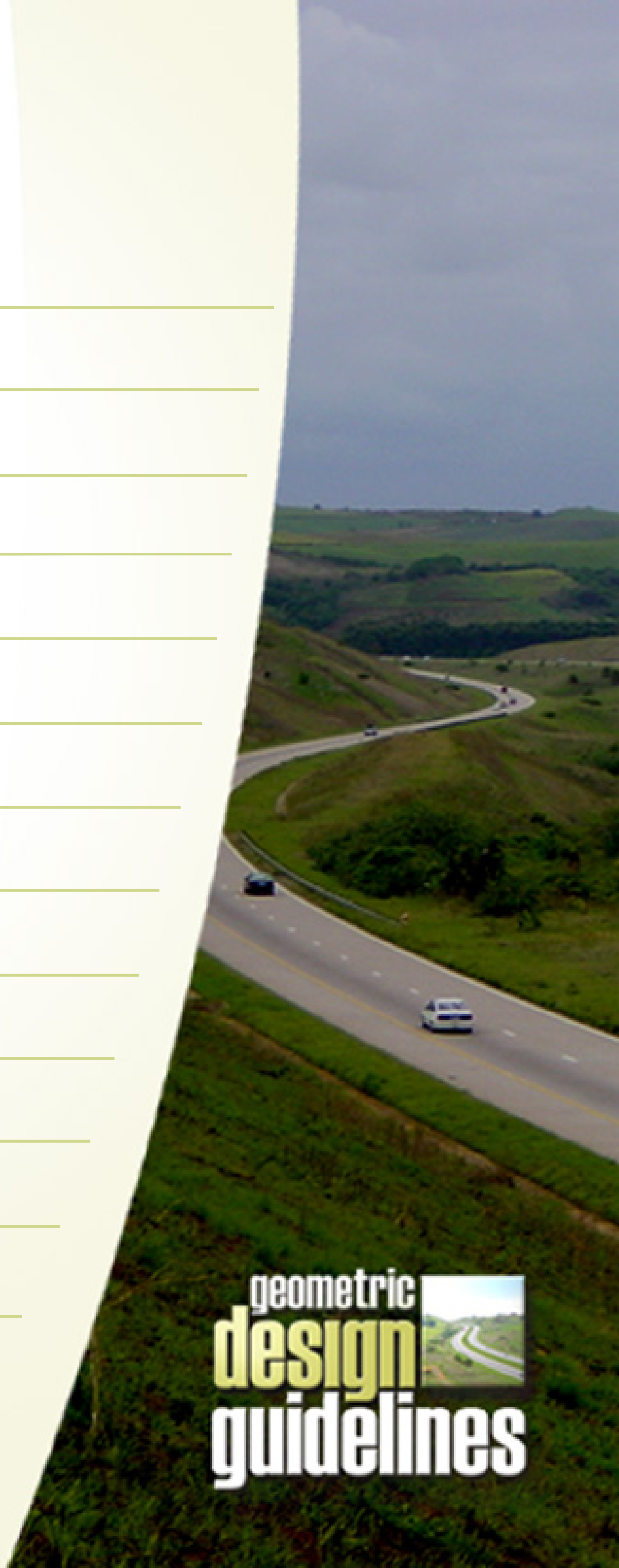
index



FOREWORD

This document describes the paradigm shift that has occurred in the field of geometric design over the last five years. Previous geometric design manuals were based on the limitations of the vehicles for which the roads were being designed and were not particularly concerned with the limitations or desires of the drivers of those vehicles. They also were prescriptive in nature, laying down rigid standards that did not permit any form of relaxation to match some or other specific circumstance. It was presumed that a road that was designed to standard was safe. Furthermore, the road reserve was perceived principally to be dedicated to the movement of vehicles, with little or no allowance being made for other users of the road reserve.

By way of contrast, this – the most recent edition of the SANRAL G2 manual – is a guideline as opposed to a prescriptive set of standards and is based on a human factors approach to design. The flexibility that results from application of the concepts of the design domain and context sensitive design is discussed.

Previous manuals did not address issues such as community sensitivities and values. In consequence, there was perhaps a tendency amongst designers to ride roughshod over the needs of those for whom the road, in fact, was being designed. While this guideline document does not spell out the methods whereby community support for the design process and its outcome can be obtained, it nevertheless stresses the need to achieve this involvement.

If mentioned at all in previous manuals, the need for coordination of the horizontal and vertical alignment was essentially aimed at driver guidance. This was achieved, for example, either by containing vertical curves within horizontal curves or remote from them. In this guideline, the development of an aesthetic product is dealt with in depth. Aesthetic design requires the elements of the road to be in harmony with each other. The road, as a complete entity, should constitute an enhancement of the landscape. It is thus stressed that design is as much art as it is science.

It must be realised that this departure from previous design practices will make the design process more rather than less complex. The designer will have to consider all users of the road whether in motorised or unmotorised vehicles or on foot and their reasons for being in the road reserve. These reasons span wider than a desire to travel from one point to another and include relaxation, recreation and social contact. Roadside vending is a matter of convenience to some road users and a source of income to others and needs to be managed to ensure the safety of all concerned. The sizing of any element will have to be a matter of careful consideration rather than simple adherence to a set of standards.

Throughout the document, reference is made to the importance of a sound philosophy of design. It is to the development, encouragement and maintenance of this philosophy of design that this guideline document is dedicated.

Nazir Alli
Chief Executive Officer
South African National Road Agency Limited

Glossary

A

Acceleration lane. An auxiliary lane used by an entering vehicle to accelerate before entering the travelled way.

Access control. The condition whereby the road agency either partially or fully controls the right of abutting landowners to direct access to and from a public highway or road.

Access interchange. An interchange providing access to a freeway from the adjacent non-freeway road network.

Arterial. Highway designed to move relatively large volumes of traffic at high speeds over long distances. Typically, arterials offer little or no access to abutting properties.

Auxiliary lane. Short lane located immediately adjacent to the basic or through lane to accommodate some or other special circumstance such as a turning movement to right or to left, acceleration to or deceleration from the speeds prevailing on the travelled way or heavy vehicles reduced to crawl speeds on a steep upgrade.

Average Daily Traffic (ADT). The number of vehicles per day passing a point on the highway during a defined period. If this period extends from 1 January to 31 December, reference is to Annual Average Daily Traffic (AADT).

Average running speed. The distance summation for all vehicles divided by the running time summation for all vehicles. Also referred to as space mean speed whereas time mean speed is simply the average of all recorded speeds.

Axis of rotation. The line about which the pavement is rotated to superelevate the roadway. This line normally maintains the highway profile

B

Barrier sight distance. The limiting sight distance below which overtaking is legally prohibited.

Boulevard. The area separating sidewalks from the through lanes.

Bridge. A structure erected with a deck for carrying traffic over or under an obstruction and with a clear span of six metres or more. Where the clear span is less than six metres, reference is to a culvert.

Broken-back curve. Two curves in the same direction with a tangent shorter than 500 metres long connecting them.

C

Camber. The slope from a high point (typically at the centre line of the highway) across the lanes of a highway. Negative camber refers to a central low point, usually with a view to drainage of a small urban street or alley.

Capacity. The maximum number of vehicles that can pass a point on a highway or in a designated lane in one hour without the density being so great as to cause unreasonable delay or restrict the driver's freedom to manoeuvre under prevailing roadway and traffic conditions.

Carriageway. Roadway forming part of a divided highway and intended for movement in one direction only – hence dual carriageway as an alternative name for divided highway.

Catchwater drain. Located above a cut face to ensure that storm water does not flow down the cut face causing erosion and deposition of silt on the roadway.

Channel grading. Where side channels are designed to gradients that differ from those of the road centreline, typically on either side of the highest points on crest curves and the lowest points on sag curves where the centreline gradient is less than 0,5 per cent.

Channelisation. The use of pavement markings or islands to direct traffic through an intersection

Clearance profile. Describes the space that is exclusively reserved for provision of the road or highway. It defines the minimum height of the soffit of any structure passing over the road and the closest approach of any lateral obstacle to the cross-section.

Cloverleaf interchange. An interchange with loop ramps in all quadrants to accommodate right turns and outer connectors for the left turns.

Collector. A road characterised by a roughly even distribution of its access and mobility functions.

Collector-Distributor road. A road used at an interchange to remove weaving from the through lanes and to reduce the number of entrances to and exits from the through lanes.

Compound curve. A combination of two or more curves in the same direction without intervening tangents between them.

Criterion. A yardstick according to which some or other quality of the road can be measured. Guideline values are specific numerical values of the criterion. For example, delay is a criterion of congestion.

Critical length of grade. The maximum length of a specific upgrade on which a loaded truck can operate without an unreasonable reduction in speed. Very often, a speed reduction of 15 km/h or more is considered "unreasonable".

Cross fall. See camber. In the case of cross fall, the high point is at the roadway edge.

Cross-over crownline. The line across which an instantaneous change of camber takes place. In the case of a normally cambered road, the centreline is a special case of the cross-over crownline. The cross-over crownline can be located anywhere on the road surface and need not even be parallel to the road centreline.

Crosswalk. A demarcated area or lane designated for the use of pedestrians across a road or street.

Crown runoff. (Also referred to as tangent runout) The rotation of the outer lane of a two-lane road from zero cross fall to normal camber (NC).

Culvert. A structure, usually for conveying water under a roadway but can also be used as a pedestrian or stock crossing, with a clear span of less than six metres.

Cut. Section of highway or road below natural ground level. Sometimes referred to in other documents as a cutting or excavation.

Cycle lane. A portion of the roadway which has been designated by road markings, striping and signing as being exclusively for the use of cyclists.

Cycle path. Also known as a bike way. A path physically separated from motorised traffic by an open space or barrier and located either within the road reserve or an independent reserve.

D

Decision sight distance. Sometimes referred to as anticipatory sight distance, allows for circumstances where complex decisions are required or unusual manoeuvres have to be carried out. As such, it is significantly longer than Stopping Sight Distance.

Density. The number of vehicles occupying a given length of road. Usually averaged over time and expressed as vehicles per kilometre.

Depressed median. A median lower in elevation than the travelled way and so designed to carry portion of the storm water falling on the road.

Design domain. The range of values of a design criterion that are applicable to a given design, e.g. lane widths of more than 3,3 metres.

Design hour. The hour in which the condition being designed for, typically the anticipated flow, is expected to occur. This is often the thirtieth highest hour of flow in the design year.

Design speed. The speed selected as the basis for establishing appropriate geometric elements for a section of road.

Design vehicle.
A compilation of the 85th percentile values of the various parameters of the vehicle type being designed for, e.g. length, width, wheelbase, overhang, height, ground clearance, etc.

Design year. The last year of the design life of the road or any other facility, often taken as twenty years although, for costly structures such as major bridges, a longer period is usually adopted.

Directional distribution (split). The percentages of the total flow moving in opposing directions, e.g. 50:50, 70:30, with the direction of interest being quoted first.

Divided highway. A highway with separate carriageways for traffic moving in opposite directions.

Driveway. A road providing access from a public road to a street or road usually located on an abutting property.

E

Eighty-fifth percentile speed. The speed below which 85 per cent of the vehicles travel on a given road or highway.

F

Footway. The rural equivalent of the urban sidewalk.

Freeway. Highest level of arterial characterised by full control of access and high design speeds.

Frontage road. A road adjacent and parallel to but separated from the highway for service to abutting properties and for control of access. Sometimes also referred to as a service road.

G

Gap. The elapsed time between the back of one vehicle passing a point on the road or highway and the nose of the following vehicle passing the same point. A lag is the unexpired portion of a gap, i.e. the elapsed time between the arrival of a vehicle on the minor leg of an intersection and the nose of the next vehicle on the major road crossing the path of the entering vehicle.

Gore area. The paved triangular area between the through lanes and the exit or entrance ramps at interchanges plus the graded areas immediately beyond the nose (off-ramp) or merging end (on-ramp).

Grade line. The line describing the vertical alignment of the road or highway.

Grade. The straight portion of the grade line between two successive vertical curves.

Grade separation. A crossing of two highways or roads, or a road and a railway, at different levels.

Gradient. The slope of the grade between two adjacent Vertical Points of Intersection (VPI), typically expressed in percentage form as the vertical rise or fall in metres/100 metres. In the direction of increasing stake value, upgrades are taken as positive and downgrades as negative.

Guideline. A design value establishing an approximate threshold, which should be met if considered practical. It is a recommended value whereas a standard is a prescriptive value allowing for no exceptions.

H

High occupancy vehicle (HOV) lane. A lane designated for the exclusive use of buses and other vehicles carrying more than two passengers.

High-speed. Typically where speeds of 80 km/h or faster are being considered.

Horizontal sight distance. The sight distance determined by lateral obstructions alongside the road and measured at the centre of the inside lane.

I

Interchange. A system of interconnecting roads (referred to as ramps) in conjunction with one or more grade separations providing for the movement of traffic between two or more roadways which are at different levels at their crossing point.

Intersection sight distance. The sight distance required within the quadrants of an intersection to safely allow turning and crossing movements.

J, K

Kerb. Concrete, often precast, element adjacent to the travelled way and used for drainage control, delineation of the pavement edge or protection of the edge of surfacing. Usually applied only in urban areas.

Kerb ramp. The treatment at intersections for gradually lowering the elevation of sidewalks to the elevation of the street surface.

K-value. The distance over which a one per cent change in gradient takes place.

L

Level of Service (LOS). A qualitative concept, from LOS A to LOS F, which characterises acceptable degrees of congestion as perceived by drivers. Capacity is defined as being at LOS E.

Low speed. Typically where speeds of 70 km/h or slower are being considered.

M

Median. The portion of a divided highway separating the two travelled ways for traffic in opposite directions. The median thus includes the inner shoulders.

Median opening. An at-grade opening in the median to allow vehicles to cross from a roadway to the adjacent roadway on a divided road.

Modal transfer station. The public facility at which passengers change from one mode of transport to another, e.g. rail to bus, passenger car to rail.

Mountainous terrain. Longitudinal and transverse natural slopes are severe and changes in elevation abrupt. Many trucks operate at crawl speeds over substantial distances.

N

Normal crown (NC). The typical cross-section on a tangent section of a two-lane road or four-lane undivided road.

O

Overpass. A grade separation where a minor highway passes over the major highway.

Outer separator. Similar to the median but located between the travelled way of the major road and the travelled way of parallel lanes serving a local function *if these lanes are contained within the reserve of the major road*. If they fall outside this reserve, reference is to a frontage road.

P

Partial Cloverleaf (Par-Clo) Interchange. An interchange with loop ramps in one, two or three (but usually only two) quadrants. A Par-Clo A Interchange has the loops in advance of the structure and Par-Clo B Interchange has the loops beyond the structure. A Par-Clo AB Interchange has its loops on the same side of the crossing road.

Passenger car equivalents (units) (PCE or PCU). A measure of the impedance offered by a vehicle to the passenger cars in the traffic stream. Usually quoted as the number of passenger cars required to offer a similar level of impedance to the other cars in the stream.

Passing sight distance. The total length visibility, measured from an eye height of 1,05 metres to an object height of 1,3 metres, necessary for a passenger car to overtake a slower moving vehicle. It is measured from the point at which the initial acceleration commences to the point where the overtaking vehicle is once again back in its own lane.

PC (Point of curvature). Beginning of horizontal curve, often referred to as the BC.

PI (Point of intersection). Point of intersection of two tangents.

PRC (Point of reverse curvature). Point where a curve in one direction is immediately followed by a curve in the opposite direction. Typically applied only to kerb lines.

PT (Point of tangency). End of horizontal curve, often referred to as EC.

PVC (Point of vertical curvature) The point at which a grade ends and the vertical curve begins, often also referred to as BVC.

PVI (Point of vertical intersection). The point where the extension of two grades intersect. The initials are sometimes reversed to VPI.

PVT (Point of vertical tangency). The point at which the vertical curve ends and the grade begins. Also referred to as EVC.

Q

Quarter link. An interchange with at-grade intersections on both highways or roads and two ramps (which could be a two-lane two-way road) located in one quadrant. Because of its appearance, also known as a Jug Handle Interchange.

R

Ramp. A one-way, often single-lane, road providing a link between two roads that cross each other at different levels.

Relative gradient. The slope of the edge of the travelled way relative to the gradeline.

Reverse Camber (RC). A superelevated section of roadway sloped across the entire travelled way at a rate equal to the normal camber.

Reverse curve. A combination of two curves in opposite directions with a short intervening tangent

Road safety audit. A structured and multidisciplinary process leading to a report on the crash potential and safety performance of a length of road or highway, which report may or may not include suggested remedial measures.

Roadside. A general term denoting the area beyond the shoulder breakpoints.

Road bed. The extent of the road between shoulder breakpoints.

Road prism. The lateral extent of the earthworks.

Road reserve. Also referred to as Right-of-way. The strip of land acquired by the road authority for provision of a road or highway.

Roadway. The lanes and shoulders excluding the allowance (typically 0,5 metres) for rounding of the shoulders.

Rolling terrain. The natural slopes consistently rise above and fall below the highway grade with, occasionally, steep slopes presenting some restrictions on highway alignment. In general, rolling terrain generates steeper gradients, causing truck speeds to be lower than those of passenger cars.

Rural road or highway. Characterised by low-volume high-speed flows over extended distances. Usually without significant daily peaking but could display heavy seasonal peak flows.

S

Shoulder. Usable area immediately adjacent to the travelled way provided for emergency stopping, recovery of errant vehicles and lateral support of the roadway structure.

Shoulder breakpoint. The hypothetical point at which the slope of the shoulder intersects the line of the fill slope. Sometimes referred to as the hinge point.

Side friction (f). The resistance to centrifugal force keeping a vehicle in a circular path. The designated maximum side friction (f_{max}) represents a threshold of driver discomfort and not the point of an impending skid.

Sidewalk. The portion of the cross-section reserved for the use of pedestrians.

Sight triangle. The area in the quadrants of an intersection that must be kept clear to ensure adequate sight distance between the opposing legs of the intersection.

Simple curve. A curve of constant radius without entering or exiting transitions.

Single point urban interchange. A diamond interchange where all the legs of the interchange meet at a common point on the crossing road.

Speed profile. The graphical representation of the 85th percentile speed achieved along the length of the highway segment by the design vehicle.

Standard. A design value that may not be transgressed, e.g. an irreducible minimum or an absolute maximum. In the sense of geometric design, not to be construed as an indicator of quality, i.e. an ideal to be strived for.

Stopping sight distance. The sum of the distance travelled during a driver's perception/reaction time and the distance travelled thereafter while braking to a stop.

Superelevation. The amount of cross-slope provided on a curve to help counterbalance, in combination with side friction, the centrifugal force acting on a vehicle traversing the curve.

Superelevation runoff. (Also referred to as superelevation development) The process of rotating the outside lane from zero crossfall to reverse camber (RC), thereafter rotating both lanes to the full superelevation selected for the curve.

Systems interchange. Interchange connecting two freeways, i.e. a node in the freeway system.

T

Tangent. The straight portion of a highway between two horizontal curves.

Tangent runoff. See crown runoff

Traffic composition. The percentage of vehicles other than passenger cars in the traffic stream, e.g. 10 per cent trucks, 5 per cent articulated vehicles (semi-trailers) etc.

Transition curve. A spiral located between a tangent and a circular curve.

Travelled way. The lanes of the cross-section. The travelled way excludes the shoulders.

Trumpet interchange. A three-legged interchange containing a loop ramp and a directional ramp, creating between them the appearance of the bell of a trumpet.

Turning roadway. Channelised turn lane at an at-grade intersection.

Turning template. A graphic representation of a design vehicle's turning path for various angles of turn. If the template includes the paths of the outer front and inner rear points of the vehicle, reference is to the swept path of the vehicle.

U

Underpass. A grade separation where the subject highway passes under an intersecting highway.

Urban road or highway. Characterised by high traffic volumes moving at relatively low speeds and pronounced peak or tidal flows. Usually within an urban area but may also be a link traversing an unbuilt up area between two adjacent urban areas, hence displaying urban operational characteristics.

V

Value engineering. A management technique in which intensive study of a project seeks to achieve the best functional balance between cost, reliability and performance.

Verge. The area between the edge of the road prism and the reserve boundary

W

Warrant. A guideline value indicating whether or not a facility should be provided. For example, a warrant for signalisation of an intersection would include the traffic volumes that should be exceeded before signalisation is considered as a traffic control option. Note that, once the warranting threshold has been met, this is an indication that the design treatment should be considered and evaluated and **not** that the design treatment is automatically required.

X, Y, Z

Yellow line break point. A point where a sharp change of direction of the yellow edge line demarcating the travelled way edge takes place. Usually employed to highlight the presence of the start of a taper from the through lane at an interchange.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

The geometric design of a highway or of one of its many elements is only one step in a multifaceted process from concept to construction. However, the constraints, which the physical elements ultimately place on the function and form of a highway, pervade every step in the process. Knowledge of the parameters, which govern planning and design together with their practical application, is thus essential. These guidelines seek to meet that need.

The emphasis previously of Geometric Design Manuals was on design standards for new construction. The South African primary road network is, however, substantially complete and new road works are largely limited to urban developments. This Manual thus deals not only with new works but also pays attention to rehabilitation, reconstruction and upgrading projects. A feature of these projects is that the designer's freedom of choice is often restricted by developments surrounding the road to be rehabilitated. In consequence, adherence to rigidly applied standards is not possible, in addition to the fact that blind adherence has never been construed as a thinking designer's approach to the problem at hand.

These geometric design guidelines are intended for use on National Roads – or on any other roads falling within the domain of the S A National Roads Agency Limited. For this reason, the guidelines address a wide range of functional uses and requirements. They will also need to cater for a multiplicity of users, and designers will be faced with competing demands from different sections of the community as they endeavour to design safe and operationally efficient roads.

A major objective of any road design guide is to ensure that designs achieve value for money without any significant deleterious effect on safety. The design philosophy, systems and techniques developed elsewhere in this document have been based on the Design Speed approach and related geometric parameters which will result in a much greater flexibility to achieve economic design in varied and sometimes difficult circumstances.

In line with this, the standards in this guideline will address a spectrum of road types, varying from multi-lane freeways carrying traffic volumes of over 100 000 vehicles per day, to single carriageway roads carrying volumes of the order of 500 vehicles per day. In respect of this latter class of road design, recommendations have been considerably extended to allow greater flexibility in design, with particular emphasis on the co-ordination of design elements to improve safety and overtaking conditions.

The guidelines distinguish between roads in rural areas and those in urban areas and also caters for situations where National Roads traverse the CBDs of smaller municipalities.

Overall, the greater flexibility in design introduced in these guidelines will enable more economic designs, reducing both the construction costs and the impact of new roads and road improvements on the environment.

TABLE OF CONTENTS

TABLE OF FIGURES

Chapter 2

DESIGN PHILOSOPHY AND TECHNIQUES

2.1 BACKGROUND

The Department of Transport completed the "Moving South Africa" project in 1999. One of the findings was that, in order to reduce congestion, a shift from private to public transport would be required. As many people are captive to public transport, it is also necessary to create an environment supportive of public transport.

This will require encouraging:

- Settlement densities supportive of public transport;
- Network layouts with geometric design standards suitable for bus and taxi routes including:
  - Safe stopping sight distance;
  - Reduced gradients;
  - Minimum horizontal curvature;
  - Intersection layouts that are simple to negotiate;
  - User friendly bus stops;
  - Terminals and modal transfer stations; and
  - High occupancy vehicle lanes.

Optimising resources requires the building of networks with the lowest possible whole-life costs. This has always been a goal but, historically, the emphasis tended to be on minimum construction costs and, more recently, on minimum combined construction and maintenance costs. A subsequent shift in emphasis caused the focus to move towards the whole-life economy of the network.

The network with the shortest overall length compatible with linking all origins and destinations would theoretically have the lowest cost. It could also represent a saving in maintenance cost provided that the attempt to reduce the network length did not adversely impact on the vertical alignment, resulting in very steep gradients or a poorly drained road, both of which could carry a maintenance and construction penalty. Assuming that maintenance is practical, it is possible that the network, short though it may be, forces circuitous travel paths, which would nullify any savings on construction and maintenance. It follows that the shape of the network is as important as its overall length in optimising the life-cycle cost. Geometric Design, which is often incorrectly construed as the selection, sizing and grouping of a set of components to create a road network, must therefore contain a strong element of Geometric Planning.

Geometric Planning includes careful selection of the cross-section. The road width and shape has a significant impact on the cost of construction but economizing on the cross-section by reducing the number and width of lanes could have a crippling effect on traffic flow and a consequential increases in road user costs. As such, classification of the various links in the road network and estimating their traffic volumes is essential for planning a truly economical road network.

Another historic emphasis was on design for mobility and accessibility. Design was specifically for passenger cars with some attention being paid to the requirements of other vehicles, particularly at intersections. However, geometric designers must now recognize that the road network, particularly in dense settlements, serves other functions in addition to mobility and accessibility. Community needs, including social interaction, relaxation and commerce, are becoming ever more important. In urban areas there is a trend towards mixed land usage. A consequence of this change is that trip lengths are shorter and modes of transport other than passenger cars and buses become a practical option. Walking and cycling can be expected to become more pervasive in the urban environment. The design process will have to make provision for these mobility options as part of the total package available to the traveller.

As there is a need to consider:

- network reconstruction and rehabilitation;
- the findings of the Moving South Africa project;
- the whole-life economy of the road network;
- the broader functionality of the road network; and
- the possibility of an increase in non-motorised transport;

it follows that the focus of geometric planning and design has to change.

2.2 FUNDAMENTAL PRINCIPLES

The laws of motion govern the interaction of the vehicle and the roadway. Isaac Newton's formulation of these laws states that "The change of motion is proportional to the motive force impressed; and is made in the direction of the right line in which that force is impressed" and also that "To every action there is always opposed an equal reaction: or, the mutual actions of two bodies upon each other are always equal, and directed to contrary parts". Professor Newton clearly understood the implications of these laws for he goes on to say "The power and use of machines consists only in this, that by diminishing the velocity we may augment the force, and the contrary."

By applying the laws of motion, together with judicious experimentation, we are able to gain a reasonable understanding of the interaction between the vehicle and the roadway, as they are essentially deterministic. In essence, this understanding describes what a vehicle moving along a road can do and not necessarily what the driver wishes to do. Therefore, to properly describe a highway operating system these laws must be integrated with the human factor, which includes the perceptions, reactions, tolerances and failures of a wide spectrum of individuals under continuously changing circumstances.

Design manuals tend to focus on vehicle dynamics, with all the frailties of the human component of the system being summed up in a single reaction time. The randomness of human behaviour is disregarded. Crash investigations often reveal, however, that it is not always the road or the vehicle but rather the human component of the system that fails under stress.

A vehicle moving along a roadway is a highly complex system with an infinite range of possibilities and outcomes. There are numerous critical elements, each with its own probability of failure. When these are factored together, the sheer number of elements ensures that the probability of failure of the system as a whole is very high indeed. We measure these failures as crashes.

According to Hauer, roads designed to published standards are neither safe nor unsafe and the linkage between standards and safety is largely unpremeditated. He illustrates his contention by reference to the vector diagram that describes the forces operating on a vehicle traversing a superelevated curve. This is Newtonian dynamics and, if it offered a proper explanation of the situation, curves should theoretically have no accidents at all or, at worst, should have exactly the same accident rate as the tangents that precede and follow them. Furthermore, vehicles leaving the road should be equally distributed between the inside and the outside of the curve. The reality of the situation is that the accident rate on curves is higher than on tangents and most vehicles leaving the road do so on the outside of the curve.

Clearly, the vector diagram is not a complete or sufficient exposition of the problem. For example, drivers sometimes steer into a curve only after they have passed its starting point and are thus obliged to follow a path with a smaller radius than that provided by the designer. If the designed curve is at minimum radius, the sub-minimum path actually being followed could have unanticipated consequences. A panic reaction under these circumstances could cause the vehicle to swerve out of control.

While reference is made to human error as the prime cause for most crashes, it is noteworthy that many drivers manage to make the same mistake at the same point along the road. While it is necessary to reconsider the role of the Newtonian models on which geometric standards are based, human factors require careful evaluation.

2.3 DESIGN PHILOSOPHY

Commonly advocated design philosophies tend towards the simplistic and are inclined to ignore the issues discussed in Section 2.1. In search of safety they place inordinate reliance on models derived exclusively from Newtonian dynamics. Current philosophy is, in short, based on the assumption that any design that accords with established geometric design policies is safe and that those that do not are unsafe. This is taken for granted by designers and often is accepted by the courts when making decisions on questions of liability.

Despite many decades of research the complex relationship between vehicle, roadway, driver; and operational safety is not always well understood. Although numerous researchers have investigated the relationships between accident rates and specific geometric design elements, the results were often not sufficiently definitive for practical use. This is due to the narrow focus of this research, which, in examining the relationship between accidents and individual design elements, fails to consider the interactive effects of other parameters, which could lead to bias and mask important relationships.

From this rather unhappy state of affairs we can only conclude that a new design philosophy is warranted.

A design philosophy should encompass two levels. In the first instance, the focus should be on

Geometric Planning, which has seldom, if ever, been discussed in Geometric Design Manuals. Geometric Planning explicitly addresses the matters discussed in Section 2.1. In a sense, it is these issues that dictate how user-friendly the ultimate design will be to both the road user and the community.

Detailed Design is about operational safety, which is the second level of geometric design. This is the level on which Manuals typically focus and the effectiveness and the safety of road elements enjoy equal attention. It is proposed that, in the new philosophy, safety should be the prime consideration. Sacrificing safety in the interests of efficiency and economy is not an acceptable practice.

A more holistic philosophy should thus be founded on the concept of reducing the probability of failure to the lowest possible level and, furthermore, should seek to minimise the consequences of those failures that do occur. To achieve this goal, designs must begin with a clear understanding of purpose and functionality. From this foundation comes the selection of appropriate design elements followed by their integration into the landform and its current and future use. The hallmark of professionalism in road design is the ability to foresee and optimize the conflicting objectives that are inherent in any project.

2.4 DESIGN TECHNIQUES

To arrive at an acceptable design there is no substitute for experience and study. There is, however, a range of useful tools and techniques at the designer's disposal. These are formalised expressions of particular objectives and include:

- Flexibility in highway design;
- Interactive highway design;
- Design domain concept;
- Safety audits;
- Economic analysis; and
- Value engineering.

2.4.1 Flexibility In highway design

A review of the standards and warrants in this manual will quickly reveal that it allows some degree of design flexibility. The degree to which this flexibility is employed in the design process is in fact, nothing more than the application of the art and science of engineering.

In an attempt to formalise the process and to guide the designer towards appropriate choices, the United States Department of Transportation published a report in 1997 entitled "Flexibility in Highway Design". It consists of three main sections: an introduction to the highway design process, general guidelines referring to the major elements of highway design, and examples of six design projects presented as case studies. The concepts described are now more commonly referred to as "context sensitive design".

The most important concept to keep in mind throughout the highway design process is that every project is unique. The setting and character of an area, the values of the surrounding community, the needs of the highway users and the associated physical challenges and opportunities are unique factors that highway designers must consider with each project. For each potential project, designers are faced with the task of balancing the need for improvement of the highway with the need to safely integrate the design into the surrounding natural and human environments.

To accomplish this, highway designers must exercise flexibility. There are a number of options available to aid in achieving a balanced road design and to resolve design issues. Among these are the following:

- Use the flexibility available within the design standards;
- Recognise that design exceptions may be required where environmental impact consequences are great;
- Be prepared to re-evaluate decisions made earlier in the project planning and environmental impact assessment phase;
- Lower the design speed where appropriate;
- Maintain the road's existing horizontal and vertical geometry and cross section where possible;
- Consider developing alternative design standards, especially for scenic or historic roads; and
- Recognise the safety and operational impacts of various design features and modifications.

In addition to exercising flexibility, a successful highway design process should include the public. To be effective, the public view should be canvassed at the outset, even before the need for the project has been defined. If the primary purpose and need for the improvement has not been agreed on, it would be extremely difficult to reach consensus on alternative design solutions later in the process. Public input can also help to assess the characteristics of the area and to determine what physical features are most valued by the community and, thus, have the greatest potential for impact. Awareness of these valued characteristics at an early state will help designers to avoid changing them during the project, reducing the need for mitigation and the likelihood of controversy.

After working with the community to define the basic project need and to assess the physical character of the area, public involvement is necessary to obtain input on design alternatives. Working with the affected community to solve design challenges as they arise is far more effective than bringing the public into the process only after major design decisions have been made. The public needs to be involved at all points in the project where there are the greatest opportunities for changes to be made in the design.

One of the major and continuing sources of conflict between highway agencies and the communities they serve relates to the topic of functional classification. In particular, the need to identify the "correct" functional classification for a particular section of highway, and a regular re-examination of functional classification as changes in adjacent land use take place, would resolve many potential design conflicts before they take place.

There are a number of other fundamental design controls that must be balanced against one another. These include:

- The design speed of the facility;
- The design-year peak-hour level of

service on the facility;

- The physical characteristics of the design vehicle;
- The performance characteristics of the design vehicle;
- The capabilities of the typical driver on the facility (i.e., local residents using low-speed neighbourhood streets versus long distance travellers on inter-urban freeways); and
- The existing and future traffic demands likely to be placed on the facility.

2.4.2 Interactive Highway System Design Model (IHSDM)

A suite of computer modules within the CAD environment is currently under development by the U.S. Federal Highway Administration. When completed, designers will have a powerful tool with which to assess the safety effects of their geometric design decisions.

As currently planned, IHSDM will be applicable to two lane highways. It is composed of six modules.

The Crash Prediction Module

This module will estimate crash potential for a design alternative, including all roadway segments and intersections. Estimates will be quantitative and will include the number of crashes for a given roadway segment or intersection as well as the percentages of fatal and severe crashes.

The module will allow the user to compare the number of crashes over a given time period for different design alternatives or to perform sensitivity analyses on a single alternative.

The Design Consistency Module

This module evaluates the operating-speed consistency of two-lane highways. The evaluation is performed using a speed-profile model that estimates 85th percentile speeds on each element along an alignment. The module generates two consistency-rating measures:

- The difference between estimated 85th percentile speeds and the design speed of the roadway, and
- The reduction in 85th percentile speed between each approach tangent-curve pair.

The module will consist of a speed-profile model and consistency rating measures that have been validated and are applicable to most two-lane, free flowing highways in the United States.

The Driver/Vehicle Module

This will consist of a Driver Performance Model linked to a Vehicle Dynamics Model. Driver performance is influenced by cues from the roadway/vehicle system (i.e., drivers modify their behaviour based on feedback from the vehicle and the roadway). Vehicle performance is, in turn, affected by driver behaviour/performance. The Driver Performance Model will estimate a driver's speed and path along a two-lane highway in the absence of other traffic. These estimates will be input to the Vehicle Dynamics Model, which will estimate measurements including lateral acceleration, friction demand, and rolling moment.

The Driver/Vehicle Module will produce the following measures of effectiveness and, where appropriate, threshold or reference values for comparison purposes:

- Lateral acceleration in comparison with discomfort, skid, and rollover threshold values;
- Friction demand in comparison with the skid threshold;
- Rolling moment in comparison with the rollover threshold;
- Estimated vehicle speed in comparison with threshold speeds for discomfort, skidding, and rollover; and
- Vehicle path (lateral placement) relative to the lane lines.

The Intersection Diagnostic Review Module

This module will be used to evaluate the geometric design of at-grade intersections on two-lane highways and to identify possible safety treatments. The Intersection Diagnostic Review Module will incorporate qualitative guidance from the American Association of State Highway and Transportation Officials document "A Policy on the geometric design of highways and streets" (generally referred to as the Green Book) and other design policies, design guidelines based on past research and design guidelines based on expert opinion. The primary focus is to identify combinations of geometric design elements that suggest potential design deficiencies, even when each element considered individually could be regarded as being within good design practice.

The Policy Review Module

This module is intended for use in all stages of highway planning and design, including design review, for both new and reconstruction projects. Design elements that are not in compliance with policy will be identified, and an explanation of the policy violated will be provided. In response to this information, the user may correct any deficiencies, analyse the design further using other IHSDM modules, and/or prepare a request for design exception. A summary of the policy review will be provided, including a listing of all design elements that do not comply with policy. The categories of design elements to be verified include: horizontal alignment, vertical alignment, cross section, intersections, sight distance, and access control/management.

The Policy Review Module will notify designers of any design elements that deviate from minima/maxima set by the AASHTO Green Book, the "Roadside Design Guide," and the "Guide for the Development of Bicycle Facilities." The Module will also have the capability of reviewing designs relative to alternative, user-specified design policies, such as State Department of Transportation design guidelines.

The Traffic Analysis Module

This module will link highway geometry data with a traffic simulation model to provide information on speed, travel time, delay, passing rates, percentage following in platoons, traffic conflicts and other surrogate safety measurements. TWOPAS, a traffic simulation model for two-lane highways, will form the basis for this module.

2.4.3 The "design domain" concept

The design domain concept recognizes that there is a range of values, which could be adopted for a particular design parameter within absolute upper and lower limits. Values adopted for a particular design parameter within the design domain would achieve an acceptable though varying, level of performance in average conditions in terms of safety, operation, and economic and environmental consequences.

Figure 2.1 illustrates the concept.

While values within the lower region of the design domain for a particular parameter are generally less safe and less operationally efficient, they are normally less costly than those in the upper region. In the upper region of the domain, resulting designs are generally "safer" and more efficient in operation, but may cost more to construct. In fact, the design domain sets the limit within which parameters should be selected for consideration within the value engineering concept.

During recent years there have been many advances in road design and in the procedures for assessment of safety and operational. These improvements, as well as initiatives in the assessing and auditing of scheme layouts, have considerably improved the design process.

It is now practical to estimate the changes in the level of service, cost and safety when the design is changed within the design domain. Where data are not available, guidance is available to the designer in the literature on the sensitivity of safety to changes in the parameter under consideration within the design domain. These evaluations are however limited in comparison to the evaluation of operational adequacy or construction costs.

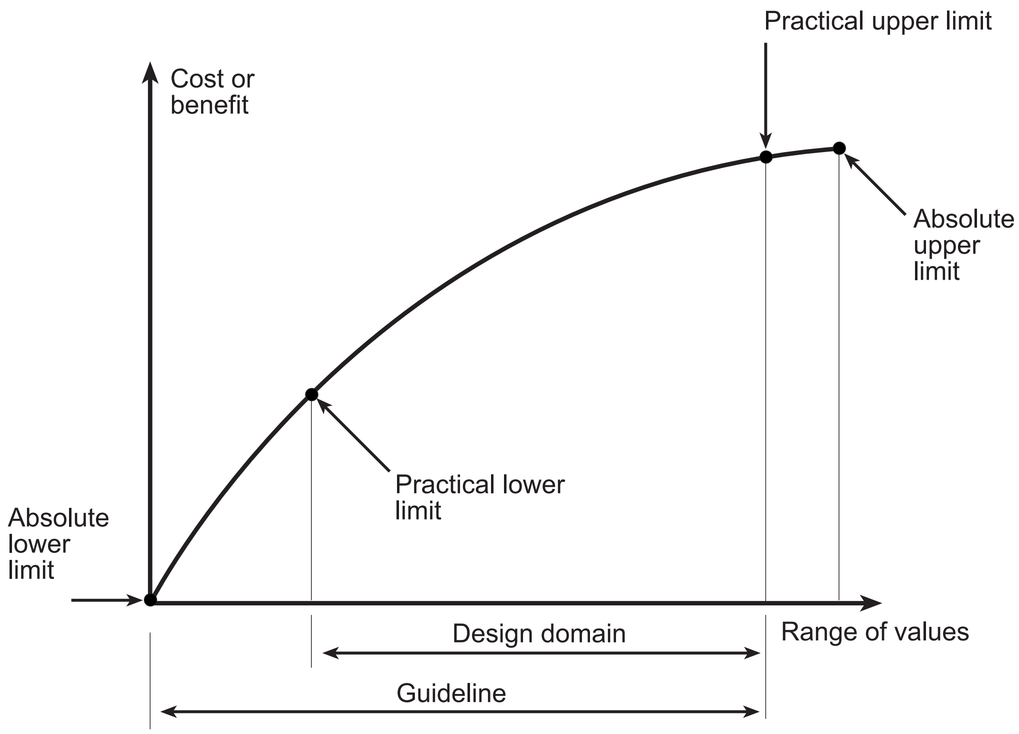


Figure 2.1 : The design domain concept

The benefits of the design domain concept are:

- It is directly related to the true nature of the road design function and process, since it places emphasis on developing appropriate and cost-effective designs, rather than on those which simply meet "standards";

- It directly reflects the continuous nature of the relationship between service, cost and safety and changes in the values of design dimensions. It thus reinforces the need to consider the impacts of trade-offs throughout the domain and not just when a "standards" threshold has been crossed, and;
- It provides an implicit link to the concept of "Factor of Safety" - a concept that isused in other civil engineering design processes where risk and safety are important.

The illustration in Figure 2.2 is an example of how different costs and benefits may vary within the design domain for a specific parameter - in this case shoulder width. The application of this concept to all design parameters will lead to an optimal project design.

Application of the concept of a design domain in practice presents practical challenges. In some cases, the concept of a design domain with upper and lower bounds, and a continuous range of values in between, may not be practical or desirable. Lane widths provide a good example of such a case. In these instances, it may only be necessary to consider a series of discrete values for the dimension in question. In other instances, there may be no upper limit to a design domain other than what is practical or economic. In these cases, the upper boundary of the design domain generally reflects typical upper level values found in practice, or the general threshold of cost-effective design.

The designer must respect controls and constraints to a greater or lesser degree, depending on their nature and significance. Often, the designer is faced with the dilemma of being

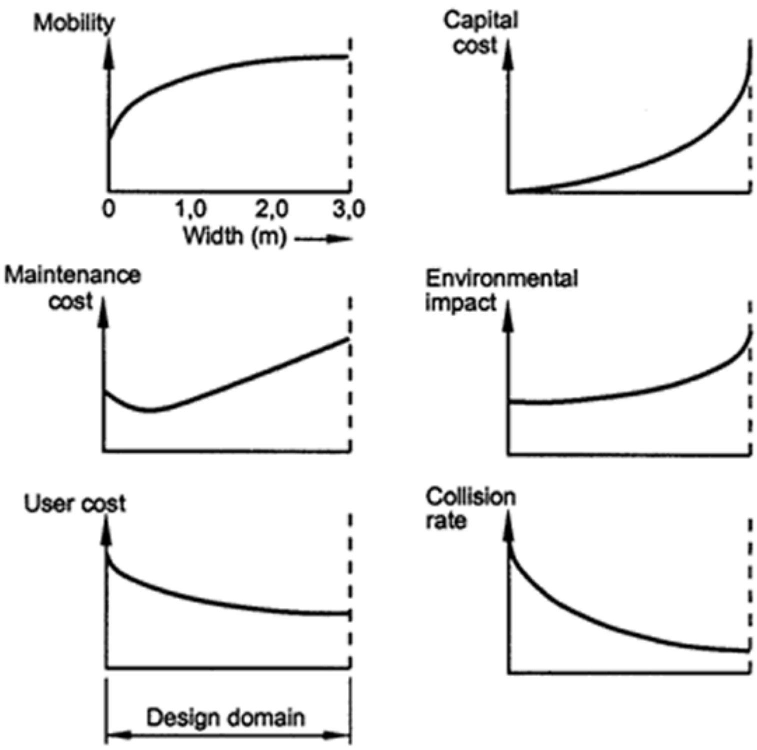


Figure 2.2: Example of design domain application - Shoulder width.

unable to choose design dimensions or criteria that will satisfy all controls and constraints, and a compromise must be reached. These are engineering decisions that call for experience, insight and a good appreciation of community values.

Some design criteria such as vertical clearance at structures are inviolate. Others are less rigid and some are little more than suggestions. Some of those chosen are for safety reasons, some for service or capacity, while others are based on comfort or aesthetic values. The choice of design criteria is very important in the design process and it is essential for the designer to have a good understanding of their origin and background. A design carefully prepared by a designer who has a good understanding, not only of the criteria, but also of their background and foundation, and who has judiciously applied the community values, will probably create the desired level of service, safety and economy.

For many elements, a range of dimensions is given and the designer has the responsibility of choosing the appropriate value for a particular application. A designer with economy uppermost in mind may be tempted to apply the minimum value, reasoning that so long as the value is within an accepted range, the design is "satisfactory". This may or may not be the case.

The designer might find it appropriate to reduce values of design criteria, which is not necessarily a poor decision. However, the consequences need to be thoroughly understood, particularly as they impacts on safety and also on the costs and benefits. Ameliorating measures, such as the use of traffic control devices, may need to be considered in the design process. If a design involves compromise, it may be more appropriate to vary several elements by a small amount than to alter one element excessively. It is important that a design be balanced.

2.4.4 Road safety audits

As the term implies, road safety auditing is a structured process that brings specialised and explicit safety knowledge to bear on a highway project so that it can be quantitatively considered. It is a formal examination of a future or existing project in which an independent, qualified examination team reports on the accident potential and safety performance of the project.

The benefits of road safety audits include:

- A reduction in the likelihood of accidents on the road network;
- A reduction in the severity of accidents on the road network;
- An increased awareness of safe design practices among traffic engineers and road designers;
- A reduction in expenditure on remedial measures; and
- A reduction in the life-cycle cost of a road.

Australian and New Zealand experience has shown that road safety audits do not add more than four per cent to the cost of a road project. It is, however, necessary to equate this cost to the potential benefits of the road safety audit, e.g.:

- A saving in time and cost by changing project details at the planning and design stage rather than by changing or removing a road element once installed;

- A reduction in the likelihood of accidents and therefore in accident costs; and
- A reduction in the cost of litigation.

The objectives of a road safety audit are;

- To identify and report on the accident potential and safety of a road project;
- To ensure that road elements with an accident potential are removed; or
- That the risk of crashes is reduced.

Road safety can be audited at any of the following six stages, however, the sooner the better:

Stage 1 Road safety audit:	Preliminary design stage
Stage 2 Road safety audit:	Draft design stage
Stage 3 Road safety audit:	Detailed design stage
Stage 4 Road safety audit:	Preconstruct-ion stage
Stage 5 Road safety audit:	Pre-opening stage
Stage 6 Road safety audit:	Existing facility

2.4.5 Economic analysis

Economic analyses form an intrinsic part of any civil engineering project where the "value for money" concept is important.

Roads are essential for mobility of people and goods. The benefits of mobility are attained at a cost. Roads cost money to build and maintain; they consume space and affect the environment; road travel consumes time, creates noise and pollution, and brings about crashes, etc. All these are the costs of mobility.

By spending more money on construction, other costs may be reduced (e.g. travel time or crashes). However, additional expenditure must create increases in benefits or reductions in other costs. Economic analyses can evaluate the trade-offs between costs and benefits.

The analysis when applied to a road can be highly complex, depending on the scope of the project. Many formal or informal evaluations may have been carried out and decisions made, before the geometric designer gets involved. In extreme cases, the designer may be so constrained by decisions already made, that there is little or no opportunity to judge many of the potential costs and benefits. It is, however, the designer's task to incorporate those judgements into planning and design wherever that freedom exists. The designer should also identify situations where policy decisions may unreasonably constrain a satisfactory design. When presented effectively, arguments made by designers may affect the timing and scope of projects and also influence changes to existing policy.

The geometric designer determines the horizontal and vertical alignment and cross section at every point on the road. In addition, special planning is required at every location where roadways intersect, to accommodate diverging, converging and conflicting traffic movements. In selecting design dimensions and layouts, the designer can directly affect some of the benefits, costs and impacts of the road, as well as allow for future expansion.

The hallmark of professionalism in road design is the ability to optimise and foresee the repercussions of design decisions on the benefits, costs and impacts of the road.

For most, if not all, road projects, the designer will have some scope for value judgements, although this will vary from place to place and from project to project, governed by policy decisions already made. Factors that the designer may be able to influence include:

- Mobility;
- Environmental impacts;
- Safety;
- Capital costs;
- Aesthetics;
- Maintenance costs and
- Vehicle operating costs.

In influencing these factors, the designer will be guided by jurisdictional policy decisions, such as the relative importance of maintenance cost versus capital cost or of fuel consumption and air pollution against capital cost.

2.4.6 Value engineering

Road design is generally carried out in an environment where a limited budget needs to be stretched as far as possible. For this reason designers are placed under considerable pressure to minimize costs.

While economy and fiscal efficiency is a key goal of all designs and should continue to be so, it is essential that changes in design should be analysed explicitly, evaluating safety in the same manner as other criteria, such as construction and maintenance costs, and environmental and operational impacts. One method is "value engineering" which is a proven management technique based on an intensive, systematic and, especially, creative study of the project to seek the best functional balance between its cost, reliability and performance.

In a road design context, this means that a value engineering exercise should be more than merely a way of minimizing construction costs, but that equal and explicit attention should also be given to the important aspects of safety, operational performance and quality. In fact, value engineering can, and sometimes does, result in increased construction costs to reduce the life-cycle costs.

More and more authorities are using the concept of value engineering to a more cost-effective design. If properly applied, this approach is a valuable input to the design process where functional balances are evaluated explicitly and quantitatively for the full range of life cycle costs and benefits and re-evaluated in response to proposed changes in design, construction sequences and practices. Only in this way can the true benefits of the value engineering process be realised.

Engineers acting independently of the design team often do value engineering. However, the concept is applicable at all times to all projects and, to do a complete job, this design team should embody value engineering in its design process. If this is done, the independent value engineering process will become less necessary.

Back to Menu

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 3

DESIGN CONTROLS

3.1 INTRODUCTION

The design of a road is that of a three-dimensional structure which should ideally be safe, efficient, functional and economical for traffic operations, and which should also be aesthetically pleasing in its finished form. However, the designer uses dimensions and related criteria within a design context that recognizes a series of design controls constraining what can be achieved. These limitations are imposed by the characteristics of vehicle and driver performance as well as by environmental factors. The designer should, therefore, relate the physical elements of the road to the requirements of the driver and vehicle so that consistency in the driver's expectations is achieved and, at the same time, ensures that environmental and other constraints are accommodated.

Good road design is the art of combining and balancing the various controls in a perceptive fashion and is not merely an exercise in three-dimensional geometry. In this chapter, the constraints and controls on the design process are discussed.

3.2 HUMAN FACTORS

3.2.1 Drivers

An appreciation of driver performance as part of the road traffic system is essential for effective road design and operation. When a design is incompatible with human capabilities (both of the driver and any other road user) the opportunities for errors and accidents increase. Knowledge of human performance, capabilities and behavioural characteristics is thus a vital input into the design task.

Road users do not all behave in the same way and designs should cater for substantial differences in the range of human characteristics and a wide range of responses. However, if the perceptual clues are clear and consistent, the task of adaptation is made easier and the response of drivers will be more appropriate and uniform. For roadway design this translates into some useful principles, viz:

- A roadway should confirm what drivers expect based on previous experience; and
- Drivers should be presented with clear clues about what is expected of them

Driver Workload and Expectations

The driver workload comprises

- Navigation: trip planning and route following;
- Guidance: following the road and maintaining a safe path in response to traffic conditions; and
- Control: steering and speed control

These tasks require the driver to receive and process inputs, consider the outcome of alternative actions, decide on the most appropriate, execute the action and observe its effects through the reception and processing of new information. There are numerous problems inherent in this sequence of tasks, arising from both the capabilities of the human driver, and the interfaces between the human and other components of the road traffic system (the road and the vehicle). These include inadequate or insufficient input available for the task at hand (e.g. during night time driving, as a result of poor sight distance, or because of complex intersection layouts). When they become overloaded, drivers shed part of the input to deal with that judged to be more important. Most importantly, drivers are imperfect decision-makers and may make errors, including in the selection of what input to shed.

The designer must provide all the information the driver needs to make a correct decision timeously, simultaneously ensuring that the information is provided at a tempo that does not exceed the driver's ability to absorb it. In the words of the American Association of State Highway and Transportation Officials: (AASHTO)

> *'A common characteristic of many high-accident locations is that they place large or unusual demands on the information-processing capabilities of drivers. Inefficient operation and accidents usually occur where the chance for information-handling errors is high. At locations where the design is deficient, the possibility of error and inappropriate driver performance increases.'*

Prior experience develops into a set of expectancies that allows for anticipation and forward planning, and these enable the driver to respond to common situations in predictable and successful ways. If these expectancies are violated, problems are likely to occur, either as a result of a wrong decision or of an inordinately long reaction time. There are three types of driver expectancy:

Continuation expectancy. This is the expectation that the events of the immediate past will continue. It results, for example, in small headways, as drivers expect that the leading vehicle will not suddenly change speed. One particularly perverse aspect of continuation expectance is that of subliminal delineation, e.g. a line of poles or trees or lights at night which suggests to the driver that the road continues straight ahead when, in fact, it veers left or right. These indications are subtle, but should always be looked out for during design.

Event expectancy. This is the expectation that events that have not happened will not happen. It results, for example, in disregard for "at grade" railway crossings and perhaps for minor intersections as well, because drivers expect that no hazard will present itself where none has been seen before. A response to this situation is more positive control, such as an active warning device at railway crossings that requires that the driver respond to the device and not to the presence of a hazard.

Temporal expectancy. This is the expectation that, where events are cyclic (e.g. traffic signals), the longer a given state prevails, the greater is the likelihood that change will occur. This, of course, is a perfectly reasonable expectation, but it can result in inconsistent responses. For example, some drivers may accelerate towards a green signal, because it is increasingly likely that it will change, whereas others may decelerate. A response to this is to ensure, to the extent possible, that there is consistency throughout the road traffic system to encourage predictable and consistent driver behaviour.

The combined effect of these expectancies is that:

- drivers tend to anticipate upcoming situations and events that are common to the road they are travelling;
- the more predictable the roadway feature, the less likely will be the chance for errors;
- drivers experience problems when they are surprised;
- in the absence of evidence to the contrary, drivers assume that they will only have to react to standard situations;
- the roadway and its environment upstream create an expectation of downstream conditions; drivers experience problems in transition areas and locations with inconsistent design or operation, and
- expectancies are associated with all levels of driving performance and all aspects of the driving situation and include expectancies relative to speed, path, direction, the roadway, the environment, geometric design, traffic operations and traffic control devices.

Driver Reaction

It takes time to process information. After a person's eyes detect and recognize a given situation, a period of time elapses before muscular reaction occurs. Reaction time is appreciable and differs between persons. It also varies for the same individual, being increased by fatigue, drinking, or other causes. The AASHTO brake reaction time for stopping has been set at 2,5 s to recognize all these factors. This value has been adopted in South Africa.

Often drivers face situations much more complex than those requiring a simple response such as steering adjustments or applying the brakes. Recognition that complex decisions are time-consuming leads to the axiom in highway design that drivers should be confronted with only one decision at a time, with that decision being binary, e.g. "Yes" or "No" rather than complex, e.g. multiple choice. Anything up to 10 seconds of reaction time may be appropriate in complex situations.

Design Response

Designers should strive to satisfy the following criteria:

- Driver's expectations are recognized, and unexpected, unusual or inconsistent design or operational situations avoided or minimized.
- Predictable behaviour is encouraged through familiarity and habit (e.g. there should be a limited range of intersection and interchange design formats, each appropriate to a given situation, and similar designs should be used in similar situations).
- Consistency of design and driver behaviour is maintained from element to element (e.g. avoid significant changes in design and operating speeds along a roadway).
- The information that is provided should decrease the driver's uncertainty, not increase it (e.g. avoid presenting several alternatives to the driver at the same time).
- Clear sight lines and adequate sight distances are provided to allow time for decision-making and, wherever possible, margins are allowed for error and recovery.

With the major response to drivers' requirements being related to consistency of design, it is worthwhile considering what constitutes consistency. Consistency has three elements that are the criteria offered for the evaluation of a road design:

Criterion I Design consistency - which corresponds to relating the design speed to actual driving behaviour which is expressed by the 85th percentile speed of passenger cars under free-flow conditions;

Criterion II Operating speed consistency which seeks uniformity of 85th percentile speeds through successive elements of the road and

Criterion III Consistency in driving dynamics - which relates side friction assumed with respect to the design speed to that demanded at the 85th percentile speed.

In the case of Criterion 1, if the difference between design speed and 85^{th} percentile speed on an element such as a horizontal curve is less than 10 km/h, the design can be considered good. A difference of between 10 km/h and 20 km/h results in a tolerable design and differences greater than 20 km/h are not acceptable.

In the case of Criterion 2, the focus is on differences in operating speed in moving from one element, e.g. a tangent, to another, e.g. the following curve. A difference in operating speed between them of less than 10 km/h is considered to be good design and a difference of between 10 and 20 km/h is tolerable. Differences greater than 20 km/h result in what is considered to be poor design.

For the third Criterion the side friction assumed for the design should exceed the side friction demanded by 0,01 or more. A difference between -0,04 and 0,01 results in a fair design. A value of less than -0,04 is not acceptable. A negative value for the difference between side friction assumed for design and the side friction demanded means that drivers are demanding more side friction than is assumed to be available - a potentially dangerous situation.

3.2.2 Other road users

Pedestrians

The interaction of pedestrians and vehicles should be carefully considered in road design, principally because 50 per cent of all road fatalities are pedestrians.

Pedestrian actions are less predictable than those of motorists. Pedestrians tend to select paths that are the shortest distance between two points. They also have a basic resistance to changes in gradient or elevation when crossing roadways and tend to avoid using underpasses or overpasses that are not convenient.

Walking speeds vary from a 15th percentile speed of 1,2 m/s to an 85th percentile of 1,8 m/s, with an average of 1,4 m/s. The 15th percentile speed is recommended for design purposes.

Pedestrians' age is an important factor that may explain behaviour that leads to collisions. It is recommended that older pedestrians be accommodated by using simple designs that minimize crossing widths and assume lower walking speeds. Where complex elements such as channelisation and separate turning lanes are featured, the designer should assess alternatives that will assist older pedestrians.

Pedestrian safety is enhanced by the provision of:

- median refuge islands of sufficient width at wide intersections, and
- lighting at locations that demand multi ple information gathering and process ing.

Cyclists

Bicycle use is increasing and should be considered in the road design process. Improvements such as:

- paved shoulders;
- wider outside traffic lanes (4,2 m minimum) if no shoulders exist;
- bicycle-safe drainage grates;
- adjusting manhole covers to the grade, and
- maintaining a smooth, clean riding surface

can considerably enhance the safety of a street or highway and provide for bicycle traffic:

At certain locations it may be appropriate to supplement the existing road system by providing specifically designated cycle paths. The design elements of cycle paths are discussed in Chapter 4.

3.3 SPEED

3.3.1 General

Drivers, on the whole, are concerned with minimising their travel times, and speed is one of the most important factors governing the selection of alternate routes to gain time savings. The attractiveness of a specific road or route is generally judged by its convenience in travel time, which is directly related to travel speed.

Various factors influence the speed of vehicles on a particular road. These include:

- Driver capability, driver culture and driver behaviour;
- Vehicle operating capabilities;
- The physical characteristics of the road and its surroundings;
- Weather;
- Presence of other vehicles, and
- Speed limitations (posted speed limits).

Speeds vary according to the impression of constraint imparted to the driver as a result of these factors.

The objective of the designer is to satisfy the road users' demands for service in a safe and economical way. This means that the facility should accommodate nearly all reasonable demands (speed) with appropriate adequacy (safety and capacity) but should not fail completely under severe load, i.e. the extremely high speeds maintained by a small percentage of drivers. Roads should, therefore, be designed to operate at a speed that satisfies most, but not necessarily all, drivers.

Various studies have shown that the 85th percentile speed generally exceeds the posted speed limit by a margin of at least 10 km/hr when weather and traffic conditions are favourable. For this reason, design speed is typically equated to the 85th percentile speed.

The relationship between road design and speed is interactive. While the designer shapes the elements of the road by the anticipated speed at which they will be used, taking into account the inherent economic trade-offs between construction and environmental costs of alternative alignments (vertical and horizontal) to match desired travel speed, the speed at which they will be used depends to a large extent on the chosen design features.

3.3.2 Speed classification

The term "speed" is often used very loosely when describing the rate of movement of road traffic. Road design recognizes various definitions or classifications of speed, all of which are interrelated. The sub-divisions are:

- Desired Speed - the speed at which a driver wishes to travel, determined by a combination of motivation and comfort.
- Design Speed - the speed selected as a safe basis to establish appropriate geometric design elements for a particular section of road and which should be a logical one with respect to topography, anticipated operating speed, the adjacent land use and the functional classification of the road.
- Operating Speed - observed speeds during free flow conditions. For an individual driver, operating speed is generally lower than desired speed since operating conditions are not usually ideal.
- Running Speed - the average speed maintained over a given route while a vehicle is in motion. The running time is the length of the road section divided by the time required for the vehicle to travel through the section. Thus, in determining the running speed, the times en route when the vehicle is at rest are not taken into account in the calculations. Running speeds are generally used in road planning and capacity and service level analyses. The difference between running speed and design speed is strongly affected by traffic volumes.
- Posted Speed - is a speed limitation set for reasons of safe traffic operations rather than for geometric design considerations and is aimed at encouraging drivers to travel at appropriate speeds for all prevailing conditions.

3.3.3 Design speed

The most important factor in geometric design is the design speed. This was previously defined as the highest continuous speed at which individual vehicles can travel with safety on the road when weather conditions are favourable, traffic volumes are low and the design features of the road are the governing condition for safety. The current definition is simply states that the design speed is the speed selected as the basis for establishing appropriate geometric elements for a section of road. These elements include horizontal and vertical alignment, superelevation and sight distance. Other elements such as lane width, shoulder width and clearance from obstacles are indirectly related to design speed.

The chosen design speed should be a logical one consistent with the road function as perceived by the driver and also one that takes into account the type of road, the anticipated operating speed, and the terrain that the road traverses. Where a difficult condition is obvious, drivers are more apt to accept a lower speed than where there is no apparent reason for it.

Other relevant factors include traffic characteristics, land costs, speed capabilities of vehicles, aesthetics, economics and social or political impacts. A highway of higher functional classification may justify a higher design speed than a less important facility in similar topography, particularly where the savings in vehicle operation and other operating costs are sufficient to offset the increased costs of right-of-way and construction. A low design speed, should not be assumed where the topography is such that drivers are likely to travel at high speeds.

When carefully selected, these factors should result in a design speed which is acceptable to all but a very few drivers. Above minimum design values should be used where feasible though consistency is essential.

When a substantial length of road is being designed, it is desirable to adopt a constant design speed to maintain consistency. Changes in terrain and other physical controls may, however, dictate a change in design speed on certain sections. Each section, however, should be relatively long, compatible with the general terrain or development through which the road passes. The justification for introducing a reduced design speed should be obvious to the driver. A case in point is where a road leaves relatively level terrain and starts traversing hilly or mountainous terrain. Moreover, the introduction of a lower or higher design speed should not be effected abruptly but over sufficient distance to encourage drivers to change speed gradually.

Where design speeds exceed 90 km/h the variation between successive speeds should be limited to 10 km/h and, below 80 km/h, this variation should be limited to 20 km/h. Where it is necessary to change the design speed, the new design speed should apply to an extended section of road. Even if properly signposted, isolated design speed variations are hazardous as they do not match driver expectations and it is always possible that the signpost may be obscured, illegible, removed or even simply not perceived by the driver. Isolated design speed changes are, therefore, to be avoided.

The need for a multilane cross-section suggests that traffic volumes are high. A design speed of at least 120 km/h should be used if the topography permits. Major roads, even if two-lane two-way roads, should also be designed to this speed if possible. Rolling terrain may, however, necessitate a reduction to 100 km/h in the design speed and, in the case of mountainous terrain, it may even be necessary to reduce the design speed to 80 km/h.

Secondary and tertiary roads may have lower design speeds than those advocated for the primary road network. However, where traffic is likely to move at relatively high speeds on these roads, higher design speeds should be selected.

There is still debate as to whether speeds greater than 120 km/h should be used for design purposes on freeways. Higher design speeds not only safeguard against early obsolescence of the highway, but also provide an increased margin of safety for those driving at high speeds. That there is some validity in this statement is reflected by the fact that the design speed of high-type roads is now at least 120 km/h as compared with 56 km/h in 1927, a change brought about by the continuing increase in vehicle performance.

The choice of a design speed for a dual carriageway is much less influenced by construction cost than that for other rural roads. In practice, lower design speeds are often accepted on single carriageway roads in order to keep construction costs within certain limits. There is danger in this philosophy since, although drivers will obviously accept lower speeds in what are clearly difficult locations, repeated studies have shown that they do not adjust their speeds to the importance of the facility. Instead they endeavour to operate at speeds consistent with the traffic on the facility and its physical limitations.

Ideally, then, design speed should be chosen to reflect the 85th percentile desired speed that is likely to materialize. This is often achievable for roads for which the primary function is mobility and where severe physical constraints do not exist. Limited studies in South Africa have shown that the 85th percentile speed exceeds 120 km/h for unhindered vehicles on a four-lane divided roadway. Use of a design speed of 130 km/h should therefore satisfy driver demands in most areas.

The selected design speed should be logical and in harmony with the topography and the functional classification of a road. Careful consideration should also be given to its relationship to other defined speeds. While no hard relationships have been established, choice of design speed can simultaneously accommodate and influence desired, operating, running and posted speeds.

Table 3.1 provides an indication of typical design speeds for different classes of roads.

Table 3.1: Typical design speeds

Road type	Design speed (km/h)
Limited Access Roads	
Freeways in urban areas	90-130
Freeways and expressways in rural areas	110-130
Expressways in urban areas	80-110
Conventional Roads	
Rural	
Flat terrain	90-120
Rolling terrain	80-100
Mountainous terrain	60-80
Urban	
Arterial streets	60-100
Arterial streets with extensive development	50-70

3.3.4 Operating speed

Operating speed is measured under free flow conditions. The term "spot speed" is sometimes used to denote operating speed. For an individual driver, operating speed is generally lower than desired speed since operating conditions are not usually ideal. When reference is made to the operating speed of all vehicles in the traffic stream, this is taken as being the 85th percentile of all observed speeds.

Operating speed has a variety of uses. It is generally used as a measure of level of service at uninterrupted flows. It can also be used to monitor the effect of flow constrictions, such as intersections or bridges. Since operating speeds at ideal sections of road are indicative of speeds desired by motorists, they can be used to guide the selection of design speed on improved or new facilities.

When the design speed is less than the desired speed, drivers should be warned to modify their speed, as studies have shown that crash rates increase as the operating speed of a particular vehicle deviates from the mean operating speed of the other vehicles on the roadway.

The typical driver can recognize or sense a logical operating speed for a given roadway based on knowledge of the system, posted speed limits, appraisal of the ruggedness of the terrain, traffic volumes and the extent, density and size of development. Studies have shown that characteristics, such as the number of access points, nearby commercial development, road width and number of lanes, have a significant influence on vehicle speeds. Based on these factors, the driver's initial response is to react to the anticipated situation rather than to the actual situation. In most instances, the two are similar enough not to create conflicts. If the initial response is incorrect, operation and safety may be severely affected.

Some agencies conduct speed surveys to determine operating speeds at various points along a section of roadway. The results can be compared with the design speed, and may lead to a policy change in the selection of design speeds.

3.3.5 Application of design speed

Consistency of design is fundamental to good driver performance, based on satisfying the driver's expectations. Design consistency exists when the geometric features of a continuous section of road are consistent with the operational characteristics as perceived by the driver. The traditional approach to achieving design consistency has been through the application of the design speed process. Once selected, the design speed is used to determine values for the geometric design elements from appropriate design domains.

However, application of this procedure alone does not guarantee design consistency. There are several limitations of the design speed concept that should be considered during design:

1. Selection of dimensions to accommodate specified design speed does not necessarily ensure a consistent alignment design. Design speed is significant only when physical road characteristics limit the speed of travel. Thus, a road can be designed with a constant design speed, yet have considerable variation in

speeds achievable and therefore to a driver appear to have a wide variation in character. For example, the radii of curves within a section should be consistent, not merely greater than the minimum value.

2. For horizontal alignments, design speed applies only to curves, not to the connecting tangents. Design speed has no practical meaning on tangents. As a result, the operating speed on a tangent, especially a long one, can often significantly exceed the design speed of the road as a whole.

3. The design speed concept does not ensure sufficient coordination among individual geometric features to ensure consistency. It controls only the minimum value of the maximum speeds for the individual features along an alignment. For example, a road with an 80 km/h design speed may have only one curve with a design speed of 80 km/h and all other features with design speeds of 120 km/h or greater. As a result, operating speeds approaching the critical curve are likely to exceed the 80 km/h design speed. Such an alignment would comply with an 80 km/h design speed, but it would violate a driver's expectancy and result in undesirable alignment.

4. Vehicle operating speed is not necessarily synonymous with design speed. Drivers normally adjust speed according to their desired speed, posted speed, traffic volumes and perceived alignment hazards. The perception of hazard presented by the alignment may vary along a road designed with a constant design speed. The speed adopted by a driver tends to vary accordingly and may exceed the design speeds. A report on studies in Australia and the US concluded that 85th percentile operating speeds consistently exceed design speeds where the design speeds are less than 100 km/h at horizontal curves on rural two-lane highways.

5. In addition, different alignment elements may have quite different levels of perceived hazard. Entering a horizontal curve too fast will almost certainly result in loss of control, so drivers adjust their speed accordingly. However, the possibility of a curtailed sight distance concealing a hazard is considered as a remote occurrence. Unfortunately drivers do not generally adjust their speed to compensate for sight distance restrictions.

To help overcome these weaknesses in the use of design speed to design individual geometric elements, speed profiles are used. A speed profile is a graphical depiction (which can be modelled) showing how the 85th percentile operating speed varies along a length of road. This profile helps to identify undesirably large differentials in the 85th percentile operating speed between successive geometric elements, e.g. a curve following a tangent.

When an existing road is being improved, actual operating speeds can be measured to create a speed profile, but interpretation of the profile can be difficult, depending on the complexity of geometric and other features that may cause drivers to change speed. For a new road, prediction of operating speeds is needed to create a speed profile model.

3.4 DESIGN VEHICLES

3.4.1 Introduction

The physical characteristics of vehicles and the proportions of the various sizes of vehicles using a road are positive controls in design and define several geometric design elements, including intersections, on- and off-street parking, site access configurations and specialized applications such as trucking facilities. It is necessary to identify all vehicle types using the facility, establish general class groupings and select hypothetical representative design vehicles, within each design class. The dimensions used to define design vehicles are not averages or maxima, nor are they legal limiting dimensions. They are, in fact, typically the 85th percentile or 15th percentile value of any given dimension. The design vehicles are therefore hypothetical vehicles, selected to represent a particular vehicle class.

The dimensions in the previous Geometric Design Manual were based on typical design vehicles in South Africa pertinent to 1965. The range of vehicle types and their operating characteristics have changed significantly since then. The vehicle size regulations have also undergone substantial revisions which have generally resulted in larger trucks on the roads as well as in an increased use of recreational utility vehicles.

3.4.2 Vehicle classifications

Three general classes of vehicles have been selected for this design guide: passenger cars, trucks and buses. The passenger car class includes compacts and subcompacts, recreational utility vehicles and all light vehicles and light delivery trucks (vans and pickups). The truck class includes single-unit trucks, truck tractor-semi trailer combinations, and trucks or truck tractors with semi trailers in combination with full trailers. Buses include single unit buses, articulated buses and intercity buses. In establishing the design dimensions for the various vehicle classes, this guide focuses on vehicles in regular operation only.

Vehicles defined in the Road Traffic Act include:

- Passenger cars and minibuses (kombis);
- Standard single unit buses;
- Articulated buses ("Bus Train");
- Two axle trucks, with and without trailers;
- Three and four axle vehicles;
- Three, four and five axle articulated trucks;
- Five and six axle articulated trucks, and
- Multi vehicle combinations.

3.4.3 Vehicle characteristics

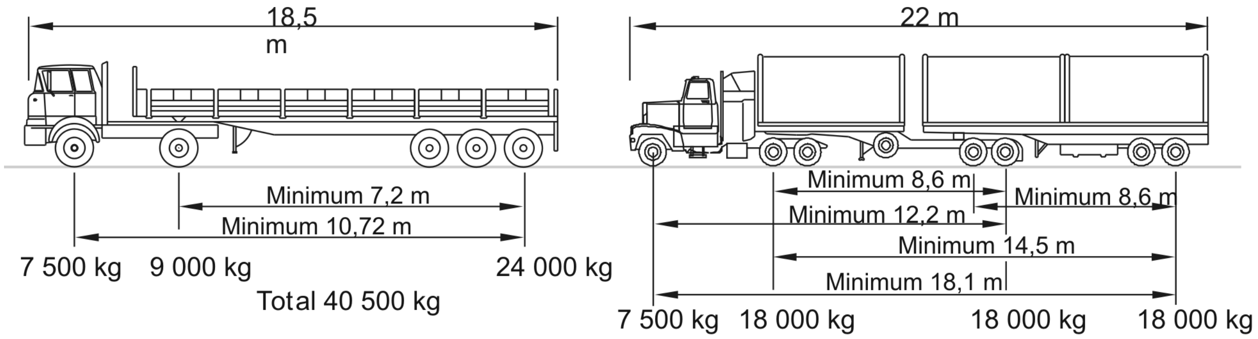
The dimensions adopted for the various design vehicles are given in Table 3.2.

The WB15 vehicle has an overall length of 17 m, whereas the regulations allow for a semi-trailer to have an overall length of 18,5 m. The multiple vehicle combination, being a semi-trailer plus trailer and typically in an Interlink configuration, can have a maximum overall length of 22 m. Examples of these vehicles are illustrated in Figure 3.1.

f these vehicles are expected to use a route with any frequency, the designer will have to carefully plan the layout of the intersections to ensure that they can be accommodated. As described below, accommodation does not necessarily require lane widths sufficient to complete a turning movement within the lane. A degree of

Table 3.2: Dimensions of design vehicles (m)				
Vehicle	**Wheel base**	**Front overhang**	**Rear overhang**	**Width**
Passenger car (P)	3,1	0,7	1,0	1,8
Single unit (SU)	6,1	1,2	1,8	2,5
Single unit + trailer (SU+T)	6,7+3,4*+6,1	1,2	1,8	2,5
Single unit bus (BUS)	7,6	2,1	2,6	2,6
Semi-trailer (WB-15)	6,1+9,4	0,9	0,6	2,5

* Distance between SU rear wheels and trailer front wheels



The gross mass of this vehicle is determined by the sum of the axle loads unless the axle group centres are less than the minimum when regulation 241 applies

Minimum power.......Regulation 239 (2) (b)...169 kW

Total gross mass is limited to 56 000 kg
The gross combination mass is determined by Regulation 241 and the 560 000 kg limit

Minimum power.......Regulation 239 (2) (b)...234 kW

Figure 3.1: Five Axle Vehicles and Multi Vehicle Combinations

encroachment on adjacent lanes is permissible depending on the frequency of occurrence.

Turning Radii

In constricted situations where the templates are not appropriate, the capabilities of the design vehicle become critical. Minimum turning radii for the outer side of the vehicle are given in Table 3.3. It is stressed that these radii are appropriate only to crawl speeds.

In terms of regulation 355 (a) of the Road Traffic Act, all vehicles must be able to describe a minimum turning radius not exceeding 13,1m.

Vehicle height

Regulation 354 in the Road Traffic Act limits the overall height of a double decker bus to 4,6 metres, and that of any other vehicle to 4,3 metres. The 15th percentile height of a passenger car has been established to be 1,3 m. This has been selected for design purposes as the

Table 3.3: Minimum turning radii	
Vehicle	**Radius (m)**
Passenger Car (P)	6,8
Single unit truck (SU)	10,0
Bus (B)	11,5
Semi-trailer (WB-15)	11,0

passenger car is also an object that has to be seen by the driver in the cases of passing and intersection sight distance.

Driver Eye Height

The passenger car is taken as the critical vehicle for driver eye height and a figure of 1,05 metres is recommended. For buses and single unit vehicles a typical value is 1,8 metres and for semi-trailer combinations the height of the eye can vary between 1,9 metres and 2,4 metres.

Vehicle Turning Paths

The designer should allow for the swept path of the selected design vehicle as it turns. The swept path is established by the outer trace of the front overhang and the path of the inner rear wheel. This turn assumes that the outer front wheel follows the circular arc defining the minimum turning radius as determined by the vehicle steering mechanism.

It is assumed that the turning movements critical to the design of roadway facilities are done at low speeds. At these speeds, the turning behaviour of vehicles is mainly determined by their physical characteristics. The effects of friction and dynamics can safely be ignored. It is also assumed that groups of evenly spaced axles mounted on a rigid bogie act in the turn as a single axle placed at the centre of the group for the purpose of measuring critical turning dimensions.

Commercially available templates and computer software define the turning envelope of vehicles in forward motion and also support plotting of the turning envelope of reversing non-articulated vehicles. Prediction of the reversing behaviour of articulated vehicles is, however, very complex, mainly because this behaviour is inherently unstable, and additional turning controls come into play.

3.4.4 Selecting a design vehicle

In general, buses and heavy vehicles should be used as the design vehicle for cross section elements, with the car as the design vehicle for the horizontal and vertical alignment. For most major intersections along arterial roads or within commercial areas, it is common practice to accommodate the semi trailer. The occasional larger vehicle may encroach on adjacent lanes while turning but not on the sidewalk.

Many authorities designate and signpost specific truck routes. The intersections of two truck routes or intersections where trucks must turn to remain on a truck route should be designed to accommodate the largest semi-trailer combination expected to be prevalent in the turning traffic stream. Where local residential roads intersect truck routes or arterials, the intersections should be specifically designed not to accommodate trucks easily, in order to discourage them from travelling through the residential area.

On major haulage routes, large tractor-trailer combination trucks are prevalent and these routes should be designed to accommodate them. Raised channelising islands are typically omitted in recognition of low pedestrian volumes and other constraints such as right of way and construction costs. The absence of raised islands also allows more manoeuvring area for large trucks.

3.5 SIGHT DISTANCE

3.5.1 General

A critical feature of safe road geometry is adequate sight distance. As an irreducible minimum, drivers must be able to see objects in the road with sufficient time to allow them to manoeuvre around them or to stop. Other forms of sight distance are pertinent. They are:

- Passing sight distance, which is required for substantial portions of the length of two-lane roads;
- Intersection sight distance, to allow a driver on the minor road to evaluate whether it is safe to cross or enter the opposing stream of traffic;
- Decision sight distance where, for example, a driver must be able to see and respond to road markings;
- Headlight sight distance, typically applied to sag vertical curves, and
- Centre line barrier sight distance.

It is also necessary to consider the terrain or obstructions on the inside of horizontal curves when evaluating adequate sight distance.

3.5.2 Deceleration rates

Although research in North America has shown that drivers can choose (or apply) a deceleration of greater than 5 m/sec^2, there is a large degree of variability in driver and vehicle capabilities and the 90th percentile deceleration is of the order of 3,4 m/sec^2. The Institute of Transportation Engineers' Traffic Engineering Handbook states that decelerations of up to 3,0 m/sec^2 are reasonably comfortable for passenger car occupants. This deceleration rate has been adopted for these guidelines.

3.5.3 Object height

The object height to be used in calculation of stopping sight distance is often a compromise between the length of the resultant sight distance and the cost of construction. Stopping is generally in response to another vehicle or large hazard in the roadway. To recognize a vehicle as a hazard at night, a line of sight to its headlights or taillights would be necessary. Larger objects would be visible sooner and provide longer stopping distances. To perceive a very small hazard, for example, a surface obstruction, a zero object height may be necessary. However, at the required stopping sight distances for high speeds, small pavement variations and small objects (especially at night) would not be easily visible. Thus, most drivers travelling at high speeds would have difficulty in stopping before reaching a small obstruction.

A driver will usually attempt to take evasive action rather than to stop for small objects on the roadway. Although not recommended as a design parameter, the time available to manoeuvre is a useful measure when examining variations of geometry in restricted situations or reconstruction projects. In this case, the appropriate object is the pavement surface.

3.5.4 Stopping sight distance

The minimum sight distance on a roadway should be sufficient to enable a vehicle travelling at the design speed on a wet pavement to stop before reaching a stationary object in its path.

Stopping sight distance is the sum of two distances:

Table 3.4: Object height design domain

Object Height (m)	Applicability
0,00	Risk of road washouts Pavement markings in critical locations
0,15	Risk of fallen trees or rocks Risk of log or construction debris fallen from truck Risk of fallen person
0,60	Vehicle tail or brake light
1,30	Passing sight distance for top of car Intersection sight distance

The designer should adopt an object height based on the probability of a particular object occurring on the roadway, as shown in Table 3.4 below. For stopping sight distance, a conservative tail light height of 0,60 m is recommended. If fallen trees or rocks are a real risk, an object height of 0,15 m is recommended. In this context, research has established that the probability of a collision involving an object of a height of 0,15 m or less is infinitesimally small. For passing sight distance, an object height of 1,30 m will allow the driver to discern the top of an oncoming car. A zero object height is recommended where road washouts are a serious risk. It is also recommended for pavement markings in situations such as at intersections or interchanges, where these provide essential guidance.

- the distance traversed by the vehicle from the instant the driver sights an obstruction to the instant the brakes are applied, and
- the distance required to stop the vehicle from the instant the brakes are applied

These are referred to as brake reaction distance and stopping distance, respectively.
These two components, using a reaction time of 2,5 seconds and a deceleration rate of 3,0 m/s^2, result in the relationship

$$s = v\,(0{,}694 + 0{,}013v)$$

where: s = stopping sight distance, m
v = initial speed, km/h

Table 3.5: Recommended stopping sight distances for design			
Design Speed (km/h)	Stopping Sight Distance (m)		
	Calculated	Recommended for Design	Green Book 2000
30	32,5	35	35
40	48,6	50	50
50	67,2	70	65
60	88,4	90	85
70	112,3	110	105
80	138,7	140	130
90	167,8	170	160
100	199,4	200	185
110	233,6	230	220
120	270,5	270	250
130	309,9	310	285

Stopping sight distances calculated using this equation are given in Table 3.5, rounded up for design purposes. Also shown in the table for general interest are the values of stopping sight distance adopted in the 2000 AASHTO Policy on the Geometric Design of Highways and Streets, the "Green Book 2000".

In the measurement of stopping sight distance, the driver's eye height is taken as being at 1,05 m and the object height is as defined in Table 3.4.

3.5.5 Effect of gradient on stopping sight distance

When a highway is on a gradient, the equation for stopping sight distance becomes

$$s = v(0{,}694 + \frac{0{,}004v}{0{,}3 \pm G})$$

in which G is the percentage gradient divided by 100, with upgrades being positive and downgrades negative and the other terms as previously stated. The brake reaction time is assumed to be the same as for level conditions. The stopping sight distance for design speeds from 30 to 130 km/h as corrected for gradient is illustrated in Figure 3.2.

The sight distance at any point on the highway is generally different in each direction, particularly on straight roads in rolling terrain. As a general rule, the sight distance available on downgrades is longer than on upgrades, more or less automatically providing the necessary corrections for grade. This is because down grades are normally followed by sag vertical curves, with the following grade also being visible to the driver

3.5.6 Variation of stopping sight distance for trucks

The recommended minimum stopping sight distance model directly reflects the operation of passenger cars and trucks with antilock braking systems. Trucks with conventional braking systems require longer stopping distances from a given speed than do passenger cars. However, AASHTO suggests that the truck driver is able to see the vertical features of the obstruction from substantially further because of the higher driver eye height. In addition, posted speed limits for trucks in South Africa are considerably lower than for passenger vehicles. Separate stopping sight distances for trucks and passenger cars are, therefore, not generally used in highway design.

There is, however, evidence to suggest that the sight distance advantages provided by the higher driver eye level in trucks do not always compensate for their inferior braking. Some reasons for the longer truck braking distances include:

- Poor braking characteristics of empty trucks. The problem relates to the sus pension and tyres that are designed for maximum efficiency under load;
- Uneven load between axles;
- Propensity of truck drivers not to obey posted speed limits;
- Inefficient brakes of articulated trucks, and
- Effect of curvature. where some of the friction available at the road/tyre interface is used to hold the vehicle in a circular path.

To balance between the costs and benefits in designing for trucks, truck stopping sight distances should be checked at potentially hazardous locations. In general, the deceleration rate for trucks is 1,5 m/s2. The driver's eye height is taken as being at 1,8 m and the object height is as defined in Table 3.4.

The designer should also consider measures such as additional signs to improve road safety

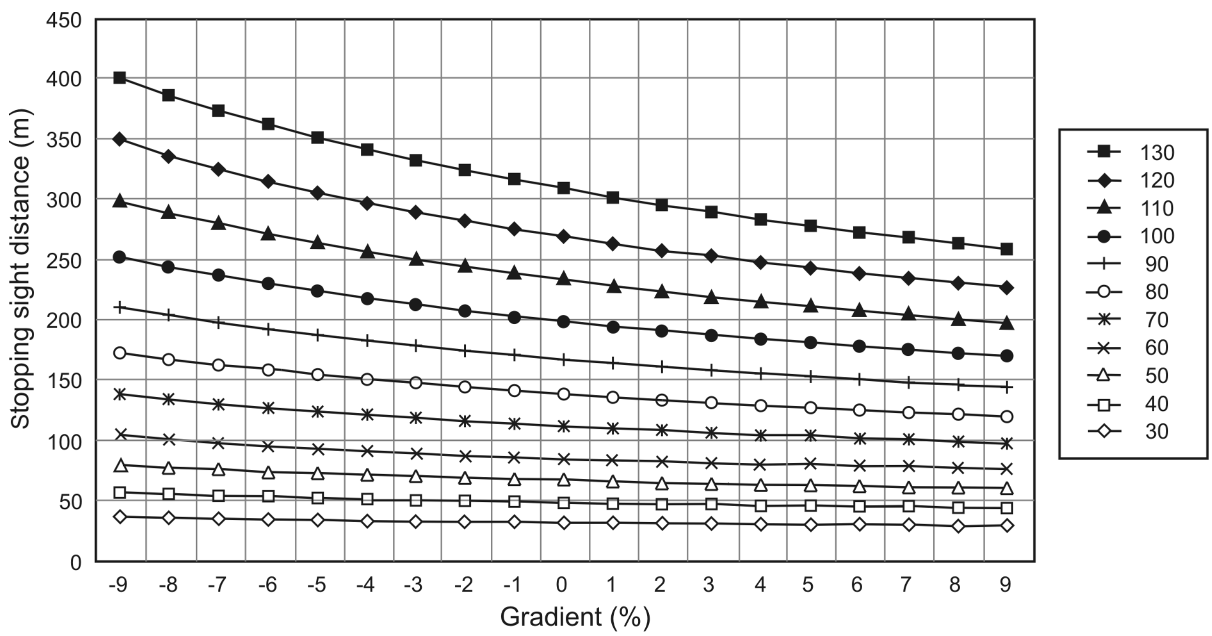


Figure 3.2: Stopping distance corrected for gradient

if stopping sight distance is found to be inadequate for trucks and it is not possible to improve the geometric design. However, it is emphasized that provision of signage is not a substitute for appropriate design practices.

3.5.7 Passing sight distance

On a two-lane rural road, the passing manoeuvre is one of the most significant yet complex and important driving tasks. The process is relatively difficult to quantify, primarily because of the many stages involved, the relative speed of vehicles and the lengthy section of road needed to complete the manoeuvre. Road safety, capacity and service levels are all affected by the passing ability of faster vehicles. This ability is influenced by a variety of factors, including traffic volumes, speed differentials, road geometry and human factors. The minimum sight distance required by a vehicle to overtake safely on two-lane single carriageway roads is the distance which will enable the overtaking driver to pass a slower vehicle without causing an oncoming vehicle to slow below the design speed.

It should be pointed out that there are a variety of models defined for the overtaking manoeuvre. The distances usually given are those required to enable an overtaking driver to complete or abort a manoeuvre already commenced, with safety. In addition to this distance, the Austroads approach introduces a distance that is needed for the driver to identify a length of road as a potential overtaking zone. This "establishment" distance is considerably longer than the overtaking manoeuvre distance.

Table 3.6 shows the minimum overtaking sight distances generally used for various design speeds. Passing manoeuvres involving trucks, particularly in South Africa, require longer distances than those indicated. Designers must take this into account for roads where significant percentages of heavy vehicles are expected in the traffic stream.

Table 3.6: Passing sight distance

Design Speed (km/h)	Absolute Minimum Passing Sight Distance (m)	Desirable Minimum Passing Sight Distance (m)
30	220	250
40	290	350
50	350	400
60	410	450
70	490	550
80	550	650
90	610	750
100	680	900
110	730	1000
120	800	1100
130	860	1200

As mentioned above, the designer should seek opportunities to introduce passing lanes on two-lane roads, particularly where the terrain limits sight distance. A report on a review and evaluation of research studies concluded that passing and climbing lane installations reduce collision rates by 25 per cent compared to untreated two-lane sections. They provide safer passing opportunities for drivers who are uncomfortable in using the opposing traffic lane and for those who become frustrated when few passing opportunities exist, owing to terrain or traffic density.

Sections with adequate passing sight distance should be provided as frequently as possible. The appropriate frequency is related to operating speed, traffic volumes and composition, terrain and construction cost. As a general rule, if passing sight distance cannot be economically provided at least once every 2 km, passing lanes should be considered. The 2+1 cross-section currently in vogue in Europe has some merit. This three-lane cross-section has two lanes in one direction and a single lane in the opposing direction. At about two to three kilometre intervals, the second lane is allocated to movement in the opposite direction. A minimum shoulder width is required as discussed in Chapter 4.

3.5.8 Decision sight distance

Stopping sight distances are usually sufficient to allow reasonably competent and alert drivers to stop under ordinary circumstances. However, these distances are often inadequate when:

- Drivers must make complex decisions;
- Information is difficult to perceive, or
- Unexpected or unusual manoeuvres are required.

Limiting sight distances to those provided for stopping may also preclude drivers from performing evasive manoeuvres, which are often less hazardous and otherwise preferable to stopping. Even with an appropriate complement of standard traffic control devices, stopping sight distances may not provide sufficient visibility for drivers to corroborate advance warning and to perform the necessary manoeuvres. It is evident that there are many locations such as exits from freeways, or where lane shifts or weaving manoeuvres are performed where it would be prudent to provide longer sight distances. In these circumstances, decision sight distance provides the greater length that drivers need. If the driver can see what is unfolding far enough ahead, he or she should be able to handle almost any situation.

Decision sight distance, sometimes termed anticipatory sight distance, is the distance required for a driver to:

- detect an unexpected or otherwise difficult-to-perceive information source or hazard in a roadway environment that may be visually cluttered;
- recognize the hazard or its potential threat;
- select an appropriate speed and path; and
- initiate and complete the required safety manoeuvre safely and efficiently.

Because decision sight distance gives drivers additional margin for error and affords them sufficient length to manoeuvre their vehicles at the same or reduced speed rather than to just stop, it is substantially longer than stopping sight distance.

Drivers need decision sight distances whenever there is likelihood for error in either information reception, decision-making, or control actions. Critical locations where these kinds of errors are likely to occur, and where it is desirable to provide decision sight distance include:

- Approaches to interchanges and inter sections;
- Changes in cross-section such as at toll plazas and lane drops;
- Design speed reductions, and
- Areas of concentrated demand where there is apt to be "visual noise", e.g where sources of information, such as roadway elements, opposing traffic, traffic control devices, advertising signs and construction zones, compete for attention.

The minimum decision sight distances that should be provided for specific situations are shown in Table 3.7. If it is not feasible to provide these distances because of horizontal or vertical curvature or if relocation is not possible, special attention should be given to the use of suitable traffic control devices for advance warning. Although a sight distance is offered for the right side exit, the designer should bear in mind that exiting from the right is in total conflict with driver expectancy and is highly undesirable. The only reason for providing this value is to allow for the remote eventuality that a right side exit has to be employed.

In measuring decision sight distances, the 1 050 mm eye height and 0 mm object height have been adopted.

3.5.9 Headlight sight distance

Headlight sight distance is typically used in establishing the rate of change of grade for sag vertical curves. At speeds above 80 km/h, only large, light coloured objects can be perceived at the generally accepted stopping sight distances. A five-fold light increase is necessary for a 15 km/h increase in speed and a 50 per cent reduction in object size.

Table 3.7: Decision sight distance

Design Speed (km/h)r	Situations				
	Interchanges : Sight distance to nose (metres)		Lane drop, closure, merge. Sight distance to taper area (metres)	Lane shift Sight distance to beginning of shift (metres)	Intersections Sight distance to turn lane
	Left Exit	Right Exit			
50	N/A	N/A	150	85	150
60	200	275	200	100	200
80	250	340	250	150	250
100	350	430	350	200	350
120	400	500	400	250	400

For night driving on highways without lighting, the length of visible roadway is that which is directly illuminated by the headlights of the vehicle. This length is typically shorter than the minimum sight distance.

When headlights are operated on low beam, the reduced candlepower at the source and the downward projection angle significantly restrict the visible length of roadway surface.

For crest vertical curves, the area beyond the headlight beam point of tangency with the roadway surface is shadowed and receives only indirect illumination. Also, a general limit of 120 to 150 metres sight distance is all that can be safely assumed for visibility of an unilluminated object on a bitumen surfacing. This corresponds to a satisfactory stopping sight distance for 80 to 90 km/h or a decision time of about 5 seconds at 100 km/h.

Since the headlight mounting height (typically about 600 mm) is lower than the driver eye height (1 050 mm for design), sight distance is controlled by the height of the vehicle headlights and a one degree upward divergence of the light beam from the longitudinal axis of the vehicle. Any object within the shadow zone must be high enough to extend into the headlight beam to be directly illuminated.

3.5.10 Barrier sight distance

Barrier sight distance is not a geometric design factor, but is rather an operational guide to the driver to promote safety on two-lane roads.

Barrier sight distance is the limit below which overtaking is legally prohibited, in order to ensure that two opposing vehicles travelling in the same lane should be able to come to a stop before impact. A logical basis for the determination of the barrier sight distance is that it should at least equal twice the stopping distance. Values given in the South Africa Road Traffic Signs Manual approximate this approach.

Barrier sight distance is measured to an object height of 1,3 metre from an eye height of 1,05 m. . The object height is the height of an approaching passenger car.

Hidden dip alignments are poor design practice, but are found on many rural roads. They typically mislead drivers into believing that there is more sight distance available than actually exists. In checking vertical alignment, designers should pay attention to areas where this deficiency exists, and ensure that drivers are made aware of any such inadequacies.

3.5.11 Obstructions to sight distance on horizontal curves

Physical features, such as a concrete barrier wall, a bridge pier, a tree, foliage, or the back slope of a cutting, can affect available sight distance. Accordingly, designs need to be checked in both the horizontal and vertical planes for obstructions.

Minimum radii of horizontal curvature are determined by application of vehicle dynamics and not through sight distance controls. It is, therefore, possible that the selected radius may not be adequate to ensure the safe stopping sight distance requirements. If the obstructions to sight distance are immovable, re- alignment may be necessary.

The problem is illustrated in Figure 3.3. The driver's eye is assumed to be at the centre of the nearside lane. The chord AB is the sight line and the curve ACB is the stopping sight distance. A zero gradient is assumed. It follows that selection of a radius for a given distance of

L	=	lane width (m)
s	=	stopping sight distance for specific gradient and design speed (m) from Figure 3.1.

Given the nature of the relationship, a trial-and-error approach to the solution is required.

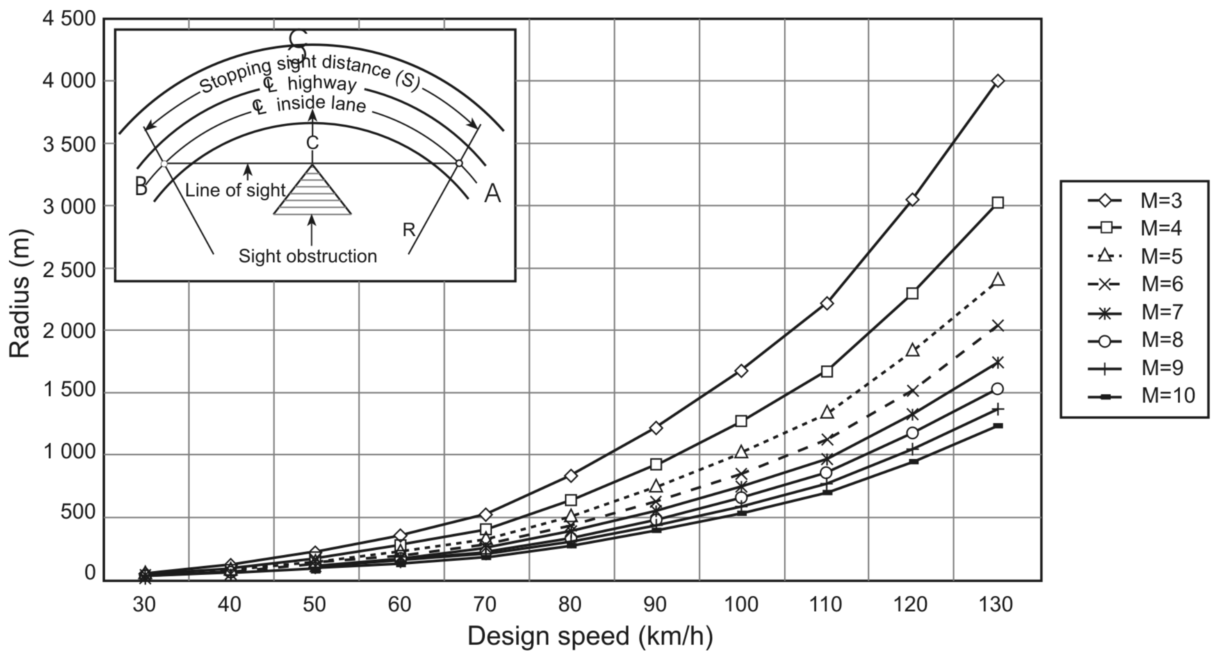


Figure 3.3: Horizontal restrictions to stopping sight distance

obstruction from the inner lane centre line will constitute an under-design if the inner lane is on a downgrade.

A radius to satisfy stopping sight distance criteria can be calculated from the following formula;

$$C = \left(R - \frac{L}{2}\right)\left\{1 - \left[1 - \tan^2\left(\frac{28{,}6s}{R - \frac{L}{2}}\right)\right]^{0{,}5}\right\}$$

where	C	=	distance from centre of inside lane to obstruction (m)
	R	=	radius of curve (m)

3.6 ENVIRONMENTAL FACTORS

A road is a key element in the modern environment with wide ranging implications. Planning for effective integration is therefore essential. In South Africa, the National Environmental Management Act of 1998 lays down prescriptions for the provision and operation of infrastructure, including roads. The designer should be aware of the constraints imposed by this law. For example, constraints may include avoiding a particular watercourse or wetland area, or accommodating prescribed mitigatory measures such as screening berms or sound fences. In carrying out their mandate to plan and design road systems, road designers should consider on the one hand, making facilities aesthetically pleasing and being "good neighbours" in the community and, on the other, providing safe and efficient transportation links to users.

3.6.1 Land use and landscape integra tion

With regard to environmental factors, the objective of route selection should be to choose a route that has both the minimum effect on landform and requires the smallest number of large earthworks. Integration with the existing landform can best be achieved by grading out cuttings and embankments to slopes that reflect the surrounding topography. This in turn may affect adjacent sites of conservation or heritage interest and, in such cases, a balance needs to be struck. A major consideration is that non-renewable resources, such as wetlands, should be avoided wherever possible.

Designs should aim to achieve the best possible use of excavated materials, thus minimizing the need for off-site spoil or borrow pits. If off-site works are necessary, they should be subject to the same good design principles as those used on site, achieved by liaison with the appropriate planning authority. Earthworks can only be integrated successfully if the new landform and its soil structure allow effective strategic rehabilitation. Restoration to agricultural use can be a particularly effective strategy.

Design objectives should be:

- To choose the route least damaging to the landscape and which respects existing landform best by avoiding disruption of major topographical features;
- To find an alignment that uses the existing landform to good effect and which minimizes the scale of earthworks;
- To design profiles which reflect existing natural slopes;
- To retain the least road footprint, by the return of land to its former use;
- To use existing landform to minimize noise and visual intrusion: for example, placing a road in a cutting or behind rising ground to protect settlements;
- To develop new landforms, including mounds and false cuttings, to screen the road from settlements, and
- To achieve a balance between horizontal and vertical alignment.

3.6.2 Aesthetics of design

Design aesthetics and attention to landform are very closely related topics. Aesthetic improvements can often be achieved without incurring additional costs, provided the designer approaches the subject in a sensitive manner. In fact, alignments that are visually pleasing are usually less hazardous than other alignments.

On any roadway, creating pleasing appearance is a worthwhile objective. Scenic values can be considered along with safety, utility, economy, and all the other factors considered in planning and design. This is particularly true of the many portions of the National Road system situated in areas of natural beauty. The location of the road, its alignment and profile, the cross section design, and other features should be in harmony with the setting. Economy consistent with traffic needs is of paramount importance, although a reasonable additional expenditure can be justified to enhance the beauty of the highway.

This topic is addressed in detail in Chapter 5.

3.6.3 Noise abatement

Noise is defined as an unwanted sound, a subjective result of sounds that intrude on or interfere with activities such as conversation, thinking, reading or sleeping. Motor vehicle noise is generated by the functioning of equipment within the vehicle, by its aerodynamics, by the action of tyres on the roadway and, in metropolitan areas, by short-duration sounds such as braking squeal, exhaust backfires, hooters and sirens.

The decrease in sound intensity with distance from the source is influenced by several factors. Measurements taken near roads show that doubling the distance results in a lowering of 3 dBA over clean, level ground and 4,8 dBA over lush growth.

Some sustained (ambient) noise is always present. In a quiet residential neighbourhood at night, it is in the 32 to 43 dBA range; the urban residential daytime limits are about 41 to 53 dBA. In industrial areas the range is 48 to 66 dBA, while, in downtown commercial locations with heavy traffic, it is 62 to 73 dBA. Increases up to 9 dBA above ambient noise levels bring only sporadic protests. Protests become widespread with increases in the 9 to 16 dBA range. At increases greater than 16 dBA, there may be community action.

A design objective is to keep noise at or below acceptable levels and this can be achieved either by absorbing the noise or deflecting it upwards. Strategies addressing noise levels include, for example, depressing and sometimes covering the roadway or by installing sound barriers of earth or masonry. However, these may also trap air pollutants.

Special sound barriers may be justified at certain locations, particularly along ground level or elevated roads through noise-sensitive areas. Concrete, wood, metal, or masonry walls are very effective in deflecting noise. One of the more aesthetically pleasing barriers is the earth berm that has been graded to achieve a natural form that blends with the surrounding topography. The feasibility of berm construction should be planned as part of the overall grading plan for the roadway. There will be instances where an effective earth berm can be constructed within the normal right-of-way or with a minimal additional right-of-way purchase. If the right-of-way is insufficient to accommodate a three metre high berm, a lower berm can be constructed in combination with a wall or screen to achieve the desired height.

Shrubs, trees or ground covers are not very efficient in shielding sound because of their permeability to the flow of air. However, almost all buffer plantings offer some noise reduction, and exceptionally wide and dense plantings may result in substantial reductions in noise levels. Even where the noise reduction is not considered significant, the aesthetic effects of the plantings will produce a positive influence.

3.6.4 Air pollution by vehicles

The highway air-pollution problem has two dimensions: the area-wide effects of primarily reactive pollutants; and the high concentrations of largely non-reactive pollutants at points or corridors along or near roads. The motor vehicle is a primary contributor to both forms, accounting for an estimated 70 per cent of the CO, 50 per cent of the HC, and 30 per cent of the NOx.

Area-wide conditions are exacerbated when temperature inversions trap pollutants near the ground surface and there is little or no wind, so that concentrations of pollutants increase. For some individuals, eyes burn and breathing is difficult. It is alleged that lives can be shortened and some deaths have actually resulted from these exposures. Also, certain kinds of vegetation are killed, stunted, or the foliage burned.

The quantity of air pollutants can be reduced by judicious design. Exhaust emissions are high while vehicle engines are operating at above-average levels of output, for example while accelerating from a stationary position or when climbing a steep hill. Smooth traffic flow at constant speeds, such as in "green wave" conditions on a signalised route, reduces exhaust emissions in addition to leading to a reduction in noise levels. By way of contrast, speed humps, which are popular as speed-reducing devices in residential areas, have the dual penalty of increased pollution and increased noise levels. In rural areas, vertical alignments should be designed with a minimum of "false rises".

In addition to being able to modify the quantity of pollutants in the atmosphere, the designer can influence the extent to which emissions impact on local communities. Temperature inversions that trap polluted air are typically associated with closed or bowl-shaped valleys. If at all possible, major routes should not traverse such areas but should rather be located on the higher ground surrounding inversion-prone valleys, with relatively low-volume road links serving developments in the valley areas. Attention should also be paid to prevailing wind directions so that routes bypassing local communities are located downwind of these settlements.

3.6.5 Weather and geomorphology

Land shape, on a broad scale, as well as prevailing weather conditions, which could influence the design, are factors over which the designer does not have any control. Certain areas of the country are prone to misty conditions and others subject to high rainfall. Both are factors that have to be taken into account in design.

Mist and rain both cause reduced visibility. Where these are a regular occurrence, they tend to lie in belts, sometimes fairly narrow, across the landscape. Designers should acquire local knowledge about the quirks of the weather patterns and seek ways to reduce their effect.

Where it is not possible to avoid a mist belt, the designer should pay particular attention to the concept of the "forgiving highway", by providing flat side slopes and avoiding alignments where short radius curves follow each other in quick succession. Steep downgrades followed by short radius horizontal curves are particularly to be avoided. A real effort should also be put into avoiding high fills. In conditions of heavy mist, vehicles will tend to move very slowly but, even at speeds significantly below the design speed of the road, the restricted visibility will lead to high levels of stress. Drivers are more likely to make incorrect decisions when under stress and designers should thus do everything possible to keep stress levels within manageable limits.

3.7 TRAFFIC CHARACTERISTICS

3.7.1 General

Factual information on expected traffic volumes is an essential input to design. This indicates the need for improvements and directly affects the geometric features and design.

Traffic flows vary both seasonally and during the day. The designer should be familiar with the extent of these fluctuations to enable him or her to assess the flow patterns. The directional distribution of the traffic and the manner in which its composition varies are also important parameters. A thorough understanding of the manner in which all of these behave is a basic requirement of any realistic design.

3.7.2 Traffic volumes

Traffic flow is measured by the number of vehicles passing a particular station during a given period of time. Typically, the flow of interest is the Average Daily Traffic (ADT). Flows may also be reported per hour, such as the "hourly observed traffic volume" or the "thirtieth-highest hour" or "hundredth-hour", which are commonly used for design purposes. Very short duration flows, such as for a five-minute period, are typically applied to studies of signalised intersections.

If hourly flows are ordered from highest to lowest, it is customary, in rural areas, to design for the thirtieth highest hourly flow, i.e that flow which is exceeded in only 29 hours of the year. This is because rural roads have very high seasonal peaks and it is not economical to have a road congestion-free every hour throughout the year. In urban areas, seasonal peaks are less pronounced and the 100th highest hourly flow is considered a realistic flow level for design purposes.

To predict hourly flows, it is necessary to know the ADT and the peaking factor, ß. The parameter, ß, is a descriptor of the traffic flow on a given road and depends on factors such as the percentage and incidence of holiday traffic, the relative sizes of the daily peaks, etc. The peaking factor can fluctuate between -0,1 and -0,4. A value of -0,1 indicates minimal seasonal peaking. This value of ß should be used in urban designs. A value of -0,4 suggests very high seasonal peaks and would normally be applied to roads such as the N3. As a general rule, a value of -0,2 could be used as being a typical value. Equation 3.1 below can be used to estimate flows between the highest and 1030th highest hour. Although not a particularly good model, flows beyond the 1030th hour can be estimated by using a straight line relationship from the 1030th flow to zero veh/hr at the 8760th or last hour of the year.

$$Q_N = 0{,}072ADT(N/1030)^{\beta}$$

where
Q_N = two-directional flow in N-th hour of year (veh/h)
ADT = average daily traffic (veh/day)
N = hour of year
ß = peaking factor.

It is interesting to note that the peak hour factor, K, quoted in the Highway Capacity Manual is often assumed for design purposes to be 0,15. Reference is commonly made to the 30th highest hour of the year as being the design hour. Applying a value of -0,2 to ß, and assuming N to be 30, Q_N according to Equation 3.1 is 0,146 x ADT for the thirtieth highest hour.

Designers need to estimate future traffic flows for a road section. It is recommended that a design period of 20 years be used for forward planning. The 30th or 100th highest flow used in the design is that occurring in the design year, typically twenty years hence. Staged construction or widening of roads over this period can be a feature of an economical design.

The capacity of rural road sections is influenced by the following key characteristics:

- Road configuration - e.g. two-lane two-way, multi-lane divided or undivided;
- Operating speed;
- Terrain;
- Lane and shoulder width;
- Traffic composition, and
- Gradients.

In the case of two-lane two- way roads, the following additional factors are important:

- Directional distribution of traffic flow; and
- Passing opportunities - sight distance, overtaking lanes, climbing lanes or slow vehicle turnout lanes.

3.7.3 Directional distribution

Directional distribution of traffic is an indication of the tidal flow during the day. In urban areas, the morning peak traffic is typically inbound towards the central business district (with relatively low outward-bound flows), whereas the afternoon peak is in the reverse direction. It is important to realize that the design flow is actually a composite and not a single value. A road must be able to accommodate the major flow in both directions.

The actual distribution to be used for design purposes should be measured in the field. If an existing road is to be reconstructed, the field studies can be carried out on it beforehand. For new facilities, measurements should be made on adjacent roads from which it is expected the traffic will be diverted and modelling techniques applied.

Directional distribution is relatively stable and does not change materially from year to year. Relationships established from current traffic movements are normally also applicable to future movements.

3.7.4 Traffic composition

Vehicles of different sizes and mass have different operating characteristics. Trucks have a higher mass/power ratio and occupy more roadway space than passenger cars. Consequently, they constitute a greater impedance to traffic flow than passenger vehicles, with the overall effect that one truck is equivalent to several passenger cars. For design purposes, the percentage of truck traffic during the peak hours has to be estimated.

For design of a particular highway, data on the composition of traffic should be determined from traffic studies. Truck traffic is normally

Table 3.8: Equivalent passenger car units				
Vehicle type	Passenger car units			
	Rural roads	Urban Streets	Roundabouts	Traffic Signals
Cars and light vans	1,0	1,0	1,0	1,0
Commercial vehicles	3,0	1,75	2,8	1,75
Buses and coaches	3,0	3,0	2,8	2,25
Motorcycles	1,0	0,75	0,75	0,33
Pedal cycles	0,5	0,33	0,5	0,2

expressed as a percentage of total traffic during the design hour in the case of a two-lane road; and as a percentage of total traffic in the predominant direction of travel in the case of a multi-lane road.

It is not practical to design for a heterogeneous traffic stream and, for this reason, trucks are converted to equivalent Passenger Car Units (PCUs). The number of PCU's associated with a single truck is a measure of the impedance that it offers to the passenger cars in the traffic stream. This topic is exhaustively addressed in the Highway Capacity Manual and is not discussed further here.

Passenger car unit equivalents have, in general, been derived from observations as illustrated in Table 3.8. The values offered serve only as a rough indication and the designer should refer to the Highway Capacity Manual for detailed calculation of PCU's appropriate to the different environments and circumstances.

3.7.5 Traffic growth

The design of new roads or of improvements to existing roads should be based on the future traffic expected to use the facilities.

It is difficult to define the life of a "road" because major segments may have different lengths of physical life. Each segment is subject to variations in estimated life expectancy because of influences not readily subject to analysis such as obsolescence and unexpected changes in land use, resulting in changes in traffic volumes, pattern and load. Regardless of the anticipated physical life of the various elements of the road, it is customary to use a single value as the "design life". In essence, the road is expected to provide an acceptable level of service for this period. Whether or not any of its various components have a longer physical life expectancy than this design life is irrelevant. For example, the alignment and, in some instances, the surfacing of roads built during Roman times are still in use today without there being any reference to a design life of 2 000 years.

A typical value of design life is twenty years. In the case of major and correspondingly expensive structures a design life of fifty years may be assumed. This should not, however, be confused with the concept of a bridge being able to withstand the worst flood in fifty years.

3.7.6 Capacity and design volumes

The term capacity is used here to define the ability of a road to accommodate traffic under given circumstances. Factors to be taken into account are the physical features of the road itself and the prevailing traffic conditions

Prevailing road conditions

Capacity figures for uninterrupted flow on highways have to be modified if certain minimum physical design features are not adhered to. Poor physical features that tend to cause a reduction in capacity are:

- Narrow traffic lanes. Lane widths of 3.65 m are accepted as being the minimum necessary for heavy volumes of mixed traffic, i.e. before capacity of the lane is reduced.
- Inadequate shoulders. The narrowness, or lack of, shoulders alongside a road cause vehicles to travel closer to the centre of the carriageway, thereby increasing the medial traffic friction. In addition, vehicles making emergency stops must, of necessity, park on the carriageway. This causes a substantial reduction in the effective width of the road, thereby reducing capacity.
- Side obstructions. Vertical obstructions such as poles, bridge abutments, retain ing walls or parked cars that are located within about 1,5 m of the edge of the carriageway contribute towards a reduction in the effective width of the outside traffic lane.
- Imperfect horizontal or vertical curvature. Long and/or steep hills and sharp bends result in restricted sight distance. As drivers then have reduced opportunities to pass, the capacity of the facility will be reduced.

In addition to the above, the capacities of certain rural roads and the great majority of urban roads are controlled by the layouts of intersections. Physical features having considerable influence are the type of intersection, i.e. whether plain, channelised, roundabout or signalised, the number of intersecting traffic lanes and the adequacy of speed-change lanes.

Unlike the physical features of the highway, which are literally fixed in position and have definite measurable effects on traffic flows, the prevailing traffic conditions are not fixed but vary from hour to hour throughout the day. Hence, the flows at any particular time are a function of the speeds of vehicles, the composition of the traffic streams and the manner in which the vehicles interact with each other, as well as of the physical features of the roadway itself.

Capacity

The term "capacity" was introduced in the USA in the Highway Capacity Manual, in which it is defined as "the maximum number of vehicles that can pass a given point on a roadway or in a designated lane during one hour without the traffic density being so great as to cause unreasonable delay, hazard, or restriction to the drivers' freedom to manoeuvre under the prevailing roadway and traffic conditions". This definition gives a reasonable method of approach but, in practice, it is necessary to choose one or more arbitrary criteria of what constitutes restriction of traffic movement, or "congestion".

The Highway Capacity Manual procedure must however be used for specific road capacity designs.

For typical South African conditions and to balance financial, safety and operational considerations it is recommended that the capacity of a two-lane rural road be taken on average as being between 10 000 and 12 000 vehicles per day while, for freeways, consideration could be given to changing from a four-lane to a six-lane freeway when the traffic flow is of the order of 35 000 to 40 000 vehicles per day.

3.8 ROAD CLASSIFICATION

The classification of roads into different operational systems, functional classes or geometric types is necessary for communication between engineers, administrators and the general public. Classification is the tool by which a complex network of roads can be subdivided into groups having similar characteristics. A single classification system, satisfactory for all purposes, would be advantageous but has not been found to be practicable. Moreover, in any classification system the division between classes is often arbitrary and, consequently, opinions differ on the best definition of any class. There are various schemes for classifying roads and the class definitions generally vary depending on the purpose of classification.

The principal purposes of road classification are to:-

- Establish logical integrated systems that, because of their particular service, should be administered by the same jurisdiction;
- Relate geometric design control and other design standards to the roads in each class, and
- Establish a basis for developing long-range programmes, improvement priori ties and financial plans.

Classification of roads by design types based on the major geometric features (e.g. freeways) is the most helpful one for road location and design purposes. Classification by route numbering is the most helpful for road traffic operational purposes, whilst administrative classification is used to denote the level of government responsible for, and the method of, financing road facilities. Functional classification, the grouping of roads by the character of service they provide, was developed for transportation planning purposes.

3.8.1 Classification criteria for South African roads

Although numerous classification criteria are used on road networks world-wide, in South Africa there are basically three criteria used to classify road types. These are:

- Functional classification;
- Administrative classification, and
- Design Type classification, based on traffic usage.

This document uses the third type for design purposes.

As a result of growing awareness of the interdependency of the various modes of transport as well as the creation of Metropolitan Transport Authorities within South Africa's major metropolitan areas there is considerable overlap between the functional and administrative classification criteria. Within these metropolitan areas, the general public is more dependent on, and understands, a route numbering or functional classification than on an administrative classification of roads within the area.

Although these guidelines are based on a design type classification, the three different approaches mentioned above are briefly described in order to provide a picture of the road system hierarchy in South Africa.

3.8.2 Functional classification concept

For transportation planning purposes, road are most effectively classified by function. The functional classification system adopted for the South African road network is illustrated in Table 3.9. This was used for the South African Rural Road Needs Study carried out during the early 1980s.

Another and less comprehensive form of functional classification was developed for the purposes of road signing as shown in the South African Road Traffic Signs Manual. As stated in SARTSM,

"There are definite limits to the number of ways in which GUIDANCE signs and specifically DIRECTION signs can be made to indicate with sufficient immediate recognition potential, the different classes into which the road network is divided for signing purposes."

Classification for signing thus differentiates mainly between numbered and unnumbered routes and, in respect of numbered routes, also draws a distinction between freeways and other roads.

Table 3.9: Road functional classification

Functional Classification	Description of Road Function or Type
1	Roads which form the principal avenues of communication between: ❑ Major regions of the RSA and/or ❑ Defined or proposed metropolitan areas and/or ❑ Major regions of the RSA and other countries and which are declared National roads or the extensions of these.
1a*	Roads which comply with the above but which are not declared National roads or extensions of these.
2	Roads, not being Class 1 or 1a, which form main avenues of communication between: ❑ Important centres** and class 1 and 1a roads and/or ❑ Important centres and/or ❑ Of an arterial nature within a town in a rural area.
3	All other surfaced roads for which the road authorities are responsible and which are not Class 1, 1a, 2 or 5 roads.
4	All Gravel Roads for which the road authorities are responsible, excluding Class 5 roads.
5***	Special Purpose Roads: Roads which provide for some particular activity or function and which are not assigned to Classes 1 to 4, e.g.: ❑ For minerals development; ❑ For strategic or defence purposes; ❑ For social need, or ❑ For agricultural or other development.

NOTES: * *Class 1a roads include declared National roads which do not follow an "N" route.*

** *Important centres are centres with population of 5 000 or more.*

*** *Class 5 roads may be either surfaced or gravel roads*

Roads have two functions: to provide mobility and to provide land access. However from a design standpoint, these functions are incompatible. For mobility, high or continued speeds are desirable and variable or low speeds undesirable; for land access, low speeds are desirable and high speeds undesirable. For example, freeways provide a high degree of mobility, with access provided only at spaced interchanges to preserve the high-speed, high-volume characteristics of the facility. The opposite is true for local, low-speed roads that primarily provide local access. The general relationship of functionally classified systems in serving mobility and land access is illustrated in Figure 3.4.

Given a functional classification, design criteria can be applied to encourage the use of the road as intended. Design features that can convey the level of functional classification to the driver include width of roadway, continuity of alignment, spacing of intersections, frequency of access points, building setbacks, alignment and grade standards, and traffic controls.

3.8.3 Administrative classification

Legislation and, in some instances, the Constitution, assigns to certain levels of government the responsibilities for providing, regulating and operating roads and streets for public use. This concept and the principles of law that support it were developed in Great Britain, and even earlier by the Romans. Within the limits of its constitutional powers, a particular road authority may delegate its authority for roads to bodies such as a Roads Board or a Toll Road Concessionaire. In South Africa, the three separate levels of government each have a roads function mandated to them. Despite this separation of authority for various classes or roads, it is essential to bear in mind that roads act as a total system or network and that the subdivision of roads into administrative classes bears no relation to the functional type of a road under the control of a specific authority. The administrative classification approach thus divides the South African road network into:

- National,
- Provincial, and
- Local Authority roads.

This administrative classification of roads also bears no relation to the design type of a road under the control of a specific authority. Thus, until the turn of the 21st century, there were sections of National Roads that were unsurfaced, with very low traffic volumes, the most heavily trafficked roads in South Africa, also up to the turn of the century, often being the responsibility of a local authority. Furthermore certain "routes" in the country comprise both National and Provincial roads, and could even include local authority roads. Administratively, a National Road, which is denoted as such by a legal proclamation, in general comprises those roads that form the principal avenues of communication between major regions of the country, and/or between major population conurbations and/or between major regions of South Africa and other countries.

3.8.4 Design type classification

The most widely accepted design type criteria are those developed by AASHTO which classify a road system into:

- Freeways;

- Arterial roads other than freeways, and
- Collector roads.

For each classification, specific design standards and criteria and access and other policies have been developed and are applied.

For use in the present document the following service or design classifications are proposed, Design designations of these specific National Roads are as follows;

Class I **Primary Roads**

Class IA *Primary Freeways in rural areas*

- Illustrative threshold ADT (with more than 12% heavy vehicles) = 15 000 veh/d.

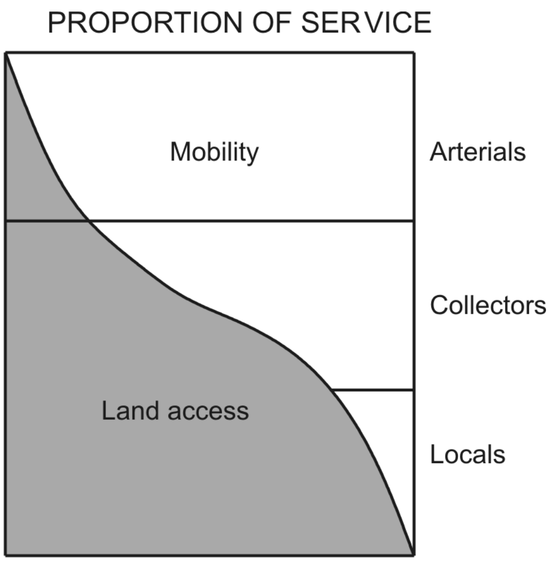


Figure 3.4: Relationship of functional road classes

related to design traffic volumes given in ADT terms. Each classification groups roads with similar functions. The factors that influence the classification of a roadway to a certain group include;

- Trip purpose (for the majority of users);
- Trip length;
- Size and type of population centre served;
- Traffic characteristics, and
- Network and system requirements

- Average travel distances on links indicated by at least 45 per cent of the trips being more than 2 hours in duration;
- May be provided for network continuity purposes, and
- Minimum design speed 130 km/h.

Class IB *Primary Freeways in metropolitan areas*

- Illustrative threshold ADT = 20 000 veh/d

- Average travel distance on links indicated by majority of trips being less than 2 hours in duration
- Form integral element of Metropolitan Road Network
- Generally extension of Rural Freeways (Class IA roads)
- Minimum design speed 130 km/h

Class II Primary Arterials, 4 lane single carriageway roads

Class IIA Primary Rural Arterials

- Generally provided when 2 lane single carriageway road reaches capacity and freeway not financially affordable
- Ilustrative threshold ADT : 8 000 - 10 000 veh/d, with bottom end of scale applicable where percentage of truck traffic exceed 15 per cent
- Minimum design speed 120 km/h

Class IIB Primary Metropolitan Arterial

- Design in context in which it operates.

Class III Secondary Rural Arterial

- Provided to address inter-regional travel demands, or providing access to tourist or National resource areas
- Provided to address inter-regional travel demands, or providing access to tourist or National resource areas
- Provided to address inter-regional travel demands, or providing access to tourist or National resource areas
- Two lane, single carriageway roads
- Class IIIA
- Design ADT greater than 4 000 veh/d
- Minimum design speed 120 km/h
- Class IIIB
- Design ADT 500 - 4 000 veh/d
- Minimum design speed 110 km/h
- Class IIIC
- Design ADT less than 500 veh/d
- Minimum design speed 100 km/h.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 4

ROAD DESIGN ELEMENTS

4.1 INTRODUCTION

Regardless of the philosophy brought to bear on the design of a road or the classification of the road in the network, the final design comprises a grouping and sizing of different elements. A vertical curve is an element. The development of super elevation is an element. The cross-section is heavily disaggregated, comprising a large group of elements. Geometric design thus comprises:

- The selection of elements to be incorporated in the design;
- The sizing of the selected elements, and
- Linking the elements into a three-dimensional sequence.

The selected, sized and linked elements in combination represent the final design of a road which, when built, must constitute a network link which will satisfactorily match the criteria of safe, convenient and affordable transportation with minimum side effects, simultaneously addressing needs other than those pertaining directly to the movement of people or freight.

A road network comprises links and nodes. The intersections and interchanges are the nodes and the roads connecting them the links. Intersections and the elements that constitute them are discussed in Chapter 6, with interchanges being dealt with in Chapter 7.

In this chapter, the various elements involved in link design are discussed.

4.2 HORIZONTAL ALIGNMENT

The horizontal alignment comprises three elements: tangents, circular curves and the transitions between tangents and curves.

Tangents (sometimes referred to as "straights") have the properties of bearing (direction or heading) and length. Circular curves also have two properties; radius and deflection (or deviation) angle, these two properties directly leading to a third property of interest, namely curve length. The basic properties of transition curves are shape and length. Design of the horizontal alignment includes selection of the values associated with each of these six properties.

4.2.1 General Controls for horizontal alignment

The horizontal alignment should be as directional as possible and consistent with the topography. However, it is equally important in terms of context sensitive design to preserve developed properties and areas of value to the community.

Winding alignments composed of short curves and tangents should be avoided if at all possible because they tend to cause erratic operation and a high consequent crash rate.

Most design manuals recommend that minimum radii should be avoided wherever possible, suggesting that flat curves (i.e. high values of radius) should be used, retaining the minimum for the most critical conditions. The concept of consistency of design, on the other hand, suggests that the difference between design speed and operating speed should ideally be held to a maximum of 10 km/h, with a 20 km/h difference still representing tolerable design. This could be construed as a recommendation that minimum curvature represents the ideal, which is wholly at variance with the historic approach to selection of curve radius. What is actually intended however is that the designer should seek to employ the highest possible value of design speed for any given circumstance.

For small deflection angles, curves should be sufficiently long to avoid the appearance of a kink. A widely adopted guideline is that, on minor roads, curves should have a minimum length of 150 metres for a deflection angle of 5^{o} and that this length should be increased by 30 metres for every 1^{o} decrease in deflection angle. On major roads and freeways, the minimum curve length in metres should be three times the design speed in km/h. The increase in length for decreasing deflection angle also applies to these roads. In the case of a circular curve without transitions, the length in question is the total length of the arc and, where transitions are applied, the length is that of the circular curve plus half the total length of the transitions.

South African practice recommends an upper limit to the length of horizontal curves. Curves to the left generally restrict passing opportunities and, furthermore, dependant on their radius, operation on long curves tends to be erratic. For this reason, it is desirable to restrict the length of superelevated curves to a maximum of 1 000 metres. If a curve radius is such that the curve either has a normal camber or a crossfall of 2 per cent, the limitation on curve length falls away and the curve can be dealt with as though it were a tangent.

Broken-back (also referred to as "flat back") curves are a combination of two curves in the same direction with an intervening short tangent. These should be avoided on the twin grounds of aesthetics and safety. Not only are broken-back curves unsightly but drivers do not always recognize the short intervening tangent and select a path corresponding to the radius of the first curve approached, hence leaving the road on the inside and part way along the tangent. On roads in areas with restrictive topography, such as mountain passes, the designer may have no option but to accept the use of broken-back curves but should, nevertheless, be aware of their undesirability. Where the intervening tangent is longer than 500 metres, the appellation "broken-back" is no longer appropriate.

Reverse curves are also a combination of two curves but in opposite directions with an intervening short tangent. These curves are aesthetically pleasant but it is important to note that the intervening tangent must be sufficiently long to accommodate the reversal of superelevation between the two curves.

Although compound curves afford flexibility in fitting the road to the terrain and other ground controls, their use should be avoided outside intersection or interchange areas. Once they are on a horizontal curve, drivers expect the radius to remain unaltered hence supporting a constant speed across the length of the curve. A compound curve is thus a violation of driver expectancy and can be expected to have a correspondingly high crash rate.

4.2.2 Tangents

Two fundamentally different approaches can be adopted in the process of route determination. In a curvilinear approach, the curves are located first and thereafter connected up by tangents. This approach can be adopted with advantage in mountainous or rugged terrain. A typical consequence of this approach is that curves tend to be long and tangents short. Because of the local topography, the approach to route determination most frequently adopted in South Africa is for the tangents to be located first and the curves fitted to these tangents thereafter.

Bearing

Economic considerations dictate that, where other constraints on route location are absent, roads should be as directional as possible. In consequence, tangents may be located on bearings that have an adverse impact on driver comfort and safety. Two conditions require consideration.

In the first instance, a combination of gradient, direction of travel and time of day may cause the driver to be dazzled by the sun. The most obvious example is the east-west bearing where the vehicle would be moving towards the rising or setting sun. No direction of travel is completely exempt from this problem. For example, when travelling at midday in mid-winter in a northerly direction up a gradient steeper than about eight per cent, the sun can present a problem. Obviously this would be of fairly limited duration in terms of the season and the likelihood of a prolonged gradient of this magnitude, whereas the east-west bearing would be a problem all year round and over an extended period of time during the day. The designer should be aware of this problem and, if possible, avoid selecting bearings that reduce visibility. If the dazzle problem cannot be avoided, warning signage may be considered.

In the second instance, a bearing at right angles to the prevailing wind direction can cause problems for empty trucks with closed load compartments, e.g. pantechnicons, or passenger cars towing caravans. On freight routes or routes with a high incidence of holiday traffic, the designer should seek to avoid this bearing. If not possible, locating the road on the lee of a hill that could then offer shelter from the wind may be an option to be explored in order to offer some relief.

Length

The minimum allowable length of tangent is that which accommodates the rollover of superelevation in a reverse curve situation.

It has been found that extremely long tangents, e.g. lengths of twenty kilometres or more, have accident rates similar to those on minimum length tangents, the lowest accident rate occurring in a range of eight to twelve kilometres. This range is recommended for consideration in fixing the maximum length of tangent on any route. This maximum is based on the assumption of a design speed of 120 km/h or more. At lower design speeds, it is necessary to consider maximum lengths considerably shorter than eight to twelve kilometres as discussed below.

As a rough rule of thumb (which is adequate for planning purposes) the maximum length of tangent in metres should not exceed ten to twenty times the design speed in kilometres/hour. If the achievable maximum length of tangent across the length of the route is regularly greater than this guideline value, thought should be given to consideration of a higher design speed. Rules of thumb have their limitations and, in this case, application should be limited to design speeds of 100 km/h or less. At a design speed of 120 km/h or higher, a maximum tangent length of 1 200 to 2 400 metres would clearly be meaningless.

As an example of the application of the rule of thumb, a design speed of 80 km/h would suggest that the maximum length of tangent should be in the range of 800 to 1 600 metres. In this range, drivers would tend to maintain a fairly constant speed of about 80 km/h. At greater lengths of tangent, drivers would accelerate to some higher speed only to decelerate at the following curve. This oscillation in speed is inherently dangerous as discussed in more detail below.

Consistency of design dictates that:

- The difference between design speed and 85th percentile speed, and
- Variations in 85th percentile speeds between successive elements

should be limited as far as possible. Research has indicated that, ideally, these differences and variations should be less than 10 km/h but that an acceptable design still results if they are less than 20 km/h. Differences in excess of 20 km/h constitute poor design so that, at a greater level of precision than the proposed rule of thumb, three possibilities arise. These are:

- Case 1 - The length of tangent is less than or equal to the distance, T_{min}, required to accelerate from the operat ing peed appropriate to one curve to that appropriate to the next curve;
- Case 2 - The length of tangent allows acceleration to a speed higher than that appropriate to the next curve but not as high as the desired speed remote from inhibiting curves, and
- Case 3 - The length of tangent, T_{max}, allows acceleration up to desired speed. These cases are intended to be applied in areas where curves are to design speeds of 100 km/h or less.

To apply the guideline values of allowable speed variations and differences, it is necessary to estimate the operating speeds on the curves preceding and following a tangent. In the absence of information specific to South African conditions, the international value of V_{85} given in Equation 4.1 will have to be employed.

$$V_{85} = 105{,}31 + 1{,}62 \times 10^{-5} \times B^2 - 0{,}064 \times B \quad (4.1)$$

where	V_{85}	=	85th percentile speed (km/h)
	B	=	Bendiness
		=	57 300/R (degrees/km)
	θ	=	Deviation angle
and	L	=	Total length of curve (metres).

The general form for bendiness, B, in the case where the circular curve is bounded by transition curves is

$$B = \frac{(L_{cl1}/2R + L_{cr}/R + L_{cl2}/2R) \times 180/\pi \times 1000}{L} \quad (4.2)$$

where $L = L_{Cl1} + L_{Cr} + L_{Cl2}$

with L_{Cl1} and L_{Cl2} being the lengths of the preceding and succeeding transition curves and L_{Cr} the length of the circular curve.

The length of tangents is to be compared with the values of Tmin and Tmax that, as suggested, are a function of the speeds achievable on the curves preceding and following these tangents. The values are calculated using an acceleration or deceleration rate of 0,85 m/s^2 (determined by car-following techniques and which also corresponds to deceleration without braking)

$$T_{MIN} = \frac{V85_1^2 - V85_2^2}{22{,}03} \qquad (4.3)$$

and

$$T_{Max} = \frac{2(V85T_{max})^2-(V85_1)^2-(V85_2)^2}{22{,}03} \qquad (4.4)$$

where $V85T_{max}$ = 85th percentile speed on long tangent, i.e. V85, (km/h)

$V85_1$ = 85th percentile speed on preceding curve (km/h)

$V85_2$ = 85th percentile speed on following curve (km/h)

A tangent has a bendiness of zero so that V85 for T_{Max} is, according to Eq. 4.1, 105,31 km/h. South African research has derived an expression for average speed as given in Eq. 4.5.

□C = 143,96 - 10,39 ln Q - 0,04 (G2 - 5,20) - 18,08 D -33,89 PT - 54,15 PS (4.5)

where □C = average passenger car speed (km/h)

Q = flow (veh/h)

G = gradient (per cent)

D = directional split as a decimal fraction

PT = number of trucks in the traffic stream as a dec imal fraction

PS = number of semi trailers in the traffic stream as a decimal fraction.

In the absence of significant volumes of traffic and on a level grade, the average speed would, according to this equation, be of the order of 120 km/h, suggesting that the 85th percentile speed estimated by Eq. 4.1 is very conservative.

If the tangent length is shorter than T_{Min} the tangent is non-independent and it is only necessary for the operating speeds of the two adjacent curves to be within the difference ranges described above to constitute good or tolerable design. In essence, acceleration to the operating speed of the following curve could take place on this curve itself. Where it is necessary to decelerate to the operating speed of the following curve, it will be necessary for the driver to brake in order to achieve the appropriate speed at the start of the curve. It follows that, on a two-lane two-way road, tangent lengths shorter than T_{Min} are potentially dangerous.

Where the tangent length is just equal to T_{Max} the vehicle will be able to accelerate from the operating speed of the preceding curve to the desired speed and then immediately decelerate to the operating speed of the following curve. In this case, the difference between operating speeds on each curve and the desired speed has to be within the allowable range.

In the case of a tangent length falling in the range $T_{Min} < T < T_{Max}$, it will be necessary to calculate the highest operating speed that can be reached by accelerating at a rate of 0,85 m/s^2 from the operating speed of the first curve, allowing for a deceleration at the same rate to the operating speed of the second. It is the difference between this maximum operating speed and the speeds on the adjacent curves that is critical. The maximum operating speed on a tangent of this length is calculated as

$$V85 = [11{,}016(T - T_{Min}) + V85_1^{\ 2}]^{0,5} \quad \text{for } V85_1 > V85_2 \qquad (4.6)$$

4.2.3 Curves

Over the years, various theoreticians have proposed a variety of polynomials as the most desirable forms of horizontal curvature, with desirability presumably being determined by the aesthetics of the end resultant and usually from a vantage point not normally available to the driver. Accident history suggests, however, that drivers have enough difficulty in negotiating simple circular curves that have the property of providing a constant rate of change of bearing. It is recommended that anything more complex than circular arcs should be avoided, the most noteworthy exception being the loop ramp on interchanges.

In the preceding section, relationships between operating speed and degree of curvature were offered, as were acceptable differences between the operating speeds observed on succeeding curves. These define, in effect, the relative design domain of horizontal curvature on any given road, i.e. the possible range of values of radius of any curve given, the radius of the preceding curve.

The safety of any curve is dictated not only by the external factors described above but also by factors internal to it, namely radius, superelevation, transitions and curve widening. Of these factors, the most significant is radius as research carried out in Washington State shows consistently that crash frequency increases as the curve radius decreases. At present, the best model shows that

$$A = (0{,}96\,L + 0{,}0245/R - 0{,}012S)\,0{,}978^{(3.3 \times W - 30)} \qquad (4.7)$$

where A = crashes/million vehicles entering from both directions

L = curve length (km)

R = curve radius (km)

S = 1, if transition curves have been provided

= 0, otherwise

W = roadway width (lanes plus shoulders) (m).

Using this relationship, the designer would be able to estimate the merits of increasing the radius of a curve. This would presumably be of great benefit in the case where an existing road is to be upgraded or rehabilitated.

It is necessary to determine the absolute boundaries of the design domain. The upper bound is obviously the tangent in the sense that it has a radius of infinite length. The lower bound is the minimum radius for the selected design speed and this is a function of the centripetal force necessary to sustain travel along a circular path. This force is developed in part by friction between the vehicle's tyres and the road surface and in part by the superelevation provided on the curve. The Newtonian dynamics of the situation is illustrated in Figure 4.1.

The side friction factor is a function of the condition of the vehicle tyres and the road surface and varies also with speed. For the purposes of design, it is desirable to select a value lower than the limit at which skidding is likely to occur and the international general practice is to select values related to the onset of feelings of discomfort. Canadian practice suggests that the

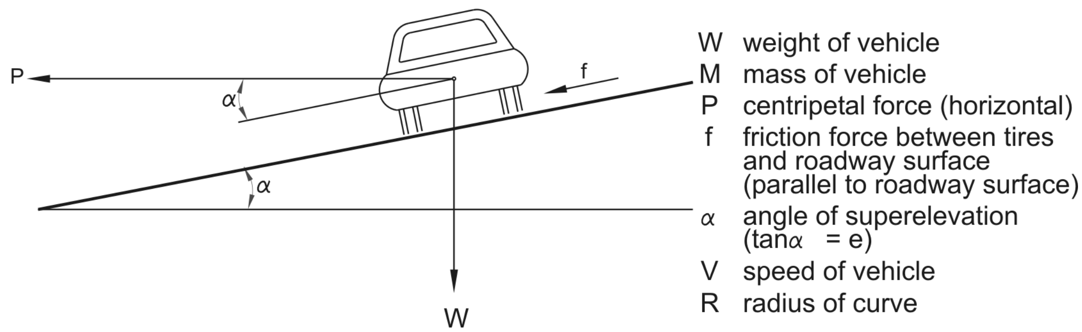


Figure 4.1: Dynamics of a vehicle on a curve

The relationship between speed, radius, lateral friction and superelevation is expressed by the relationship:

$$e + f = V^2/127\,R \qquad (4.8)$$

where e = superelevation(taken as positive when the slope is downward towards the centre of the curve)

f = lateral or side friction factor

V = speed of vehicle (km/h)

R = radius of curvature (m).

This equation is used to determine the minimum radius of curvature that can be traversed at any given speed.

side friction factor be taken as

$$f = 0{,}21 - 0{,}001 \times V \qquad (4.9)$$

where V = vehicle speed (km/h).

For any given speed, it is thus only necessary to select the maximum rate of superelevation, e_{max}, in order to determine the minimum allowable radius of horizontal curvature for that speed. This selection is based on considerations of the design domain as discussed in the following section. In practice, four values of emax are used, being 4, 6, 8, and 10 per cent. The minimum radius of curvature appropriate to design speeds in the range of 40 km/h to 130 km/h for each of these values of e_{max} is given in Table 4.1.

Table 4.1: Minimum radii for various values of e_{max} (m)				
Design speed (km/h)	4 %	6 %	8 %	10 %
40	60	55	50	50
50	100	90	80	80
60	150	130	120	110
70	200	190	170	150
80	280	250	230	210
90	380	340	300	280
100	490	440	390	360
110	680	600	530	480
120	870	750	670	600
130	1100	950	830	740

Guidelines are offered in the following section for the selection of e_{max}

4.2.4 Superelevation

The selection of the appropriate value of e_{max} is at the discretion of the designer in terms of the design domain concept. The higher values of e_{max} are typically applied to rural areas and the lower values to the urban environment.

The spatial constraints in urban areas will very often preclude the development of high values of superelevation. Because of congestion and the application of traffic control devices, the speeds achieved at any point along the road can fluctuate between zero and the posted speed - or even higher depending on the local level of law enforcement. Negotiating a curve with a superelevation of 10 per cent at a crawl speed can present a major problem to the driver. As a general rule, urban superelevations should not exceed 6 per cent although, in the case of an arterial, this could be taken as high as 8 per cent, provided that this value of superelevation is used only between intersections and that the superelevation is sufficiently remote from the intersections for full run-off to be achieved prior to reaching the intersection area.

In rural areas, the range of observed speeds is relatively limited and adequate distance to allow for superelevation development and runoff is usually available. Climatic conditions may, however, impose limitations on the maximum value of superelevation that can be applied. Icing of the road surface is not a typical manifestation of the South African climate but has been known to occur in various high-lying parts of the country. Heavy rainfall reduces the available side friction and relatively light rain after a long dry spell also reduces side friction. This applies particularly to areas where the road surface is polluted by rubber and oil spills, as is the case in urban areas and the immediately surrounding rural areas. Where any of these circumstances are likely to occur, a lower value of e_{max} is recommended.

A lower value of e_{max} should also be considered in a road where steep gradients occur with any frequency. A superelevation of 10 per cent would present trucks with some difficulties when they are climbing a steep grade at low speeds. As shown in Table 7.3, the combination of a superelevation of 10 per cent and a gradient of 8 per cent has a resultant of 12,8 per cent at approximately 45^{O} to the centreline.

Whatever the value selected for e_{max}, this value should be consistently applied on a regional basis. Its selection governs the rate of superelevation applied to all radii above the minimum. Variations in e_{max} result in curves of equal radius having different rates of superelevation. Drivers select their approach speeds to curves on the basis of the radius that they see and not on the degree of superelevation provided. A lack of consistency with regard to superelevation would almost certainly lead to differences in side friction demand with possibly critical consequences. Recommended rates of e_{max} are offered in Table 4.2.

Table 4.2: Design domain for e_{max}

Domain	Suggested range for e_{max}
Rural roads	8 % to 10 %
High-speed urban roads	6 % to 8 %
Minor urban roads	4 % to 6 %

Distribution of e and f

There are a number of methods of distributing e and f over a range of curves flatter than the minimum for a given design speed. Five methods are well documented by AASHTO. These are:

- Method 1: Both superelevation and side friction are directly proportional to the inverse of the radius;
- Method 2: Side friction is first applied to sustain lateral acceleration down to radii requiring f_{max} followed by increasing e with reducing radius until e reaches e_{max}. In short, first f and then e are increased in inverse proportion to the radius of curvature;
- Method 3: Effectively the reverse of Method 2 with first e and then f increased in inverse proportion to the radius of curvature;
- Method 4: As for Method 3, except that design speed is replaced by average running speed, and
- Method 5: Superelevation and side friction are in curvilinear relations with the inverse of the radius of curvature, with values between those of Methods 1 and 3.

These methods of distribution are illustrated in Figure 4.2.

In terms of the design domain concept, Method 2 has merit in the urban environment. As pointed out earlier, provision of adequate superelevation in an environment abounding in constraints such as closely spaced intersections and driveways is problematic. It is thus sensible to make as much use as possible of side friction before having to resort to the application of superelevation. It also should be noted that drivers operating at relatively low speeds in an urban environment are prepared to accept higher values of side friction than they would at high speeds on a rural road.

Method 5 is recommended for adoption in the case of rural and high-speed urban roads. In practice it represents a compromise between Methods 1 and 4. The tendency for flat to intermediate curves to be overdriven is accommodated by the provision of some superelevation. The superelevation provided sustains nearly all lateral acceleration at running speeds (assumed to be about 80 per cent of design speed) with considerable side friction available for greater speeds. On the other hand, Method 1, which avoids the use of maximum superelevation for a substantial part of the range of curve radii, is also desirable. Method 5 has an unsymmetrical parabolic form and represents a practical distribution for superelevation over the range of curvature.

Tables 4.3, 4.4, 4,5 and 4,6 provide values of e for a range of horizontal radii and values of e_{max} of 4, 6, 8, and 10 per cent respectively.

Superelevation runoff

In the case of a two-lane road, superelevation runoff (or runout) refers to the process of rotating the outside lane from zero crossfall to reverse camber (RC) thereafter rotating both lanes to full superelevation. Tangent (or crown) runoff refers to rotation of the outside lane from

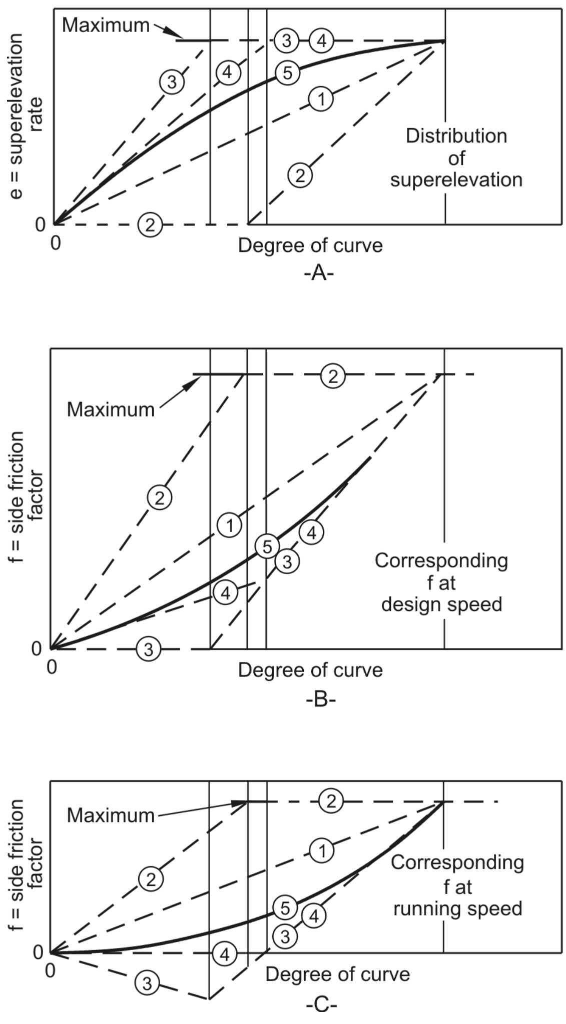


Figure 4.2: Methods of distributing of e and f

zero crossfall to normal camber (NC). Rotation is typically around the centreline of the road although constraints, such as driveway entrances or drainage in the urban environment, may require rotation to be around the inside or outside edge. These latter alternatives result in the distortion of the vertical alignment of the road centreline and a severe slope on the road edge being rotated, with a potentially unaesthetic end result. In the case of dual carriageway cross-sections, rotation is typically around the outer edges of the median island.

The designer should be sensitive to the fact that zero crossfall implies a lack of transverse drainage resulting in the possibility of storm water ponding on the road surface. A further cause of ponding could be where the centreline gradient is positive and equal to the relative gradient. In this case, the inner edge of the road would have zero gradient over the entire length of the superelevation runoff.

Ponding is extremely dangerous for two reasons. The more obvious danger is that it can cause a vehicle to hydroplane, causing a total loss of traction and steering ability. If the front wheels are pointing in any direction other than straight ahead when the vehicle moves out of the ponded water, the sudden availability of fric-

Table 4.3: Values of superelevation for above min radii of curvature (%): e_{max} = 4 %

RADIUS (m)	DESIGN SPEED (km/h)									
	40	50	60	70	80	90	100	110	120	130
7000	NC	NC	NC	NC	NC	NC	NC	A four per cent superelevation is appropriate only to low speed urban roads. Design speeds in excess of 100 km/h are not suitable for this design domain.		
5000	NC	NC	NC	NC	NC	NC	NC			
4000	NC	NC	NC	NC	NC	NC	NC			
3000	NC	NC	NC	NC	NC	NC	RC			
2000	NC	NC	NC	NC	RC	RC	RC			
1500	NC	NC	NC	RC	RC	RC	2,2			
1400	NC	NC	NC	RC	RC	RC	2,3			
1300	NC	NC	NC	RC	RC	2,0	2,4			
1200	NC	NC	NC	RC	RC	2,1	2,6			
1000	NC	NC	RC	RC	2,0	2,5	3,0			
900	NC	NC	RC	RC	2,1	2,7	3,2			
800	NC	NC	RC	RC	2,3	2,9	3,4			
700	NC	RC	RC	2,0	2,6	3,2	3,6			
600	NC	RC	RC	2,3	2,9	3,5	3,9			
500	NC	RC	2,1	2,6	3,3	3,8	4,0			
400	RC	RC	2,5	3,1	3,7	4,0				
300	RC	2,3	3,1	3,6	4,0					
250	RC	2,6	3,4	3,8						
200	2,1	3,1	3,8	4,0						
180	2,3	3,3	3,9							
160	2,5	3,5	4,0							
140	2,8	3,7								
120	3,1	3,9								
100	3,4	4,0								
90	3,6									
80	3,8									
70	4,0									
60	4,0									
50										

Note: NC denotes Normal Camber, i.e. 2 per cent fall from the centreline to either edge of the travelled way

RC denotes Reverse Camber, i.e. 2 per cent crossfall from the outer edge of the travelled way to the inner edge

tion can lead to a sharp swerve and subsequent loss of control. The other possibility is that of one front wheel striking the water before the other, in which case the unbalanced drag could also lead to the vehicle swerving out of control. The designer should therefore endeavour to avoid the combination of zero longitudinal gradient and zero crossfall. A pond depth of 15 mm is sufficient to cause hydroplaning. In the case of worn tyres, a lesser depth will suffice.

The length of the superelevation runoff section is selected purely on the basis of appearance, which suggests that there is a maximum acceptable difference between the gradients of the axis of rotation and the pavement edge. Experience indicates that relative gradients of 0,8 and 0,35 per cent provide acceptable runoff lengths for design speeds of 20 km/h and 130 km/h respectively. Interpolation between these values provides the relative gradients shown in Table 4.7.

Many States of the United States have opted for a standard relative gradient of 1:200, whereas Canada has elected to use a relative gradient of

Table 4.4: Values of superelevation for above min radii of curvature (%): e_{max} = 6 %

RADIUS (m)	DESIGN SPEED (km/h)									
	40	50	60	70	80	90	100	110	120	130
7000	NC	NC	NC	NC	NC	NC	NC	NC	RC	RC
5000	NC	NC	NC	NC	NC	NC	NC	RC	RC	RC
4000	NC	NC	NC	NC	NC	NC	RC	RC	RC	2,3
3000	NC	NC	NC	NC	NC	RC	RC	2,0	2,4	3,6
2000	NC	NC	NC	RC	RC	2,3	2,2	2,9	3,4	4,0
1500	NC	NC	RC	RC	2,4	2,9	2,9	3,6	4,2	4,9
1400	NC	NC	RC	RC	2,5	3,1	3,1	3,8	4,4	5,1
1300	NC	NC	RC	2,1	2,6	3,2	3,3	4,1	4,6	5,3
1200	NC	NC	RC	2,3	2,8	3,3	3,4	4,3	4,9	5,5
1000	NC	RC	2,1	2,7	3,2	3,7	4,0	4,8	5,4	5,8
900	NC	RC	2,3	2,9	3,4	3,9	4,3	5,1	5,7	6,0
800	NC	RC	2,5	3,1	3,6	4,2	4,6	5,4	5,9	
700	NC	2,1	2,7	3,4	3,9	4,5	5,0	5,8	6,0	
600	NC	2,4	3,0	3,7	4,2	4,8	5,4	6,0		
500	RC	2,7	3,4	4,1	4,6	5,2	5,9			
400	2,3	3,1	3,8	4,5	5,1	5,7	6,0			
300	2,8	3,7	4,4	5,1	5,7	5,9				
250	3,1	4,0	4,8	5,5	6,0	6,0				
200	3,6	4,5	5,2	5,9						
180	3,8	4,7	5,4	6,0						
160	4,0	4,9	5,6							
140	4,3	5,2	5,9							
120	4,6	5,5	6,0							
100	4,9	5,8								
90	5,1	6,0								
80	5,4									
70	5,6									
60	5,9									
50	6,0									

Note: NC denotes Normal Camber, i.e. 2 per cent fall from the centreline to either edge of the travelled way

RC denotes Reverse Camber, i.e. 2 per cent crossfall from the outer edge of the travelled way to the inner edge

Table 4.5: Values of superelevation for above min radii of curvature (%): e_{max} = 8 %

RADIUS (m)	DESIGN SPEED (km/h) 40	50	60	70	80	90	100	110	120	130
7000	NC	NC	NC	NC	NC	NC	NC	RC	RC	RC
5000	NC	NC	NC	NC	NC	NC	RC	RC	RC	2,1
4000	NC	NC	NC	NC	NC	NC	RC	RC	2,1	2,6
3000	NC	NC	NC	NC	NC	RC	RC	2,3	2,8	3,3
2000	NC	NC	NC	RC	2,1	2,6	2,6	3,3	4,0	4,7
1500	NC	NC	RC	2,1	2,7	3,2	3,3	4,2	5,0	5,8
1400	NC	NC	RC	2,2	2,8	3,4	3,5	4,5	5,3	6,1
1300	NC	NC	RC	2,4	3,0	3,6	3,7	4,7	5,6	6,4
1200	NC	NC	RC	2,6	3,2	3,8	4,0	5,0	5,9	6,7
1000	NC	RC	2,3	2,9	3,6	4,3	4,7	5,7	6,6	7,4
900	NC	RC	2,5	3,2	3,9	4,6	5,1	6,2	7,1	7,8
800	NC	2,0	2,7	3,5	4,2	4,9	5,5	6,6	7,5	8,0
700	NC	2,3	3,0	3,8	4,6	5,3	6,1	7,2	7,9	
600	RC	2,6	3,4	4,2	5,0	5,8	6,7	7,7	8,0	
500	2,1	3,0	3,9	4,8	5,6	6,4	7,3	8,0		
400	2,5	3,5	4,5	5,4	6,3	7,1	8,0			
300	3,1	4,2	5,3	6,3	7,2	8,0				
250	3,5	4,7	5,9	6,9	7,8					
200	3,9	5,4	6,5	7,5	8,0					
180	4,4	5,7	6,8	7,8						
160	4,7	6,0	7,2	8,0						
140	5,1	6,4	7,6							
120	5,5	6,9	8,0							
100	6,1	7,4								
90	6,4	7,7								
80	6,7	8,0								
70	7,1									
60	7,5									
50	8,0									

Table 4.6: Values of superelevation for above minradii of curvature (%): e_{max} = 10 %

RADIUS (m)	DESIGN SPEED (km/h) 40	50	60	70	80	90	100	110	120	130
7000	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC
5000	NC	NC	NC	NC	NC	NC	NC	NC	RC	RC
4000	NC	NC	NC	NC	NC	NC	RC	RC	RC	2,4
3000	NC	NC	NC	NC	NC	RC	2,4	2,1	2,5	2,7
2000	NC	NC	NC	NC	RC	2,2	2,7	3,1	3,6	4,0
1500	NC	NC	NC	RC	2,4	2,9	3,5	4,1	4,8	5,3
1400	NC	NC	RC	2,1	2,6	3,1	3,8	4,3	5,1	5,7
1300	NC	NC	RC	2,3	2,8	3,3	4,0	4,6	5,5	6,1
1200	NC	NC	RC	2,4	3,0	3,6	4,3	5,0	5,9	6,6
1000	NC	RC	2,2	2,9	3,5	4,2	5,1	5,9	7,0	7,9
900	NC	RC	2,5	3,2	3,9	4,6	5,6	6,4	7,7	8,7
800	NC	RC	2,7	3,5	4,3	5,1	6,2	7,1	8,5	9,7
700	RC	2,3	3,1	4,0	4,8	5,8	6,9	8,0	9,5	10,0
600	RC	2,7	3,6	4,5	5,5	6,5	7,8	9,0	10	
500	2,3	3,1	4,2	5,3	6,4	7,6	8,9	10,0		
400	2,8	3,8	5,0	6,3	7,5	8,8	9,8			
300	3,6	4,8	6,3	7,8	9,0	9,9	10,0			
250	4,2	5,6	7,1	8,7	9,7	10,0				
200	5,0	6,6	8,2	9,6	10,0					
180	5,5	7,1	8,7	9,9						
160	6,0	7,6	9,2	10,0						
140	6,4	8,1	9,7							
120	7,0	8,8	10,0							
100	7,7	9,5								
90	8,2	9,8								
80	8,6	10,0								
70	9,1									
60	9,6									
50	10,0									

Note: NC denotes Normal Camber, i.e. 2 per cent fall from the centreline to either edge of the travelled way

RC denotes Reverse Camber, i.e. 2 per cent crossfall from the outer edge of the travelled way to the inner edge

Table 4.7: Maximum relative gradients		
Design speed (km/h)	Maximum relative gradient (%)	Equivalent maximum relative slope 1:
40	0,72	140
50	0,68	147
60	0,64	156
70	0,60	167
80	0,56	179
90	0,52	192
100	0,48	208
110	0,44	227
120	0,40	250
130	0,35	286

1:400 in the calculation of the length of the superelevation runoff. Other widely used options include adopting the distance travelled in four seconds and previous editions of the AASHTO policy suggested the distance travelled in 2 seconds. As can be seen, there is a large degree of arbitrariness attaching to determination of the length of superelevation runoff. The designer can thus vary the relative gradient to accommodate other elements of the design, such as the distance between successive curves or the distance to the following intersection. It is, however, suggested that the relative gradients offered in Table 4.7 should provide a pleasing appearance and the designer should at least attempt to achieve relative gradients of a similar magnitude.

The gradient of the tangent runout is simply a continuation of whatever relative gradient was adopted for the superelevation runoff.

If the relative gradient approach to determination of runoff length is adopted, this length is calculated as

$$L = \frac{(wn)e_d}{\Delta} . b \qquad 4.10$$

where

L	=	length of superelevation runoff (m)
w	=	width of one traffic lane, (m)
n	=	number of lanes rotated
e_d	=	superelevation rate (per cent)
Δ	=	relative gradient, (per cent)
b	=	adjustment factor for number of lanes rotated

Adjustment factor for number of lanes

If the above relationship is applied to cross-sections wider than two lanes, the length of the

Table 4.8: Lane adjustment factors		
Number of lanes rotated, n	Adjustment factor, b	Length increase relative to one lane rotated
1	1	1,00
1,5	0,83	1,25
2	0,75	1,50
2,5	0,70	1,75
3	0,67	2,00
3,5	0,64	2,25
For other values of n, use the equation b = [1 + 0,5(n-1)]/n		

superelevation runoff could double or treble and there may simply not be enough space to allow for these lengths. On a purely empirical basis, it is recommended that the calculated lengths be adjusted downwards by the lane adjustment factors offered in Table 4.8.

Location of superelevation runoff

The two extremes of runoff location are:

- Full superelevation attained at the beginning of the curve (BC), and
- Only tangent runout attained at the BC.

Both alternatives result in high values of lateral acceleration and are thus considered undesirable. The preferred option would be to have a portion of the runoff located on the tangent and the balance on the curve. Experience has shown that having about 2/3 of the runoff on the tangent produces the best result in terms of limiting lateral acceleration. If circumstances demand, deviation by about 10 per cent from this ratio is tolerable.

The superelevation runoff and tangent runout are illustrated in Figure 4.3. Figures 4.4 and 4.5 show possible treatments for superelevation runoff on reverse and broken back curves. In these cases, the superelevation runoff terminates at a crossfall of two per cent rather than the more customary zero camber on the outside lane.

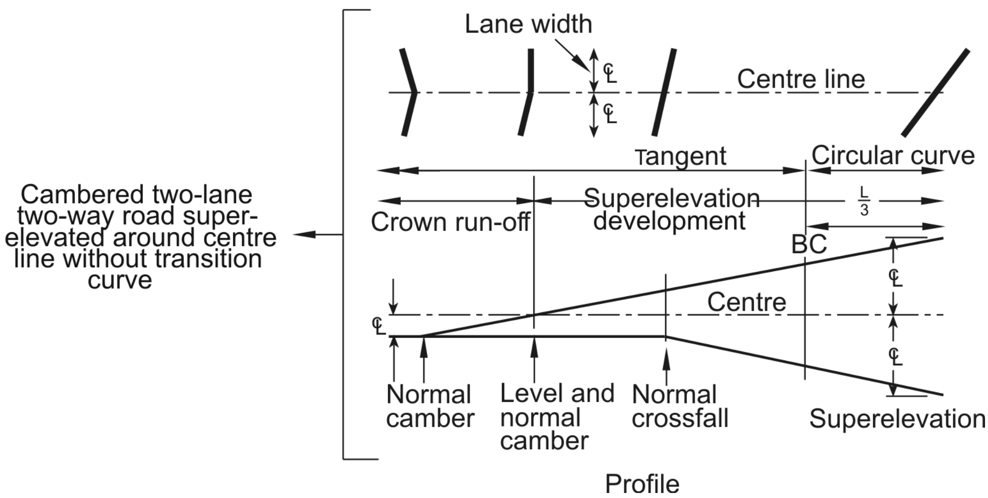


Figure 4.3: Attainment of superelevation

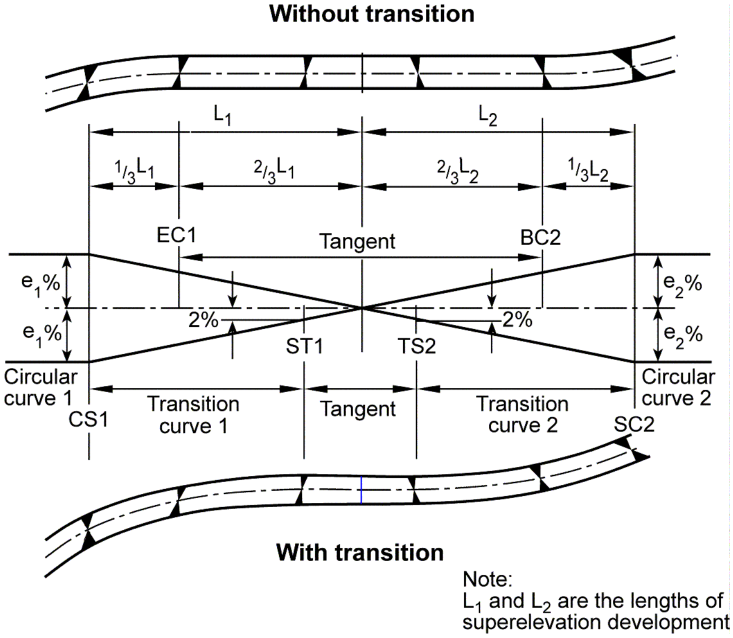


Figure 4.4: Superelevation runoff on reverse curves

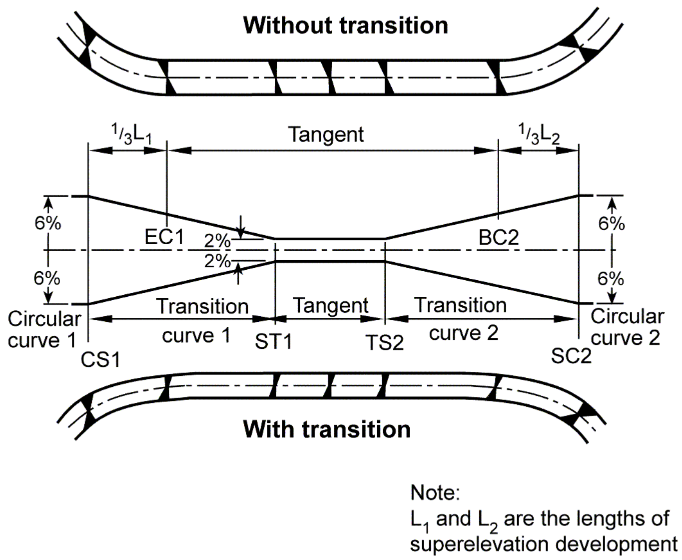


Figure 4.5: Superelevation runoff on broken-back curves

If the circular curve is preceded by a transition curve, all of the superelevation runoff should be located on the transition curve.

4.2.5 Transition curves

Any vehicle entering a circular curve does so by following a spiral path. For most curves, this transition can be accommodated within the limits of normal lane width. At minimum radii for the design speed, longer transition paths are followed and, if these occur on narrow lanes, the shift in lateral position may even lead to encroachment on adjacent lanes. Under these circumstances, it may be convenient to shape the horizontal alignment such that it more accurately reflects the path actually followed by a vehicle entering the circular curve.

Various curves can be used to provide a transition from the tangent to the circular curve. Whatever form is used, it should satisfy the conditions that:

- It is tangential to the straight;
- Its curvature should be zero (i.e. infinite radius) on the straight;
- The curvature should increase (i.e. radius decrease) along the transition;
- Its length should be such that, at its junction with the circular curve, the full super elevation has been attained;
- It should join the circular arc tangentially, and
- The radius at the end of the transition should be the same as that of the circular curve.

Candidate curves are the lemniscate, the cubic parabola (also known as Froude's spiral), and the clothoid (also known as Euler's spiral). The cubic parabola achieves a maximum value and then flattens out again and is thus not a true spiral and the lemniscate requires an unacceptable length of arc to achieve the desired radius. The clothoid, which has the relationship whereby the radius, R, at any point on the spiral varies with the reciprocal of the distance, L, from the start of the spiral, is thus the preferred option. Expressed mathematically, this relationship is

$$R = A^2/L \qquad (4.11)$$

where A is a constant called the spiral parameter and has units of length.

The length of a transition curve may be based on one of three criteria. These are:

- Rate of change of centripetal acceleration, essentially a comfort factor, varying between 0,4 m/s^3 and 1,3 m/s^3;
- Relative slope as proposed in Table 4.3; or
- Aesthetics.

Since relative slope is applied to curves where transitions are not provided, its use also for transitioned curves would be a sensible point of departure. For the various design speeds, a radius corresponding to a specified centripetal acceleration can be calculated. These radii are listed in Table 4.9 for an acceleration of 1,3 m/s^2. There is little point in applying transition curves to larger radii where the centripetal acceleration would be lower.

Setting out of transitions

Application of the spiral has the effect that the circular curve has to be offset towards its centre. It is thus located between new tangents that are parallel to the original tangents but shifted from

Table 4.9: Maximum radii for use in spiral transition curves	
Design Speed (km/h)	Maximum Radius (m)
40	100
50	150
60	210
70	290
80	380
90	480
100	590
110	720
120	850
130	1000

them by an amount, s, known as the shift. The value of the shift is given by

$$s = L^2/2R \qquad (4.12)$$

where L = selected length of tran sition curve (m)

R = radius of circular curve (m)

The starting point of the spiral is located at a distance, T, from the Point of Intersection (PI) of the original tangents with

$$T = (R+s) \tan \theta/2 + L/2 \qquad (4.13)$$

where θ = deviation angle of the circular curve

The most convenient way to set out the spiral is by means of deflection angles and chords and the deflection angle for any chord length, l, is given as

$$a = l^2/6RL \times 57{,}246 \qquad (4.14)$$

4.2.6 Lane widening

When vehicles negotiate a horizontal curve, the rear wheels track inside the front wheels. In the case of semi trailers with multiple axles and pivot points, this off-tracking is particularly marked. The track width of a turning vehicle, also known as the swept path width, is the sum of the track width on tangent and the extent of off-tracking, with the off-tracking being a function of the radius of the turn, the number and location of pivot points and the length of the wheelbase between axles. The track width is calculated as

$$U = u + R - (R^2 - \Sigma L_i^2)^{0,5} \qquad (4.15)$$

where U = track width on curve (m)

u = track width on tangent (m)

R = radius of turn (m)

L_i = wheel base of design vehicle between successive axles and pivot points (m).

Strictly speaking, the radius, R, should be the radius of the path of the midpoint of the front axle. For ease of calculation, however, the radius assumed is that of the road centreline.

The front overhang is the distance from the front axle of the vehicle to the furthest projection of the vehicle body in front of the front axle. In the case of the turning vehicle, the width of the front overhang is defined as the radial distance between the path followed by the outer front edge of the vehicle and the tyre path of the outer front wheel. The width of the front overhang is calculated as

$$F_A = [R^2 + A(2L + A)]^{0,5} - R \quad (4.16)$$

where F_A = width of front overhang (m)

R = radius of curve (m)

A = front overhang (m)

L = wheel base of single unit or tractor (m)

The width of the rear overhang is the radial distance between the outside edge of the inner rearmost tyre and the inside edge of the vehicle body. In the case of a passenger car this dis-

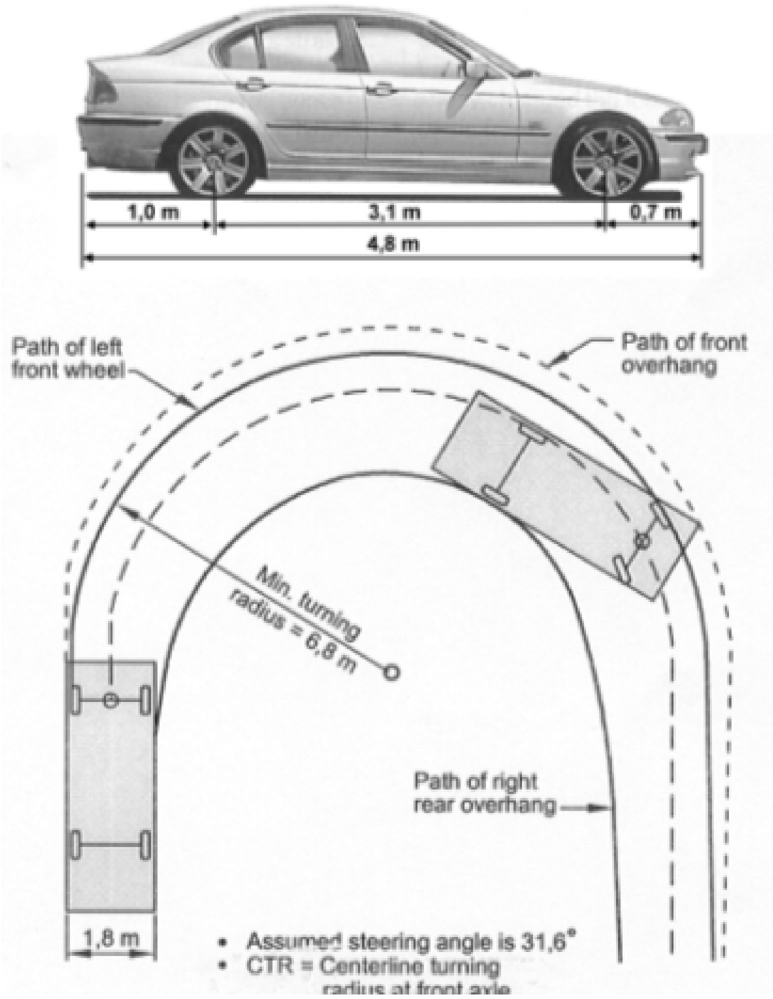


Figure 4.6: Typical turning path

tance is typically less than 0,15 m. The width of truck bodies is usually the same as the wheel-base width so that the width of the rear over-hang is zero.

A typical turning path is illustrated in Figure 4.6. Turning paths for numerous vehicles are provid-ed in the 2000 edition of the AASHTO Policy on geometric design of highways and streets and the designer is directed towards Exhibits 2-3 to 2-23 of that document.

It is necessary to provide an allowance, C, for lateral clearance between the edge of the road-way and the nearest wheel path, and for the body clearance between passing vehicles. Typical values of C are:

- 0,60 m for a travelled way width of 6,0 m;
- 0,75m for a travelled way width of 6,6 m, and
- 0,90 m for a travelled way width of 7,4 m.

A further allowance, Z, is provided to accommo-date the difficulty of manoeuvring on a curve and the variation in driver operation. This addi-tional width is an empirical value that varies with the speed of traffic and the radius of the curve. It is expressed as

$$Z = 0,1(V/R^{0,5}) \qquad (4.17)$$

where V = design speed of the road (km/h)

By combining Eqs 4.12, 4.13 and 4.14 with the clearance allowances, C and Z, the width of the travelled way can be calculated as

$$W_C = N(U + C) + F_A(N - 1) + Z \qquad (4.18)$$

where N = number of lanes

and the other variables are as previously defined.

As a general rule, values of curve widening, being (W_C - W) where W is the width of the trav-elled way on tangent sections, that are less than 0,6 m are disregarded. Lane widening is thus generally not applied to curves with a radius greater than 300 metres, regardless of the design speed or the lane width.

Widening should transition gradually on the approaches to the curve so that the full addi-tional width is available at the start of the curve. Although a long transition is desirable to ensure that the whole of the travelled way is fully usable, this results in narrow pavement slivers that are difficult, and correspondingly expen-sive, to construct. In practice, curve widening is thus applied over no more than the length of the super elevation runoff preceding the curve. For ease of construction, the widening is normally applied only on one side of the road. This is usually on the inside of the curve to match the tendency for drivers to cut the inside edge of the travelled way.

In terms of usefulness and aesthetics, a tangent transition edge should be avoided. A smooth graceful curve is the preferred option and can be adequately achieved by staking a curved transition by eye. Whichever approach is used, the transition ends should avoid an angular break at the pavement edge.

Widening is provided to make driving on a curve comparable with that on a tangent. On older roads with narrow cross-sections and low design speeds and hence sharp curves, there was a considerable need for widening on curves. Because of the inconvenience attached to widening the surfacing of a lane, it follows that the required widening may not always have been provided. Where a road has to be rehabilitated and it is not possible to increase the radius of curvature, the designer should consider the need for curve widening.

In the case of an alignment where curves in need of widening of the travelled way follow each other in quick succession, the inconvenience associated with the application of curve widening can be avoided by constructing the entire section of road, including the intervening tangents, to the additional width.

4.3 VERTICAL ALIGNMENT

Vertical alignment comprises grades (often referred to as tangents) and vertical curves.

Grades have the properties of length and gradient, invariably expressed as a percentage, representing the height in metres gained or lost over a horizontal distance of 100 metres.

Curves may be either circular or parabolic, with South African practice favouring the latter. The practical difference between the two forms is insignificant in terms of actual roadway levels along the centreline. The parabola has the property of providing a constant rate of change of gradient with distance, which is analogous to the horizontal circular curve, which provides a constant rate of change of bearing. It thus has a certain academic appeal.

The general equation of the parabola is
$y = ax^2 + bx + c$,
from which it follows that the gradient, dy/dx, at any point along the curve is expressed as $2ax + b$ and the rate of change of gradient, d^2y/dx^2, is $2a$. This has the meaning of extent of change over a unit distance. Normal usage is to express the rate of change in terms of the distance required to effect a unit change of gradient. This expression is referred to as the K-value of the curve and is equal to $1/2a$. It follows that the length of a vertical curve can be conveniently expressed as being

$$L = A \times K \qquad (4.19)$$

where L = curve length (m)
A = algebraic difference between the gradients on either side of the curve
K = rate of change

4.3.1 General controls for vertical alignment

On rural and high-speed urban roads, a smooth grade line with gradual changes, which are consistent with the class of the road and the character of the terrain, is preferable to an alignment with numerous breaks and short lengths of grades and curvature. A series of successive, relatively sharp crest and sag curves creates a roller coaster or hidden dip profile which is aesthetically unpleasant. Hidden dips can be a safety concern, although, at night, the loom of approaching headlights may provide a visual clue about oncoming vehicles. Such profiles occur on relatively straight horizontal alignments where the road profile closely follows a rolling natural ground line.

A broken-back grade line, which consists of two vertical curves in the same direction with a short length of intervening tangent, is aesthetically unacceptable, particularly in sags where a full view of the profile is possible. A broken plank grade line, where two long grades are connected by a short sag curve, is equally unacceptable. As a general rule, the length of a curve (in metres) should not be shorter than the design speed in km/h. In the case of freeways, the minimum length should not be less than twice the design speed in km/h and, for preference, should be 400 metres or longer to be in scale with the horizontal curvature. The broken-back and broken-plank curves are the vertical counterparts of the horizontal broken-back curve and the long tangent/small radius curve discussed earlier. The only difference between them is that these forms of vertical alignment are, at least, not dangerous.

In theory, vertical curves in opposite directions do not require grades between them. In practice, however, the outcome is visually not successful. The junction between the two curves creates the impression of a sharp step in the alignment, downwards where a sag curve follows a crest and upwards where the crest curve follows the sag. A short length of grade between the two curves will create the impression of continuous, smoothly flowing vertical curvature. The length of the intervening grade in metres need not be more than the design speed in km/h to achieve this effect.

Where the total change of gradient across a vertical curve is very small, e.g. less than 0,5 per cent, the K-value necessary to achieve the minimum length of curve would be high. Under these circumstances, the vertical curve could be omitted altogether without there being an adverse visual impact.

The vertical alignment design should not be carried out in isolation but should be properly coordinated with the horizontal alignment as discussed later. In addition to the controls imposed on the grade line by the horizontal alignment, the drainage of the road may also have a major impact on the vertical alignment. The top of a crest curve and the bottom of a sag imply a zero gradient and the possibility of ponding on the road surface. Where water flow off the road surface is constrained by kerbs, the gradient should be such that longitudinal flow towards drop inlets or breaks in the kerb line is supported. On lower-speed urban roads, drainage design may often control the grade design.

There are, to date, no specific guidelines on consistency of vertical alignment in terms of the relative lengths of grades and values of vertical curvature, as is the case in horizontal alignment. However, where grades and curves are of approximately equal lengths, the general effect of the grade line tends to be pleasing.

4.3.2 Grades

The convention adopted universally is that a gradient that is rising in the direction of increasing stake value is positive and a descending gradient negative. Gradient and length are interlinked in that steep gradients have an adverse effect on truck speeds and hence on the operating characteristics of the entire traffic stream. This effect is not limited to upgrades because truck operators frequently adopt the rule that speeds on downgrades should not exceed those attainable in the reverse direction.

It is desirable that truck speeds should not decrease too markedly. Apart from the operational impact of low truck speeds, it has also been established that there is a strong correlation between crash rates and the speed differential between trucks and passenger cars. American research indicates that crash rates for speed reductions of less than 15 km/h fluctuate between 1 and 5 crashes per million kilometres of travel increasing rapidly to of the order of 21 crashes per million kilometres of travel for a speed reduction of 30 km/h. In the absence of South African research, it is presumed that a similar trend would manifest itself locally, although probably at higher crash rates. For these reasons, reference is made to the "critical length of grade" typically taken as being the distance over which a speed reduction of 15 km/h occurs. For a given gradient, lengths less than the critical length result in acceptable operation in the desired range of speeds. Where the desired freedom of operation is to be maintained on grades longer than the critical length, it will be necessary to consider alleviating measures such as local reductions of gradient or the provision of extra lanes.

Local research indicates that the 85th percentile mass/power ratio is of the order of 185 kg/kW. The performance of the 85th percentile truck is illustrated in Figure 4.7. The critical lengths of grade for a speed reduction of 15 km/h are derived from these performance curves and are shown in Table 4.10.

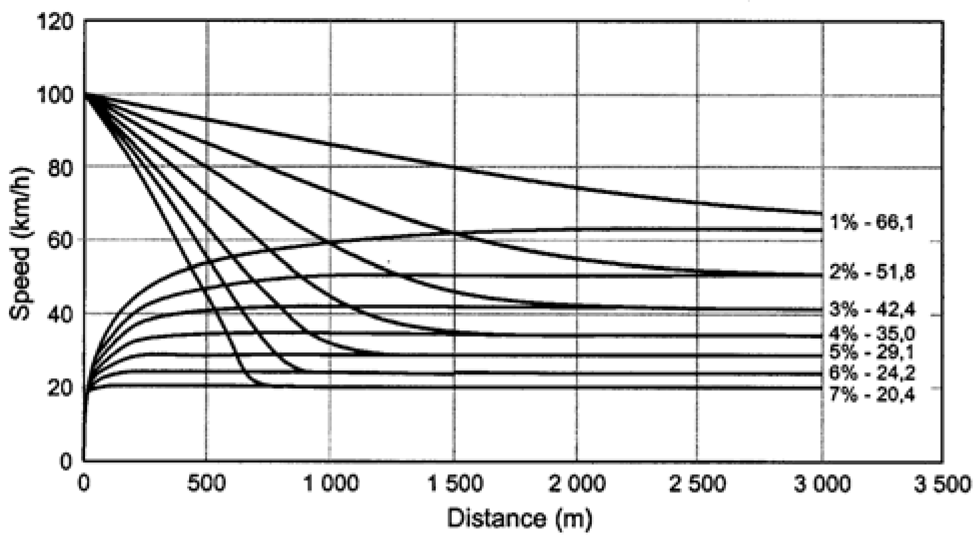


Figure 4.7: Truck speeds on grades

As suggested earlier, grades longer than those given in Table 4.10 may require some form of alleviating treatment. One such treatment is the provision of climbing lanes. This is discussed in

Table 4.10: Lengths of grade for 15 km/h speed reduction	
Gradient (%)	**Critical length of grade (m)**
2	550
3	380
4	300
5	240
6	180
7	140
8	100

Section 4.4.2. Stepping the grade line, i.e. by inserting short sections of flatter gradient, as an alleviating treatment is sometimes offered as relief to heavy trucks at crawl speeds on steep gradients. In practice, this has proved to be ineffective because drivers of heavy trucks simply maintain the crawl speed dictated by the steeper gradient in preference to going through the process of working their way up and down through the gears.

Maximum acceptable gradients shown in Table 4.11 are dictated primarily by the topography and the classification of the road. Topography is described as being flat, rolling or mountainous which is somewhat of a circular definition in the sense that what is really being described is not the topography itself but the gradients that can readily be achieved under the three sets of circumstances. Other factors that should be borne in mind in selecting a maximum gradient include:

- traffic operations, where high volumes would suggest a reduction in maximum gradient in order to maintain an accept able Level of Service;
- costs, being the whole-life cost of the road and not merely its initial construction cost;
- property, where relatively flat gradients in a rugged environment may result in high fills or deep cuts necessitating the acquisition of land additional to the normal road reserve width;
- environmental considerations; and
- adjacent land use in heavily developed or urban areas

Table 4.11: Maximum gradients			
Design speed (km/h)	**Topography**		
	Flat	**Rolling**	**Mountainous**
	Gradients (%)		
60	6	7	8
80	5	6	7
100	4	5	6
120	3	4	5

It is the designer's responsibility to select a maximum gradient appropriate to the project being designed. The values offered in Table 4.11 are thus only intended to provide an indication of gradients appropriate to the various circumstances.

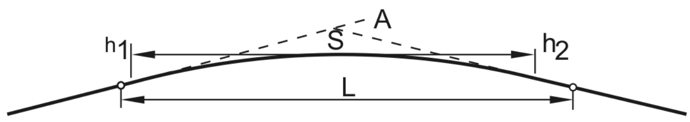
Maximum gradients on freeways should be in the range of three to four per cent regardless of the topography being traversed. Lower order roads have been constructed to gradients as steep as twenty per cent but it is pointed out that compaction with a normal 12/14 tonne roller is virtually impossible on a gradient steeper than about twelve per cent. It is recommended that this be considered the absolute maximum gradient that can be applied to any road.

The minimum gradient can, in theory, be level, i.e. zero per cent. This could only be applied to rural roads where storm water would be removed from the road surface by the camber and allowed to spill over the edge of the shoulder. If used on a road that is kerbed, channel grading would have to be employed. Given the limits of accuracy to which kerbs and channels can be constructed, the flattest gradient that should be considered is of the order of 0,5 per cent. It is recommended that, even without kerbing, gradients should not be less than 0,5 per cent. If the grade is longer than 500 metres, increasing the camber to 2,5 or 3,0 per cent should be considered. The latter value of camber should only be considered in areas subject to heavy rainfall because it may give rise to problems related to steering and maintaining the vehicle's position within its lane.

Crest curves are often in cut and, for the minimum value of K for a design speed of 120 km/h, the gradient would be at a value of less than 0,5 per cent for a distance of 55 metres on either side of the crest. Over this distance, channel grading should be applied to the side drains. On sag curves, the distance over which the longitudinal gradient is less than 0,5 per cent is 26 metres on either side of the lowest point at the minimum K-value.

Varying the camber between 2 per cent and 3 per cent over a distance of about 80 metres will provide an edge grading at 0,5 per cent in the case where the centreline gradient is flat. As an alternative means of achieving adequate

Crest curves



L length of vertical curve in metres
S minimum sight distance in metres
L > S
h_1 height of driver's eye, 1,05 m
h_2 height of object, described in following section
A algebraic difference in grades, percent

$K = \frac{L}{A}$

Figure 4.8: Sight distance on crest curves

drainage, this is more useful in theory than in practice because shaping and compacting the road surface to have this wind in it is extremely difficult. This method is not unknown but is not recommended because of the construction problem.

4.3.3 Curves

As the parameter, K, has been described as the determinant of the shape of the parabolic curve, it follows that some or other value of K can be determined such that it provides adequate sight distance across the length of the curve. Sight distance is measured from the driver eye height, h_1, to a specified object height, h_2. In the case of a crest curve, the line of sight is taken as being a grazing ray to the road surface, as illustrated in Figure 4.7.

Vertical curvature for stopping sight distance

The required value of K is derived from the equation of the parabola as indicated in Equation 4.20.

$$K = \frac{S^2}{200(h_1^{0,5} + h_2^{0,5})^2} \qquad (4.20)$$

where K = Distance required for a 1 % change of gradient (m)

S = Stopping sight distance for selected design speed (m)

h_1 = Driver eye height (m)

h_2 = Object height (m)

A = Algebraic difference in gradient between the approaching and depart ing grades (%).

This relationship applies to the condition whereby the required sight distance is contained within the length of the vertical curve.

If the curve length is shorter than the required sight distance, lesser values of K can be employed as indicated by Equation 4.21.

$$K = \frac{2S}{A} - \frac{200(h_1^{0,5} + h_2^{0,5})^2}{A^2} \qquad (4.21)$$

where K = Distance required for a 1 % change of gradient (m)

S = Stopping sight distance for selected design speed (m)

h_1 = Driver eye height (m)

h_2 = Object height (m)

A = Algebraic difference in gradient between the approaching and departing grades (%)

The values of K offered in Table 4.12 are based on a curve length longer than the required stopping distance with a driver eye height of 1,05 metres and various heights of object as discussed in Chapter 3.

Vertical curvature for passing sight distance

Similar calculations can be carried out based on passing sight distance. High values of K result so that, in the situation where the crest of the curve is in cut, the increase in volumes of excavation will be significant. Although the designer should seek to provide as much passing sight distance as possible along the length of the road, it may be useful to shorten the crest curve in order to increase the lengths of the grades on either side rather than to attempt achieving passing sight distance over the crest curve itself.

Table 4.12: Minimum values of k for crest curves				
Design speed (km/h)	Stopping sight distance (m)	K-Value for height of object (m) equal to		
		0,0	0,15	0,6
40	50	12	6	4
50	70	25	12	8
60	90	40	20	12
70	110	60	30	18
80	140	90	50	30
90	170	140	70	45
100	200	190	100	60
110	230	250	130	80
120	270	350	180	110
130	310	460	240	150

Vertical curvature for barrier sight distance

On undivided roads, barrier sight distance (also referred to as non-striping sight distance) indicates whether no-passing pavement markings are required. Barrier sight distance is shorter than passing sight distance. The designer should attempt to provide barrier sight distance or more wherever possible because passing manoeuvres can often be completed in less than the calculated passing sight distance specifically in cases where the actual speed differentials between the overtaking and the overtaken vehicles are greater than those applied in the derivation of passing sight distances. K values of crest curvature corresponding to barrier sight distance are offered in Table 4.13.

Sag curves

During the hours of daylight or on well-lit streets at night, sag curves do not present any problems with regard to sight distance. Under these

Table 4.13: Minimum k-values for barrier sight distance on crest curves		
Design speed (km/h)	Barrier Sight Distance (m)	K-Value
40	100	10
50	140	20
60	180	34
70	220	50
80	280	85
90	340	125
100	400	170
110	460	225
120	540	310
130	620	410

circumstances, the value of K is determined by considerations of comfort, specifically the degree of vertical acceleration involved in the change in gradient. The maximum comfortable vertical acceleration is often taken as 0,3 m/s^2.

Where the only source of illumination is the vehicle's headlights, the line of sight is replaced by a line commencing at headlight height, taken as being 0,6 metres, and with a divergence angle of 1° relative to the grade line at the position of the vehicle on the curve. This situation is illustrated in Figure 4.9.

Although not a frequent occurrence, sight distance on a sag curve may be impaired by a structure passing over the road. Checking the available sight distance at an undercrossing is best made graphically on the profile. In this case, the line of sight is a grazing ray to the soffit of the structure. The selected clearance is thus of interest. Clearances are typically taken as 5,2 metres measured at the lowest point on the soffit. In the case of pedestrian bridges, the clearance is 5,9 metres while, in the case of existing pedestrian bridges, a clearance of 5,6 metres may be accepted. This is suggested because pedestrian overpasses are relatively light structures that are unable to absorb severe impact and are more likely to collapse in such an event. The increased vertical clearance reduces the probability of damage to the structure and improves the level of safety for pedestrians using it.

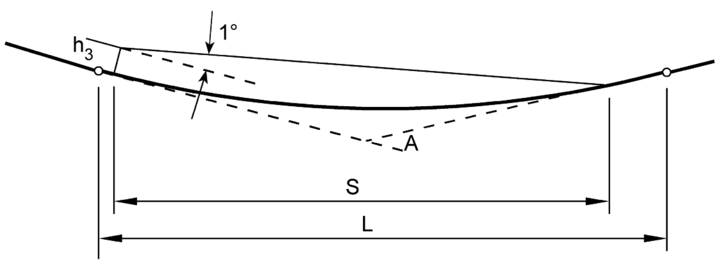
The criteria of comfort, sight distance and vertical clearance lead to a minimum desirable length of sag curve. A further criterion is that of drainage, whereby a minimum distance with a gradient of less than 0,5 per cent is desired.

Where the stopping sight distance is less than the length of the curve, the value of K is given by Equation 4.22.

$$K = \frac{S^2}{120 + 3{,}5S} \qquad (4.22)$$

where S = stopping distance (m)

Sag curves



L length of vertical curve in metres
S minimum stopping distance in metres
L > S
h_3 height of head lamps, 0,6 m
1° angle of light beam upward from plane of vehicle
A algebraic difference in grades, percent

$K = \frac{L}{A}$

Figure 4.9: Sight distance on a sag curve

The criterion of comfort, as expressed in Equation 4.23, provides K-values roughly half of those dictated by considerations of stopping sight distance.

$$K = V^2 / 395 \qquad (4.23)$$

where V = design speed (km/h)

K-values for sag curves, as determined by headlight distance and comfort, for the case where the sight distance is less than the length of the curve, are given in Table 4.14

In both crest and sag curves, K-values dictated by sight distance, where the curve length is greater than the sight distance, can be used throughout. The alternative case, where the sight distance is longer than the curve length, generally offers a lower K-value.

K-values appropriate to headlight distance should be used in rural areas and also where street lighting is not provided. Where street lighting is provided, the lower K-values associated with the comfort criterion may be adopted.

4.4 CROSS-SECTIONS

The prime determinants of cross-section design are:

- The function that the road is intended to serve;
- The nature and volume of traffic to be accommodated; and
- The speed of the traffic.

Road function refers to a spectrum of needs ranging from accessibility to mobility. Furthermore, a road can be classified on a variety of different bases. A common feature of these considerations is that they largely concern addressing the needs of the occupants of a moving vehicle. These needs may find expression in a desire for ready access to or from a property adjacent to the road, freedom to manoeuvre in a terminal or intersection area or for high-speed long distance travel. High-speed traffic requires more space than relatively slow-moving traffic. Space takes the form of wider lanes, wider shoulders and (possibly) the inclusion of a median in the cross-section.

Table 4.14: Minimum k-values for sag curves

Design speed (km/h)	K-Value	
	Headlight distance	Comfort
40	8	4
50	14	6
60	20	9
70	25	12
80	30	16
90	40	20
100	50	25
110	60	30
120	70	36
130	80	43

All these needs have to be met in terms of overall objectives of safety, economy, convenience and minimum side effects.

In urban areas, road functions also have to include considerations of living space. People enjoy casual encounters, meeting people on neutral territory, as it were, without the obligation of having to act as host or hostess in the home. The sidewalk café, the flea market and window-shopping all have to be accommodated within the road reserve. All of these activities impact on the cross-section, which has to be designed accordingly.

Traffic does not exclusively comprise motorised vehicles. In developing areas, it may be necessary to make provision for animal-drawn vehicles and, in this context, developing areas are not necessarily exclusively rural. The volume of motorised vehicles will have an impact on the design of the cross-section with regard to the number of lanes that have to be provided. High volumes of moving vehicles will generate a need for special lanes such as for turning, passing, climbing or parking.

In urban areas, the presence of large numbers of pedestrians will require adequate provision to be made in terms of sidewalk widths. Pedestrians are also to be found on rural roads. On rural roads, speeds are high so that crashes involving pedestrians are inevitably fatal. It is thus sensible to make at least modest provision for pedestrians on rural roads, even though their numbers may be low.

Cyclists can often be accommodated on the normal travelled lanes but, when the number of cyclists increases, it may be necessary to widen these lanes or, as a further development, to provide cycle paths adjacent to or, for preference, removed from the travelled lanes.

Although the horizontal and vertical alignments are disaggregated in the sense that they are a combination of tangents and curves, the cross-section is heavily disaggregated, comprising a multitude of individual elements. These elements are illustrated in Figure 4.10. Design is thus concerned primarily with the selection of elements that have to be incorporated within the cross-section, followed by sizing of these individual elements.

In spite of this disaggregated approach to design, there are numerous combinations of elements that occur frequently. Cases in point are:

- Two-lane two-way roads;
- Two-plus-one roads;
- Four-lane undivided and divided roads, and
- Four, six (or more) lane freeways.

Each of these composite cross-sections leads to the development of a standard road reserve width that often is enshrined in legislation. In this section, the individual elements rather than the composite cross-section will be discussed.

4.4.1 General controls for cross-sections

Safety is a primary consideration in the design of the cross-section. The safety of the road user refers to all those within the road reserve, whether in vehicles or not.

Wide lanes supposedly promote the safety of the occupants of vehicles although current evidence suggests that there is an upper limit beyond which safety is reduced by further increases in lane width. The reverse side of the coin is that wide lanes have a negative impact on the safety of pedestrians attempting to cross the road or street. South Africa has a particularly bad record in terms of pedestrian fatalities, which account for approximately half of the total number of fatalities. In devising safe cross-sections, it is therefore necessary to consider the needs of the entire population of road users and not just those in vehicles.

In urban areas, it is necessary to make provision for boarding and alighting public transport passengers, disabled persons and other non-vehicular users of the facility in addition to accommodating pedestrians and cyclists. In these areas, design speed usually plays a lesser role in the design of the cross-section.

Land availability is of particular importance in urban areas. Land for the road reserve may be restricted because of the existence of major buildings or high cost of acquisition or some or other protected land use. During the planning process, it is prudent to attempt to acquire additional road reserve width to allow for improvements to traffic operations, auxiliary lanes, wider pedestrian areas, cycle paths as well as for provision of utilities, streetscaping and maintenance considerations. This should, however, not be taken to the extreme, where large tracts of land are unnecessarily sterilised in anticipation of some or other future eventuality.

The location of existing major utilities, which may be either above or below ground, and difficult or costly to relocate is a fairly common design control in urban areas. In particular the location of aboveground utilities in relation to clear zone requirements should be carefully considered.

In the discussion that follows, dimensions are essentially discussed in terms of new or "green field" designs. Very often, however, the designer does not have this level of freedom of choice. For example, rehabilitation projects often require the creation of additional lanes without

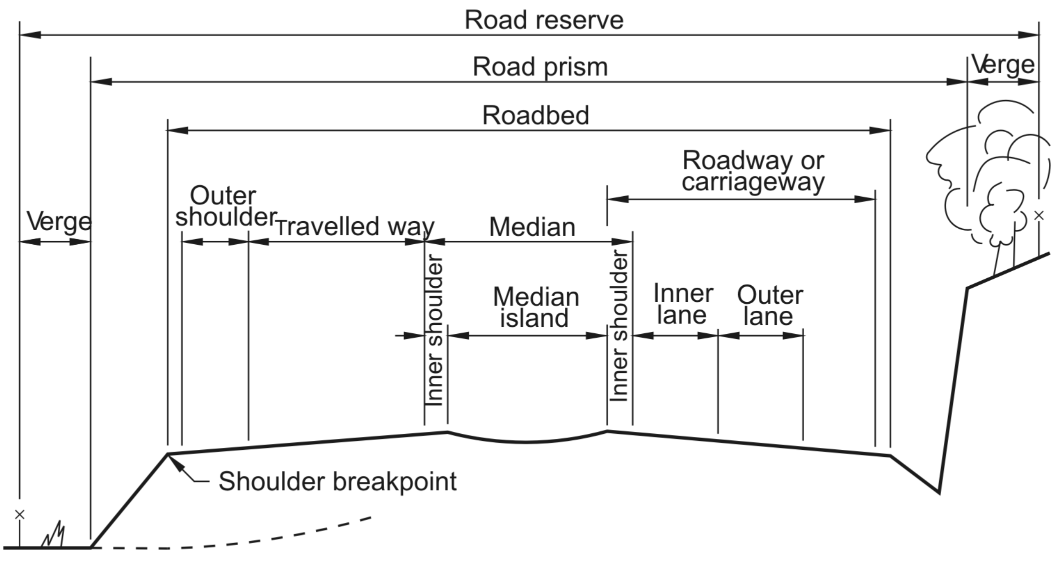


Figure 4.10: Cross-section elements

the funding or space available for additional construction. The choice then is one of either accepting a lane or shoulder width that is narrower than would be desired or foregoing the additional lanes.

As discussed below, the preferred cross-section for a two-lane two-way road has a total width of 13,4 metres. This comprises lane widths of 3,7 metres and usable shoulder widths of 2,5 metres plus a 0,5 metre allowance for shoulder rounding. If a cross-section of this width already exists and it is deemed necessary to incorporate a climbing lane without incurring extra construction costs, this can be achieved by accepting a climbing lane width of 3,1 metres with the adjacent shoulder having a width of 1,0 metres. The through lanes have to be reduced to a width of 3,4 metres each and the opposite shoulder to a width of 1,5 metres thus allowing for the shoulder rounding of 0,5 metres on both shoulders.

Many of the older cross-sections have a total paved width between shoulder breakpoints of 12,4 metres, comprising lanes with a width of 3,4 metres, shoulders with a width of 2,3 metres and 0,5 metres of shoulder rounding on both sides. Three lanes with a width of 3,1 metres each would allow for shoulders that are approximately 1 metre wide and 1 metre of shoulder rounding. It is suggested that the dimensions offered for the three-lane cross-section constitute an irreducible minimum and can be considered as a palliative measure only over a relatively short distance, e.g. to accommodate a climbing lane. Accommodating a third lane on cross-sections narrower than 12,4 metres between shoulder breakpoints should not be essayed.

Similarly, a fourth lane on a 13,4 metre wide cross-section would imply lane widths of 3,1 metres and a zero width of shoulder if the shoulder rounding of 0,5 metres is to be maintained. The traffic volumes necessitating consideration of a four-lane cross-section would render such a configuration of lanes and shoulders highly undesirable.

4.4.2 Basic Lanes

Basic lanes are those that are continuous from one end of the road to the other. The number of lanes to be provided is largely determined by traffic flow and the desired Level of Service that the road is to provide. Reference is thus to the Highway Capacity Manual.

The anticipated traffic speed offers an indication of the required width of lane. Lane widths typically used are 3,1 metres, 3,4 metres and 3,7 metres. Research has indicated that the crash rate starts to show a marginal increase above a lane width of 3,6 metres. It is accordingly recommended that widths significantly greater than 3,7 m should not be employed. Some local authorities use lane widths of 5,4 metres on major routes. The intention is apparently to make provision for parking or cycle traffic without demarcating the lanes as such. In practice, passenger cars are very inclined to use these lanes as though they were two unmarked lanes 2,7 metres in width. Although this is a very economical method for providing adequate capacity, the safety record of such cross-sections may be suspect and their use should be discouraged.

The narrowest width recommended for consideration (3,1 metres) allows for a clear space of

300 mm on either side of a vehicle 2,5 metres wide. This width would only be applied to roads where traffic volumes and/or speeds are expected to be low.

Where traffic volumes are such that a multi-lane or divided cross-section is required, 3,7 metres would be a logical width to adopt. High speeds would also warrant this width because, on a narrower lane, momentary inattention by a driver could easily cause a vehicle to veer into the path of another. Intermediate volumes and speeds are adequately accommodated by a lane width of 3,4 metres.

In the case of a rehabilitation or reconstruction project, it may be necessary to add lanes to the cross-section. The cost of the additional earthworks may however be so prohibitive that the funds required to upgrade the road to full 3,7 metre lanes and 3,0 metre shoulders may simply not be available. The options available to the road authority are then either not to upgrade the road at all or to accept some lesser widths of lanes and/or shoulders. It is suggested that a stepwise approach to the problem, first reducing the available width of shoulder and thereafter considering reductions in lane width, could be adopted by. A 3,4 metre lane is reasonably safe, even at a speed of 120 km/h, but is not a comfortable solution when applied over an extended distance. A reduction in the posted speed limit may be desirable if the lesser lane width is applied over several kilometres. At still lower lane widths, it may be advisable to provide a posted speed limit of 100 km/h even for relatively short sections of road.

Travelled lanes have either a camber, being a slope from the centreline towards the outside shoulder, or a cross fall, being a slope from one edge of the travelled way to the other edge. These slopes are provided to ensure drainage of the road surface and are typically at two per cent although, in areas where high rainfall intensities are likely to occur, the slope could be increased to as much as three per cent.

4.4.3 Auxiliary lanes

Auxiliary lanes are located immediately adjacent to the basic lanes. They are generally short and are provided only to accommodate some or other special circumstance.

Auxiliary lanes are often used at intersections. They can be turning lanes, either to the left or to the right, or through lanes. The turning lanes are principally intended to remove slower vehicles, or stopped vehicles waiting for a gap in opposing traffic, from the through traffic stream hence increasing the capacity of the through lanes. Through auxiliary lanes are often employed at signalised intersections to match the interrupted flow through the intersection area with the continuous flow on the approach lanes. The application and design of these lanes are discussed in depth in Chapter 6.

Auxiliary lanes are also employed at interchanges. These are through lanes and are intended to achieve lane balance where turning volumes are sufficient to warrant multi-lane on- and off-ramps. They may also be used between closely spaced interchanges to support weaving, principally replacing a merge-diverge with a Type A weave. Auxiliary lanes at interchanges are discussed in Chapter 7.

Climbing lanes and passing lanes are auxiliary lanes employed on network links, i.e. other than at intersections and interchanges. Climbing lanes are often referred to as truck lanes, crawler lanes, overtaking lanes or passing lanes. The function of climbing lanes is, however, very different from that of passing lanes. Climbing lanes remove slower vehicles from the traffic stream and have the effect of reducing the number of passenger car equivalents (PCE's) in the stream. If the reduction is sufficient, the Level of Service (LOS) on the grade will match that on the preceding and succeeding grades. Passing lanes also remove slower traffic from the stream but, in this case, the objective is platoon dispersal, thus supporting an increase in the capacity of the road. In short, the climbing lane seeks to match the LOS on a steep grade to that on the preceding and following flatter grades, whereas the passing lane improves the capacity of the road as a whole. The ultimate demonstration of the latter circumstance is the four-lane road, which could be described as a two-lane road with continuous passing lanes in both directions.

From a safety point of view, it is important that drivers are made aware of the start and, more particularly, the end of an auxiliary lane. The basic driver information requirements in the latter case are:

- Indication of the presence of a lane drop;
- Indication of the location of the lane drop; and
- Indication of the appropriate action to be undertaken.

It has been observed that, without adequate signposting and road markings indicating the presence of a lane drop to drivers, erratic and last-second manoeuvres occur. In the case of merging, such manoeuvres can be extremely hazardous. Reference should be made to the SADC Road and Traffic Signs Manual for the recommended signage and road markings.

The lane drop should not be located so far from the end of the need for the auxiliary lane that drivers accept the increase in number of lanes and do not expect the reduction. Furthermore, the location should not be such that the lane drop is effectively concealed from the driver. Concealment can arise from location immediately beyond a crest curve. An "architectural" approach which locates the lane drop on a horizontal curve and selects curve radii that result in a smooth transition from two lanes to one across the length of the curve is a particularly subtle form of concealment that should be avoided at all costs. To simplify the driving task, the lane drop should be highly visible to approaching traffic and drivers should not be subjected to successive decision points too quickly implying that lane drops should be removed from other decision points.

It is possible that the shoulder adjacent to an auxiliary lane could be very narrow. However, for emergency use it is recommended that a three metre wide surfaced shoulder be provided along and extending downstream of the taper for a distance equal to the stopping sight distance for the design speed of the road.

Climbing lanes

Four types of warrants for climbing lanes are in use. These are:

- Reduction of truck speed through a given amount or to a specified speed;

- Reduction in truck speed in association with a specified volume of traffic;
- Reduction in LOS through one or more levels, and
- Economic analysis.

The first three focus on the performance of trucks and infer some or other impact on passenger vehicles. The fourth directly quantifies the effect of slow moving vehicles (which need not necessarily be exclusively truck traffic) on the traffic stream in terms of delay over the design life of the climbing lane and compares the benefit of the removal of delay with the cost of providing and maintaining the climbing lane.

Speed reductions adopted internationally vary in the range of 15 km/h to 25 km/h and are usually intended to be applied to a single grade. The most widely occurring value of speed reduction is 15 km/h, based on considerations of safety. The Australian approach bases the need for climbing lanes on examination of a considerable length of road. The justification for climbing lanes is based on traffic volume, percentage of trucks and the availability of passing opportunities along the road. Speed reduction is to 40 km/h and not by 40 km/h.

South African practice, as described in TRH17, uses a combination of the speed and traffic volume as a warrant, requiring both to be met before the climbing lane is warranted. The speed reduction applied is 20 km/h from an initial 80 km/h. The volume warrant is given in Table 4.15 below.

Truck speed reductions without reference to traffic volumes have merits in terms of safety. Their principal benefit lies in reduction in the speed differential in the through lane, thereby reducing the probability of the occurrence of a crash. It is, however, theoretically possible that a climbing lane would be considered warranted merely because it would lead to the required truck speed reduction even if total traffic volumes were very low with virtually no trucks in the traffic stream. The addition of a volume warrant increases the likelihood of a reasonable economic return on the provision of a climbing lane. It is conceded that performance-based warrants such as those described above are not intended to be - or are ever likely to be - fully economic.

Level of Service is a descriptor of operational characteristics in a traffic stream. An important feature is that it is purely a representation of the driver's perception of the traffic environment and is not concerned with the cost of modifying that environment. The warrants suggested by the Highway Capacity Manual are:

Table 4.15: Warrant for climbing lanes

Gradient (%)	Traffic volume in design hour	
	5 % trucks	10 % trucks
4	632	486
6	468	316
8	383	243
10	324	198

- A reduction of two or more Levels of Service in moving from the approach segment to the grade, or
- Level of Service E existing on the grade.

The Highway Capacity Manual warrant is severe and does not, in any event, match the speed reduction warrant. The provision of a climbing lane at a specific site is thus dependent on the type of warrant selected. It is accordingly suggested that, if a designer wishes to apply this warrant, the reduction considered should be of one or more Levels of Service, and not two or more.

In view of the economic restraints on new construction, a compromise between convenience and cost effectiveness is required. The compromise proposed is that, while delay - seen as a major criterion of Level of Service - is employed in determination of the need for provision of a climbing lane, the delay considered would not be that suffered by the individual vehicle but rather by the entire traffic stream. Commercially available software calculates the value of the time saved during the design life of the climbing lane and relates this to input from the designer in respect of construction and maintenance costs of the lane.

In mountainous terrain, where trucks are reduced to crawl speeds over extended distances and relatively few opportunities for overtaking exist, the cost of construction of climbing lanes may be prohibitive. Under these circumstances, an alternative solution to the operational problem may be to construct short lengths of climbing lane as opposed to a continuous lane over the length of the grade. These short climbing lanes are variously referred to as passing bays, turnouts or partial climbing lanes and are typically 100 to 200 metres long. Because vehicles entering the turnout do so at crawl speeds, the tapers can be very short, e.g. twenty to thirty metres long, corresponding to taper rates of 1:6 to 1:10 in the case of 3,1 metre wide lanes.

Climbing lanes usually have the same width as the adjacent basic lane. In very broken terrain, a reduction in width to as little as 3,1 metres can be considered because of the low speeds of vehicles in the climbing lane. On the same grounds, the shoulder width may also be reduced but to not less than 1,5 metres. If the shoulders elsewhere on the road are three metres wide, the additional construction width required to accommodate the climbing lane and reduced shoulder is thus only 1,6 metres.

While the decision whether or not to provide the climbing lane could be based on the economics of the matter, the location of its terminals is dependent purely on safety based on the operational characteristics of truck traffic. The warrant for the provision of climbing lanes in terms of truck speed reduction is set at 20 km/h as described previously. It is suggested, as a safety measure to allow for variation in the hill climbing capabilities of individual trucks, that, having established that a climbing lane is warranted, a speed reduction of 15 km/h be used to determine the location of the climbing lane terminals. Table 4.10 shows critical lengths of grade in terms of a truck speed reduction of 15 km/h for various gradients. It is recommended that the full width of the climbing lane be provided at or before the end of the critical length, with this length being measured on the preceding sag curve from a point halfway between the Vertical Point of Intersection (VPI) and the end of the vertical curve (EVC). The full width of the climbing lane should be maintained until the point is reached where truck speed has once again increased to be 15 km/h less than the normal speed on a level grade.

The length of the entrance taper should be such that a vehicle can negotiate the reverse curve path with the benefit of a 2 per cent crossfall on the first curve followed by a negative superelevation of 2 per cent on the second curve with a short intervening tangent to allow for the reversal of curvature. Ideally, it should not be necessary for the vehicle path to encroach on the shoulder. These conditions can be met by a taper which is about 100 metres long, corresponding to a taper rate of 1 : 27. This is approximately half the taper rate applied to interchange off-ramps, which is appropriate as, in this case, the taper addresses a reverse path and not a single change of direction.

As stated above, the exit terminal of the climbing lane should be the point at which trucks have accelerated back to a speed that is, at most, 15 km/h slower than the truck operating speeds on the basic lanes. If there is a barrier line at this point, the lane should be extended to the point at which the barrier line ends. The reason for this is that a vehicle entering the basic lane may inadvertently force an overtaking vehicle into the opposing lane. This is a potentially dangerous situation and the designer must ensure that there is sufficient sight distance to support appropriate decisions by the drivers involved.

The terminal may take one of two possible forms. One option is to provide a right-to-left taper merging the basic lane with the climbing lane, followed by a left-to-right taper back to the two-lane cross-section. The motivation for this layout is that faster-moving vehicles find it easy to merge with slower-moving traffic. This is its only advantage. Should a vehicle not manage to complete the merge before the end of the right-to-left taper, its only refuge is the painted island, thereafter being confronted by an opposing lane situation. Furthermore, in this layout, the basic lane terminates at the end of the climbing lane with the climbing lane thereafter becoming the basic lane. This is a contradiction of the fundamental definitions of the basic and auxiliary lanes. This option is not recommended for two-lane roads, although it could possibly be used on multilane or divided cross-sections.

The alternative is to provide a simple taper, dropping the climbing lane after it has served its purpose. A vehicle that cannot complete the merging manoeuvre at the end of the climbing lane has the shoulder as an escape route. This is a safer option than that described above.

As described with regard to the entrance taper, a vehicle exiting from the climbing lane could negotiate a taper that is 100 metres in length. However, a flatter taper would allow time to find a gap in the opposing traffic. A taper rate of 1:50 is suggested for on-ramps at interchanges where vehicles are required to negotiate only a single change of direction. The reverse curve path followed in exiting from an auxiliary lane may require a still flatter taper rate for which reason a rate of the order of 1:70 is suggested, leading to a taper that is approximately 200 metres in length.

Passing lanes

The procedure followed in the design of the alignment of a road should seek, in the first instance, to provide the maximum possible passing opportunity. Thereafter, the need for climbing lanes should be evaluated and decisions taken on where climbing lanes are to be provided and what the lengths of these climbing lanes should be. At this stage, it is possible to relate the total passing opportunity to the overall length of the road.

The analysis of two-lane roads as described in the Highway Capacity Manual is based on two criteria, being percentage time spent following and average travel speed. Both of these are adversely influenced by inadequate passing opportunities. For example, a road with sixty per cent passing opportunity demonstrates a 19 per cent increase in time spent following by comparison with a road with similar traffic volumes and unlimited passing opportunities. This assumes a 50/50 directional split. The increase in time spent following is greater with unbalanced flows. If the flow is as low as 800 vehicles per hour, passing opportunities limited to 60 per cent result in a reduction of the order of three per cent in average speeds by comparison with the speeds on roads with unlimited passing opportunity. It is accordingly necessary for the designer to carry out an analysis based on the Highway Capacity Manual to establish whether or not additional passing opportunities should be provided in order to maintain the desired Level of Service.

Passing lanes are normally provided in areas where construction costs are low and where there is an absence of passing opportunities. They are aimed at platoon dispersion and local research has demonstrated that a passing lane length of about one kilometre is adequate for this purpose. Numerous short passing lanes are preferable to few long passing lanes and it is recommended that they be located at two, four and eight kilometre spacings. Where traffic volumes are low, the longest spacing can be used and, as traffic volumes increase, the intervening lanes can be added in a logical manner.

With one-kilometre long passing lanes provided at two-kilometre intervals, the next level of upgrading would be a Two + One cross-section. In this case, the road is effectively provided with a three-lane cross-section from end to end with the centre lane being alternately allocated to each of the opposing directions of flow. In keeping with the spacings discussed above, the switch in the direction of flow in the centre lane should be at about two-kilometre intervals.

Unlike climbing lanes, passing lanes tend to operate at the speeds prevailing on the rest of the road. Reductions in lane width are thus not recommended and passing lanes should have the same width as the basic lanes.

As recommended for climbing lanes, the entrance taper to a passing lane could be 100 metres in length and the length of the exit taper double this to allow adequate time for merging vehicles to find a gap in the through flow. Seeing that both the entrance and the exit tapers signal a change in operating conditions on the road, it is recommended that decision sight distance should be available at these points.

High occupancy vehicle (HOV) lanes

As described above, auxiliary lanes are short and are intended only to deal with a specific circumstance. As soon as this circumstance changes, the auxiliary lane is dropped. HOV lanes, on the other hand, form part of the rapid transit system of a city and can thus be provided over a substantial distance. Whether they should be considered as auxiliary lanes could thus be debated.

HOV lanes are typically applied on commuter routes with a view to encouraging the use of public transport or lift clubs hence reducing congestion. The average occupancy of passenger cars is of the order of 1,5 persons per vehicle, whereas a municipal bus can convey 80 passengers effectively replacing 50 or more passenger cars in the traffic stream. In view of the fact that buses can be 2,6 metres wide, narrow lane widths are inappropriate to HOV lanes that, ideally, should not be narrower than 3,6 metres, i.e. allowing a clear space of 0,5 metres between the sides of the vehicle and the lane markings.

HOV lanes have to be policed to ensure that only vehicles qualifying for the privilege use them. Signalisation can be employed to give vehicles in HOV lanes priority over other road users. This is described in the South African Road Traffic Signs Manual in some detail. The combination of policing and priority usage underpins the effectiveness of HOV lanes but these operational issues are normally outside the terms of reference of geometric design.

It is important that the designer draws a distinction between the basic lanes and the HOV lanes in the cross-section. The essential point of difference lies in the fact that the passenger car is usually taken as the design vehicle for basic lanes whereas the HOV lanes are designed to accommodate buses. Kerb radii at intersections and the width of the turning lanes on bus routes should be such that buses can negotiate these curves without encroaching on the adjacent lanes or, more importantly, on the sidewalks.

Bus routes typically converge on the CBD. The HOV lanes, which, as part of the normal cross-section, may have served well in the outlying or suburban areas, could prove inadequate to accommodate the increased volume of bus traffic in the CBD. It may then be necessary to designate various of the streets in the CBD as exclusive bus roads or, alternatively, to consider a system similar to the O-Bahn routes.

4.4.4 Kerbing

Kerbs are raised or near-vertical elements that are located adjacent to the travelled way and are usually used for:

- Drainage control;
- Delineation of the pavement edge; and
- Reduction in maintenance operations by providing protection for the edge of surfacing.

Kerbing is normally only applied in urban areas where vehicle speeds are relatively low.

In rural areas, the drainage function is normally accommodated by channels or open drains of various forms. Delineation is usually by means of an edge line or a contrasting colour on the shoulder. Protection of the edge of surfacing can be by means of buried edge blocks or, more typically, by means of a thickened edge. A thickened edge is simply a bitumen-filled V-shaped groove cut into the base course.

Kerbs may be barrier or semi-mountable or mountable. Barrier and semi mountable kerbs normally are accompanied by a channel (or gutter) whereas the mountable kerb is, in effect, a channel itself.

Barrier kerbs are intended primarily to control drainage as well as access and can inhibit slow-moving vehicles from leaving the roadway. When struck at high speeds, barrier kerbs can result in loss of control and damage to the vehicle. In spite of the name, barrier kerbs are inadequate to prevent a vehicle from leaving the road after a high-speed impact. In addition, a barrier kerb can lead to a high-speed errant vehicle vaulting over a barrier or guardrail. For this reason, barrier kerbing is not generally used on urban freeways and is considered undesirable on expressways and arterials with design speeds higher than 70 km/h. Barrier kerbs are never used in conjunction with rigid concrete barrier systems.

Semi-mountable kerbs have a face slope of 25 mm/m to 62,5 mm/m and are considered mountable under emergency conditions. They are typically used on urban freeways and arterials and also in intersections areas as a demarcation of raised islands.

Mountable kerbs have a relatively flat sloping face of 10 mm/m to 25 mm/m and can be crossed easily by vehicles. They are particularly useful as a form of lane demarcation on high-speed roads but are not effective as a form of drainage. Perhaps their widest application is to be found in residential areas, where vehicles can drive off the travelled way to park on the verge.

Channels are usually about 300 mm wide, thus automatically providing an offset between the kerb and the edge of the travelled way. Where channels are not provided, the offset should still be maintained for reasons of safety.

4.4.5 Shoulders

Shoulders are the usable areas immediately adjacent to the travelled way and are a critical element of the roadway cross-section. They provide:

- A recovery area for errant vehicles;
- A refuge for stopped or disabled vehicles;
- An area out of the travel lanes for emergency and maintenance vehicles; and
- Lateral support of the roadway structure.

In addition, shoulders support use of the road by other modes of transport, for example cyclists and pedestrians.

Regulation 298 of the regulations promulgated in terms of National Road Traffic Act (Act 93/1996) prohibits driving on the shoulder except that this is permitted:

- On a two-lane road;
- Between the hours of sunrise and sunset,
- While being overtaken by another vehicle.

provided this can be done without endangering the vehicle, other vehicles, pedestrians or property and if persons and vehicles on the road are clearly discernable at a distance of at least 150 metres.

Considering the above applications of the shoulder, a stopped vehicle can be accommodated on a shoulder that is three metres wide. There is no merit in adopting a shoulder width greater than this. The shoulder should, however, not be so narrow that a stopped vehicle could cause congestion by forcing vehicles travelling in both directions into a single lane. A partly blocked lane is acceptable under conditions of low speed and low traffic volume. Assuming the narrowest width of lane, i.e. 3,1 metres, it would be possible for two vehicles to pass each other next to a stopped car if the shoulder were not less than 1,0 metres wide. Hazards, including the edges of high fills, cause a lateral shift of vehicles if closer to the lane edge than 1,5 metres. Allowing for shoulder rounding of 0,5 metres, the usable shoulder is thus 1,0 metres wide and this should be considered the irreducible minimum width of shoulder.

Where the traffic situation demands a dual-carriageway cross-section, the greatest width of shoulder, i.e. three metres, is called for. This width would apply to the outer shoulder. The inner shoulder need only be one metre wide:

- to protect the integrity of the pavement-layers;
- to avoid drop-offs at the lane edge; and
- provide space for roadmarkings

provided the median island is not kerbed, thus allowing a disabled vehicle to be moved clear of the adjacent lane. If a barrier, such as kerbing or a guardrail, makes the median island inaccessible, the full shoulder width should be provided in the case of a six-lane cross section because negotiating two lanes to reach the safety of the outside shoulder (with a disabled vehicle) could be difficult.

Between the two extremes of 3,0 metres and 1,0 metres, shoulder widths of 1,5 or 2,5 metres could be used in the case of intermediate traffic volumes and speeds. These alternative shoulder widths would not normally be used for the inner shoulders of a dual carriageway road. Table 4.16 illustrates the application of the various shoulder widths on undivided rural roads.

Paved widths of between 1,5 and 2,5 metres should be avoided. The presence of the paving may tempt a driver to move onto the shoulder to allow another vehicle to overtake, but these widths cannot accommodate a moving vehicle with any safety.

These shoulder widths are recommended for adoption for new construction. In the case of rehabilitation or reconstruction projects, there may not be sufficient width of cross-section to accommodate the desirable widths and some lesser width will have to be considered. As pointed out in Section 4.4.2, it may be advisable to first reduce the shoulder width before considering reductions of lane width.

The shoulder breakpoint is usually about 500 mm beyond the edge of the usable shoulder to allow for shoulder rounding.

Where guardrails or other roadside appurtenances have to be provided, these are located 300 to 500 mm beyond the usable shoulder. The shoulder breakpoint should be a further 500 mm beyond these appurtenances, as a lesser distance will not provide the support needed by a guardrail when hit by an errant vehicle.

The surfacing of shoulders is recommended:

- For freeways;
- In front of guardrails;
- Where the total gradient, being the resultant of the longitudinal gradient and the camber or superelevation, exceeds six per cent;
- Where the materials with which the shoulders are constructed are readily erodible, or where the availability of material for maintenance of the shoulders is limited;
- Where heavy vehicles would tend to use the shoulder as an auxiliary lane;
- In mist belts; or
- Where significant usage by pedestrians occurs.

A patchwork of surfaced shoulders would be both unsightly and unsafe. Where the intervening lengths of unsurfaced shoulders are short, it is suggested that they also be surfaced. As a guideline, it is proposed that if surfacing sixty or more per cent of the shoulder is warranted, this should constitute adequate grounds for full surfacing of the shoulders.

Full surfacing implies continuous surfacing along the length of the road and not necessarily across the full width of the shoulder, although this is the desirable option. It is suggested above that the minimum recommended width of shoulder is 1,0 metres. If it were considered necessary to surface the shoulder at all, there would be little or no operational advantage in surfacing a lesser width than this. In the case of new construction, the designer has the option of considering the economic merits of a relatively narrow surfaced shoulder vis-à-vis a wide unsurfaced shoulder. In the case of rehabilitation projects, it may be decided to retain the full 3,0 metre shoulder but, as a cost-saving measure, to surface only half of the total width.

The cross fall on surfaced shoulders is normally an extension of that on the travelled lanes. Where shoulders are not surfaced, the cross fall

Table 4.16: Shoulder widths for undivided rural roads

Design speed (km/h)	Design hour volume (veh/h)		
	<250	250 – 450	>450
	Width of shoulders (m)		
50	1,0		
60	1,5	1,5	
70	1,5	2,5	
80	2,5	2,5	2,5
90	2,5	2,5	3,0
100	2,5	2,5	3,0
110		3,0	3,0
120		3,0	3,0
130			3,0

Note: Shoulder widths are not quoted for unlikely combinations of speed and volume is normally one per cent steeper than that on the lane to allow for the rougher surface and the consequently slower rate of flow of storm water off the road surface.

At night or during inclement weather it is important that the driver should be able to distinguish clearly between the shoulder and the lane. This can be accomplished by the use of a shoulder material of a contrasting colour or texture. Edge marking is a convenient way of indicating the boundary between the lane and the shoulder. Rumble strips can also be used and have been shown to reduce the rate of run-off-road incidents by twenty per cent or more. Rumble strips can be raised or grooved. Being intended to provide the driver with an audible warning, the noise level they generate is unacceptable in urban areas and should therefore only be used in rural areas.

4.4.6 Medians

The median is the total width between the inner edges of the inside traffic lanes and includes the central island and the median shoulders. As long ago as the early 1930s it was proposed to "separate the up-traffic from the down-traffic", a function which the median fulfils to this day. This separation is intended to reduce the probability of head-on crashes and also to reduce the nuisance of headlight glare (usually by the planting of shrubs on the central island). The reduction in head-on crashes is achieved through selection of a suitable width of median or the use of median barriers. Shrubs can also serve as a barrier to prevent cross-median accidents but the stems of the shrubs should not grow so thick as to become a further hazard. A maximum stem thickness of 175 to 200 mm, corresponding to the diameter of a guardrail post, is recommended.

A further application of medians refers to access management, where right turns into or out of local land uses are often discouraged on high-speed roads.

From the above it can be inferred that medians are typically applied in the case of high speed or high volume roads with a basic function of mobility.

Median islands can be as narrow as one metre, which is sufficient to contain a median barrier comprising back-to-back guardrails. This suggests that, including minimum width median shoulders, the minimum width of the median should be not less than three metres. Inner shoulders are often not provided in the urban cross-section but kerbing would require an offset of about 300 mm. The minimum width of an urban median should thus be 1,6 metres.

Research has found that few out-of-control vehicles travel further than nine metres from the edge of lane, so that this width of median would be sufficient to avoid most head-on crashes.

In urban areas, medians often contain right-turning lanes. In intersection designs, the inner shoulder is invariably replaced by kerbing so that the median would be the sum of the lane width plus the width required to provide a pedestrian refuge. Pedestrians do not feel safe on median islands narrower than about two metres, suggesting that the median should have a width of the order of 5,5 to 6,0 metres. This width is adequate to accommodate pedestrians as well as the right-turning lane. If intersections are closely spaced, it may be necessary to apply this width to the full length of the median, whereas with widely spaced intersections, e.g. 500 metres or more between intersections, a lesser width can be applied between the intersections with the median being flared out by means of active tapers at the intersections.

Medians with a width of nine metres or more allow for individual grading of the two carriageways, which can be useful in rolling terrain. In addition, these medians lend themselves to landscaping and to the creation of a park like environment. Unfortunately, they create problems at intersections by virtue of the long travel distances that they impose on turning vehicles. The incidence of crashes at intersections increases with increasing width of median and, at widths of 20 metres, the intersection should be designed as two intersections back-to-back, with traffic control on the roadway crossing the median. The wide rural median does not translate well to the urban environment so that roads on the outskirts of urban areas should be designed with medians appropriate to a future urban characteristic.

Medians may be either depressed or raised. Depressed medians are normally used in rural areas and raised medians in urban areas. This differentiation between rural and urban areas arises for two reasons: drainage and safety.

Storm water drainage in rural areas is generally above the surface for ease of maintenance whereas, in urban areas, a well-developed piped drainage system is generally available. The depressed median allows the roadbed to drain into the median, specifically on curves where water from the outer carriageway is prevented from draining across the inner carriageway. In the urban context, kerbing includes drop inlets directing storm water into the underground system.

Urban median islands are usually narrower than their rural counterparts and do not normally have barriers. The barriers have to be terminated at every intersection and at some entrances so that the safety offered by the barrier is more than offset by the hazard of the barrier ends. Kerbing offers a modest degree of protection to pedestrians who may be on the median while crossing the road. In addition, kerbing can, to a limited extent, redirect errant vehicles back into their own lanes.

The speeds on rural roads make kerbing inappropriate in this environment, as the driver of a vehicle striking a kerb at high speed would almost certainly lose control of the vehicle, with this problem being compounded by the inevitable damage to the front wheels of the vehicle.

Depressed medians generally have flat slopes and a gently rounded bottom so that the driver of a vehicle leaving the road has an opportunity to regain control, minimising occupant injury and vehicle damage. Overturning crashes are more frequent on slopes steeper than 1 : 4 for median widths of six to twelve metres. This should, therefore, be considered the steepest allowable slope, with slopes of 1 : 6 or flatter being preferred.

4.4.7 Outer separators

The outer separator is the area between the edges of the travelled way of the major road and the adjacent parallel road or street. It comprises the left shoulder of the major road, an island and the right shoulder of the adjacent road or street. The outer separator serves as a buffer between through traffic and local traffic on a frontage or service road. It is typically applied where the corridor has to serve the two functions of long distance travel and local accessibility. An arterial passing through a local shopping area is an example of this application.

If travel on the frontage road is one way and in the same direction as that on the adjacent through lane, the outer separator can be as narrow as three metres or, if barriers are provided, two metres.

At night, drivers on the through lane would find an opposing direction of flow on the frontage road very confusing, being confronted by headlights both to the right and to the left. Under these circumstances, the width of the outer separator should be substantially increased, preferably doubled, to minimise the effect of the approaching traffic, particularly headlight glare on non-illuminated sections of the road. Plantings or dazzle screens on the outer separator are recommended for the same reason.

On rural freeways, the outer separator should be at least nine metres wide, based on the distance that an out-of-control vehicle is able to move away from the edge of the through lane. Reference in the literature is to outer separator widths of twenty to thirty-five metres in rural areas.

At an intersection, the frontage road should either be terminated or moved a substantial distance away from the through lanes. This is intended to safeguard the operation of the intersection because vehicles attempting to turn from the through road to a frontage road could very easily generate a queue that backs up onto the through lanes. Not only is this operationally undesirable but it could also be unsafe.

Where it is anticipated that a road will have to be widened at some time in the future, the width of the outer separator should be such that it can accommodate the additional lane, hence minimising the extent of damage to the rest of the road cross-section.

4.4.8 Boulevards

Boulevards are only used in urban areas and are similar to outer separators with regard to their function and location. The principal difference is that they separate a sidewalk and not a frontage road from the through lanes. Boulevards are a desirable feature because:

- The separation between the sidewalk and the vehicular traffic provides increased safety for pedestrians and children at play;
- The probability of a pedestrian/vehicle collision is reduced as the sidewalk is placed some distance from the kerb;
- Pedestrians are less likely to be splashed by passing vehicles in wet weather;
- Space is provided for street furniture and streetscaping as well as for surface and underground utilities, and
- Changes to the cross-slope of the sidewalk to provide for appropriate driveway gradients are minimised using the

boulevard area to effect the gradient change.

The verge, showing the location and dimensions of the boulevard, is illustrated in Figure 4.11.

Aesthetic considerations in the urban environment are important, particularly when major streets pass through or are adjacent to parkland and residential areas. Desirably, the positive aesthetic qualities of the adjacent land use are carried over into the verge and boulevard areas of the street cross-section. As a feature of the urban landscape, boulevards are usually grassed or landscaped. If the boulevards are narrower than 1,5 metres, they are surfaced rather than grassed because of the maintenance difficulties associated with narrow strips.

The entire area from the reserve boundary to the road edge is normally sloped towards the road to assist drainage, not only of this area but of the adjacent development as well. Because of the impedance offered by grass to overland flow, the slope of the boulevard should be at least four per cent. Local circumstances may require steeper slopes but, where driveways have to be provided across boulevards, the variation in slope should not be so drastic that vehicles cannot traverse the area without scraping their undersides on the ridge between the boulevard and the sidewalk.

Boulevards are as wide as the road reserve allows. Ideally, they should not be narrower than two to three metres.

4.4.9 Bus stops and taxi lay-byes

Bus stops and lay-byes can be located within the width of the boulevard. In this case, grassing of the boulevard is discontinued and the area surrounding the bus stop is paved as an extension of the sidewalk to provide users of public transit with all-weather access to buses.

Pedestrian accidents often occur at bus stops. This can be attributed to the fact that buses frequently stop too close to the road edge, thus obstructing oncoming drivers' view of pedestrians crossing the road. Under these circumstances, a pedestrian stepping out from behind the bus would be moving directly into the path of an oncoming vehicle.

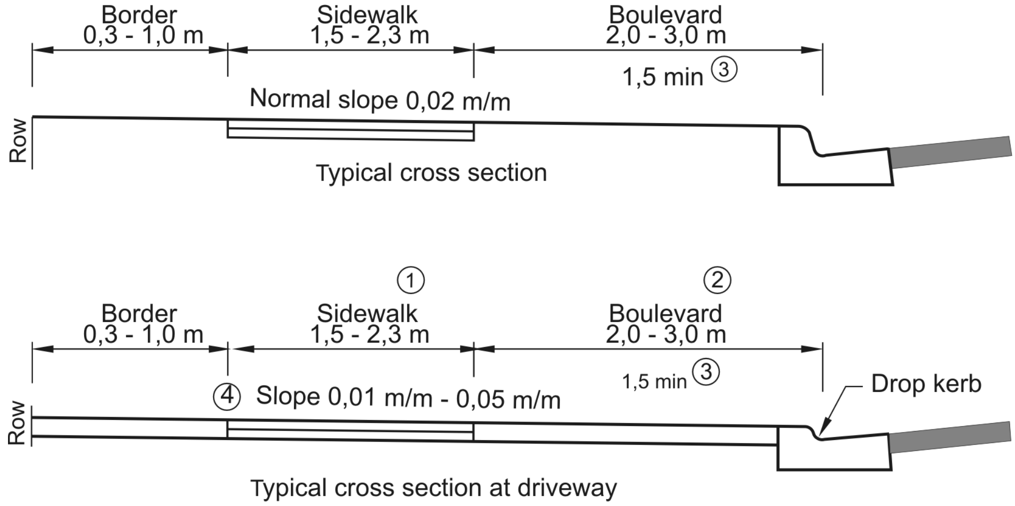


Figure 4.11: Verge area indicating location of boulevard

Two approaches can be adopted to minimise this problem. If the bus stop is provided with adequate entrance and exit tapers, it is easy for buses to move well clear of the travelled way. If space permits, a painted island can be provided between the bus stop and the travelled way so that the stop is, in effect, a short length of auxiliary lane. In addition to an approach aimed at the physical dimensions of the bus stop, a further safety measure could be the provision of barriers preventing bus passengers from crossing the road until they have moved clear of the bus stop itself.

The location of bus stops can have an adverse impact on safety. A bus at a stop located immediately in advance of intersections would force left-turning vehicles into a situation of heavily reduced sight distance. Furthermore, while pulling out of the stop, the bus could seriously influence the operation of the intersection as a whole.

Best practice suggests that bus stops should be located beyond intersections. However, they should not be located more than about fifty metres from the nearest intersection.

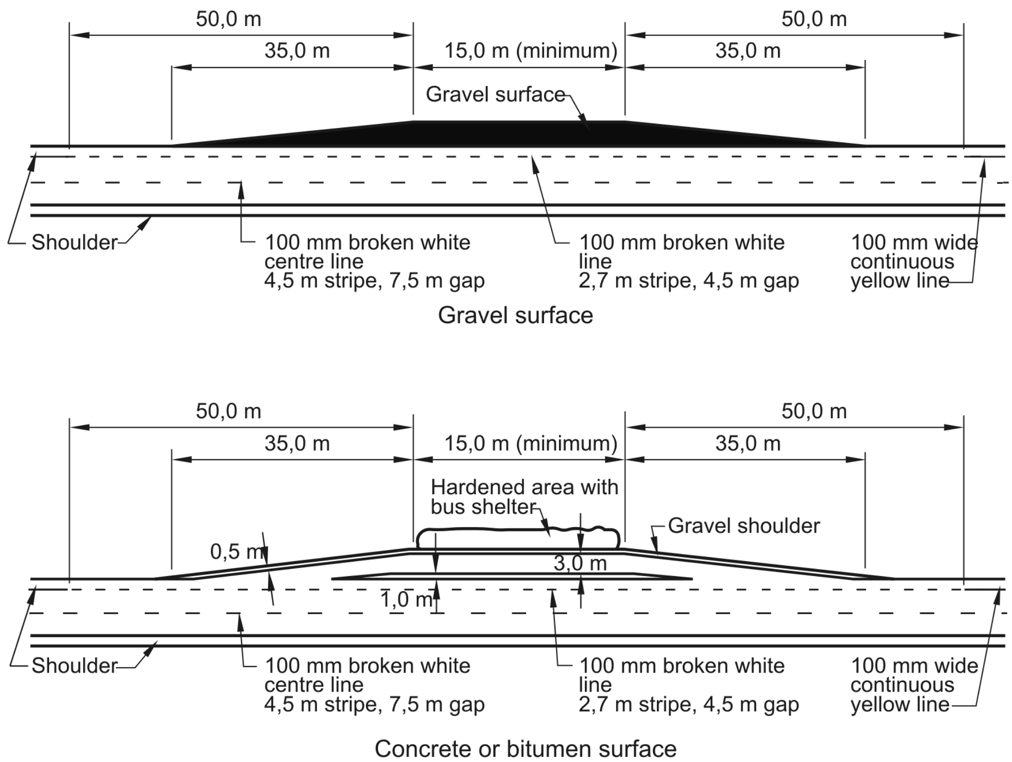


Figure 4.12: Typical layout of a bus stop

A proposed typical layout of a bus stop is illustrated in Figure 4.12. If the frequency of service on a particular road is high, e.g. where two or more bus routes have converged upstream of the bus stop, the length of the bus stop should be increased to 25 metres to accommodate two buses. If necessary, the tapers can be reduced to not less than 1 : 3 for design speeds of 70 km/h or less.

Destinations for bus passengers may be on the bus route itself but are more likely to be to one side or the other of the route. The close proximity of the bus stop to an intersection offers passengers a convenient route to their final destination. However, it is suggested that a bus stop should not be located closer than about 15 metres from the kerb line of the intersecting road or street. A lesser spacing would make it difficult for a left-turning bus to enter the bus stop and, furthermore, may result in encroachment on the sight triangle required by a driver on the intersecting road or street.

4.4.10 Sidewalks

Pedestrian traffic is not encouraged in the road reserves of freeways, expressways or other high-speed arterials and accommodation of pedestrian traffic is usually handled elsewhere. On all other urban streets, pedestrian traffic can be expected and it is necessary to provide sidewalks.

In commercial areas or areas where the road reserve width is restricted, sidewalks may extend from the kerb to the road reserve boundary. As discussed above, there is distinct merit in placing a boulevard between the sidewalk and the travelled way, but pedestrian volumes may be so high that the entire available width would have to be utilised for the sidewalk. The width of a sidewalk should not be less than 1,5 metres and a minimum width of two metres should be provided near hospitals and old-age homes where wheelchair traffic could be expected. If the sidewalk is immediately adjacent to the kerbing, the minimum width should be increased by about 0,6 to 1 metre. This is to make provision for fire hydrants, street lighting and other road furniture. It also allows for the proximity of moving vehicles and the opening of car doors.

The normal cross-slope on a sidewalk is 2 per cent. Cross-slopes steeper than this present a problem to people with walking impairments or who are in wheel chairs. Sidewalks crossing driveway entrances may have to have a steeper cross-slope than this to match the gradient of the driveways but should not exceed a cross-slope of five per cent.

Kerbs, raised medians and channelising islands can be major obstructions to the elderly and people with disabilities, particularly those in wheelchairs. The most common method for minimising the impact of these obstacles is to provide ramps, also referred to as kerb cuts, dropped kerbs or pram dips. Ramps should have a slope of not more than about six per cent. A kerb height of 150 millimetres would thus require a ramp length of 2,5 metres. There should be a clear sidewalk width of 1,5 metres beyond the top end of the ramp so that, where a ramp is provided, the overall sidewalk width should be not less than four metres. Wheel chairs may be 0,75 metres wide so that two wheelchairs passing each other on the ramp would require a ramp width of the order of 2 to 2,5 metres. If it is not possible to provide this width, a width of not less than 1,5 metres should be considered.

The designer should be aware that, in accommodating one group of pedestrians with disabilities, a different group might be disadvantaged in the process. Visually impaired pedestrians would have trouble in locating the kerb face in the presence of a ramp. As a vertical face across the sidewalk would be unexpected, even by sighted pedestrians, the sides of the ramp should also be sloped.

Sidewalks are not normally provided in rural areas. It should, however, be noted that approximately half the fatalities on the South African road network are pedestrians, with many of

Table 4.17: Warrants for pedestrian footways in rural areas			
Footway	Average daily traffic (veh/d)	Pedestrian flow per day	
		Design speed < 80 km/h	Design speed > 80 km/h
On one side of road	400 to 1 400	300	200
	> 1 400	200	120
On both sides of road	700 to 1 400	1 000	600
	> 1 400	600	400

these fatalities occurring in rural areas. Provision should therefore be made for pedestrian safety outside urban areas. Paved footways could be considered under the warranting conditions listed in Table 4.17.

Footways can be as little as one metre wide, but a width of 1,8 metres would allow two people to walk side by side.

The safest location for footways is at the edge of the road reserve. This location is not popular with pedestrians because the footway then follows all the variations in the natural ground level. In rolling or mountainous terrain through cuts and fills, such a footway would not make for comfortable walking. Even if this footway were provided, pedestrians would almost certainly prefer to walk on the more level surface of the shoulder. In level terrain, the footway should, if possible, be situated at least three metres away from the travelled way. This corresponds to a location immediately outside the edge of the usable shoulder in the case of a high volume high-speed road.

In cases where footways are not warranted but where a large number of pedestrians walk alongside the road, the road shoulder should be upgraded to cater for them. The minimum width of these shoulders should be three metres. If not surfaced, they should be bladed and compacted regularly to provide pedestrians with a hard surface to walk on. In high rainfall areas, a portion at least of the shoulder should be paved, with this paved area being at least 1,5 metres wide. Furthermore, the road shoulder should be well drained to prevent the accumulation of water, which would force pedestrians to walk on the carriageway.

4.4.11 Cycle paths

Changes from single to multiple land usage will result in shorter trip lengths, making the bicycle a more popular form of transport. In addition, people are becoming conscious of the need for exercise and of the bicycle as means of exercise. Finally, unlike the motor vehicle, bicycles are environmentally friendly. If adequately planned, designed and maintained, cycle paths can play an important role in the transportation system. It is important to realise that cyclists need sufficient space to operate with safety and convenience rather than simply being assigned whatever space is left over after the needs of vehicular traffic has been accommodated.

The basic requirements of cyclists are:

- Space to ride;
- A smooth surface;
- Speed maintenance, and

- Connectivity.

The bicycle design envelope and clearances are illustrated in Figure 4.13. The one metre envelope allows for the width of the bicycle as well as for erratic tracking. Adequate clearances to fixed objects and passing vehicles should be provided, in addition to the one metre envelope.

Bicycles have narrow tyres inflated to high pressures and have no suspension system to speak of. A smooth surface is therefore desirable for bicycles to be used effectively, comfortably and safely. The surface of a cycle path should not deviate from a three-metre straightedge by more than 5 mm and should also be shaped to existing features to within 5 mm.

For bicycles to be effective as a means of transportation, cyclists must be able easily to maintain a steady speed with ease. Cyclists typically travel at speeds of twenty to thirty km/h sometimes reaching 50 km/h on downgrades. Once slowed or stopped it takes considerable time and effort to regain the desired operating speed. Bicycle routes should thus be designed for con-

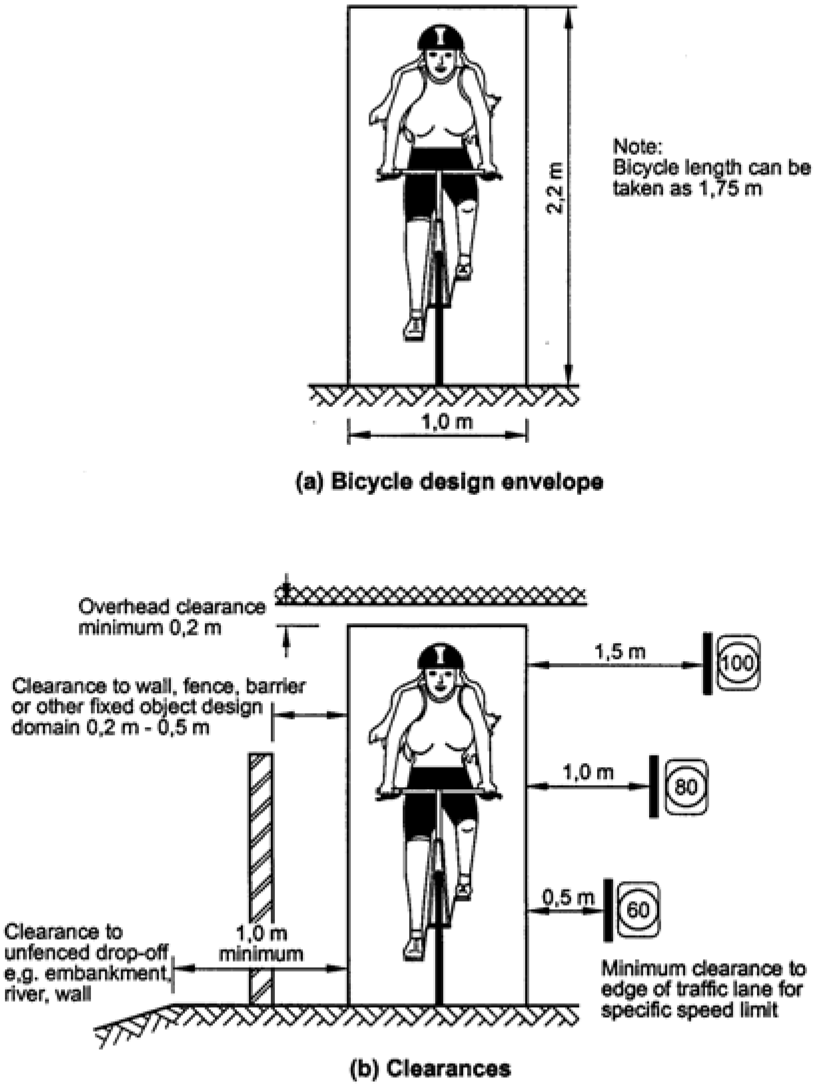


Figure 4.13: Bicycle envelope and clearances

Table 4.18: Cycle lane widths	
Classification	**Lane width**
Two-way exclusive path	2,5 – 3,0
Two-way path shared with pedestrians	3,0 – 4,0
One-way exclusive path or lane	1,5 – 2,0
One-way path shared with pedestrians	2,0 – 3,0
Shared roadway/cycle lane for ADT of	Roadway lane width =
1 – 1 000	4,0
1 000 – 3 000	4,3
3 000 – 6 000	4,0 – 4,5
>6 000	4,3 – 4,8
Note: If ADT exceeds 6 000 and trucks exceed 10 per cent, add 0,5 m If road speed is 100 km/h or greater, add 0,5 m	

tinuous movement, avoiding steep gradients, rough surfaces, sharp corners, intersections or the need to give way to other road users.

It should be possible to undertake and complete journeys by bicycle. Bicycle routes on roadways or separate paths should form a connected network on which bicycle trips can be made effectively and conveniently. Connectivity is an important aspect of effective bicycle routes and should be given careful consideration during the planning process.

Facilities for bicycles can take the form of a:

- Shared roadway/cycle lane where motor vehicles and bicycles travel in a common lane;
- Cycle lane, which is part of the travelled way but demarcated as a separate lane;
- Shoulder lane, which is a smooth, paved portion of the shoulder, properly demarcated by pavement markings or traffic signs. As the shoulder provides a useful area for cycling with few conflicts with fast-moving motor vehicles, this facility is very useful in rural areas.
- Cycle path, which is physically removed from the roadway and from which all motor traffic, with the exception of maintenance vehicles, is excluded.

Cycle paths may be located within the road reserve or in an independent reserve.

Cycle lanes should have the widths indicated in Table 4.18.

The geometry dictated by motor vehicles is generally adequate for bicycles except that bicycles have a longer stopping sight distance.

In the case of separate cycle paths, the bicycle becomes the design vehicle, with dimensions, performance, stopping sight distance, minimum horizontal and vertical curvature, and clearances.

4.4.12 Slopes

The slopes of the sides of the road prism are, like those of medians, dictated by two different conditions. Shallow slopes are required for safety and a slope of 1:4 is the steepest acceptable slope for this purpose. The alternative is to accept a steeper slope and provide for safety by some other means, such as barriers. In this case the steepest slope that can be used is dictated by the natural angle of repose and erodibility of the construction material.

Non-cohesive materials require a slope of 1:2, whereas cohesive soft materials can maintain a slope of 1:1,5. Cuts in firm cohesive materials such as stiffer clays can be built to a slope of 1:1. Rock cuts can be constructed to a slope of 1:0,25 (4:1) provided that the material is reasonably unfissured and stable.

It is stressed that the slopes suggested are only an indication of normally used values. The detailed design of a project should therefore include geotechnical analysis, which will indicate the steepest slopes appropriate to the construction or in-situ material. Economic analysis will indicate the height of fill above which a slope of 1:4 should be replaced by a steeper slope and alternative provision made for safety. As a rule of thumb, the transition from the flat slopes to slopes dictated by the materials typically occurs at a fill height of about 3 m.

4.4.13 Verges

The verge is defined as the area between the longitudinal works and the road reserve boundary. The limit of the longitudinal works in the case of the rural cross-section is often at the top of cut or the toe of fill. Where drainage works, such as side drains or catch water drains, are required, these form part of the longitudinal works, which may thus be wider than the actual road prism. In rural areas the verge simply represents the difference in width between a statutory road reserve and the width actually needed to provide the road and its appurtenant works. Utilities not directly connected with the road, e.g. telephone or power lines are normally located in the verge.

In urban areas, the area between the edge of the travelled way and the road reserve boundary provides space for a variety of elements that, for convenience, are summarised in Table 4.19.

Some of these elements have been discussed previously. The intention is to provide the designer with a checklist of elements that should be accommodated. Some elements will be mutually exclusive. For example, the use of barrier kerbs indicates that mountable kerbs cannot be present. Others may represent a temporary change in cross-section, for example the boulevard being replaced by a bus embayment. Yet others may overlap, for example the driveway approach that crosses a sidewalk.

Elements that are most likely to be accommodated in the verge are;

- Berms intended to function as barriers protecting the surrounding development from visual intrusion or noise;
- Cut and fill slopes;
- Driveway approaches, and
- Underground services.

Even in the unlikely event that none of these elements have to be provided, there has to be a clear space between the edge of the travelled way and the road reserve boundary. This space would provide sight distance in the case of horizontal curvature and also allow for emergency stopping. Furthermore, there should be some flexibility to accommodate future unknowns. It

Table 4.19: Typical widths of roadside elements	
Element	**Width (m)**
Barriers, guardrails	0,5
Berm 1,5 metres high	6,0
Boulevard	3,0
Bus stop/Taxi lay-bye embayment	3,0 – 4,5
Bus passenger queue	0,7 to 1,4
Cycle paths	1,5 to 3,0
Drainage, drop inlet/manhole	1,5
Driveway approaches	5,0
Fill or cut slopes	4,0
Kerb barrier	0,15
Kerb, mountable	0,3
Kerb, semi mountable	0,15
Parallel parking	2,5
Sidewalk	1,5 to 3,0
Street lighting	0,5
Traffic signals or signs	0,6 to 2,0
Trench width for underground services	1,0 minimum

is suggested that this clear space should ideally be a minimum of five metres wide with an absolute minimum width of three metres.

4.4.14 Clearance profiles

The clearance profile describes the space that is exclusively reserved for provision of the road. It defines the lowest permissible height of the soffit of any structure passing over the road and also the closest approach of any lateral obstacle to the road cross-section. Clearance profiles are described in detail in Chapter 10.

4.4.15 Provision for utilities

Both surface and underground utilities are often located within the road reserve. Utilities convey the sense of services not directly related to the road itself. However, to the driver the presence of a pole is a hazard to be avoided and whether the pole is carrying a power line or a streetlight is a matter of indifference. Given this broader approach to utilities, surface utilities typically located in the reserve include:

- Electrical transmission lines;
- Telephone lines;
- Street lighting;
- Traffic signal poles, and
- Fire hydrants.

Underground utilities include:

- Storm and foul water sewers;
- Water reticulation;
- Buried telephone lines;
- Gas pipelines, and
- Power transmission cables.

Most urban authorities have guidelines for the

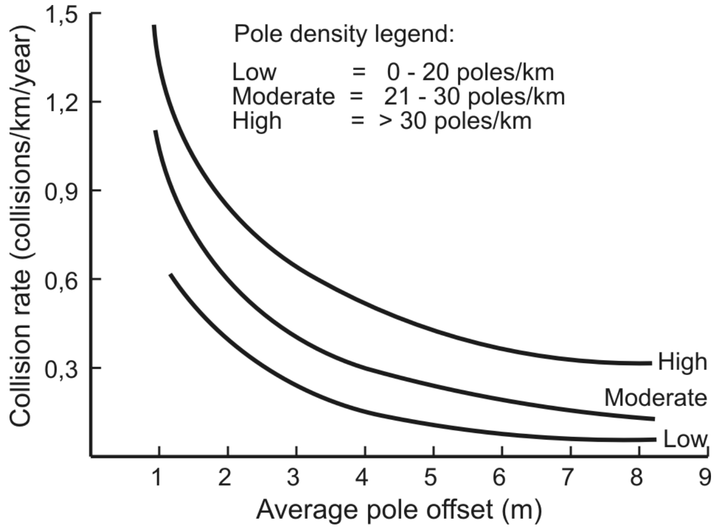


Figure 4.14: Collision rate

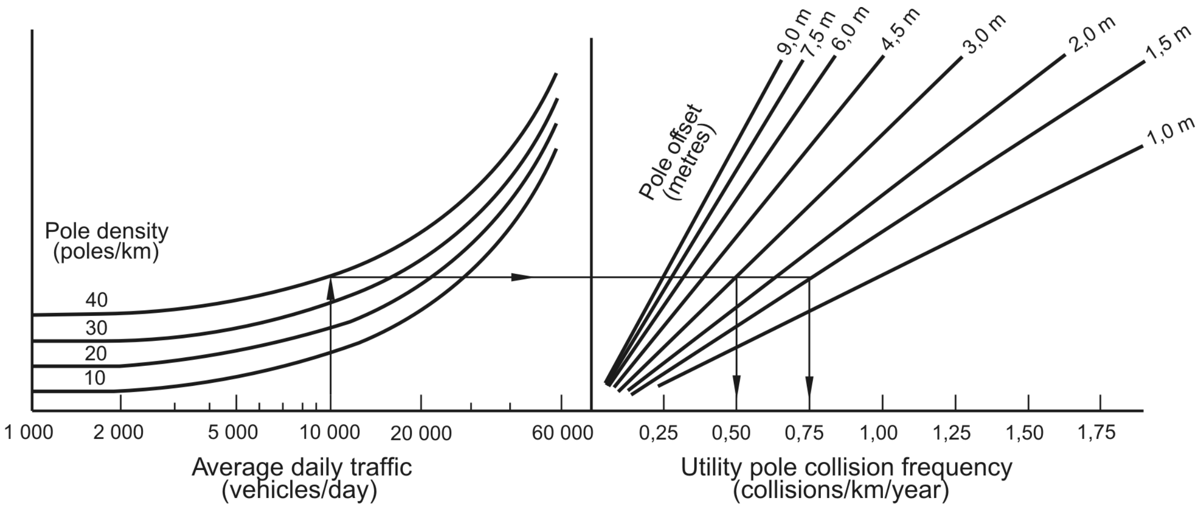


Figure 4.15: Prediction of utility pole crashes

placement of utilities. The use of an integrated process in the planning and location of roadways and utilities is encouraged in order to avoid or at least to minimise conflicts.

As a rule, underground utilities should be located in the verges or boulevards. Access to these utilities is by manholes and open manholes are an unnecessary hazard to pedestrians. In older municipal areas, services were sometimes located under the roadway itself. This practice should be actively discouraged as it places both maintenance workers and passing vehicles at risk. Every time the road is resurfaced, it is necessary to remove the manholes and replace them at the new level. This operation carries an element of risk but, if not carried out, the manhole is at a lower level than the road surface and the drop could be sufficient for a driver to lose control of the vehicle. The lower level of the manhole would, during rainy weather, lead to the creation of a pond of water that could slowly drain into the conduits of the buried utility, possibly leading to disruption of the service being provided.

The problem with surface utilities is that they are carried on poles that can be hit by errant vehicles. Research has indicated that the frequency of crashes is a function of the pole density in poles per kilometre and the average pole offset from the travelled way. The crash frequency is typically of the order of 0,1 crashes per kilometre per year with a pole spacing of less than 20 poles per kilometre and an offset of eight metres. When the pole density is higher than 30 poles per kilometre and the offset less than one metre, the collision rate climbs to a high of 1,5 crashes per kilometre per year. This is illustrated in Figure 4.14.

A nomograph predicting utility pole crashes is given in Figure 4.15.

The example shows that a road with an ADT of 11 000 vehicles and a pole density of 40 poles per kilometre will experience 0,75 crashes per kilometre per year if the pole offset is 1,5 metres. If the designer were to increase the pole offset to 3 metres, the crash rate would reduce to 0,5 crashes per kilometre per year, an improvement of 33 per cent.

4.4.16 Drainage elements

The process of design of storm water drainage systems is exhaustively discussed in the South African Roads Agency Drainage Manual, 1986.

In this section, discussion is limited to the elements that the designer should incorporate in the cross-section to ensure adequate drainage of the road reserve and the adjacent land.

A distinction is drawn between rural and urban drainage. Rural drainage focuses largely on the swift removal of storm water from the travelled way onto the verge and on its movement to a point where it can be taken from the upstream to the downstream side of the road. In an urban environment, the road reserve serves as the principal conduit of storm water from surrounding properties and its conveyance to a point where it can be discharged into natural watercourses. Rural drainage is, in short, the removal of water from the road reserve whereas urban drainage attracts water to the reserve.

In both rural and urban environments, storm water drainage is aimed both at the safety of the road user and the integrity of the design layers of the road. Previously, rural drainage was exclusively directed towards safeguarding the design layers. It was previously common practice to recommend a minimum depth of drain. The safety of the road user dictates rather that a maximum depth of drain be specified. The recommended maximum depth is 500 mm. The volume of water to be conducted by a drain thus indicates the required width of the drain rather than its depth, since the need to keep the design layers unsaturated has not changed.

Rural drainage

Rural drainage is normally by means of unpaved open drains, which may either silt or scour, depending on the flow speeds in them. Both silting and scouring of a drain increase the hazard to the road user. Scour would lead to the creation of a deep channel that would be impossible to traverse with any degree of safety. It may also cause erosion of the shoulder and ultimately threaten the integrity of the travelled way itself. Silting may block the drain, so that water that should have been removed would be discharged onto the road surface.

The effectiveness of the drain depends on water speed, which is a function of longitudinal slope, as well as of other variables. There is a range of slopes over which the flow velocity of water on in-situ materials will be so low that silting occurs, and another range where the flow velocity will be high enough to cause scour. On slopes between these two ranges neither silting nor scouring will occur and unpaved drains will be effective.

Paving solves some of the problems caused by both silting and scouring. Paving generally has a lower coefficient of roughness than in-situ materials, so that water speeds are higher in a paved drain than in an unpaved drain with the same slope. Furthermore, it is possible to force higher speeds in the paved drain by selection of the channel cross- section. The problem of silting can be resolved, at least partially, by paving the drain.

The flow velocity below which silting is likely to occur is 0,6 m/s. Flow velocities above which scour is likely to occur are given in Table 4.20. Conventional open-channel hydraulics will, in conjunction with Table 4.20, indicate when either silting or scouring is likely, and hence whether it is necessary to pave a drain or not.

As a rough guide to longitudinal slopes, it is suggested that unpaved drains should not be steeper than 2 per cent, or flatter than 0,5 per cent. Paved drains should not be flatter than 0,3 per cent. Practical experience indicates that it is difficult to construct a paved drain accurately to the tolerances demanded by a slope flatter than 0,3 per cent, so that local imperfections may cause silting of an otherwise adequate drain.

Where the longitudinal slope is so flat that self-cleansing water speeds are not achieved, even with paving, it will be necessary to consider a piped drainage system.

As an alternative to lining a material subject to scour, it is possible to reduce flow velocity by constructing weirs across an unpaved drain. The drain will then in effect become a series of stilling basins at consecutively lower levels. While this could be an economical solution in terms of construction cost, it has the disadvantage that an area of deep localised erosion, immediately followed by a stone-pitched or concrete wall, would confront an errant vehicle. If this alternative is to be considered at all, it should be restricted to roads with very low traffic volumes and the weirs should be spaced as far apart as possible.

Drains constructed through in-situ materials generally have flat inverts so that, for a given flow, the flow velocity will be reduced. The flat inverts reduce the possibility of scour and are easy to clear if silting occurs. Paved drains, not being susceptible to scour, have a V-profile. Self-cleansing velocities are thus achieved at relatively small flows and the need for maintenance is reduced.

reduce the likelihood of a vehicle digging its front bumper into the far side of the drain and somersaulting.

Typical drain profiles are illustrated in Figure 4.16.

Table 4.20: Scour velocities for various materials

Material	Maximum permissible velocity (m/s)
Fine sand	0,6
Loam	0,9
Clay	1,2
Gravel	1,5
Soft shale	1,8
Hard shale	2,4
Hard rock	4,5

The sides of the drain should not be so steep as to be dangerous to the road user; a maximum slope of 1:4 is recommended. Ideally, both sides of the drain should be designed to this slope or flatter. Where space for the provision of the drain is restricted, the slope closest to the road should remain at 1:4 and the outer slope steepened. This has the effect of positioning the drain as far as possible from the path of vehicles. One example of this is a side drain in a cut, where the outer slope of the side drain forms an extension of the cut face. These slopes, in combination with the flat invert, give the trapezoidal profile of an unpaved drain.

It is recommended that the bottom of a lined V-profile and the junctions between the sides and bottom of an unlined trapezoidal profile be slightly rounded. The rounding will ease the path of an errant vehicle across the drain, and

(a) Side drains

Side drains are located beyond the shoulder breakpoint and parallel to the centre line of the road. While usually employed in cuts, they may also be used to run water along the toe of a fill to a point where the water can conveniently be diverted, either away from the road prism or through it, by means of a culvert. When used in conjunction with fills, side drains should be located as close to the edge of the reserve boundary as is practicable to ensure that erosion of the toe of the fill does not occur. Side drains are intended as collectors of water and the area that they drain usually includes a cut face and the road surface.

(b) Edge drains

Edge drains are intended to divert water from fill slopes that may otherwise erode either because

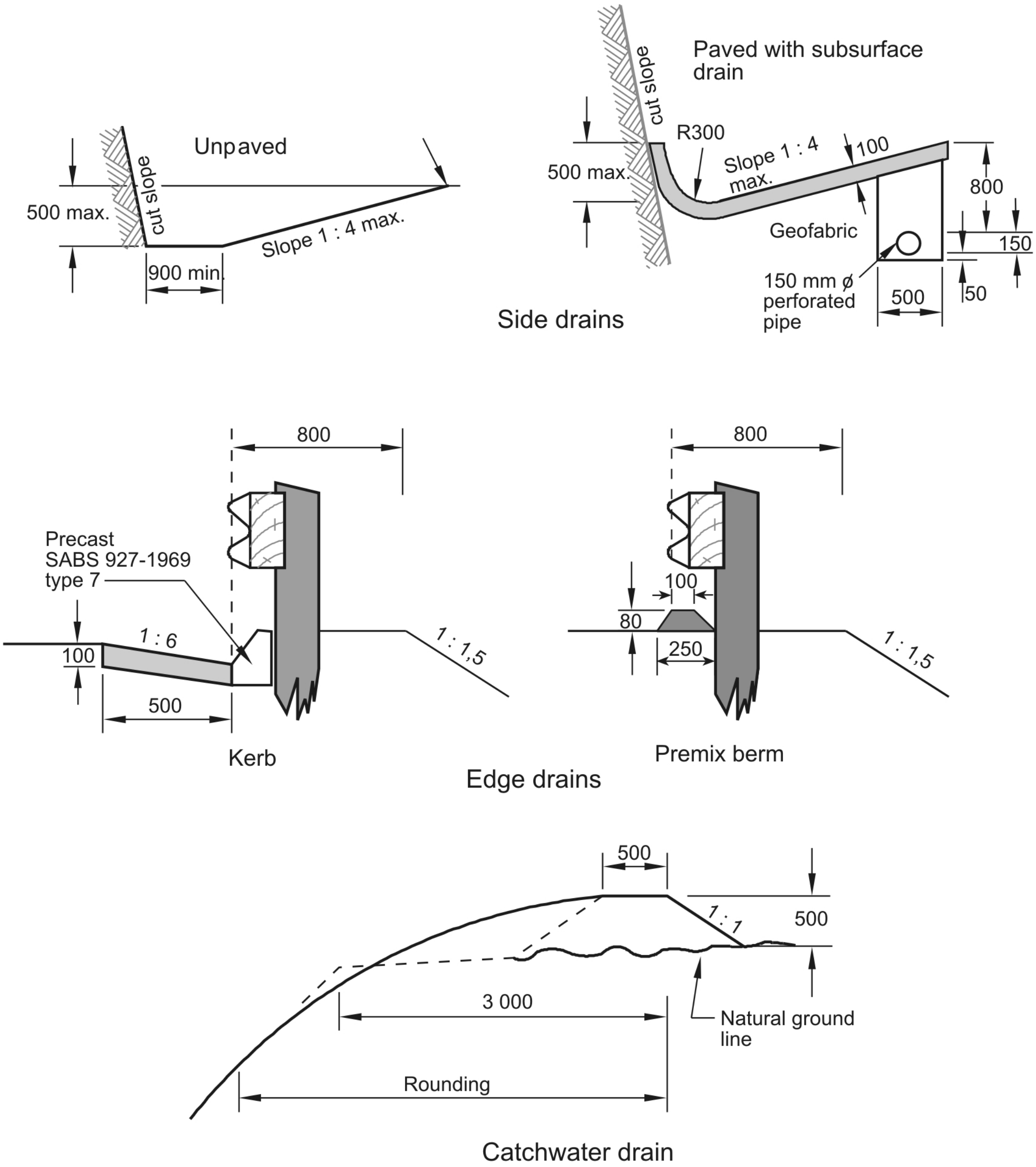


Figure 4.16: Typical drain profiles

of the erodibility of the material or because they are subjected to concentrations of water and high flow velocities. Guardrail posts tend to serve as points of concentration of water, so that, as a general rule, edge drains are warranted when the fill material is erodible or when guardrails are to be installed.

Edge drains should preferably be raised rather than depressed in profile. A depressed drain located almost under a guardrail would heighten the possibility that a vehicle wheel might snag under the guardrail.

Edge drains are constructed of either concrete or premixed asphalt. Premix berms normally have a height of 75 to 80 mm, and are trapezoidal in profile with a base width of 250 mm and a top width of 100 mm. Concrete edge drains are normal barrier kerbs and channels. These require a properly compacted backing for stability and are, therefore, not as easy to construct as premix berms.

(c) Catch water drains

The catch water drain, a berm located at the top of a cut, is to the cut face what the edge drain is to a fill. It is intended to deflect overland flow from the area outside the road reserve away from the cut face. Even if the cut is through material which is not likely to scour, the catch water drain serve to reduce the volume of water that would otherwise have to be removed by the side drain located at the bottom of the cut face.

Catch water drains are seldom, if ever, lined. They are constructed with the undisturbed topsoil of the area as their inverts and can readily be grassed as a protection against scour. Transverse weirs can also be constructed to reduce flow velocities, since the restrictions previously mentioned in relation to weirs do not apply to catch water drains. The cut face and the profile of the drain reduce the probability of a vehicle entering the drain but, should this happen, the speed of the vehicle will probably be low.

(d) Median drains

Median drains do not only drain the median but also, in the case of a horizontal curve, prevent water from the higher carriageway flowing in a sheet across the lower carriageway. The space available for the provision of median drains makes it possible to recommend that the transverse slopes should be in the range of 1:4 to 1:10. If the narrowest median recommended is used, a transverse slope flatter than 1:10 may make it difficult to protect the design layers of the road. Unlike side drains, median drains, whether lined or not, are generally constructed with a shallow V-profile with the bottom gently rounded.

(e) Chutes

Chutes are intended to convey a concentration of water down a slope that, without such protection, would be subject to scour. They may vary in size from large structures to half-round pre-cast concrete products, but they are all open channels. Flow velocities are high, so that stilling basins are required if down-stream erosion is to be avoided. An example of the application of chutes is the discharge of water down a fill slope from an edge drain. The entrances to chutes require attention to ensure that water is deflected from the edge drain into the chute, particularly where the road is on a steep grade. It is important that chutes be adequately spaced to remove excess water from the shoulders of the road. Furthermore, the dimensions of the chutes and stilling basins should be such that these drainage elements do not represent an excessive risk to errant vehicles. Generally, they should be as shallow as is compatible with their function and depths in excess of 150 mm should be viewed with caution.

Because of the suggested shallow depth of chutes, particular attention should be paid to their design and construction to ensure that the highly energised stream is not deflected out of the chute. This is a serious erosion hazard that can be obviated by replacing the chute with a pipe.

(f) Mitre banks

As their name implies, these banks are constructed at an angle to the centre line of the road. They are intended to remove water from a drain next to the toe of a fill and to discharge it beyond the road reserve boundary. Several mitre banks can be constructed along the length of a drain, as the concentration of water in the drain should ideally be dispersed and its speed correspondingly reduced before discharge. Speed can be reduced not only by reducing the volume, and hence the depth, of flow but also by positioning the mitre bank so that its toe is virtually parallel to the natural contours. The upstream face of a mitre bank is usually protected by stone pitching, since the volume and speed of flow of water that it deflects may cause scour and ultimately lead to breaching of the mitre bank.

(g) Rural underground systems

The geometric designer is not directly concerned with the underground system, except for its inlets. These should be hydraulically efficient and correctly positioned to ensure that water does not back up onto the road surface or saturate the design layers. To restrict the hazard to the road user, inlets that are flush with the surface drain invert are preferable to raised structures.

Underground reticulation is costly both to provide and to maintain. The designer should therefore, without violating the principles discussed above, attempt to reduce the use of underground drainage as far as possible by the discerning use of surface drainage.

Urban drainage

Urban drainage entails the provision of protection from major and minor storms and the basic requirements for these forms of protection are usually conflicting. For major storms, the runoff should be retarded to reduce flood peaks and, for minor storms, the runoff is best handled by rapid removal. The solution is to provide two separate but allied drainage systems.
ajor storms involve considerations such as attempting to achieve rates of runoff that do not exceed pre-development levels. This is achieved in part through the application of detention and retention ponds. Furthermore, the layout of road patterns can be coordinated with the run-off requirements of the system to increase the time of concentration, hence reducing the risk of flood hazards

Accommodation of minor storms is achieved with kerbs and channels, drop inlets and underground reticulation. The runoff is initially collected in channels until the flooded width of the road reaches a specified limit and then discharged into the underground system, which is connected to an outfall point - typically a natural watercourse. In the case of high-speed routes, such as freeways, expressways and major arterials, no encroachment of storm water onto the travelled lanes can be considered. Minor arterials and collectors should have one clear lane of 3,0 metres minimum width in each direction and local streets need only have one clear lane of 3,0 metres minimum width. In all cases, the 100-year storm should not cause a barrier kerb to be overtopped. The designer should ensure, therefore, that the resultant of gradients and crossfall is sufficiently steep and the spacing of drop inlets sufficiently short to ensure that the recommended lateral spreads of water are not exceeded.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 5

ALIGNMENT DESIGN

5.1 INTRODUCTION

Most of the chapters in this guide assist the designer by describing the principles on which design parameters are based and by offering ranges and guidance in the selection of suitable standards. In this section an attempt is made to stand away from the detail and discuss highway design holistically.

South Africans are blessed with a beautiful country. There are thus boundless opportunities to develop highway location and design as an art form in this country's varied landscape from mountains and plains to deserts and seas. When properly applied, there can be benefits both to the user and to the landscape.

Creating a harmonious alignment is architecture. It requires the ability to visualize the final form both from the perspective of the driver and the outside observer as well as a grounding in engineering principles The ability to capture the final form in the mind's eye and translate the design into two-dimensional representations defining the layout represents creativity at a high level. Most texts on the subject of the aesthetics of alignment design stress the role of experience. Undoubtedly, we all learn as we grow in the profession. However, in our society at present we do not have the luxury of widely available, experienced engineers. Designers often have to tackle difficult problems for the first time and at short notice. In what follows we have therefore concentrated on describing what it is important to know and on giving some guidance as to what is, and what is not, good practice.

5.2 VISION

The eye is a truly remarkable instrument sensitive enough for the dark-adapted eye to detect a single photon. It is also backed up by enormous computing power and a memory bank of visual images. Although it is capable of processing fine detail in real time, it remains subject to the laws of physics.

Despite the difference in scale, the stars overhead can be brighter than the spark from a campfire. This demonstrates better than words that seeing depends on the energy of the light received by the eye. Also there has to be a reasonable contrast between the source and the background. Stars are not visible during the day. In daylight, a contrast in brightness of five per cent can be perceived. At twilight, much greater differences are required. The cones of the eye's retina which perceive colour also require significant light energy and do not operate at low levels at night. The yellow and green colours at the centre of the spectrum appear significantly brighter than the reds and blues which have longer and shorter wavelengths respectively.

Although very quick, our visual response is not immediate. Approximately 0,2 of a second is required to fix on an object. However, in a changing environment we can at best process twelve images per second. Changing focus from the dashboard clock to a road sign takes about one second. From there, reaction time requires 1,3 seconds thus leading to a final safe reaction/response time of 2,5 seconds. A complex environment, where the desired response is not immediately obvious, may require a significantly longer reaction time as discussed in Chapter 3.

The driver sees the road and surrounds as if he were stationary in a 3D movie. Nearby objects speed past in a blur. Objects in the foreground can only be seen briefly. Only objects at the infinity focus can be scanned at leisure.

Under conditions of good lighting, people with good eyesight can observe an object that subtends an angle of one arc minute or 29 millimetres at 100 metres. Lines can be perceived more acutely. However, they should subtend an angle of at least 4 arc seconds to be visible. People with good eyesight are not the norm and allowance has to be made for lower acuity when designing objects and messages that are important to the driver in a high-speed environment and a constantly changing visual field.

Because of the ability of both the head and eyes to swivel, it is difficult to place boundaries on the driver's field of view. However, the sensitive part of the retina at the centre of the visual field has a relative arc of 2,5 degrees. The peripheral vision lies outside this cone where little detail can be perceived. The visual field also diminishes the more finely we focus.

From these general principles of vision, J R Hamilton and L L Thurston (in a paper entitled "Human Limitations in Automobile Driving" published in 1937) enunciated five propositions that are applicable to the highway environment. These were summarized in "Man Made America" by Tunnard and Pushkarev (1963), which should be required reading for all highway designers and are further summarized below:

1. As speed increases, the number of objects and incidences that must be reacted to increases proportionately. Concentration focuses on the approaching road and traffic in the immediate vicinity. Observing irrelevant objects outside of the necessary area of attention becomes more and more dangerous. It follows that the driver will only be able to see interesting objects in the centre of his visual field and therefore the road should aim the eye towards objects of interest and create variety through curvature.Planes that stand perpendicular to the road are prominent while parallel ones are not.
2. As speed increases the eyes seem to focus further and further ahead. Drivers anticipate the distance ahead that they will require to respond to emergencies.At 70 km/h the focal point of the eye is approximately 400 metres ahead while at 120 km/h the focal point can be up to 1 000 metres ahead. From the principle of visual acuity, it follows that anything that has to be brought to the driver's attention must lie close to the axis of vision and also be large enough to be recognized at a long distance.
3. As the level of concentration increases, the total visual field decreases with the result that the peripheral vision diminishes as speed increases. This is sometimes referred to as tunnel vision and, unless the point of concentration is made to move through an arc by means of a curving roadway, driving along a straight

and uneventful highway can become-hypnotic.

4. As speed increases, foreground features begin to fade because the driver is not able to see clearly except at a distant focus. Foreground detail is greatly diminished at 80 km/h, and beyond 100 km/h reception of the foreground is negligible. Thus, only at a distance of between 50 metres and 100 metres does vision become adequate at 100 km/h. It follows that emphasizing elaborate detail is meaningless for the driver. Only large simple shapes are usually perceived and particularly the geometry of the paved road at the centre of the driver's vision. Outside the road, only the general outline and form of the land together with objects on the horizon are distinguishable.
5. Space perception also becomes impaired as speed increases. This is a complex subject and is related to the fact that we cannot perceive small relative changes in objects at long distances. A person requires clues from the surrounding landscape to perceive motion. The movements of objects travelling parallel and closest to the axis of vision cannot be perceived beyond 250 metres on either side. As speed increases the time interval between first discerning movement and passing the object reduces. It follows therefore that the highway should offer as many clues as possible to allow the driver to judge his speed and remain in tune with reality as the space around him changes continuously.

When taken together, these principles confirm that, at the high speeds on main and freeway networks, drivers have to concentrate in order to survive. They must focus as far ahead as possible to anticipate the approaching vehicles and changes in alignment. They cannot cope with sudden events in the foreground and both their peripheral vision and space perception are impaired.

With modern virtual reality techniques it is also possible to show that, at 120 km/h, the paved area and median across a 30 metre wide freeway takes up thirty per cent of the visual field, the roadside about fifteen per cent, and the sky dominates at fifty five per cent. When the roadway is only 15 metres wide at somewhat lower speeds the roadway takes up about fifteen per cent, the sky approximately twenty five per cent, and the roadside the remaining sixty per cent.

This leads us to the conclusion that our approach to aligning a divided freeway should be very different from that for a two lane main road. For freeways, the pavement and its median dominate and the designer should focus his attention on the architecture of the alignment as it is moulded into the land form. On the other hand, for the narrow pavements typical of much of our network, the opportunity to use the road as a platform for viewing the environment can be used to dramatic effect in road architecture.

5.3 DESIGN APPROACH

In engineering terms, the geometric elements of the highway are simple. We use tangents, horizontal curves, grades and vertical curves. However, the use of these elements in combination can be nearly infinite: The result can be pleasing or discordant; economical or expensive; safe or dangerous.

The two main factors that dominate alignment design are the purpose of the road and the site. The standards for a particular facility are usually determined at the early planning stage and are based on the highway's position in the road hierarchy and on the budget available. This leads to a definition of standards in terms of width, design speed, road surface, maximum gradient, accessibility and drainage standards as input to the initial design.

With these standards in mind, it is then the task of the designer to understand the site as holistically and in as much detail as possible. We need to know about:

1. The landform and how it was sculptured by nature;
2. The underlying geomorphology;
3. The engineering properties of the surface materials including erodibility;
4. The cadastral layout and ownership patterns;
5. The use of the land and the natural vegetation;
6. The normal and extreme weather cond - ions;
7. The drainage patterns, streams and peak discharges of the important catchments including their silt load;
8. Sensitive environmental areas that require protection or preservation;
9. The location of road building material;
10. The need for access to and through the road including an understanding of local circulation; and
11. The history, competency and funding of road maintenance in the region.

In short, context-sensitive information is required which should be collated and summarized on plans and in tables to build up an understanding of the site and used to underpin the design trade-offs that will invariably follow. Decisions such as which side of a ridge to follow, where and at what angle to cross a river, where material can be borrowed to support a grade separation structure, where intersections or interchanges can be safely located, are all commonplace when a road alignment is designed. As they are also usually inter-related, comprehensive information is essential to effective and efficient design.

With the purpose of the road established and the opportunities and constraints understood, the next step is to identify broad corridors or avenues for further investigation both from the information gathered and from on-site reconnaissance. Within each corridor, sketch-planning techniques can be employed, ideally at a scale of 1:5 000 to enable the alternatives to more closely examined. The ranking of the preferred options that flow from this examination should be rigorous and scientific. It should deal with all issues, including aesthetics, engineering, socio-political issues and the life cycle cost. This process, with its plethora of feedback loops and consultation, invariably leads to the selection of one preferred corridor, or at most two, in which an alignment can be designed at a scale suitable for final construction drawings. This process should be followed, even if there is only one obvious and preferred corridor, as it is part of the process of understanding the site.

5.4 FITTING THE ROAD TO THE LANDSCAPE

The appearance of the road to the traveller and its appearance in the landscape depend on how we string the tangents, grades and curves together. These elements form an inclined plane, a simple geometric form common to all engineering design. However, nature did not create the surface of the earth using pencil and paper. The inclined plane is not nature's way of folding the landscape, forming escarpments and incising the rivers. A road with its rigid geometric form made up of vectors and circular arcs is thus a discordance in nature. It is the designer's task to minimise that discord. This can be done by the judicious use of curvature.

Returning to the earlier discussion on what the eye sees at speed and how vehicles behave, we concluded that there is a vast scale difference between a divided freeway and a two-lane road. The former can easily be 35 to 40 metres wide while the other is typically 13,7 metres wide or at most 15 if there are sidewalks or surfaced drains. On a freeway, the paved ribbon dominates the driver's view, whereas, on a two-lane road, the roadside takes up most of the visual field. It follows, therefore, that the design approaches for these two cases need to be very different indeed.

This difference in approach is captured in the twin concepts of internal and external harmony of the alignment.

Internal harmony describes the drivers' view of the road itself as the centreline rises and falls or changes direction and the road edges rise and fall relative to the centreline. The notion is thus that of an abstract ribbon in space, which the designer should be able to visualise from study of the survey plan, the longitudinal section and the cross-section. As a rule, the most pleasing appearance results when the horizontal and vertical curves are of approximately equal length and in phase with one another.

External harmony describes the match or mismatch between the road and its environment. The achievement of external harmony results in the road being an enhancement of the landscape rather than a scar across it. An example of the latter is a long tangent at right angles to the natural contours and with a succession of short crest and sag curves closely following the ground line. This results in a roller coaster appearance totally at odds with the form of the landscape.`

5.4.1 Freeway design

In locating a freeway or divided roadway, the appearance of the road as a ribbon in the landscape from the viewpoint of the driver is crucial. To create a continuous and homogenous appearance, sudden breaks, kinks or abrupt changes in the alignment must be avoided. This invariably requires the use of above minimum standards for horizontal and vertical curves and the use of the clothoid spiral to form transitions between tangents and curves. To create interest in a changing landscape the designer should strive for the curvilinear or 'splined' alignment. Figure 5-1 illustrates a range of acceptable curve lengths for varying deflections and radii that are visually desirable. Curve radii of up to 18 000 metres with lengths in excess of 4 000 metres have been used successfully in South Africa.

Visualising the roadway from the three dimensional viewpoint of the driver is important. The road that flows through the landscape, avoids major obstacles and is in scale with the terrain should be the designer's objective.

The co-ordination of the vertical and horizontal alignment ensures that the scale of the plan and profile view are in harmony. However, it is not always possible to ensure that horizontal and vertical curves coincide. When these elements are out of phase or out of scale the designer should take particular care to avoid unpleasant effects. Several examples of good and bad practice are illustrated and discussed in Section 5.4.3.

Although the aesthetics of the 3-D alignment of the road are enhanced by having the vertical curves contained within the horizontal curves, the designer must be aware of the optical illusion that results from this combination. A crest curve causes a horizontal curve to appear to have a larger radius than it has in reality and a sag curve causes the horizontal curve to appear sharper than it really is. The lower the K-value of the vertical curvature, the more pronounced the effect is. If the horizontal curve has a minimum radius, the crest curve could tempt a driver to maintain a speed higher than is safe whereas the sag curve could lead to unnecessary and sharp braking which would not necessarily be anticipated by following drivers. Correct phasing of the vertical and horizontal alignment should thus be accompanied by the use of horizontal radii that are well above the minimum for the design speed of the road.

A useful tool for analysing the curvilinear nature of an alignment is the 1/R diagram in which the inverse of the radius is plotted against the centreline distance. In the diagrams shown in the following figures tangents have a zero value and

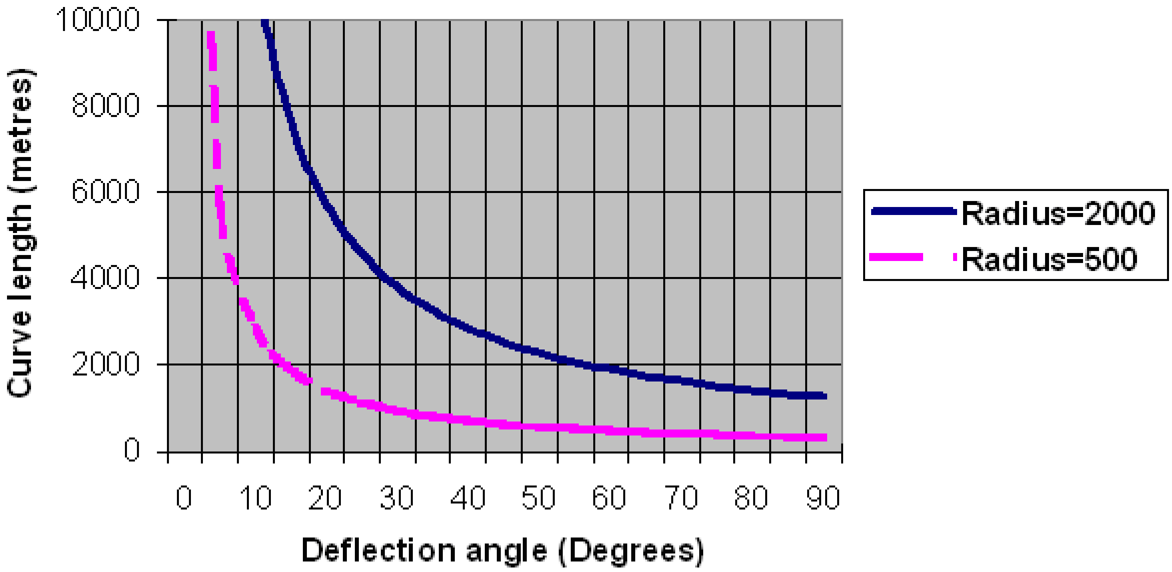


Figure 5-1: Radii for acceptable curve lengths

the curves a constant value that is inversely proportional to the radius. Transitions appear as sloping lines.

The examples in Figures 5-2 to 5-4 illustrate how 1/R diagrams are used to visualise the curvilinear nature of an alignment. Not only is the disjointed nature of the alignment shown in Figure 5-2 immediately apparent but the total area enclosed by the 1/R line and the horizontal

Figure 5-5 illustrates alternative 1/R diagrams for a group of successive tangents employing:

1. Minimum radius curves;
2. Curves approximately equal in length to the adjacent tangents; and
3. Long curves with the intervening tangents only long enough to accommodate superelevation development

A fourth 1/R diagram illustrates the effect of

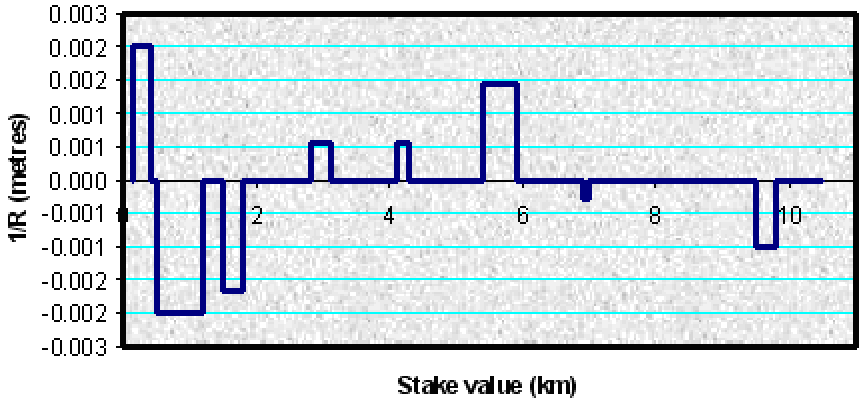


Figure 5-2: 1/R diagram for short curve/long tangent alignment

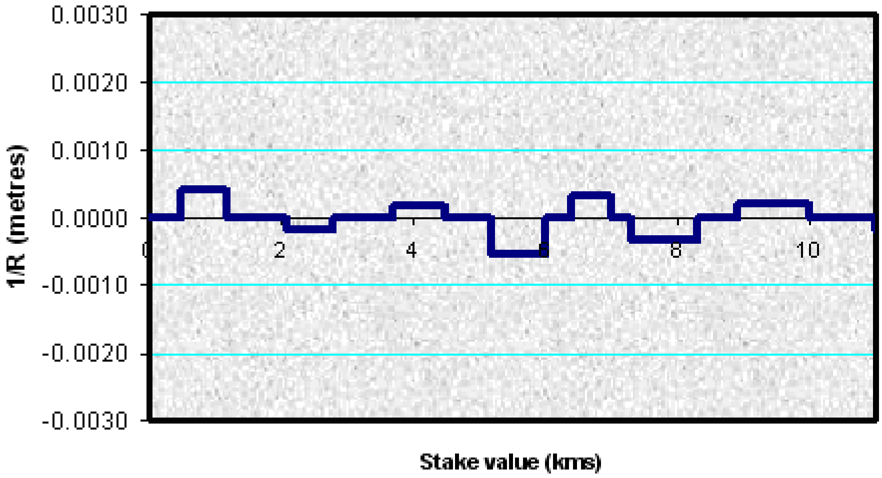


Figure 5-3: 1/R diagram for long curve/short tangent alignment

axis in Figure 5-4 is significantly less than the same areas in Figures 5-2 and 5-3.

removing the broken-back curve between SV 1 800 and SV 6 000. Depending on topographic

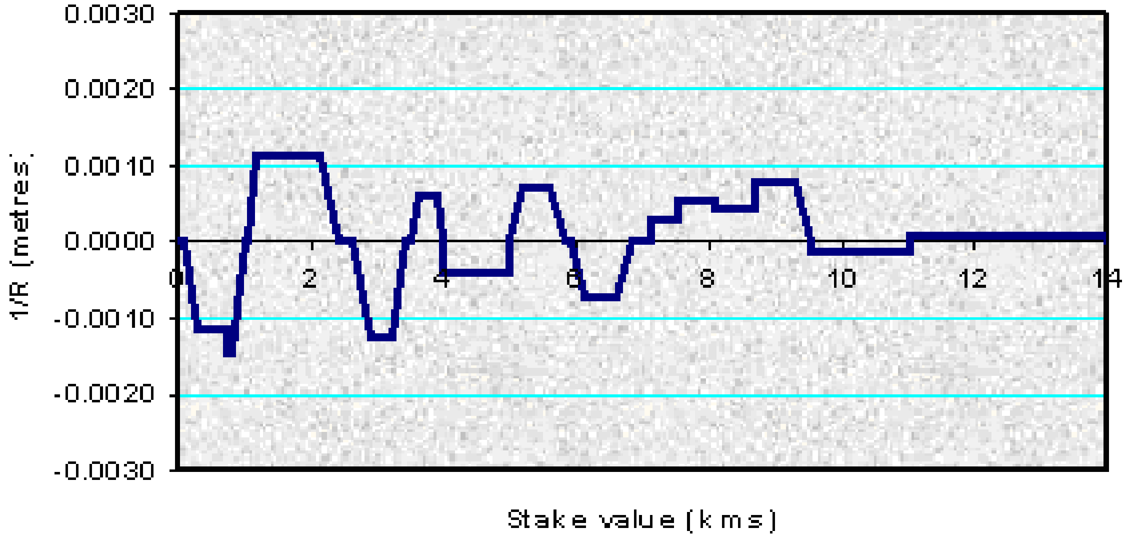


Figure 5-4: 1/R diagram for curvilinear alignment with transitions

restraints, this alignment would be the preferred option. It illustrates two sets of true S-curves, being reverse curves with equal radii. The first S-curve has radii of approximately 3 000 metres and the other radii of approximately 1 100 metres.

Relating the areas to the deflection angles and the remaining tangents by simply making the curves as long as possible is not, of course, sufficient for aesthetic design. It is essential that the road flow with the landscape. When approaching an escarpment from rolling terrain the curve radii and curve length should be gradually decreased in the transition zone until the road enters the pass. This continuity of alignment enhances both the aesthetic and operating characteristics by reducing the element of surprise. It is under these conditions that co-ordi-

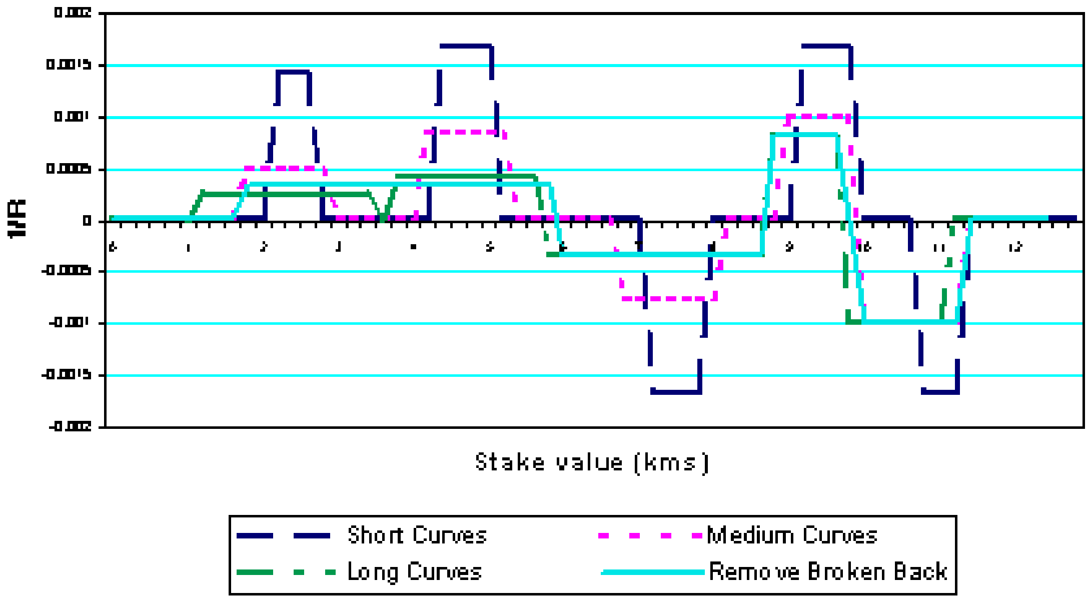


Figure 5.5: 1/R diagrams for alternative alignments

nating the horizontal and vertical alignment is particularly important, as the gradients will also change in keeping with the horizontal alignment.

Treatment of the median can play an important role in the design of a divided highway. Medians that are less than 10 metres wide should never be narrowed and should usually be treated as unifying the two carriageways. When median widths of 15 metres or more are used, the carriageways can be treated independently and some variation of median width or even separation of the carriageways can assist in fitting the highway to the landscape.

5.4.2 Single carriageway roads

The advantages of curvilinear, co-ordinated alignments apply equally well to single carriageway roads. However, because the carriageway is not as wide as that of a dual carriageway, geometric standards that are well above the minimum are not demanded by the scale of the road but are certainly not precluded.

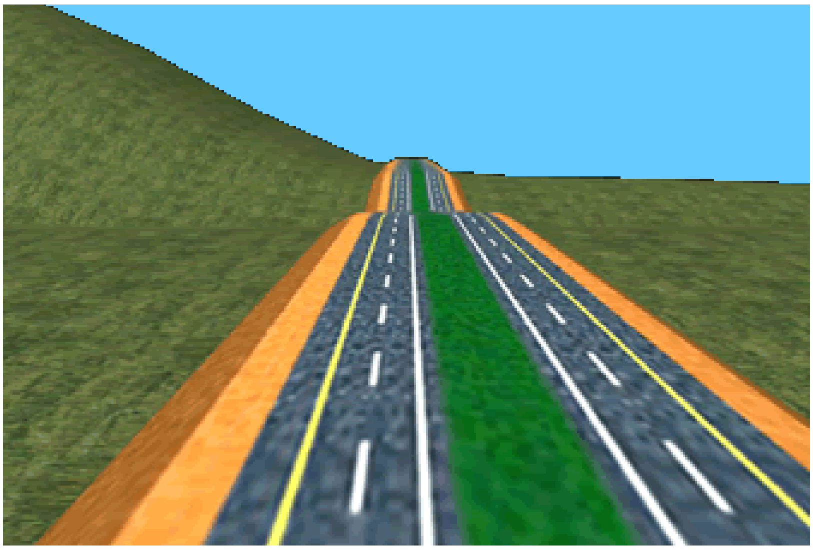
In designing the road using curves to fit the landscape, the chosen radii should allow, wherever possible, for passing sight distance. This may require the day lighting of shallow cuts. With sufficiently long radii, proper sight distance can be achieved. The alternative method of achieving an adequate percentage of passing sight distance is by reducing the radius and hence the length of the horizontal curve enabling provision of passing sight distance on the adjacent tangents. The result may provide adequate sight distance but possibly at the cost of aesthetics.

As the roadside takes up more than half the driver's visual field when driving on a single carriageway road, there is significant benefit to be gained by maintaining the grade line above ground level for as much of the alignment as possible. Roadways elevated in this way create greater interest and improve overall visibility in addition to the obvious benefit of having sufficient height to accommodate drainage structures.

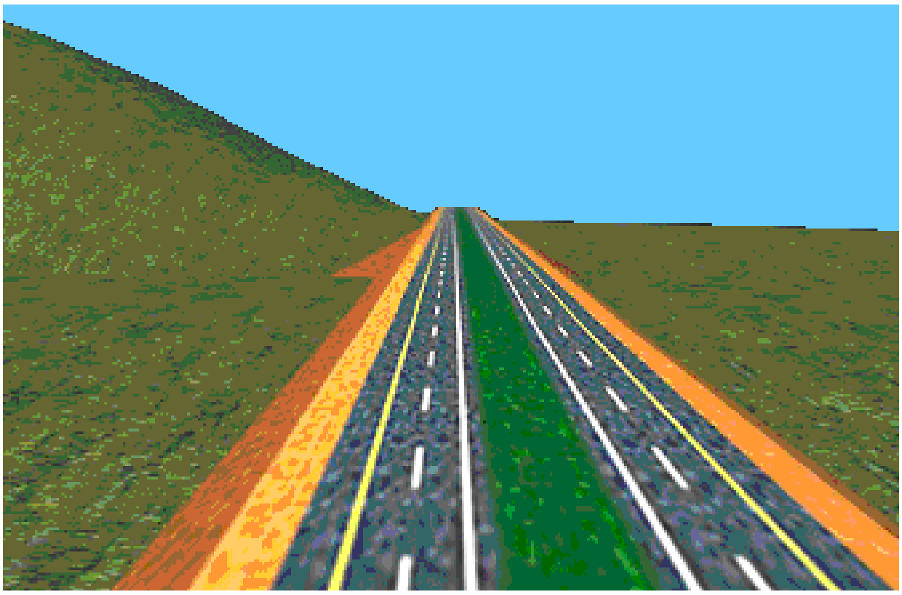
5.4.3 Perspectives in road design

To illustrate the advantages of visualising the alignment in three dimensions and to guide the design towards good practice, a number of alignment combinations are shown in Figures 5-5 to 5-17.

Figure 5-5 shows the advantage of maintaining a constant, uniform grade for as long as possible. Local dips to minimise earthworks that result in a disjointed alignment will be there for the life of the road.



A: Local dip on long grade



B: Local dip eliminated on long grade

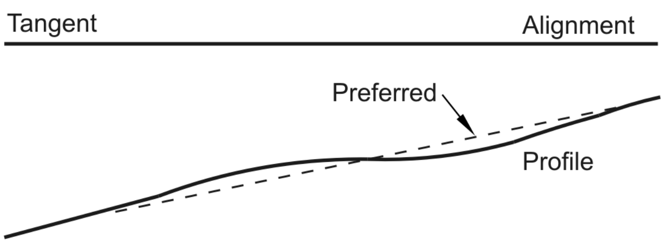
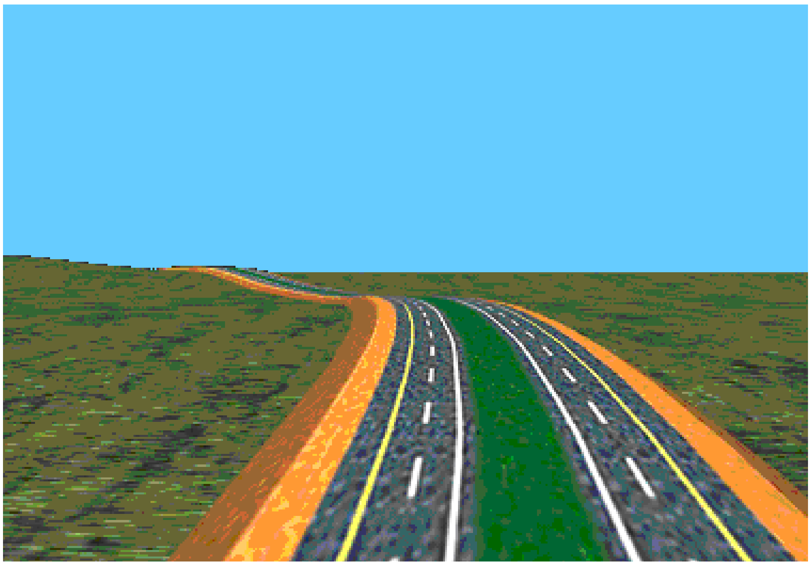


Figure 5-5: Short sag curve on long tangent

Short crests and sags should also be avoided on horizontal curves, as shown in Figure 5-6. Maintaining a constant grade is the preferred option.



A: Short humps on long horizontal curve

B: Removal of humps on horizontal curve

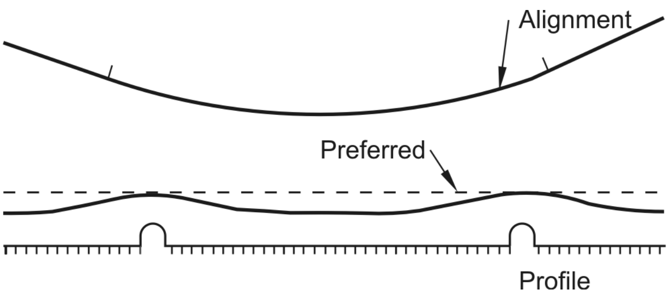
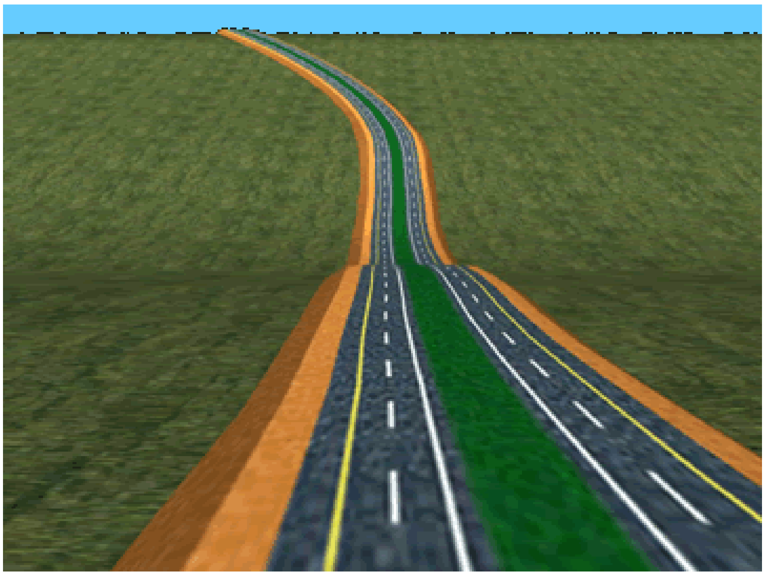
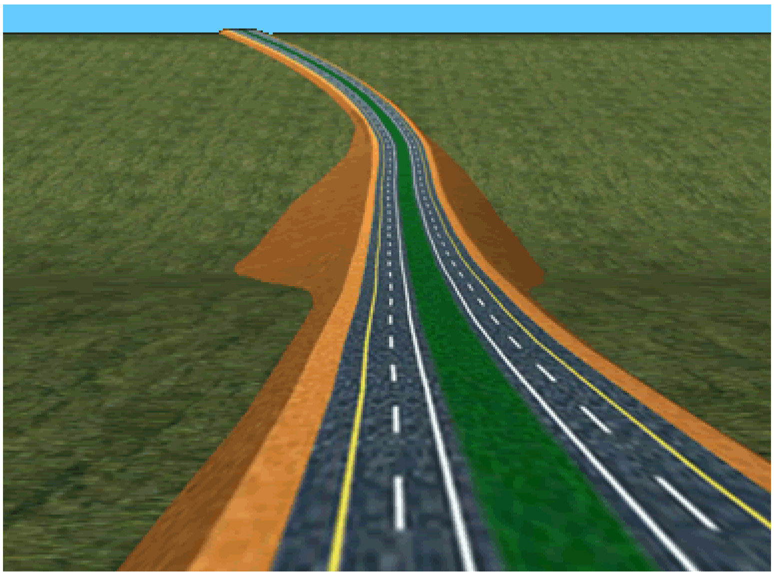


Figure 5-6: Short humps on long horizontal curve

A short discontinuity or dip in the alignment preceding a horizontal curve creates a particularly discordant view. Eliminating the crest curves in advance and following the sag curve improves the appearance, as shown in Figure 5-7.



A: Short hump and dip preceding horizontal curve



B: Long sag curve linking into horizontal curve

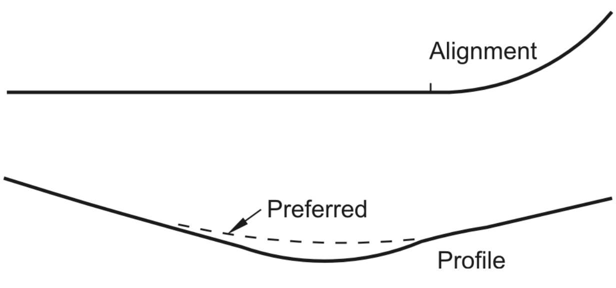


Figure 5-7: Short vertical curves preceding a long horizontal curve

A common fault in road alignment is illustrated in Figure 5-8. The roadway is often unnaturally curved to cross a small stream or grade separation at right angles.

The advantages in the alignment aesthetics of a skew crossing often far outweigh the savings deriving from a square crossing

A: Distorted alignment to create square river crossing

B: Skew crossing improves horizontal alignment

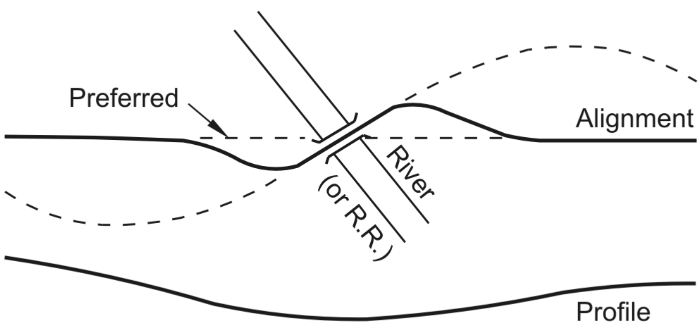
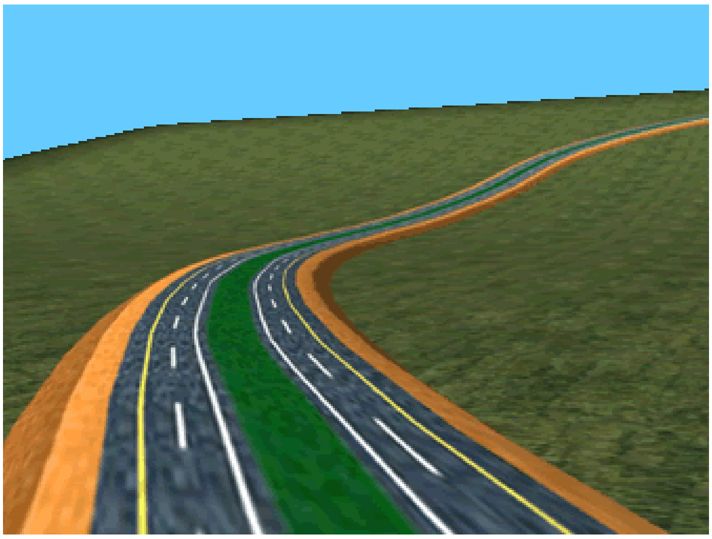
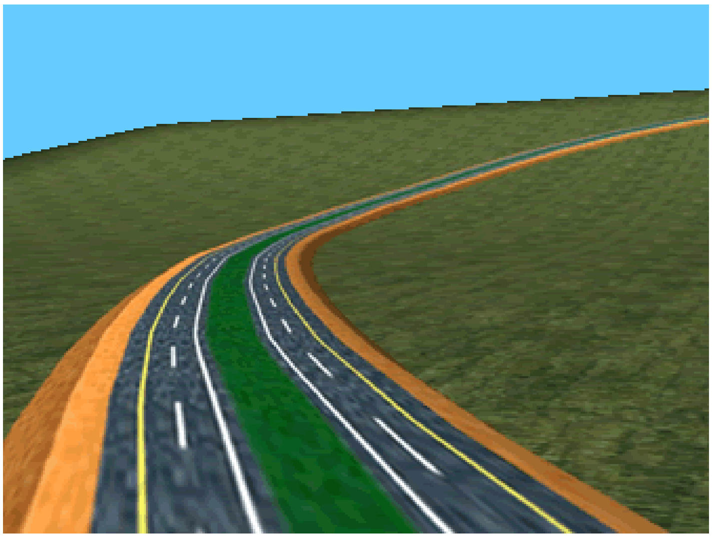


Figure 5-8: Distorted alignment at bridge crossing

Figure 5-9A illustrates the broken-back horizontal curve, or two curves in the same direction separated by a short tangent. The sag curve on the separating tangent intensifies the broken-back effect. The advantages of using a single radius curve throughout are illustrated in Figure 5-9.B.



A: Broken-back curve



B: Replacement of broken-back curve by single radius long curve

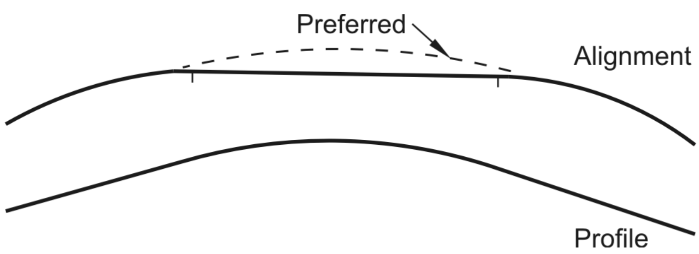
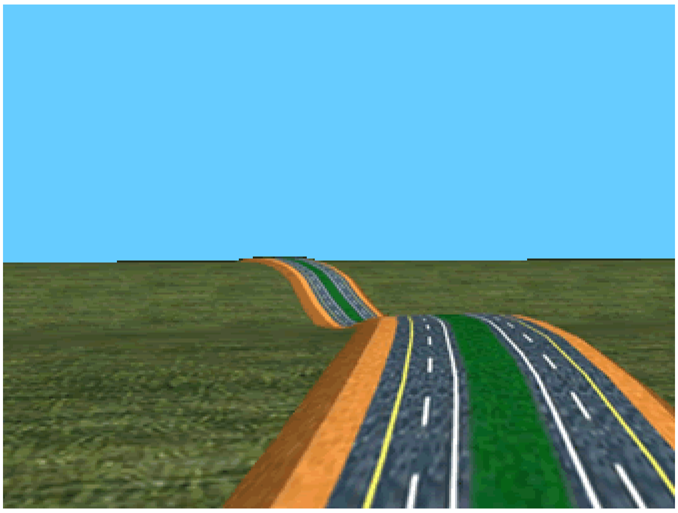


Figure 5-9: Broken-back curve

A sag curve at the start of a horizontal curve has the effect of enhancing the sharp angle appearance as shown in Figure 5-10, and should be avoided. Raising the preceding grade will move the sag curve downstream. A longer radius on the horizontal curve would cause it to start earlier. Applying both remedial measures should result in a better phasing of the horizontal and vertical alignments.

Figure 5-10: Out-of-phase vertical and horizontal alignments

Minor changes in grade or rolling of the vertical alignment as shown in Figure 5-11 should be avoided on long horizontal curves.



A: Rolling gradeline

B: Elimination of rolling gradeline

Figure 5-11 B illustrates the advantages of co-ordinating the horizontal and vertical alignments.

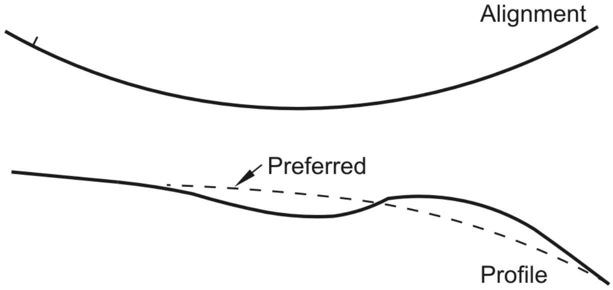
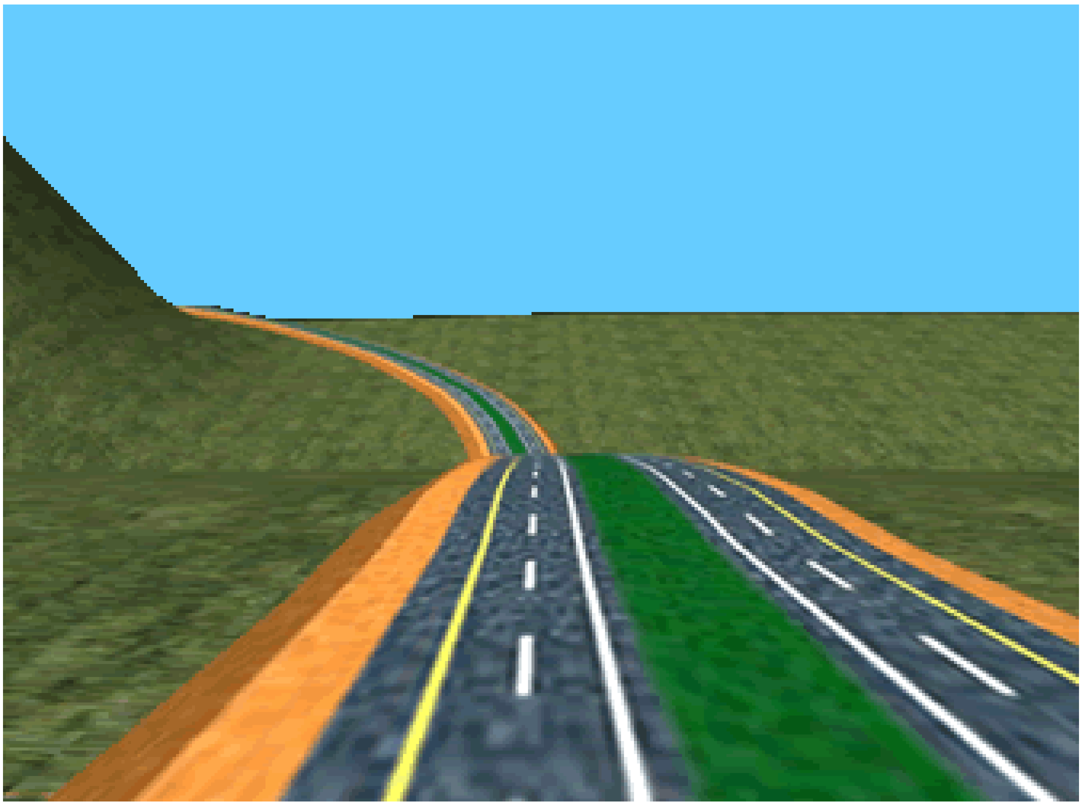


Figure 5-11: Minor rolling on long horizontal curve

Figure 5-12 shows the effect when the start of a horizontal curve is hidden by an intervening crest and the continuation of the curve is visible in the distance. The road appears disjointed.



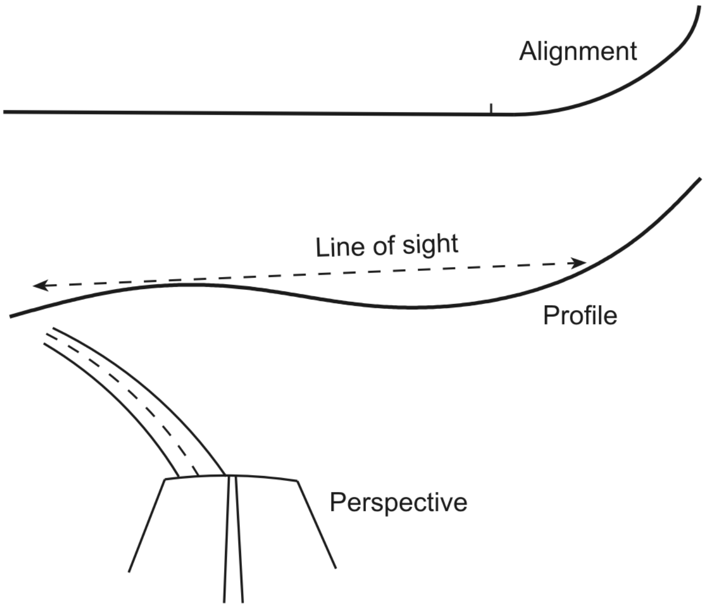


Figure 5-12: Break in horizontal alignment

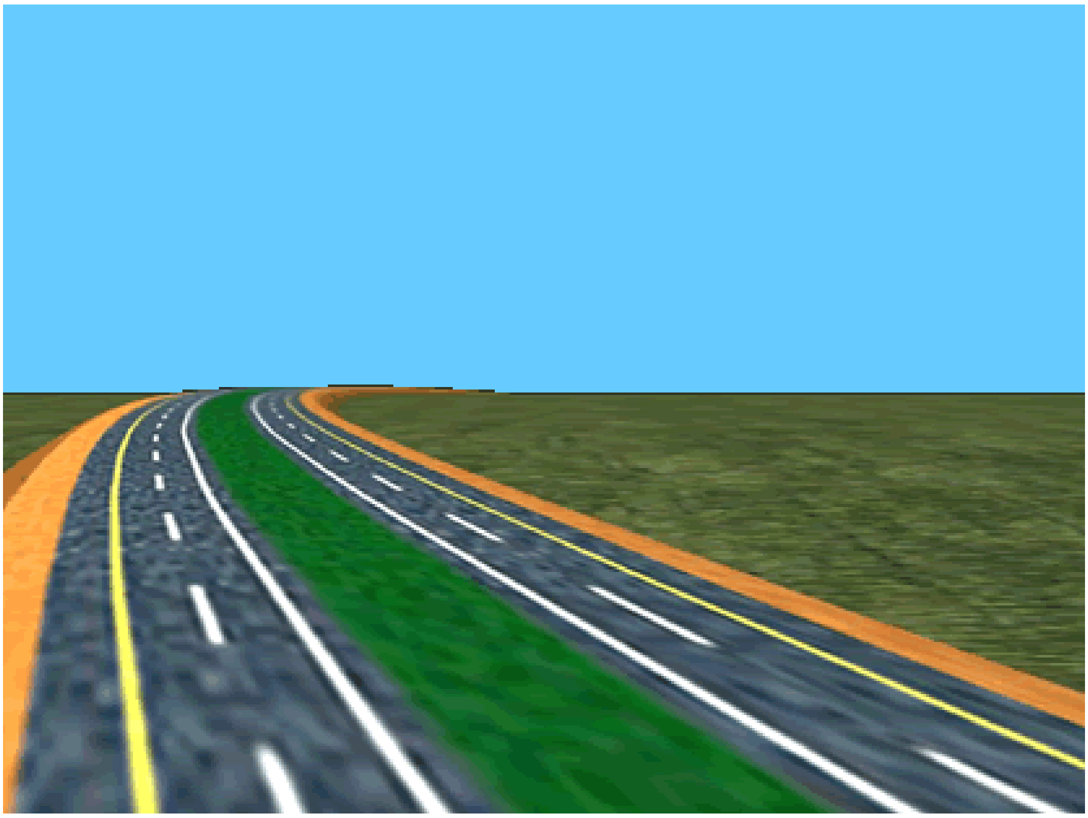


Figure 5-13: Well-coordinated crest and horizontal curves

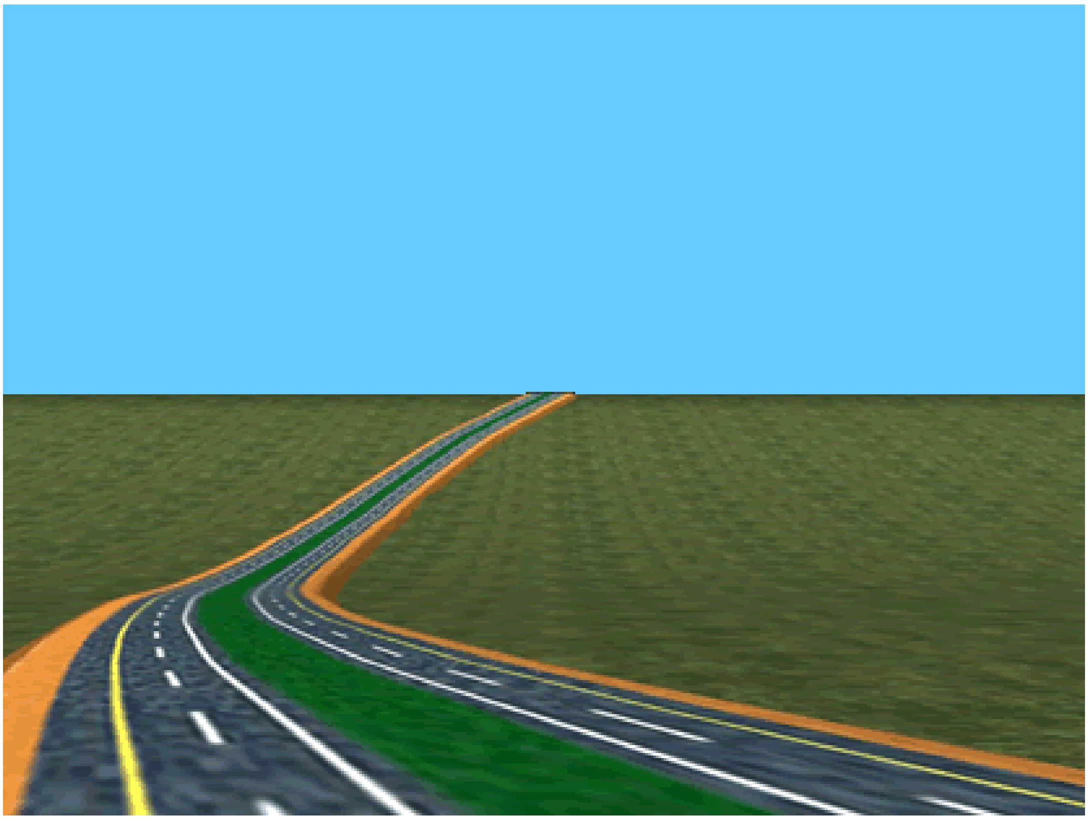


Figure 5-14: Well coordinated sag and horizontal curves

Figure 5-13 and Figure 5-14 illustrate the advantages of co-ordinating the horizontal and vertical alignment. In each case the vertical curve is contained within the horizontal curve.

Figure 5-15: N2 North: Curvilinear alignment in a valley

Figures 5-15 to 5-17 illustrate examples where excellent aesthetic designs have been achieved in practice.

Figure 5-16: N2 North - Median widening to accommodate stream

Figure 5-17: N2 - Combining horizontal and vertical alignment

5.5 INTEGRATION WITH THE ENVIRONMENT

A road is never the exclusive preserve of the user as it affects everybody in its surroundings. The impacts on the non-user are usually a combination of social, economic and environmental factors that act together in complex ways. They can be negative as in the scarring of a hillside, the diversion of a stream or the closing of an access. They can also be positive:

- When the space used by the road is orderly;
- The alignment imaginative;
- It appears to belong where it is; and
- It serves the transportation needs of the community.

In locating a road, it is necessary to synthesise land use planning and transportation planning. The traffic engineer sees the road as a conduit for goods and people. Estimates are made from the factors governing trip generation and trip origins and destinations. These ultimately result in expected and future traffic volumes. The land use planner on the other hand is concerned with the organisation of the areas and their relationship to one another. Although making a necessary contribution to communication, roads, particularly those of an arterial or freeway nature, are viewed as a barrier. They should never cut across homogenous areas of whatever type.

The challenge of reconciling the conflicts between the disciplines can best be tackled by a team approach where all participants strive to understand the viewpoints and constraints of the various disciplines. Integration with the social environment should be the primary objective. Of the social values, the aesthetic and visual impacts are often the most important. It is the detail of embankments, road signs, bridges, guardrails. Flat, rounded slopes also blend more readily into the natural contour of the land.

The general rule is that the lower the cut or fill the flatter the slope should be. On fills with a height of 8 to 10 metres, slopes of 1 : 1,1/2 to1 : 2 appear acceptable. For heights of 4 to 5 metres slopes of 1 : 4 should be the goal.

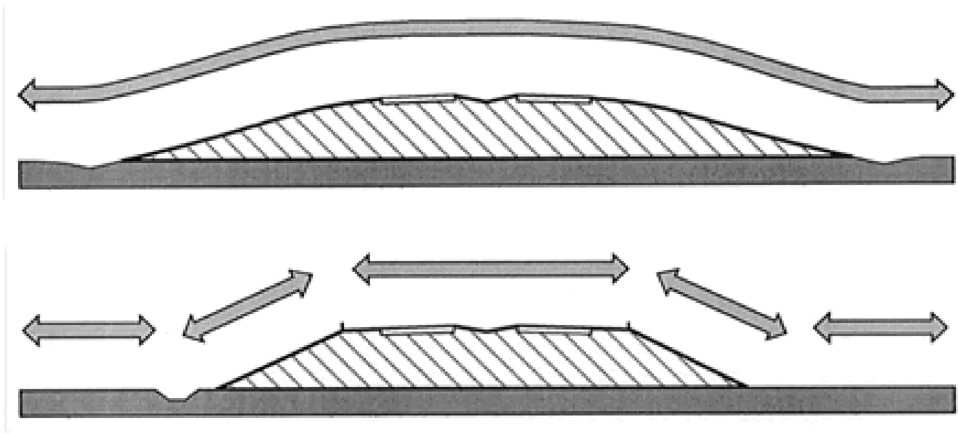


Figure 5-18: Blending of fills into landscape

the planting of the median and a host of other design features that attract the eye. All these should be integrated with the immediate environment of the road.

Of all these, the form of the cross section is the main element under the control of designer that can be manipulated to soften the impact of the road. Flattening the slopes of cut and fill and rounding the changes of slopes to create smooth contours has many benefits. It reduces soil erosion, minimises the risk of injury when a vehicle leaves the road and reduces the warrants for

Minor cuts and fills should be blended into the landscape so that they are hardly noticeable. This concept is shown in Figure 5-18. In the first sketch, the continuity of the space is preserved in the cross-section whereas, in the second, the space is chopped up into discontinuous elements.

The use of contour plans to grade cut and fill slopes are particularly useful when designing interchanges to ensure a pleasing layout of what are often large ground areas.

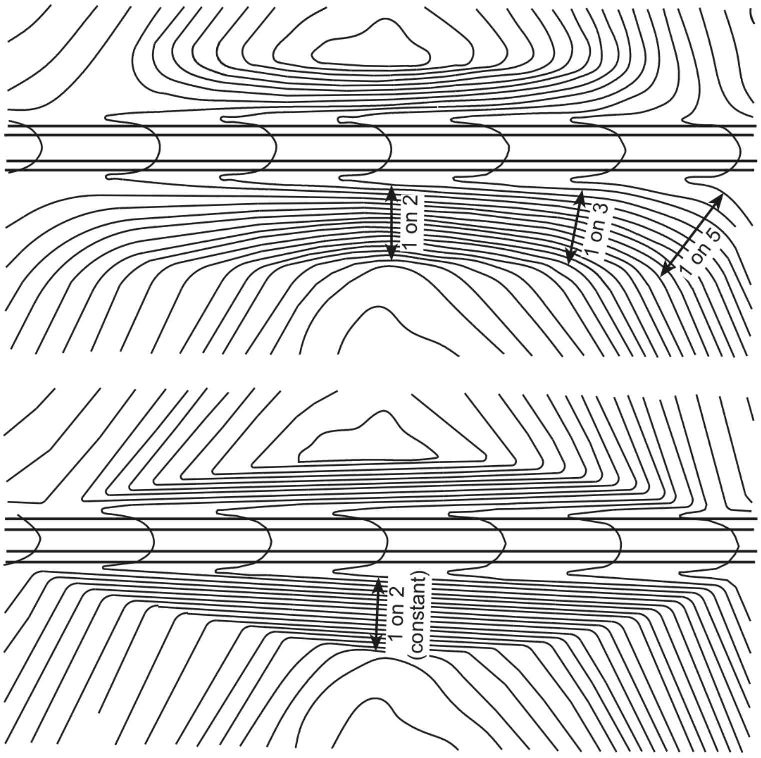
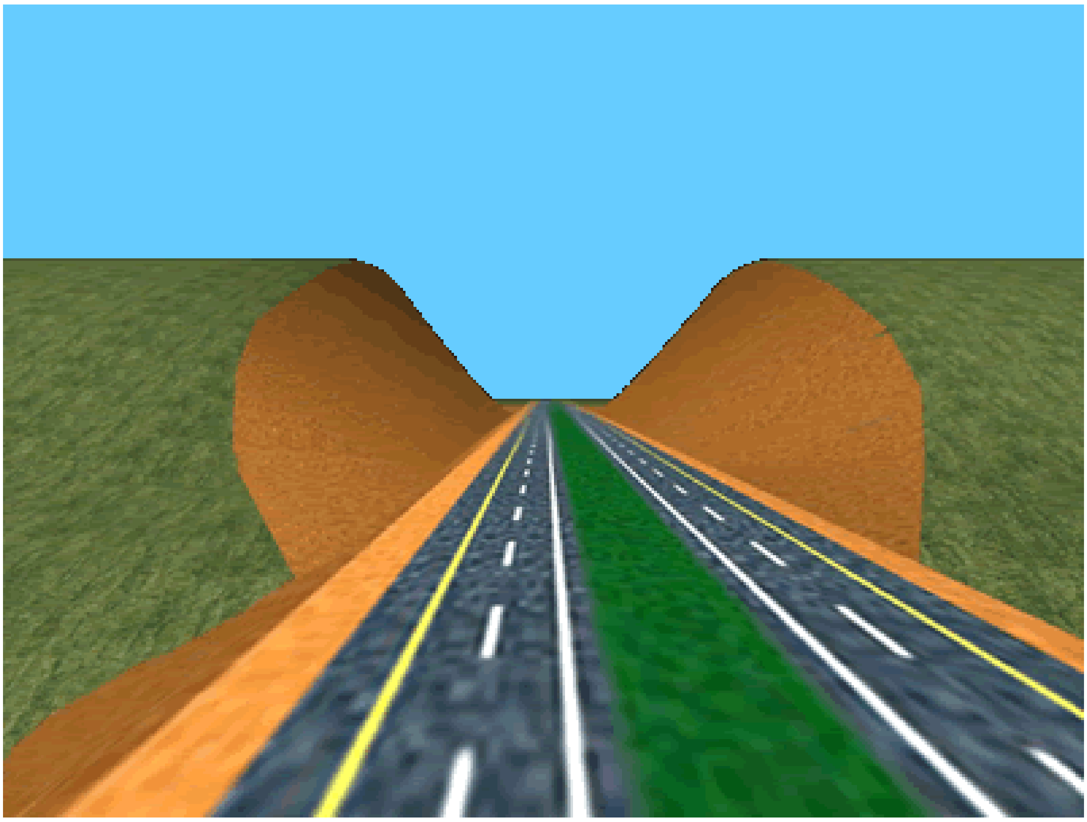
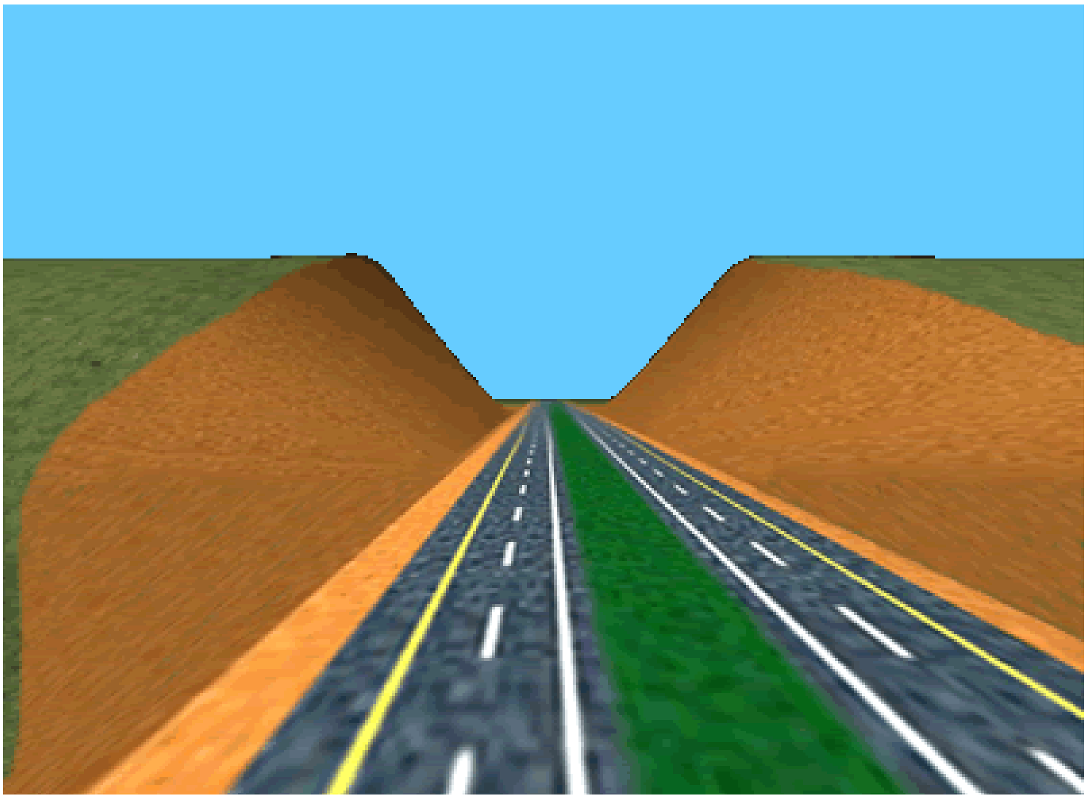


Figure 5.19: Contour plans of cut slopes

Figures 5-19 and 5-20 illustrate the advantages of transitioning the slopes into a deep cut. The approach shown in the first figure could lead to the appearance of the road being located through a natural feature whereas the second illustrates what, at best, would be a scar on the landscape.



A: Cut with constant cut batters



B: Cut with varying cut batters

Figure 5.20: Cuts with constant or varying cut batters

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 6

INTERSECTION DESIGN

6.1 INTRODUCTION

The design of an at-grade intersection requires understanding of the principles of both traffic and highway engineering. The operation of an intersection is influenced by its capacity, queue lengths and delays, accident potential, vehicle operating characteristics and traffic control. The physical layout of an intersection is defined by its horizontal and vertical alignment, roadway cross-sections, surface texture and drainage.

The successful integration of all these factors is required for good design, which must overcome

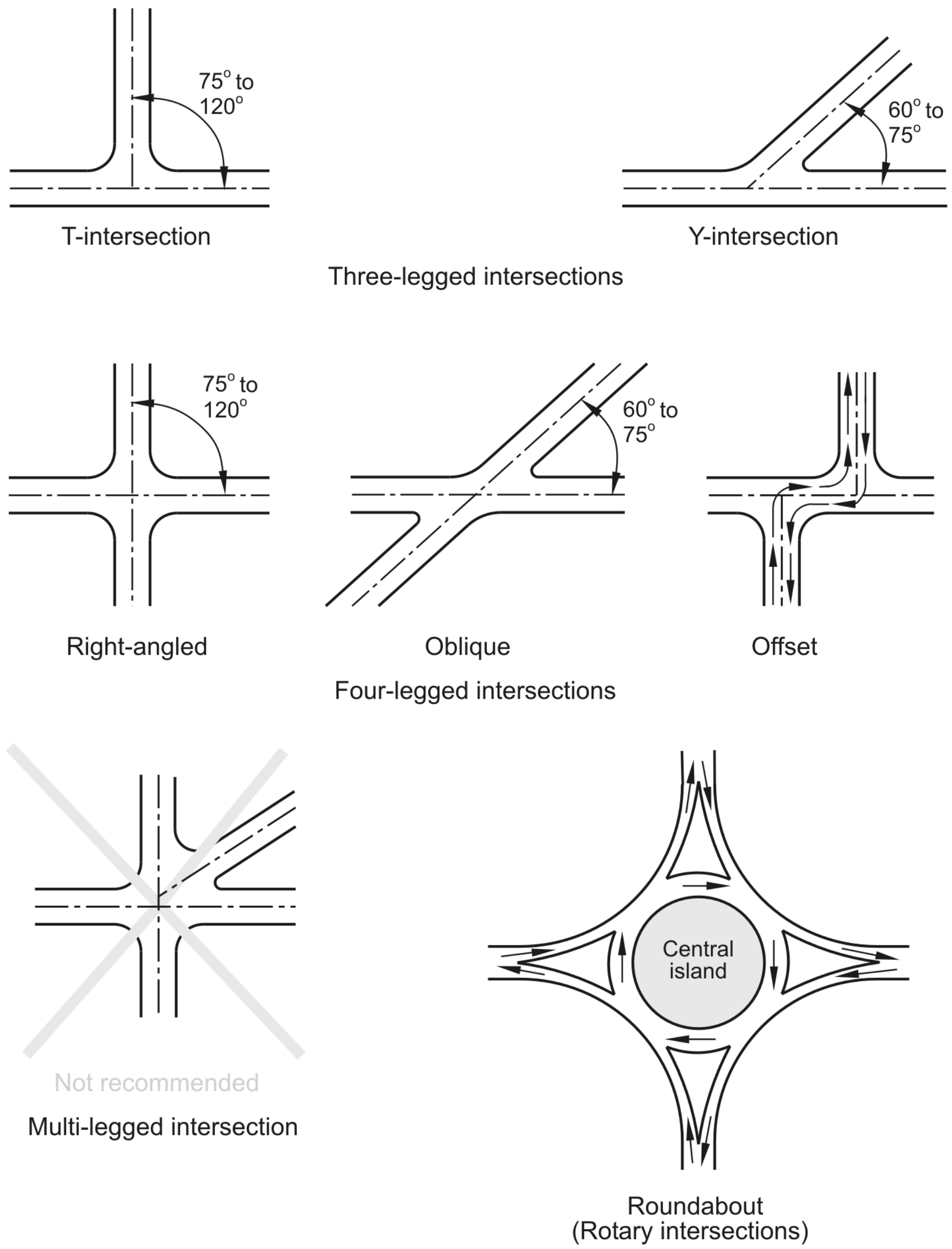


Figure 6.1: Typical intersection types

the potential safety and operation conflicts that are inherent when traffic streams interact at intersections.

Although some guidance on capacity and traffic control is offered, the focus of this chapter is on application of the geometric principles that govern the physical layout and location of an intersection.

6.2 DESIGN PRINCIPLES

6.2.1 General

The unique characteristic of intersections is that vehicles, pedestrians and bicycles travelling in many directions, must share a common area, often at the same time. The mitigation of the resulting conflicts is a major objective of intersection design. This conflict resolution is, in turn, influenced by construction and maintenance costs, environmental factors and the ease of implementation.

6.2.2 Elements affecting design

Human factors

The driver's response, discussed in sub-section 3.2.1, is a major factor in intersection design. The recommended perception and reaction time of 2,5 seconds is normally used as an input to models determining intersection sight distance. However, because of heightened awareness at controlled intersections, circumstances may indicate a lower acceptable value of 2,0 seconds at busy urban intersections.

The concept of driver expectancy is crucial in the evaluation of drivers' response and tasks within intersections.

Vehicle Characteristics

The size and manoeuvrability of vehicles is a governing factor in intersection design, particularly when channelisation features are being selected. Because of the importance of vehicle characteristics in the operation of an intersection, the selection of an appropriate vehicle, as described in sub-section 3.4.4, will influence the elements in the above table. In selecting an appropriate vehicle, the designer should carefully evaluate the expected traffic mix in context.

Table 6.1: Values of human factors appropriate to intersection design

Human factor	Design values	Design elements affected
Perception/reaction time	2,0 – 4,0 seconds	Intersection sight distance
Gap acceptance	5,5 - 7,5 seconds	Intersection sight distance
Driver height of eye	1,05 metres	Sight distance
Pedestrian walking speeds	1-1,5 metres/second	Pedestrian facilities

Various vehicle characteristics and their influence on the design of channelised intersections are described in Table 6.2.

Environmental Influences

The type of highway and area, surrounding land

Table 6.2 : Vehicle characteristics applicable to design of channelised intersections	
Vehicle Characteristics	**Intersection Design Elements Affected**
Physical Characteristics:	
Length	Lengths of storage lanes
Width	Widths of lanes Widths of turning roadways
Height	Placement of overhead signals and signs
Wheelbase	Nose placement Corner radius Widths of turning roadways
Operational characteristics	
Acceleration capability	Acceleration tapers and lane lengths Gap acceptance
Deceleration and braking capability	Lengths of deceleration lanes and tapers Stopping sight distance

use and the prevailing climate all have an influence on the type of design selected. Flexibility of approach is essential and the concepts of context-sensitive design as outlined in Chapter 2 should be applied

Functional classification is a key to applying the appropriate design standards. Primary arterials carry high traffic volumes, operate at high speeds and are often used by drivers unfamiliar with them. Large trucks and buses are common and there is a driver expectancy for route continuity and a high level of service. Intersection design should reflect and make provision for the operating characteristics of drivers and their expectations in the various classes of roads. Channelisation should accommodate the expected vehicles in a simple and direct manner. Decision sight distance is an important element and traffic control devices and pavement markings should be placed with care.

The type of area and adjacent land use governs the selection of an appropriate intersection. In urban areas, pedestrian flows, on-street parking and bus and taxi activity are commonplace. In residential areas, bicycles and school crossings need to be considered. They are usually absent in rural areas, where utility and delivery vehicles are more common.

Local climate can influence design decisions. Where the presence of mist is a frequent occurrence, sight distance would be reduced. Heavy rainfalls can obscure signs and road markings and reduce pavement friction.

6.2.3 Traffic manoeuvres and conflicts

Typical manoeuvres that result in vehicle conflict at intersections are:

- Crossing;
- Merging;

- Diverging; and
- Weaving.

Diverging and merging may be either to the left or right, mutual or multiple. Crossings may be direct, if the angle of skew is between 75° and 120°, or oblique if the angle is in the range of 60° to 75°. Oblique skews should be avoided if at all possible. If the angle of skew is less than 60°, the possibility of replacing the skew by a staggered intersection should be considered. Angles of skew greater than 120° should be replaced by relocation of the intersection to an angle of skew closer to 90°.

Traffic volumes are the most important determinant of intersection accidents. This is not surprising as the accident potential caused by conflict increases as traffic on the approach legs increases.

The type of traffic control also influences accidents. More rear-end and sideswipe accidents tend to occur at signalised intersections than at other types of control. Stop and Yield controls tend to increase the frequency of crossing accidents. Table 6.3 lists many of the condition that can lead to accidents and also the geometric and control measures that are used to mitigate poor accident experience.

Table 6.3: Features contributing to accidents at intersections and remedial measures

Geometric features or conditions contributing to adverse accident experience at intersections	Traffic engineering actions that reduce accident experience or severity
Poor approach sight distance	Addition/installation of exclusive turn-lanes
Poor corner sight distance	Upgrading of traffic control scheme
Steep grades at intersections	Improvement of sight distance
Inappropriate traffic control	Installation of lighting
Multiple approaches	Removal of fixed objects
Presence of curves within intersection	Increasing of corner radii
Number of adjacent driveways or access points	Application of channelisation
Inappropriate curb radii	
Narrow lanes	
Poor drainage and skid resistance	Improve drainage paths and roughen surface

Weaving is a combination of merging and diverging traffic moving in the same direction. It may be simple or complex.

The conflict at intersections created by the various manoeuvres leads to a unique set of operational characteristics. Understanding these is central to intersection designs and the most important characteristics are safety and capacity.

6.2.4 Capacity

To operate successfully, an intersection must be able to handle peak traffic demands. The analysis of capacity is based on the operational characteristics of conflicting vehicles separated by the time constraints imposed by traffic control devices. The measuring and forecasting of traffic flows and capacity analysis is a specialised subject and designers should refer to the manuals and references commonly used. The following is a brief summary of capacity as it relates to design.

Signalised intersections

The idealised flow rate through an intersection is known as the saturation flow rate per hour of green time. Initial driver reaction, vehicle acceleration and the behaviour of following vehicles all affect this flow rate. The capacity of an approach or leg of an intersection is proportional to the green time for that approach within the signal cycle in accordance with

	C	=	s x g/c
where	C	=	capacity (veh/h)
	s	=	saturated flow rate (veh/h)
	g/c	=	the ratio of green time to signal cycle time.

The important factors affecting saturation flow are:

- Number of lanes;
- Widths of lanes;
- Proportion of heavy vehicles;
- Gradients in excess of 3%;
- On-street parking;
- Pedestrian activity; and
- Type and phasing of signals.

The critical factors are the total number of lanes and the need for exclusive turning lanes at each approach.

Unsignalised intersections

The capacity of the major road at Stop- and Yield-controlled intersections is not affected by the presence of the intersection. The capacity of the minor road is dependent on the distribution of gaps in the major road traffic and the gap acceptance of the minor road traffic. Gap acceptance is dependent on the reaction/response time, vehicle acceleration and vehicle length. It is not a function of approach speed on the major road. Gap acceptance times used in determining capacity are usually somewhat shorter than those used to compute intersection sight distance as described in Section 6.4

Factors affecting capacity include:

- Operational speed of the major road;
- Intersection sight distance;
- Radii of turning roadways;
- Intersection layout and number of lanes;
- Type of area; and
- Proportion of heavy vehicles

The critical factors are intersection sight distance and the number and arrangement of traffic lanes.

6.2.5 Intersection types and selection

There are four basic types or classes of intersection: three-legged T and Y intersections; four legged intersections with a defined crossing path; multi-legged intersections; and roundabouts, with the last-mentioned encompassing all three of the previous types. These are shown schematically in Figure 6.1. The designer's selection of the basic intersections type is normally predicated on the design context, as intersections can vary greatly in scope, shape, degree of channelisation and traffic control measures.

Important factors to be considered in the selection of an intersection type are:

- Cost of construction;
- Type of area;
- Land use and land availability;
- Functional classes of the intersecting roads;
- Approach speeds;
- Proportion of traffic on each approach; and
- Volumes to be accommodated.

Careful consideration of these factors, together with the warrants for selecting appropriate traffic control devices, will lead to an appropriate choice or to a limited number of alternatives from which to make the final selection. The critical factors are cost and capacity.

A worldwide review of intersection design practice reported that, "typically the cheapest intersection type providing the required level of service is chosen". This cost is usually the sum of the design, construction and right-of-way costs. This view is consistent with South African experience, where the cost to the road authority is often the governing factor in the choice of intersection type.

At intersections carrying light crossing and turning volumes the capacity figures for uninterrupted flow generally apply. Table 6.4 below is a guide to the maximum traffic volumes that these intersections can accommodate.

When volumes exceed the above, the capacity of the intersection should be analysed in detail.

As safety at intersections is of key importance, the following summary of the interrelationship between intersections and accidents can act to guide the selection of an intersection type and layout:

- The U.S. National Safety Council estimated that 56 per cent of all urban accidents and 32 per cent of all rural accidents occur at intersections.
- The number of accidents is proportional to the volume and distribution of traffic on the major and minor roads.
- Roundabouts have considerable safety advantages over other types of at grade intersections.
- Poor sight distance leads to significantly higher injury and total accident rates. However, on roundabout approaches, accidents may actually increase with increasing sight distance.
- Medians should be as wide as practical at rural, unsignalised intersections but not wider than necessary at signalised intersections.
- Channelisation is usually beneficial but

Table 6.4: Typical maximum traffic volumes for priority intersections

Road Type	**Design volume (two-way veh/h)**		
Major road	500	1 000	1 500
Minor road	500	250	100

kerbed islands in the major road may be hazardous in rural areas.

- The hazard of an intersection increases as the approach speed increases.

6.2.6 Location of intersections

Given the fact that intersections are the most dangerous part of any road network, it follows that their location deserves serious attention by the designer. It is necessary to minimise both the likelihood of crashes occurring and the consequences of the crashes that do occur. There are thus various restraints on the location of intersections that should be considered.

The need for drivers to discern and perform the manoeuvres necessary to pass safely through an intersection demands that decision sight distance be available on the major road approaches. The driver on the minor road requires adequate intersection sight distance, as well as the sight triangles described in Section 6.4, in order either to merge with traffic on the major road or to cross safely. It may be necessary to modify the alignment of either the major or the minor road, or both, to ensure that adequate sight distance is available. If this is not possible, the options available to the designer are to:

- relocate the intersection;
- provide all-way Stop control; or
- provide a Jug handle or Quarter link interchange, as described in Section 7.6.4.

Where heavy earthworks, possibly beyond the normal limits of the road reserve, are required to provide adequate sight distances, relocation may be an option.

The location of an intersection on a curve can create problems for the drivers on both legs of the minor road. Drivers on the minor road leg on the inside of the curve will have difficulty in seeing approaching traffic because this traffic will be partly behind them. This constitutes, in effect, an artificial angle of skew. Furthermore, there is a possibility that a portion of the sight triangle may fall outside the limits of the road reserve, which could hamper efforts to obtain a clear line of sight for the driver on the minor

Table 6.5: Minimum radii for location of intersections on curves

Design speed (km/h)	Radius (m)
40	250
50	375
60	550
70	750
80	1 000
90	1 220
100	1 500
110	1 850
120	2 200
130	2 600

road. For these, reasons, intersections should not be located on horizontal curves with radii less than those indicated in Table 6.5.

Drivers on the outside of a curve typically have little or no difficulty in seeing opposing vehicles on the major road. The opposing vehicles are partly in front of them and they have the additional height advantage caused by the superelevation of the curve. However, they have to negotiate the turn onto the major road against an adverse superelevation. The risks involved in sharp braking during an emergency should also be borne in mind when an intersection is located on a curve. In general, an intersection should not be located on a curve with a superelevation greater than 6 per cent.

Stopping sight distances increase with steepening negative gradient. Stopping sight distance required on a gradient of -3 per cent is approximately 6 per cent longer than that on a level grade, whereas, on a -6 per cent gradient, it is approximately 16 per cent longer. It is suggested that, as a safety measure, intersections should not be located on gradients steeper than three per cent. The gradient is more critical on the minor road than on the major road because all vehicles on the minor road have to stop or yield.

One of the consequences of a collision between vehicles is that either or both may leave the road. It is therefore advisable to avoid locating intersections on high fills if at all possible. The obstruction of sight distance by bridge parapets should also be borne in mind. In the case of the crossing road ramp terminal at a narrow diamond interchange, both these problems may arise.

The location of an intersection may require modification to improve the angle of skew between the intersecting roads, as discussed in Section 6.3.1. If the angle of skew is less than 60^{o}, the intersection can be replaced by two relatively closely spaced T-intersections. A vehicle on the minor road would thus follow the route comprising a right turn onto the major road, followed by a left turn off it. Any delay to the minor road vehicle would occur clear of the high-speed traffic on the major road. If the angle of skew is greater than 120^{o}, relocation should be to a four-legged intersection in preference to the two T-intersections, because these would result in the minor road traffic following the route comprising a left turn onto the major road and a right turn off it with consequent delays to the major road traffic and increased risk to the vehicle waiting to complete the right turn off the major route. A right-left stagger or offset is, in short, to be preferred to a left-right stagger.

A further limitations on the location of intersections - being the spacing of successive intersections - is discussed in the following section.

6.2.7 Spacing of intersections

Designers seldom have influence on the spacing of roadways in a network as it is largely predicated by the original or developed land use. Nevertheless, the spacing of intersections impacts significantly on the operation, level of service and capacity of a roadway. It follows that intersection spacing should, inter alia, be based on road function and traffic volume. The principles described in the National Guidelines for Road Access Management in South Africa should therefore play a role in the determination of the location of individual intersections. This is of particular concern when the provision of a new intersection on an existing road is being considered.

Access management is aimed at maintaining an effective and efficient transportation system for the movement of people and goods, simultaneously supporting the development of the adjacent land use. Increasingly intensive land usage generally leads to demands for improved road infrastructure and the improved infrastructure makes access to it very attractive. Allowing access simply on the basis of its meeting some or other minimum geometric requirement results in increasing traffic conflicts and reduction in capacity so that the benefit of the original improvement is lost. This then leads to demands for further road improvements.

This cycle can only be broken by the development of a proper Access Management Plan by the local authority concerned. This plan specifies where intersections may be located. Furthermore, it defines the class of intersection that may be considered. Three classes of intersections are defined in the National Guidelines. These are:

- Full access, which allows for all possible movements at an intersection or access;
- Partial access, which allows left-in, left-out and right-in movements to and from a development or access road; and
- Marginal access, which allows only left-in and/or left-out movements to and from a development or access road.

These are illustrated in Figure 6.2.

Along signalised arterials, intersection spacing should be consistent with the running speed and signal cycle lengths, which are variables in themselves. If the spacing of the intersections is based on acceptable running speeds and cycle lengths, signal progression and an efficient use of the roadway can be achieved.

In Figure 6.2, these three variables are combined in a chart allowing the selection of a suitable intersection spacing. From this figure it transpires that the minimum spacing on arterial roads should be at least 400m. Where spacings closer than the minimum exist, a number of alternative actions can be considered, for example two-way flows can be converted to one-way operation or minor connecting roads can be closed or diverted, and channelisation can be used to restrict turning movements.

Where the crossing road of an interchange is an arterial, the suggested minimum distance along the arterial from the ramp terminal to the next intersection is 200m in the case of a collector road. If the next intersection is with an arterial the spacing between the ramp and the intersection should be increased to 600 metres for a Class 3 arterial and 800 metres for a Class 2 arterial.

The spacing between successive unsignalised intersections is measured by the separation between them with separation being defined as the distance between their reserve boundaries. Recommended access separations are provided in Table 6-5.

The left-in left-out class of access would typically only be applied under circumstances where

Table 6.5: Recommended minimum access separations		
Design speed (km/h)	Upstream access class	
	Unsignalised marginal	All other access types
40	20	80
50	35	110
60	50	130
70	70	175
80	100	200
100	170	300
120	250	350

speeds or traffic volumes or both are high. This is in order to minimise the disruption that would be caused by right-turning vehicles. As such, it would normally be provided with acceleration and deceleration lanes using taper rates as listed in Table 6-12 and lengths listed in Table 7-5 or 7-6

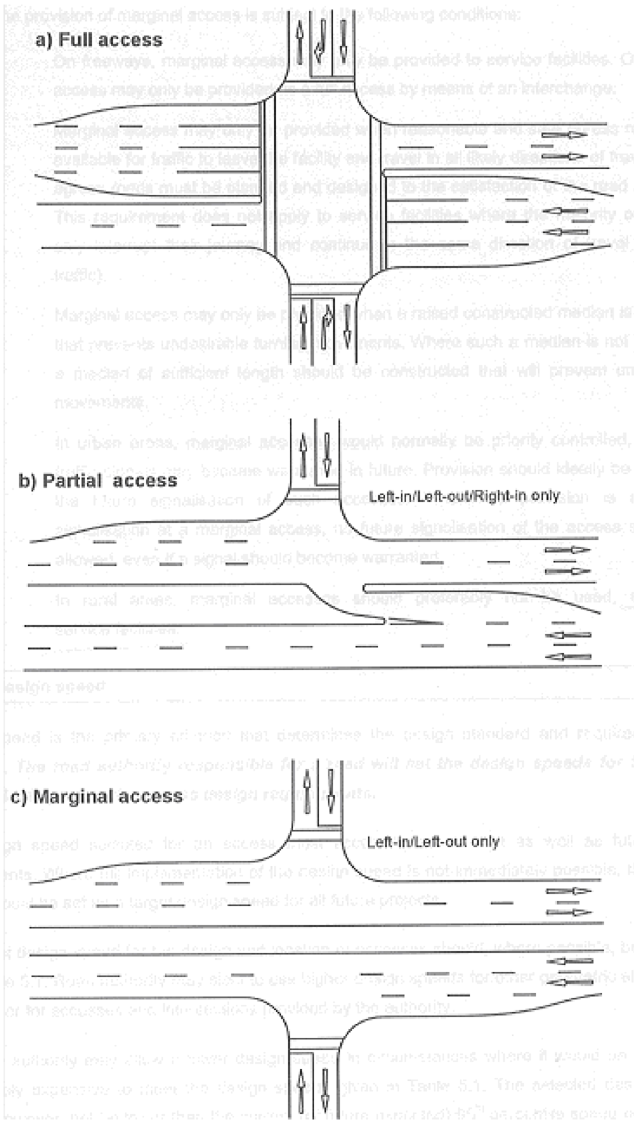


Figure 6-2: Classes of intersections

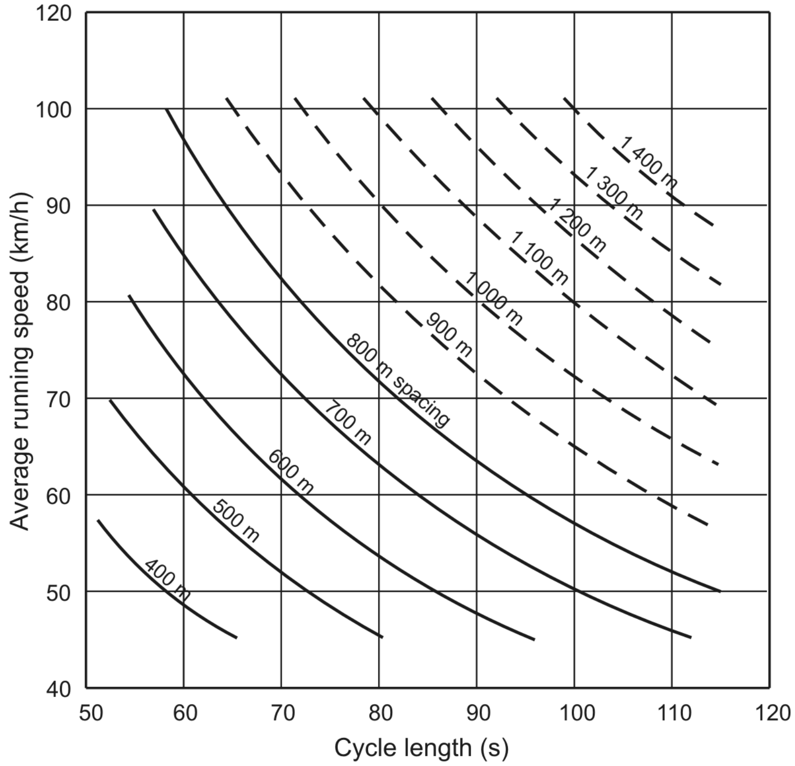
6.2.8 Channelisation

The purpose of channelisation is to achieve safe and efficient operation by managing the conflicts that are inherent to intersections. NCHRP Report 279 reports that the objectives of good intersection design are:

- Reduction of the number of points of potential conflict to the minimum compatible with efficient operation;
- Reduction of the complexity of conflict areas whenever possible;
- Limitation of the frequency of actual conflicts; and
- Limitation of the severity of those conflicts that do occur.

To achieve these objectives there are nine principles of channelisation design:

- Undesirable or wrong-way movements should be discouraged or prohibited;
- Vehicle paths should be clearly defined;
- The design should encourage safe vehicle speeds;
- Channelisation should separate points of conflict whenever possible;
- Traffic streams should cross at close to right angles and merge at flat angles;
- High priority flows should have the greater degree of freedom;
- The design should be in the context of the traffic control scheme;
- Decelerating, slow-moving or stopped vehicles should be separated from higher-speed through lanes; and
- Refuge for pedestrians and the handicapped should be provided where appropriate.



Note:
Dashed lines indicate reduced platooning and signal progression benefits for signal spacing greater than 800 m

Figure 6.3: Desirable signal spacings

The tools available to apply these principles are:

- Defining and arranging traffic lanes;
- Traffic islands of all sizes and types;
- Median islands;
- Corner radii;
- Horizontal and vertical approach geometry;
- Pavement tapers and transitions; and
- Traffic control devices.

The first six elements are a range of physical features while traffic control devices are an integral part of any intersection. These six elements are discussed in Section 6.5.

6.3 GEOMETRIC CONTROLS

6.3.1 Angle of intersection

The angle of intersection of two roadways influences both the operation and safety of an intersection. Large skews increase the pavement area and thus the area of possible conflict. Operationally they are undesirable because:

- Crossing vehicles and pedestrians are exposed for longer periods;
- The driver's sight angle is more constrained and gap perception becomes more difficult;
- Vehicular movements are more difficult and large trucks require more pavement area; and
- Defining vehicle paths by channelisation is more difficult.

For new intersections the crossing angle should preferably be in the range 75° to 120°. The absolute minimum angle of skew is 60° because drivers, particularly of trucks with closed cabs, have difficulty at this angle of skew in seeing vehicles approaching from their left. The designer should be able to specifically justify using an angle of skew less than 75°. In the remodelling of existing intersections, the accident rates and patterns will usually indicate whether a problem exists and provide evidence on any problems related to the angle of skews

6.3.2 Horizontal and vertical alignment

The horizontal and vertical alignments through and approaching an intersection are critical features. Simple alignment design allows for early recognition of the intersection and timely focus on the intersecting traffic and manoeuvres that must be prepared.

The following are specific operational requirements at intersections:

- The alignments should not restrict the required sight distance;
- The alignments should allow for the frequent braking and turning associated with intersections; and
- The alignments should not require undue direct attention to be detracted from the intersection manoeuvres and conflict avoidance.

As a general guide, horizontal curve radii at intersections should not be less than the desirable radius for the design speed on the approach roads.

For high-speed roads with design speeds in excess of 80km/h, approach gradients should not be greater than - 3 per cent. For low-speed roads in an urban environment this can be increased to - 6 per cent.

distance at the intersection to above-minimum requirements.
For new intersections, the gradient on the minor roadway is normally adjusted to form a smooth profile, as suggested in Figure 6.4.

Where major roadways intersect, the profiles of both roads are usually adjusted in approximately equal manner. When significant channelisation is introduced in association with complex gradients, intersections should be designed on an elevation plan to avoid discontinuities and ensure free drainage.

6.3.3 Lane widths and shoulders

Where intersecting roadways have shoulders or sidewalks, the main road shoulder should be continued through the intersection. Lane widths should be 3,7 m for through lanes and 3,6 metres for turning lanes. Where conditions are severely constrained, lane widths as low as 3,3 metres can be considered provided that approach speeds are below 80 km/h. In constricted urban conditions on low speed-roadways, lane widths of 3,0 metres should be the minimum adopted. Offsets from the edge of the turning roadway to kerb lines should be 0,6 to 1,0 metres.

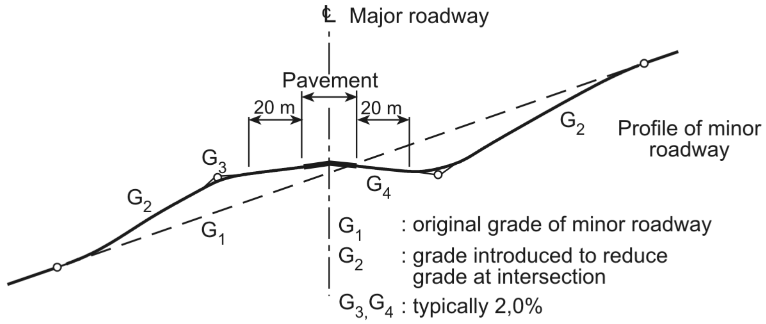
6.4 SIGHT DISTANCE

6.4.1 General Considerations

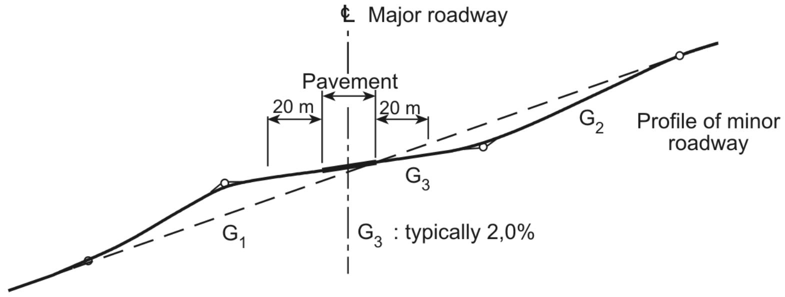
The provision of adequate sight distances and appropriate traffic controls is essential for safe intersection operation.

Stopping sight distance should be provided continuously on all roadways including at the approaches to intersections. However, in rural areas or when approach speeds are in excess of 80 km/h, the decision sight clearance set out in Section 3.5.8 should be provided on all approaches to intersections for safe operation, particularly where auxiliary lanes are added to the intersection layout to accommodate the turning movements.

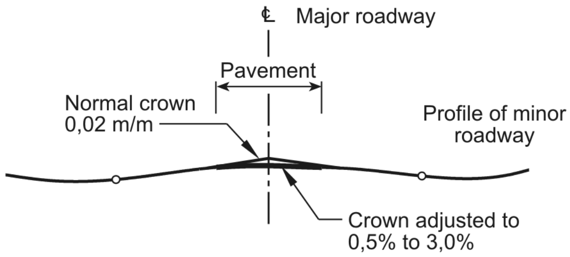
In addition to these forms of sight distance, it is necessary to provide Intersection Sight Distance (ISD). This is the sight distance required by drivers entering the intersection to enable them to establish that it is safe to do so and then to carry out the manoeuvres necessary either to join or to cross the opposing traffic streams. Previously, values of ISD were derived from elaborate models based on assumptions of reaction times, speeds and acceleration rates of turning vehicles and the deceleration rates of the opposing vehicles, etc. The distances offered in sub-section 6.4.3 are derived from research into gap acceptance as reported in NCHRP Report 383 "Intersection Sight Distance".



Minor roadway profile changed to fit crown of major roadway
(preferred, if minor roadway has stop or yield control)



Crown removed on major roadway to accommodate minor roadway
(typical consideration, if intersection may be signalised)



Crown reduced on major roadway to accommodate minor roadway

Figure 6.4: Fitting minor road profiles to major road cross-sections

6.4.2 Sight Triangles

Each quadrant of an intersection should contain a clear sight triangle free of obstructions that may block a driver's view of potentially conflicting vehicles on the opposing approaches.

Two different forms of sight triangle are required. In the first instance, reference is to approach sight triangles. The approach triangle will have sides with sufficient lengths on both intersecting roadways such that drivers can see any potentially conflicting vehicle in sufficient time to slow, or to stop if need be, before entering the intersection. For the departure sight triangle, the line of sight described by the hypotenuse of the sight triangle should be such that a vehicle just coming into view on the major road will, at the design speed of this road, have a travel time to the intersection corresponding to the gap acceptable to the driver of the vehicle on the minor road. Both forms of sight triangle are required in each quadrant of the intersection.

The line of sight assumes a driver eye height of 1,05 metres and an object height of 1,3 metres.

The approach and departure sight triangles are illustrated in Figure 6.5. The areas shown shaded should be kept clear of vegetation or any other obstacle to a clear line of sight. To this end, the road reserve is normally splayed to ensure that the entire extent of the sight triangle is under the control of the road authority. Furthermore, the profiles of the intersecting roads should be designed to provide the required sight distance. Where one or other of the approaches is in cut, the affected sight triangles may have to be "day lighted", i.e. the natural material occurring within the sight triangles may have to be excavated to ensure intervisibility between the relevant approaches.

Sight distance values are based on the ability of the driver of a passenger car to see an approaching passenger car. It is also necessary to check whether the sight distance is adequate for trucks. Because their rate of acceleration is lower than that of passenger cars and as the distance that the truck has to travel to clear the intersection is longer, the gap acceptable to a truck driver is considerably greater than that required by the driver of a passenger car. For design purposes, the eye height of truck drivers is taken as 1,8 metres for checking the availability of sight distance for trucks.

6.4.3 Intersection control

The recommended dimensions of the clear sight triangles vary with the type of traffic control used at an intersection because different types of control impose different legal constraints on drivers resulting in different driver behaviour. Sight distance policies for intersections with the following types of traffic control are presented below:

- Intersections with no control (Case A);
- Intersections with Stop control on the minor road (Case B);
    - o Right turn from the minor road (Case B1);
    - o Left turn from the minor road (Case B2);
    - o Crossing manoeuvre from the minor road (Case B3);
- Intersections with Yield control on the minor road (Case C);
    - o Crossing manoeuvre from the minor road (Case C1);
    - o Left or right turn from the minor road (Case C2);
- Intersections with traffic signal control (Case D); and
- Intersections with all-way Stop control (Case E).

A sight-distance policy for stopped vehicles turning right from a major road (Case F) is also presented.

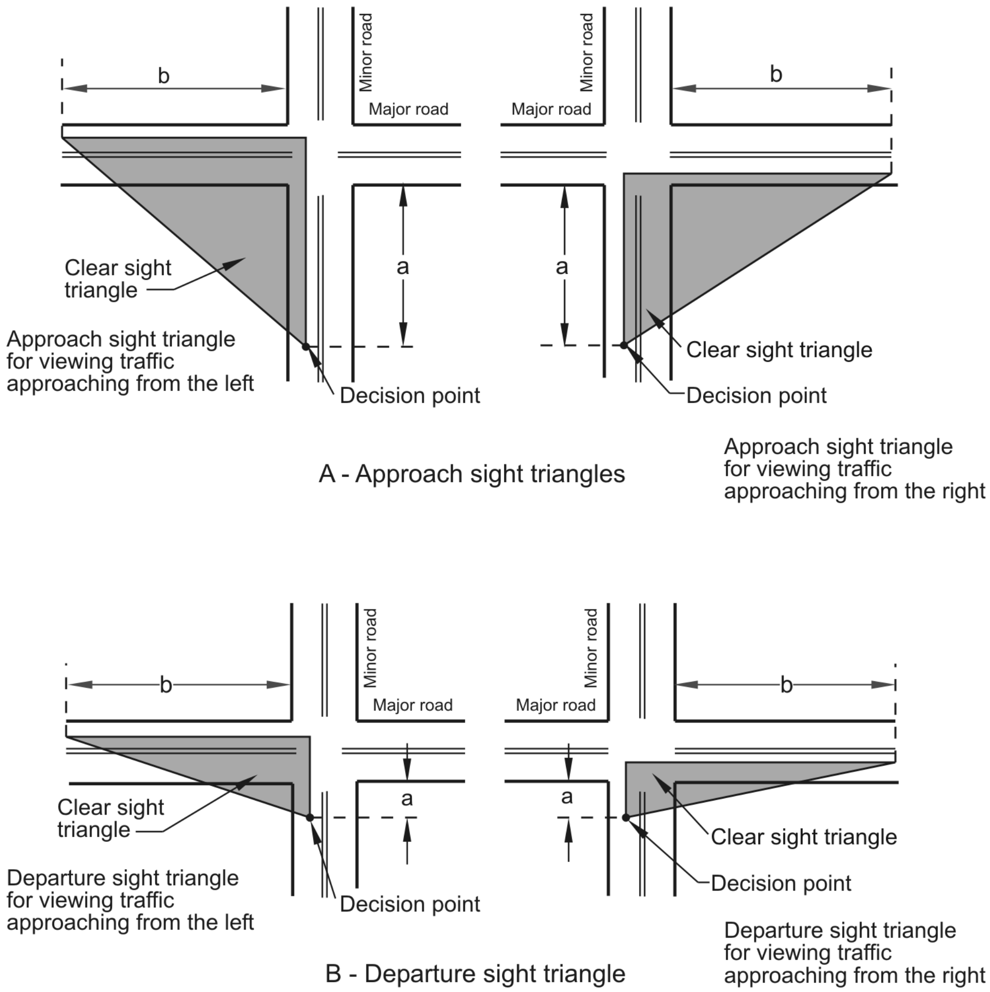


Figure 6.5: Sight triangles

Intersections with no control (Case A)

Uncontrolled intersections are not used in conjunction with the main road network, but are common in rural networks and access roads to rural settlements. In these cases, drivers must be able to see potentially conflicting vehicles on intersecting approaches in sufficient time to stop safely before reaching the intersection. Ideally, sight triangles with legs equal to stopping sight distance should be provided on all the approaches to uncontrolled intersections.

If sight triangles of this size cannot be provided, the lengths of the legs on each approach can be determined from a model that is analogous to the stopping sight distance model, with slightly different assumptions.

Field observations indicate that vehicles approaching uncontrolled intersections typically slow down to approximately 50 per cent of their normal running speed. This occurs even when no potentially conflicting vehicles are present, typically at deceleration rates of up to 1.5m/s.

Table 6.5: Recommended sight distances for intersections with no traffic control (Case A)	
Design speed (km/h)	Sight distance (m)
20	20
30	25
40	30
50	40
60	50
70	65
80	80
90	95
100	120
110	140
120	165

Braking at greater deceleration rates, which can approach those assumed in the calculation of stopping sight distances, begins up to 2,5 seconds after a vehicle on the intersecting approach comes into view. Thus, approaching vehicles may be travelling at less than their normal running speed during all or part of the perception-reaction time and can brake to a stop from a speed less than the normal running speed.

Table 6.5 shows the distance travelled by an approaching vehicle during perception, reaction and braking time as a function of the design speed of the roadway on which the intersection approach is located. These distances should be

Table 6.6: Adjustment factors for approach sight distances based on approach gradient										
Approach gradient (%)	Design speed (km/h)									
	30	40	50	60	70	80	90	100	110	120
-6	1,1	1,1	1,1	1,1	1,1	1,2	1,2	1,2	1,2	1,2
-5	1,0	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,2	1,2
-4	1,0	1,0	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1
-3 to +3	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0
+4	1,0	1,0	1,0	0,9	0,9	0,9	0,9	0,9	0,9	0,9
+5	1,0	1,0	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,9
+6	1,0	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,9

Note: Based on ratio of stopping sight distance on specified approach grade to stopping sight distance on level terrain.

used as the legs of the sight triangles shown in Figure 6.4.

Where the gradient of an intersection approach exceeds three per cent, the leg of the clear sight triangle along that approach should be adjusted by multiplying the sight distance in Table 6.5 by an adjustment factor in Table 6.6.

If these sight distances cannot be provided, advisory speed signing to reduce speeds or installing Stop signs on one or more approaches should be investigated.

Uncontrolled intersections do not normally require departure sight triangles because they typically have very low traffic volumes. If a motorist finds it necessary to stop at an uncontrolled intersection because of the presence of a conflicting vehicle, it is unlikely that another potentially conflicting vehicle will be encountered as the first vehicle departs the intersection.

Intersections with Stop control on the minor road (Case B)

Departure sight triangles for intersections with Stop control on the minor road should be considered for three situations:

- Right turns from the minor road (Case B1);
- Left turns from the minor road (Case B2); and
- Crossing the major road from the minor-road approach (Case B3).

Approach sight triangles, as shown in Figure 6.4(A), need not be provided at Stop-controlled intersections because all minor-road vehicles should stop before entering or crossing the major road.

Vehicles turning right from the minor road have to cross the stream of traffic approaching from the right and then merge with the stream approaching from the left. Left-turning vehicles need only merge with the stream approaching from the right. As the merging manoeuvre requires that turning vehicles should be able to accelerate approximately to the speed of the stream with which they are merging, it necessitates a gap longer than that for the crossing manoeuvre.

Right turn from the minor road (Case B1)

A departure sight triangle for traffic approaching from the left, as shown in Figure 6.4(B), should be provided for right turns from the minor road onto the major road for all Stop-controlled approaches.

Field observations of the gaps accepted by the drivers of vehicles turning to the right onto the major road have shown that the values in Table 6.7 provide sufficient time for the minor-road vehicle to accelerate from a stop and merge with the opposing stream without undue interference. These observations also revealed that major-road drivers would reduce their speed to some extent to accommodate vehicles entering from the minor road. Where the gap acceptance values in Table 6.7 are used to determine the length of the leg of the departure sight triangle along the major road, most major-road drivers need not reduce speed to less than 70 percent of their initial speed.

Table 6.7 applies to passenger cars. However, for minor-road approaches from which substantial volumes of heavy vehicles enter the major road, the values for single-unit trucks or semi-trailers should be applied.

Table 6.7 includes adjustments to the acceptable gaps for the number of lanes on the major road and for the approach gradient of the minor road. The adjustment for the gradient of the minor-road approach need be made only if the rear wheels of the design vehicle would be on an upgrade steeper than 3 per cent when the vehicle is at the stop line of the minor-road approach.

The length of the sight triangle along the major road (distance "b" in Figure 6.4) is the product of the design speed of the major road in metres/second and the critical gap in seconds as listed in Table 6.7.

If the sight distances along the major road based on Table 6.7 (including the appropriate adjustments) cannot be provided, consideration should be given to the installation of advisory speed signs on the major-road approaches.

Dimension "a" in Figure 6.4 (b) depends on the context within which the intersection is being designed. In urban areas, drivers tend to stop their vehicles immediately behind the Stop line, which may be located virtually in line with the edge of the major road. A passenger car driver would, therefore, be located about 2,4 metres away from the Stop line. In rural areas, vehicles usually stop at the edge of the shoulder of the major road. In the case of a three metre wide shoulder the driver would thus be approximately 5,4 metres away from the edge of the travelled way.

Table 6.7: Travel Times Used to Determine the Leg of the Departure Sight Triangle along the Major Road for Right and Left Turns from Stop-Controlled Approaches (Cases B1 and B2).

Design vehicle	Travel time (s) at design speed of major road
Passenger car	7,5
Single-unit truck	9,5
Semi-trailer	11,5
Adjustment for multilane highways: For right turns onto two-way highways with more than two lanes, add 0,5 seconds for passenger cars or 0,7 seconds for trucks for each additional lane, in excess of one, to be crossed by the turning vehicle. For left turns, no adjustment is necessary	
Adjustment for approach gradients: If the approach gradient on the minor road exceeds +3 per cent: ❑ Add 0,1 seconds per percent gradient for left turns ❑ Add 0,2 seconds per percent gradient for right turns	

Where the major road is a dual carriageway, two departure sight triangles have to be considered: a sight triangle to the right, as for the crossing movement (Case B3) and one using the acceptable gap as listed in Table 6.7 for vehicles approaching from the left. This presupposes that the width of the median is sufficient to provide a refuge for the vehicle turning from the minor road. If the median width is inadequate, the adjustment in Table 6.7 for multilane major roads should be applied with the median being counted as an additional lane.

The departure sight triangle should be checked for various possible design vehicles because the width of the median may be adequate for one vehicle type and not for another so that two different situations have to be evaluated.

Left turn from the minor road (Case B2)

A departure sight triangle for traffic approaching from the right, as shown in Figure 6.4, should be provided for left turns from the minor road. The lengths of the legs of the departure sight triangle for left turns should generally be the same as those for the right turns used in Case B1. Specifically, the length of the leg of the departure sight triangle (dimension "b") along the major road should be based on the travel times in Table 6.7, including appropriate adjustment factors.

Dimension "a" depends on the context of the design and can vary from 2,4 metres to 5,4 metres.

Where sight distances along the major road based on the travel times from Table 6.7 cannot be provided, it should be kept in mind that field observations indicate that, in making left turns, drivers generally accept gaps that are slightly shorter than those accepted in making right turns. The travel times in Table 6.7 can be decreased by 1,0 to 1,5 seconds for left turn manoeuvres, where necessary, without undue interference with major-road traffic. When the recommended sight distance for a left-turn manoeuvre cannot be provided, even with a reduction of 1,0 to 1,5 seconds, consideration should be given to the installation of advisory speed signs and warning devices on the major-road approaches.

Crossing manoeuvre from the minor road (Case B3)

In most cases it can be assumed that the departure sight triangles for right and left turns onto the major road, as described for Cases B1 and B2, will also provide more than adequate sight distance for minor-road vehicles crossing the major road. However, it is advisable to check the availability of sight distance for crossing manoeuvres:

- Where right and/or left turns are not per mitted from a particular approach and crossing is the only legal manoeuvre;
- Where the crossing vehicle has to cross four or more lanes; or
- Where substantial volumes of heavy vehicles cross the highway and where there are steep gradients on the departure roadway on the far side of the inter section that might slow the vehicle while its rear is still in the intersection.

Table 6.8 presents travel times and appropriate adjustment factors that can be used to determine the length of the leg of the sight triangle along the major road to accommodate crossing manoeuvres.

At divided highway intersections, depending on the width of the median and the length of the design vehicle, sight distance may be needed for crossing both roadways of the divided highway or for crossing the near lanes only and stopping in the median before proceeding.

or cross the major road without stopping. The sight distances needed by drivers on Yield-controlled approaches exceed those for Stop-controlled approaches because of the longer travel time of the vehicle on the minor road.

For four-legged intersections with Yield control on the minor road, two separate sets of approach sight triangles as shown in Figure 6.5(A) should be provided: one set of approach sight triangles to accommodate right and left turns onto the major road and the other for crossing movements. Both sets of sight triangles should be checked for potential sight obstructions.

Table 6.8: Travel times used to determine the leg of the departure sight triangle along the major road to accommodate crossing manoeuvres at stop-controlled intersections (Case B3)

Design vehicle	Travel time (s) at Design speed of major road
Passenger car	6,5
Single-unit truck	8,5
Semi trailer	10,5
Adjustment for multilane highways: For crossing a major road with more than two lanes, add 0,5 seconds for passenger cars and 0,7 seconds for trucks for each additional lane to be crossed. In the case of dual carriageways with inadequate width of median for refuge, count the median as another lane to be crossed.	
Adjustment for approach grades: If the approach gradient of the minor road exceeds +3 %, add 0,2 seconds per percent gradient in excess of 3 %.	

a For minor-road approach gradients that exceed +3 per cent, multiply by the appropriate adjustment factor from Table 6.6.

b Travel time applies to a vehicle that slows before crossing the intersection but does not stop.

Intersections with Yield control on the minor road (Case C)

Vehicles entering a major road at a Yield-controlled intersection may, because of the presence of opposing vehicles on the major road, be required to stop. Departure sight triangles as described for Stop control must therefore be provided for the Yield condition. However, if no conflicting vehicles are present, drivers approaching Yield signs are permitted to enter

Crossing manoeuvres (Case C1)

The lengths of the leg of the approach sight triangle along the minor road to accommodate the crossing manoeuvre from a Yield-controlled approach (distance "a" in Figure 6.5(A) are given in Table 6.9. The distances in Table 6.9 are based on the same assumptions as those for Case A control except that, based on field observations, minor-road vehicles that do not stop are assumed to decelerate to 60 per cent of the minor-road design speed rather than to 50 per cent. The distances and times in Table 6.9 should be adjusted for the gradient of the minor road approach, using the factors in Table 6.6.

The length of the leg of the approach sight triangle along the major road to accommodate the crossing manoeuvre (distance "b" in Figure 6.5(A)) should be calculated using the following equations:

$$t_c = t_a + \frac{w + L_a}{0{,}167 V_{minor}} \qquad 6.1$$

$$b = 0{,}278 V_{major} t_c \qquad 6.2$$

where:

t_c	=	travel time to reach and clear the major road in a crossing manoeuvre (sec)
b	=	length of leg of sight triangle along the major road (m)
t_a	=	travel time to reach the major road from the decision point for a vehicle that does not stop (sec) (use appropriate value for the minor-road design speed from Table 6.9, adjusted for approach grade, where appropriate)
w	=	width of intersection to be crossed (m)
L_a	=	length of design vehicle (m)
V_{minor}	=	design speed of minor road (km/h)
V_{major}	=	design speed of major road (km/h)

These equations provide sufficient travel time for the major road vehicle, during which the minor-road vehicle can:

(1) Travel from the decision point to the intersection, while decelerating at the rate of 1.5m/s² to 60 per cent of the minor-road design speed; and then

(2) Cross and clear the intersection at the same speed.

Field observations did not provide a clear indication of the size of the gap acceptable to the driver of a vehicle located at the decision point on the minor road. If the required gap is longer than that indicated by the above equations, the driver would, in all probability, bring the vehicle to a stop and then select a gap on the basis of Case B. If the acceptable gap is shorter than that indicated by the above equations, the sight distance provided would, at least, provide a margin of safety.

If the major road is a divided highway with a median wide enough to store the design vehicle for the crossing manoeuvre, then only crossing of the near lanes need be considered and a departure sight triangle for accelerating from a stopped position in the median should be provided, based on Case B1.

Left and right-turn manoeuvres (Case C2)

To accommodate left and right turns without stopping (distance "a" in Figure 6.5(A)), the length of the leg of the approach sight triangle along the minor road should be 25 metres. This distance is based on the assumption that drivers making right or left turns without stopping will slow to a turning speed of 15 km/h. The length

Table 6.9: Leg of approach sight triangle along the minor road to accommodate crossing manoeuvres from yield-controlled approaches		
Design speed (km/h)	**Distance along minor road (m)**	**Travel time from decision point to major road (t_a)[a,b]**
30	30	3,4
40	40	3,7
50	50	4,1
60	65	4,7
70	85	5,3
80	110	6,1
90	140	6,8
100	165	7,3
110	190	7,8
120	230	8,6

a For minor-road approach gradients that exceed +3 per cent, multiply by the appropriate adjustment factor from Table 6.6.

b Travel time applies to a vehicle that slows before crossing the intersection but does not stop.

of the leg of the approach sight triangle along the major road (distance "b" in Figure 6.5(B)) is similar to that of the major-road leg of the departure sight triangle for Stop-controlled intersections in Cases B1 and B2.

For a Yield-controlled intersection, the travel times in Table 6.7 should be increased by 0,5 seconds. The minor-road vehicle requires 3,5 seconds to travel from the decision point to the intersection. These 3,5 seconds represent additional travel time that is needed at a Yield-controlled intersection (Case C). However, the acceleration time after entering the major road is 3,0 seconds less for a Yield sign than for a Stop sign because the turning vehicle accelerates from 15 km/h rather than from a stop. The net 0,5 seconds increase in travel time for a vehicle turning from a Yield-controlled approach is the difference between the 3,5 second increase in travel time on approach and the 3,0 second reduction in travel time on departure explained above.

Since approach sight triangles for turning manoeuvres at Yield-controlled are larger than the departure sight triangles used at Stop-controlled intersections, no specific check of departure sight triangles at Yield-controlled intersections should be necessary.

Intersections with traffic signal control (Case D)

In general, approach or departure sight triangles are not needed for signalised intersections. Indeed, signalisation may be an appropriate accident countermeasure for higher volume intersections with restricted sight distance and a history of sight-distance related accidents.

However, traffic signals may fail from time to time. Furthermore, traffic signals at an intersection are sometimes placed on two-way flashing operation under off-peak or night time conditions. To allow for either of these eventualities, the appropriate departure sight triangles for

Case B, both to the left and to the right, should be provided for the minor-road approaches.

Intersections with all-way Stop control (Case E)

At intersections with all-way Stop control, the first stopped vehicle on each approach would be visible to the drivers of the first stopped vehicles on each of the other approaches. It is thus not necessary to provide sight distance triangles at intersections with All-way Stop control. All-way Stop control may be an option to consider where the sight distance for other types of control cannot be achieved. This is particularly the case if signals are not warranted.

Right turns from a major road (Case F)

Right-turning drivers need sufficient sight distance to enable them to decide when it is safe to turn right across the lane(s) used by opposing traffic. At all locations, where right turns across opposing traffic are possible, there should be sufficient sight distance to accommodate these manoeuvres. Since a vehicle that turns right without stopping needs a gap shorter than that required by a stopped vehicle, the need for sight distance design should be based on a right turn by a stopped vehicle.

The sight distance along the major road to accommodate right turns is the distance that would be traversed at the design speed of the major road in the travel time for the appropriate design vehicle given in Table 6.10. This table also contains appropriate adjustment factors for the number of major-road lanes to be crossed by the turning vehicle.

If stopping sight distance has been provided continuously along the major road and, if sight distance for Case B (Stop control) or Case C (Yield control) has been provided for each minor-road approach, sight distance should generally be adequate for right turns from the major road. However, at intersections or driveways located on or near horizontal or vertical curves on the major road, the availability of adequate sight distance for right turns from the major road should be checked. In the case of dual carriageways, the presence of sight obstructions in the median should also be checked.

Table 6.10: Travel times used to determine the sight distance along the major road to accommodate right turns from the major road (Class F)

Design vehicle	**Travel time (s) at design speed of major road**
Passenger car	5,5
Single-unit truck	6,5
Semi trailer	7,5
Adjustment for multilane highways: For right turns that have to cross more than one opposing lane, add 0,5 s for passenger cars and 0,7 s for trucks for each additional lane to be crossed. In the case of dual carriageways where the median is not sufficiently wide to provide refuge for the turning vehicle, the median should be regarded as an additional lane to be crossed.	

At four-legged intersections, opposing vehicles turning right can block a driver's view of oncoming traffic. If right-turn lanes are provided, offsetting them to the right, to be directly opposite one other, will provide right-turning drivers with a better view of oncoming traffic.

Effect of skew on sight distance

When two highways intersect at an angle outside the range of 75° to 120° and where realignment to increase the angle of intersection

area within each sight triangle should be clear of sight obstructions, as described above.

At skew intersections, the length of the travel paths for crossing manoeuvres will be increased. The actual path length for a crossing manoeuvre can be calculated by dividing the total width of the lanes (plus the median width, where appropriate) to be crossed by the sine of the intersection angle and adding the length of the design vehicle. The actual path length divided by the lane width applied to the major road

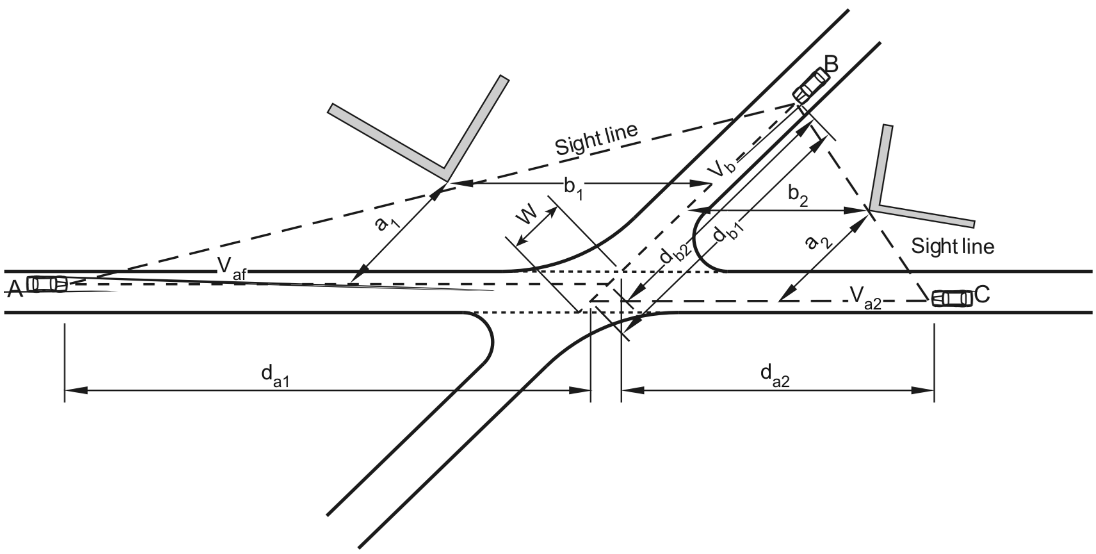


Figure 6.6: Effect of skew on sight distance at intersections

is not justified, some of the factors for determination of intersection sight distance will need adjustment.

Each of the clear sight triangles described above is applicable to oblique-angle intersections. As shown in Figure 6.6, the legs of the sight triangle will lie along the intersection approaches and each sight triangle will be larger or smaller than the corresponding sight triangle would be at a right-angle intersection. The

cross-section gives the equivalent number of lanes to be crossed. This is an indication of the number of additional lanes to be applied to the adjustment factor shown in Table 6.8 for Case B3.

The sight distances offered for Case B can, regardless of the form of control, also accommodate turning movements from the minor road to the major road at skew intersections. In the obtuse angle, drivers can easily see the full sight triangle and, in addition, often accelerate from the minor road at a higher rate than when they have to negotiate a ninety-degree change of direction. In the acute-angle quadrant, drivers are often required to turn their heads considerably to see across the entire clear sight triangle. For this reason, it is suggested that Case A should not be applied to oblique-angle intersections. Stop or Yield control should be applied and the sight distances appropriate to either Case B or Case C provided. Even in a skew intersection it is usually possible for drivers to position their vehicles at approximately 90° to the major road at the Stop line, offering added support for the application of Case B for skew intersections.

When driving through a deflection angle greater than 120°, the right turn to the minor road may be undertaken at crawl speeds. Allowance could be made for this by adding the time, equivalent to that required for crossing an additional lane, to the acceptable gap.

6.5 CHANNELISATION ELEMENTS

At-grade intersections with large paved areas, such as those with large corner radii or with angles of skew differing greatly from 90°, permit unpredictable vehicle movements, require long pedestrian crossings and have unused pavement areas. Even at a simple intersection there may be large areas on which vehicles can wander from natural or expected paths. Under these circumstances it is usual to resort to channelisation.

As stated in sub-section 6.2.8, the fundamental function of channelisation is to manage the conflicts that are inherent in any intersection. In that sub-section, nine principles of channelisation are listed. In essence, channelisation is the process whereby a vehicle can be guided safely through the intersection area from an approach leg to the selected departure leg. Guidance is offered by lane markings that clearly define the required vehicle path and also indicate auxiliary lanes for turning movements. A variety of symbols is also used as road markings to indicate inter alia that turns, either to the left or to the right, from selected lanes are mandatory. At intersections that are complex or have high volumes of turning traffic, it is usually necessary to reinforce the guidance offered by road markings by the application of:

- Channelising islands;
- Medians and median end treatments;
- Corner radii;
- Approach and departure geometry;
- Pavement tapers and transitions;
- Traffic control devices including signs and signals; and
- Arrangement and position of lanes

6.5.1 Channelising islands

Islands are included in intersections for one or more of the following purposes:

- Separation of conflicts;
- Control of angle of conflict;
- Reduction of excessive pavement areas;
- Regulation of traffic and indication of the proper use of the intersection;
- Arrangement to favour a predominant turning movement;
- Protection of pedestrians;
- Protection and storage of turning vehicles; and
- Location of traffic control devices.

The three main functions of channelising islands are thus:

- Directional - to control and direct movements, usually turning;
- Division - which can be of opposing or same direction, usually through, movements; and
- Refuge - either of turning vehicles or of pedestrians.

Typical island shapes are illustrated in Figure 6.7.

The designer should bear in mind that islands are hazards and should be less hazardous than whatever they are replacing.

Islands may be kerbed, painted or simply non-paved. Kerbed islands provided the most positive traffic delineation and are normally used in urban areas to provide some degree of protection to pedestrians and traffic control devices. Painted islands are usually used in suburban areas where speeds are low, e.g. in the range of 50 km/h to 70 km/h and space limited. In rural areas, kerbs are not common and, at the speeds prevailing in these areas, typically 120 km/h or more, they are a potential hazard. If it is necessary to employ kerbing at a rural intersection, the use of mountable kerbing should be considered. As an additional safety measure, a kerbed island could be preceded by a painted

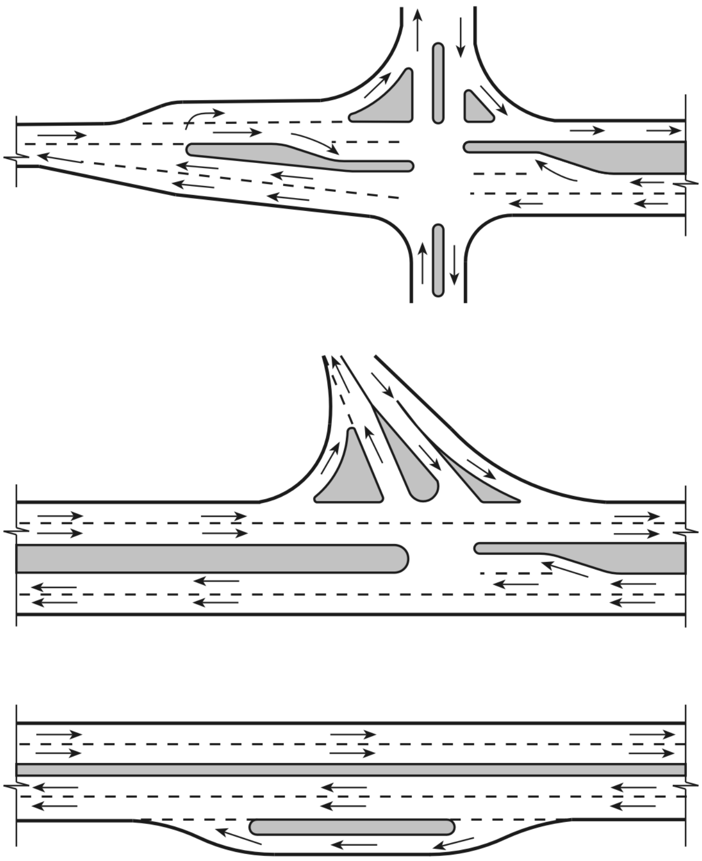


Figure 6.7 General types and shapes of islands

island. Non-paved islands are defined by the pavement edges and are usually used for large islands at rural intersections. These islands may have delineators on posts and may be landscaped.

Islands are generally either long or triangular in shape, with the circular shape being limited to application in roundabouts. They are situated in areas not intended for use in vehicle paths.

Drivers tend to find an archipelago of small islands confusing and are liable to select an incorrect path through the intersection area. As a general design principle, a few large islands are thus to be preferred to several small islands. Islands should not be less than about 5 square metres in area to ensure that they are easily visible to approaching drivers.

Directional islands are typically triangular with their dimensions and exact shape being dictated by:

- The corner radii and associated tapers;
- The angle of skew of the intersection; and
- The turning path of the design vehicle.

A typical triangular island is illustrated in Figure 6.8. The approach ends of the island usually have a radius of about 0,6 metres and the offset between the island and the edge of the travelled way is typically 0,6 to 1,0 metres to allow for the effect of kerbing on the lateral placement of moving vehicles. Where the major road has

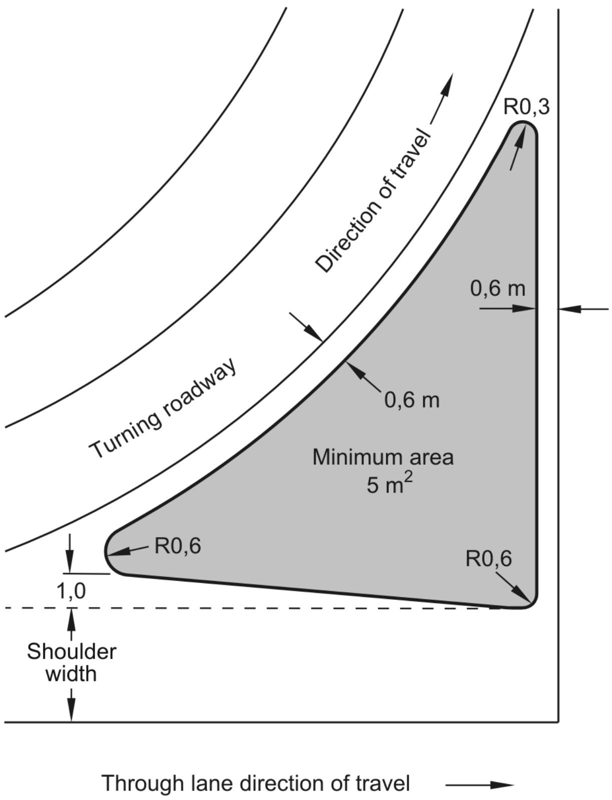


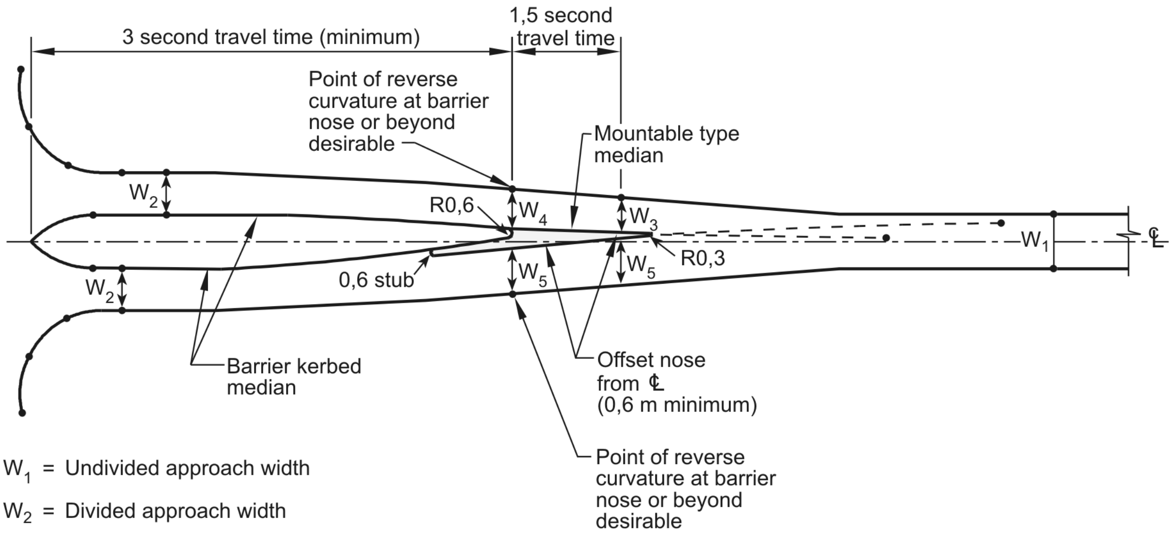
Figure 6.8: Typical directional island

shoulders, the nose of the island is offset about one metre from the edge of the usable shoulder, the side adjacent to the through lane being tapered back to terminate at the edge of the usable shoulder, thus offering some guidance and redirection. A kerbed cross-section on the major road suggests that the nose of the island should be offset by about 1,6 metres from the edge of the travelled way, with the side adjacent to the through lane being tapered back to terminate 0,6 metres from the edge of the through lane.

Dividing, or splitter, islands usually have a teardrop shape. They are often employed on the minor legs of an intersection where these legs have a two-lane, two-way or four-lane undivided cross-section. The principle function of a dividing island is to warn the driver of the presence of the intersection. This can be achieved by the left edge of the island being, at the widest point of the island, in line with the left edge of the approach leg. To the approaching driver, it would thus appear as though the entire lane had been blocked off by the island. If space does not permit this width of island, a lesser blocking width would have to be applied but it is doubted

Note:

1. Width of island controlled by nose placement as determined by control radii and angle of intersection or added right turn lane.
2. Vertical and horizontal alignment on approach to provide minimum stopping sight distance.
3. Control dimension, W, should be from edge to edge of pavement and should not include channel.



W_1 = Undivided approach width

W_2 = Divided approach width

$W_3 = \frac{W_1}{2}$ or 4,3 m whichever is larger

$W_4 = \frac{W_3 + W_2}{2}$ desirable

$W_5 = W_2 + 0,3$ m

Figure 6.9: Typical divisional (splitter) island

whether anything less than half of the approach lane width would be effective. The taper that can be employed to achieve this effect safely is discussed in Section 6.5.3. A typical dividing island is illustrated in Figure 6.9.

The shape of the splitter island discussed in Section 6.6.5 is derived from the need to redirect vehicles entering a roundabout through an angle of not more than 30°. Although it serves to create the illusion of about a half of the approach lane being blocked off, its true function is to achieve the desired extent of deflection. Furthermore, this form of splitter island does not accommodate vehicles turning from the left. In effect, it has a teardrop shape, albeit distorted by its abutting a curving roadway with a relatively short radius rather than a straight road.

The kerb height should ensure that the island would be visible within normal stopping sight distance. However, it may be advisable to draw the driver's attention to the island by highlighting the kerbs with paint or reflective markings.

As in the case of the triangular island, the nose of the dividing island should be offset by about one metre but, in this case, to the right of the centreline of the minor road. Dividing islands are usually kerbed to enhance their visibility and the offset between the kerbing and the edge of the travelled way should thus be 0,6 metres as discussed above. For the sake of consistency, the radius of the nose should be of the order of 0,6 metres.

The balance of the shape of the island is defined by the turning paths of vehicles turning to the right, both from the minor road to the major road and from the major road to the minor.

Median islands are discussed in Section 4.4.6 and outer separator islands in Section 4.4.7. At intersections, the end treatment of median islands is important. The width of the opening between two median ends should match the width of the minor road, including its shoulders, or, where the minor road is kerbed, the opening should not be narrower than the surfaced width of the minor road plus an offset of 0,6 to 1,0 metres.

The median end treatment is determined by the width of the median. Where the median is three metres wide or less, a simple semicircle is adequate. For wider medians, a bullet nose end treatment is recommended. The bullet nose is formed by arcs dictated by the wheel paths of turning vehicles and an assumed nose radius of 0,6 to 1,0 metres. It results in less intersection pavement area and a shorter length of opening than the semicircular end. Above a median width of five metres, the width of the minor road controls the length of the opening. A flattened bullet nose, using the arcs as for the conventional bullet nose but with a flat end as dictated by the width of the crossing road and parallel to the centreline of the minor road, is recommended. These end treatments are illustrated in Figure 6.10.

The bullet nose and the flattened bullet nose have the advantage over the semicircular end treatment that the driver of a right turning vehicle has a better guide for the manoeuvre for most of the turning path. Furthermore, these end treatments result in an elongated median, which is better placed to serve as a refuge for pedestrians crossing the dual carriageway road. An additional disadvantage of the use of the semicircular end treatment for wide medians is that, whereas the bullet nose and the flattened bullet nose both guide the vehicle towards the left of the centreline of the minor road, the semi-circular end treatment tends to direct the vehicle into the opposing traffic lane of the minor road.

envisaged. Reference is typically to three types of operation, being:

Case 1	One-lane one-way travel with no provision for passing stopped vehicles;
Case 2	One-lane one-way travel with provision for passing a stopped vehicles; and
Case 3	Two-lane one-way operation

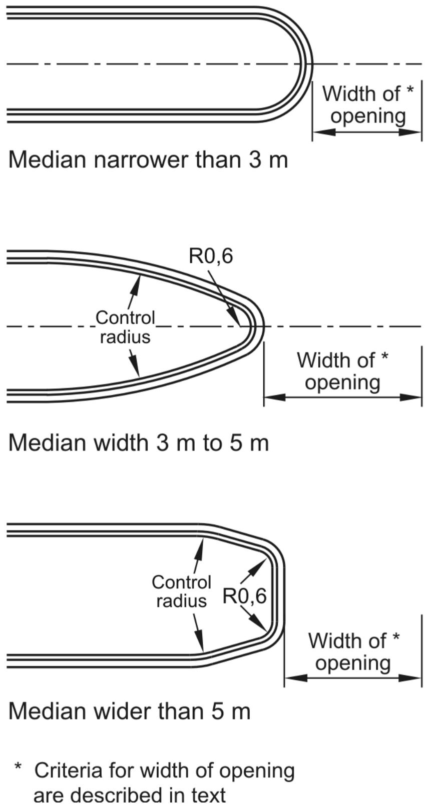


Figure 6.10: Typical median end treatments

6.5.2 Turning roadway widths

Directional islands are bounded by the major and minor roads and by a short length of one-way, typically one-lane, turning roadway. The width of the turning roadway is defined by the swept area of the design vehicle for the selected radius of curvature and the type of operation

Three traffic conditions should also be considered, being:

Condition A	Insufficient SU vehicles in the turning traffic stream to influence design;
Condition B	Sufficient SU vehicles to influence design; and
Condition C	Sufficient semi-trailers to

Table 6.11: Turning roadway widths										
Design speed (km/h)	Inner radius (m)	Case 1			Case 2			Case 3		
		Condition			Condition			Condition		
		A	B	C	A	B	C	A	B	C
20	15	4,0	5,5	7,9	6,1	8,8	13,4	7,9	10,7	15,2
25	20	4,0	5,2	6,7	5,8	8,2	11,0	7,6	10,1	12,8
30	30	4,0	4,9	6,4	5,8	7,6	10,4	7,6	9,4	12,2
35	40	3,7	4,9	6,4	5,5	7,3	8,8	7,3	9,1	10,7
40	60	3,7	4,9	5,2	5,5	7,0	8,2	7,3	8,8	10,1
50	80	3,7	4,6	5,2	5,5	6,7	7,6	7,3	8,6	9,4
60	100	3,7	4,6	4,9	5,2	6,7	7,3	7,0	8,4	9,1
70	150	3,7	4,3	4,6	5,2	6,7	7,3	7,0	8,2	9,1
	Tangent	3,7	4,0	4,3	5,2	6,4	7,0	7,0	8,0	8,2

influence design

Turning roadways are short so that design for Case 1 is usually adequate. It is reasonable to assume, even in the absence of traffic data, that there will be enough trucks in the traffic stream to warrant the application of Condition B to the design. Turning roadway widths are listed in Table 6.11

Three-centred curves are an effective alternative to the single radius curves listed in Table 6.11. These curves typically have a ratio of 3:1:3 between the successive radii. However, asymmetric combinations, e.g. 2:1:4, have also proved very useful in the past. These curves closely follow the wheel path of a vehicle negotiating the turn thus enabling the use of narrower lanes than with a single radius curve. In addition, three-centre curves allow the use of smaller central radii than do the equivalent single curves. Under Case C conditions, a 55:20:55 metre radius three-centred curve is the equivalent of a thirty metre single radius curve and permits the use of a 6,0 metre wide turning roadway compared to the 6,4 metre lane width of the single curve. The three-centred curve is particularly useful for Case C conditions because semi-trailers require an inordinate width of turning roadway. For example, the required width of turning roadway for Case 1, Condition C and a design speed of 20 km/h is 7,9 metres whereas, under the same circumstances, passenger cars require only 4,0 metres. Drivers of passenger cars could thus quite easily perceive the turning roadway as being intended for two-lane operation.

It is not possible to list all the possible alternative three-centred curve combinations. When three-centred curves are considered, the designer should determine the required roadway width by the use of templates.

6.5.3 Tapers

There are two types of taper, each with different geometric requirements. These are:

- The active taper, which forces a lateral transition of traffic; and

- The passive taper, which allows a lateral transition of traffic.

The active taper is used to narrow a roadway or lane or as a lane drop, i.e. when two lanes merge into one. The passive taper either widens or adds a lane. Active tapers constitute a hazard insofar as that a driver that fails to perceive the change in circumstances, may either drive off the travelled way or hit the adjacent kerbing. Passive tapers, on the other hand, create space on the travelled way and, thus, are not hazards. Consequently, active tapers should be long and passive tapers may be short.

If a turning roadway is preceded by an auxiliary lane to allow for deceleration or followed by an auxiliary lane allowing for acceleration, these lanes will be added or dropped by means of passive and active tapers respectively. In the absence of the auxiliary lanes, the turning roadway can, in the case of a left turn, be created by a passive taper from the left edge of the through lane to the left edge of the turning roadway.

The turning roadway may be terminated by an active taper connecting its left edge to the left edge of the through lane. Apropos the suggestion that the angle of skew of the intersection should not fall outside the range of 75° to 120°, it is important to note that a very short active taper may result in the creation of a local angle of skew of less than 75°. This would make it very difficult for the driver on the turning roadway to observe opposing traffic on the through lane. At an angle of skew of less than 5° the driver should, using the rear view mirror, be able to observe opposing traffic comfortably. In short, tapers in the range between 1:10 and 1: 0,3 should be avoided (the latter corresponding to an angle of skew of 75°).

Acceptable tapers rates are suggested in Table 6.12.

In entering a deceleration lane, a vehicle follows a reverse or S-curve alignment, which is effectively a passive taper immediately followed by an active taper. Four different combinations of taper can be employed. These are:

- A straight-line taper;
- A partial tangent taper;
- A symmetrical reverse curve; or
- An asymmetrical reverse curve.

In urban areas, short straight-line tapers appear to offer better targets for the approaching driver. Urban intersections operate at slow speeds during peak periods and, particularly for right-turn-

Table 6.12: Taper rates

Design speed (km/h)	Active tapers (1:)		Passive tapers (1:)
	Painted taper	Kerbed taper	
30	20	10	5
50	25	15	10
60	35	20	15
80	40	25	20
100	45	30	25

ing traffic, the need for storage may be more important than the ability to enter the deceleration lane at relatively high speeds. Tapers could therefore be as sharp as 1:2, which is about the limit of manoeuvrability of a passenger car at crawl speeds.

As speeds are higher in rural than in urban areas, the other forms of taper listed above may warrant consideration. The partial tangent taper is a straight line taper preceded by a short curved section with a radius such that the desired taper rate is achieved at a point about one third of the way across the width of the auxiliary lane. The symmetrical reverse curve taper has curved sections of equal radius at either end. The guideline suggested for the partial tangent, i.e. the curve traversing one third of the lane width, applies to the entry and the exit curves of the taper. The asymmetrical reverse curve usually has an entry curve radius about twice that of the exit curve.

6.6 ROUNDABOUTS

6.6.1 Introduction

Traffic circles constructed in the 1930s and 1940s were intended to operate in a weaving mode. As such, the diameters of the circles were large. No clear guidance was offered regarding priority of one vehicle over another and, in consequence, accident rates at traffic circles tended to be higher than those at conventional intersections. In an effort to rectify this situation, it was decided (in the situation of driving on the left) that vehicles should yield to those on their left. In effect, this meant that vehicles already in the circle had to give way to those wishing to enter it. Not surprisingly, gridlock resulted at heavy flow rates. Ultimately, traffic circles fell into disfavour and were replaced by conventional three- and four-legged intersections.

Modern roundabouts differ from traffic circles in their uniform characteristics and operation. Internationally, roundabouts operate on the "Yield on entry" rule so that, where vehicles drive on the left, vehicles yield to the right and vice versa. South Africa applies the same rule except that, in the case of the mini-roundabout, the rule is slightly modified by the use of the R2.2 sign which *"....indicates to the driver of a vehicle approaching a traffic circle that he shall yield right of way to any vehicle which will cross any yield line at such junction before him and which, in the normal course of events, will cross the path of such driver's vehicle."*

Drivers are thus inclined to adopt the approach that the rule of first-come-first-served applies at mini-roundabouts except that, in the case of simultaneous arrivals at the mini-roundabout, drivers will yield to the vehicle on the right.

A number of geometric elements are incorporated in the design of roundabouts and all elements appear in all roundabouts. These elements are illustrated in Figure 6.11.

6.6.2 Operation of roundabouts

Roundabouts operate by deflecting the vehicle path so as to slow traffic and promote yielding. The roadway entry is usually flared to increase capacity.

Delays at roundabouts are usually less than at conventional intersections and, in consequence, capacity is higher. The proviso is that the combined intersection flow should be less than 3 500 veh/h. Reduced delays improve vehicle operating costs.

Roundabouts have less potential conflict points than conventional intersections. In the case of the four-legged intersection, 32 conflict points are replaced by 8, as illustrated in Figure 6.12. In both roundabouts and conventional intersections, the diverge is also counted as being a conflict point. The safety performance of roundabouts is often superior to that of most conventional intersections and the reduced number of conflict points at roundabouts result in an observable reduction in accident rates.

In spite of their undoubted advantages, roundabouts are not appropriate to every situation. They may be inappropriate:

- Where spatial restraints (including cost of land), unfavourable topography or high construction costs make it impossible to provide an acceptable geometric design;
- Where traffic flows are unbalanced, with high flows on one or more approaches;
- At intersections of major roads with minor roads, where roundabouts would cause serious delays to traffic on the major roads;
- Where there are substantial pedestrian flows;
- As an isolated intersection in a network of linked signalised intersections;

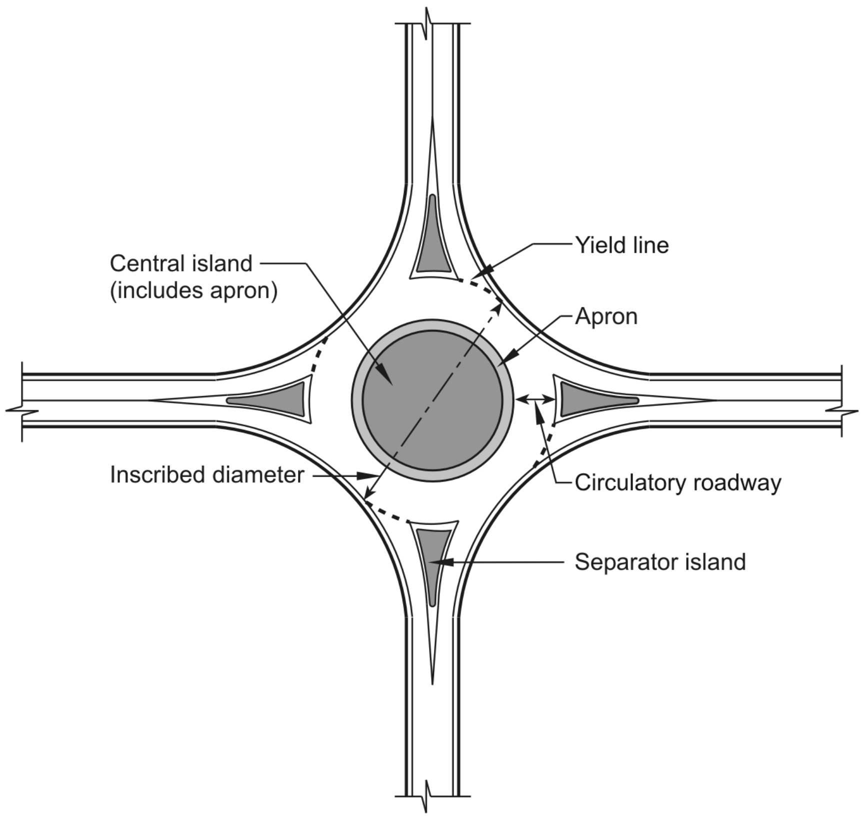


Figure 6.11: Elements of roundabouts

- In the presence of reversible lanes;
- Where semi-trailers and/or abnormal vehicles are a significant proportion of the total traffic passing through the intersection and where there is insufficient space to provide the required layout; and
- Where signalised traffic control down stream could cause a queue to back-up through the roundabout.

Roundabouts can be considered when:

- Intersection volumes do not exceed 3 000 veh/h at three-legged or 4 000 veh/h at four-legged intersections;
- The proportional split between the volumes on the major and minor road does not exceed 70/30;
    - o where, on three-legged intersections, the intersection flow is less than 1 500 veh/h or,
    - o on four-legged intersections, is less than 2 000 veh/h;
- The proportional split between the vol umes on the major and minor road does not exceed 60/40 where;
    - o on three-legged intersections, the intersection flow is greater than 1 500 veh/h or;
    - o on four-legged intersections, is greater than 2 000 veh/h;
- One major flow has a predominant through movement that is:
    - o Between 50 and 80 per cent of the approach volume; or
    - o Between 25 and 40 per cent of the intersection volume; and
    - o High volumes of right-turning movements, i.e. more than 25 per cent of the approach volume, occur and which experience long delays (e.g. 15 seconds per vehicle) and a high incidence of right-angled accidents.

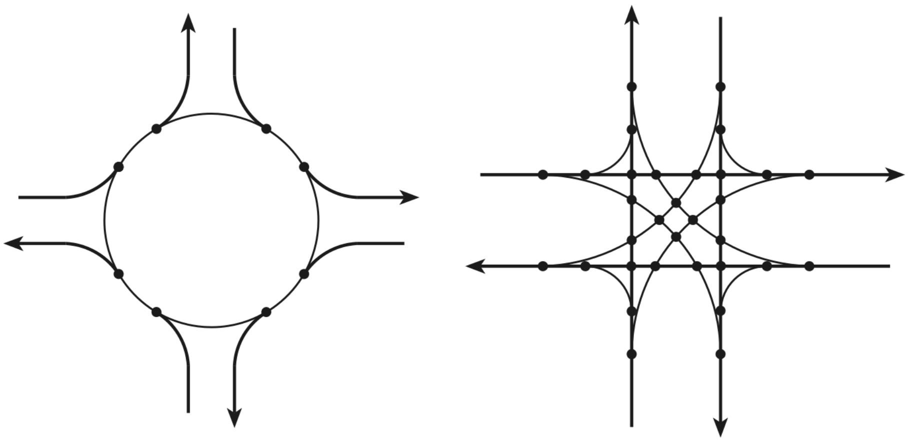


Figure 6.12: Intersection conflict points

6.6.3 Design speed

The design speed within the roundabout should ideally range between 40 to 50 km/h. Unfortunately, this suggests a radius of between 60 and 80 metres hence requiring an overall diameter of the roundabout of the order of 150 metres. Very often, the space for this size of intersection will simply not be available and some lesser design speed will have to be accepted.

Where the design speeds on the approaches are high, e.g. more than 15 km/h faster than the design speed within the facility, it may be necessary to consider forcing a reduction in vehicle speed. This could be by means of horizontal reverse curvature. The ratio between the radii of successive curves should be of the order of 1,5:1.

Speed humps should not be employed as speed-reducing devices on major roads or on bus routes. Where design speeds are of the order of 100 km/h or more, the speed hump would have to be long and the height low to ensure that the vertical acceleration caused by the speed hump does not cause the driver to lose control. It is suggested that, in practice, a suitable profile would be difficult to construct. Bus passengers, particularly those sitting in the rear overhang of the vehicle, would find traversing a speed hump distinctly uncomfortable. Speed humps should thus only be used at roundabouts in residential areas or where the intention is to apply traffic calming.

6.6.4 Sight distance

Site visibility is important in the design of roundabouts. Specifically, approaching drivers should have a clear view of the nose of the splitter or separator island. At the yield line and while traversing the roundabout, they should have an uninterrupted view of the opposing legs of the intersection at all times. This requirement suggests that the elaborate landscaping schemes sometimes placed on the central islands of roundabouts are totally inappropriate to the intended function of the layout.

Decision sight distance for intersections as described in Table 3.5 should be provided on each approach to the roundabout to ensure that drivers can see the nose of the splitter island. It follows that roundabouts should not be located on crest curves.

6.6.5 Components

The various components of a roundabout are illustrated in Figure 6.11.

Deflection

A very important component is the deflection forced on vehicles on the approach to the roundabout. The intention is to reduce the speed of vehicles so that, within limits, the greater the deflection the better. The limit is that the minimum acceptable angle of skew at an intersection is 60°, as discussed previously in this chapter. This corresponds to a deflection on entry of 30°. The approach radius should not exceed 100 metres, which corresponds to the recommended design speed of 40 to 50 km/h.

Entries and exits

The widths of single lane approaches to roundabouts are typically of the order of 3,4 to 3,7 metres. The entry width is one of the most important factors in increasing the capacity of the roundabout and can be increased above the width of the approach by flaring, i.e. by providing a passive taper with a taper rate of 1:12 to 1:15. The recommended minimum width for a single-lane entry is 5 metres.

If demanded by high approach volumes, the flaring could add a full lane to the entry to increase capacity. The width of a two-lane entry should be of the order of 8 metres. A variation on the two-lane entry is to have, in effect, a single-lane circulatory roadway with an auxiliary lane provided for the benefit of vehicles turning left at the roundabout. The auxiliary lane is shielded from traffic approaching from the right by moving the end of the splitter island forward to provide a circulatory road width adequate only for single-lane travel. This approach could be adopted with advantage when left-turning traffic represents 50 per cent or more of the entry flow or more than 300 veh/h during peak hours.

Circulatory roadway

The circulatory roadway width is a function of the swept path of the design vehicle and of the layout of the exits and entries and generally should be either equal to or 1,2 times the width of the entries. The width should be constant throughout the circle.

In the construction of the swept path of the design vehicle, it should be noted that drivers tend to position their vehicles close to the outside kerbs on entry to and exit from the roundabout and close to the central island between these two points. The vehicle path, being the path of a point at the centre of the vehicle, should thus have an adequate offset to the outside and inside kerbs. For a vehicle with an overall width of 2,6 metres, the offset should thus be not less than 1,6 metres with 2,0 metres being preferred. To ensure that vehicles do not travel faster than the design speed, the maximum radius on the vehicle path should be kept to 100 metres or less.

As a general guideline, the circulatory roadway should be sufficiently wide to allow a stalled vehicle to be passed but without sufficient trucks in the traffic stream to influence design (normally described as Case 2, Condition A in the case of turning roadways). The minimum roadway width for single-lane operation under these circumstances would be of the order of 6,5 metres between kerbs. Two-lane operation would require a roadway width of about 8 metres. If trucks are present in the traffic stream in sufficient numbers to influence design, the circulatory road width should be increased by 3 metres both in the single-lane and in the two-lane situation. A significant proportion of semi-trailers would require the width of the circulatory road width to be increased to 13 metres and 16 metres in the single-lane and the two-lane situation respectively.

A circulatory road width of 13 metres would make it possible for passenger cars to traverse the roundabout on relatively large radius curves and at correspondingly high speeds. To avoid this possibility, the central island should be modified as discussed below.

The cross-slope on the roadway should be away from the central island and equal to the camber on the approaches to the intersection.

Central island

The central island consists of a raised non-traversable area, except in the case of mini-roundabouts where the central island may simply be a painted dot. The island is often landscaped but it should be ensured that the landscaping does not obscure the sight lines across the island. Historically, central islands were often square or, if they had more than four entries, polygonal. Negotiating the right angle bends was only possible at crawl speeds and this led to substantial delays and congestion. It is now customary to provide circular islands.

While, for semi-trailers, the width of the circulatory road between kerbs would have to be 13 metres in a single-lane configuration, all other vehicles could be served by a road width of 9,5 metres. A mountable area or apron could thus be added to the central island to accommodate this difference. The apron should have crossfall steeper than that of the circulatory road, principally to discourage passenger vehicles from driving on it and a crossfall of 4 to 5 per cent is recommended.

Splitter islands

Splitter islands should be provided on the approaches to roundabouts to:

- Allow drivers to perceive the upcoming roundabout and to reduce entry speed;
- Provide space for a comfortable deceleration distance;
- Physically separate entering and exiting traffic;
- Prevent deliberate and highly dangerous wrong-way driving;
- Control entry and exit deflections; and
- Provide a refuge for pedestrians and cyclists and a place to mount traffic signs.

The sizes of splitter islands are dictated by the dimensions of the central island and inscribed circle. As a general guideline, they should have an area of at least 10 square metres so as to ensure their visibility to the oncoming driver. The length of splitter islands should be equal to the comfortable deceleration distance from the design speed of the approach to that of the roundabout.

Ideally, the nose of the splitter island should be offset to the right of the approach road centre-line by about 0,6 to 1 metre.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 7

INTERCHANGES

7.1 INTRODUCTION

7.1.1 General

The principal difference between interchanges and other forms of intersection is that, in interchanges, crossing movements are separated in space whereas, in the latter case,they are separated in time. At-grade intersections accommodate turning movements either within the limitations of the crossing roadway widths or through the application of turning roadways whereas the turning movements at interchanges are accommodated on ramps. The ramps replace the slow turn through an angle of skew that is approximately equal to 90° by high-speed merging and diverging manoeuvres at relatively flat angles.

Grade separations, discussed further in Chapter 10, provide spatial separation between the crossing movements but do not make provision for turning movements. They, therefore, do not qualify for consideration as a form of intersection.

The first interchange ever built was at Woodbridge, New Jersey, and provided (in the context of driving on the right) loops for all left turns and outer connectors for all right turns thus creating the Cloverleaf Interchange. Since then, a variety of interchange forms has been developed. These include the:

- Diamond;
- Par-Clo (from Partial Cloverleaf); and the
- Directional

The various types of interchange configuration are illustrated in Section 7.6. Each basic form can be divided into sub-types. For example, the Diamond Interchange is represented by the narrow diamond, the wide diamond and the split diamond. The most recent development in the Diamond interchange form is the Single Point Diamond Interchange. This form is also referred to as the Urban Interchange.

Historically, the type of interchange to be applied at a particular site would be selected as an input into the design process. In fact, like the cross-section, the interchange is the aggregation of various elements. A more sensible approach is thus to select the elements appropriate to a particular site in terms of the topography, local land usage and traffic movements and then to aggregate them into some or other type of interchange.

7.1.2 Design principles

Manoeuvres in an interchange area occur at high speeds close to the freeway and over relatively short distances. It is therefore important that drivers should experience no difficulty in recognising their route through the interchange irrespective of whether that route traverses the interchange on the freeway or diverts to depart from the freeway to a destination that may be to the left or the right of the freeway. In following their selected route, drivers should be disturbed as little as possible by other traffic. These requirements can be met through the application of the basic principles of interchange design.

The driver has a number of tasks to execute successfully to avoid being a hazard to other traffic. It is necessary to:

- select a suitable speed and accelerate or decelerate to the selected speed within the available distance;
- select the appropriate lane and carry out the necessary weaving manoeuvres to effect lane changes if necessary; and
- diverge towards an off-ramp or merge from an on-ramp with the through traffic.

To maintain safety in carrying out these tasks, the driver must be able to understand the operation of the interchange and should not be surprised or misled by an unusual design characteristic. Understanding is best promoted by consistency and uniformity in the selection of types and in the design of particular features of the interchange.

Interchange exits and entrances should always be located on the left. Right-hand side entrances and exits are counter to driver expectancy and also have the effect of mixing high-speed through traffic with lower-speed turning vehicles. The problem of extracting turning vehicles from the median island and providing sufficient vertical clearance either over or under the opposing freeway through lanes is not trivial. The application of right-hand entrances and exits should only be considered under extremely limiting circumstances. Even in the case of a major fork where two freeways are diverging, the lesser movement should, for preference, be on the left.

Route continuity substantially simplifies the navigational aspects of the driving task. For example, if a driver simply wishes to travel on a freeway network through a city from one end to the other it should not be necessary to deviate from one route to another.

Uniformity of signing practice is an important aspect of consistent design and reference should be made to the SADC Road Traffic Signs Manual.

Ideally, an interchange should have only a single exit for each direction of flow with this being located in advance of the interchange structure. The directing of traffic to alternative destinations on either side of the freeway should take place clear of the freeway itself. In this manner, drivers will be required to take two binary decisions, (Yes/No) followed by (Left/Right), as opposed to a single compound decision. This spreads the workload and simplifies the decision process, hence improving the operational efficiency of the entire facility. Closely spaced successive off-ramps could be a source of confusion to the driver leading to erratic responses and manoeuvres.

Single entrances are to be preferred, also in support of operational efficiency of the interchange. Merging manoeuvres by entering vehicles are an interruption of the free flow of traffic in the left lane of the freeway. Closely spaced entrances exacerbate the problem and the resulting turbulence could influence the adjacent lanes as well.

From the standpoint of convenience and safety, in particular prevention of wrong-way movements, interchanges should provide ramps to serve all turning movements. If, for any reason, this is not possible or desirable, it is nevertheless to be preferred that, for any travel movement from one road to another within an interchange, the return movement should also be provided.

Provision of a spatial separation between two crossing streams of traffic raises the problem of which to take over the top - the perennial Over versus Under debate. The choice of whether the crossing road should be taken over or under the freeway depends on a number of factors, not the least of which is the matter of terrain and construction costs. There are, however, a number of advantages in carrying the crossing road over the freeway. These are:

- Exit ramps on up-grades assist deceleration and entrance ramps on downgrades assist acceleration and have a beneficial effect on truck noise.
- Rising exit ramps are highly visible to drivers who may wish to exit from the freeway.
- The structure has target value, i.e. it provides advance warning of the possibility of an interchange ahead necessitating a decision from the driver whether to stay on the freeway or perhaps to change lanes with a view to the impending departure from the freeway.
- Dropping the freeway into cut reduces noise levels to surrounding communities and also reduces visual intrusion.
- For the long-distance driver on a rural freeway, a crossing road on a structure may represent an interesting change of view.
- The crossing road ramp terminals may include right and left turn lanes, traffic signals and other traffic control devices. Not being obstructed by bridge piers and the like, these would be rendered more visible by taking the crossing road over the freeway.

The other design principles, being continuity of basic lanes, lane balance and lane drops are discussed in Section 7.5 as matters of detailed design.

7.2 INTERCHANGE WARRANTS

7.2.1 Traffic volumes

With increasing traffic volumes, a point will be reached where all the options of temporal separation of conflicting movements at an at-grade intersection have been exhausted. One of the possible solutions to the problem is to provide an interchange.

The elimination of bottlenecks by means of interchanges can be applied to any intersection at which demand exceeds capacity and is not necessarily limited to arterials. Under these circumstances, it is necessary to weigh up the economic benefits of increased safety, reduced delay and reduced operating and maintenance cost of vehicles against the cost of provision of the interchange. The latter includes the cost of land acquisition, which could be high, and the cost of construction. As the construction site would be heavily constricted by the need to accommodate traffic flows that were sufficiently heavy to justify the interchange in the first instance, the cost of construction could be significantly higher than on the equivalent green field site.

7.2.2 Freeways

The outstanding feature of freeways is the limitation of access that is brought to bear on their operation. Access is permitted only at designated points and only to vehicles travelling at or near freeway speeds. As such, access by means of intersections is precluded and the only permitted access is by way of interchanges. Crossing roads are normally those that are high in the functional road hierarchy, e.g. arterials, although, if these are very widely spaced, it may be necessary to provide an interchange serving a lower order road, for example a collector.

It follows that the connection between two freeways would also be by means of an interchange, in which case reference is to a systems interchange as opposed to an access interchange.

7.2.3 Safety

Some at-grade intersections exhibit high crash rates that cannot be lowered by improvements to the geometry of the intersections or through the application of control devices. Such situations are often found at heavily travelled urban intersections. Crash rates also tend to be high at the intersections on heavily travelled rural arterials where there is a proliferation of ribbon development.

A third area of high crash rates is at intersections on lightly travelled low volume rural locations where speeds tend to be high. In these cases, low-cost interchanges such as the Jug-handle layout may be an adequate solution to the problem.

7.2.4 Topography

The topography may force a vertical separation between crossing roads at the logical intersection location. As an illustration, the through road may be on a crest curve in cut with the crossing road at or above ground level. If it is not possible to relocate the intersection, a simple Jug-handle type of interchange as illustrated in Figure 7.13 may be an adequate solution to the problem.

7.3 WEAVING

The Highway Capacity Manual (2000) defines weaving as the crossing of two or more traffic streams travelling in the same general direction without the aid of traffic control devices but then goes to address the merge-diverge as a separate issue. However, the merge-diverge operation, associated with successive single-lane on- and off-ramps where there is no auxiliary lane, does have two streams that, in fact, are crossing. Reference to weaving should thus include the merge-diverge.

Three types of weave are illustrated in Figures 7.1, 7.2 and 7.3. A Type A weave requires all weaving vehicles to execute one lane change. Type B weaving occurs when one of the weaving streams does not have to change lanes but the other has to undertake at most one lane change. Type C weaving allows one stream to weave without making a lane change, whereas the other stream has to undertake two or more lane changes.

The Type B weave is, in essence, a Type A weave but with the auxiliary lane extending

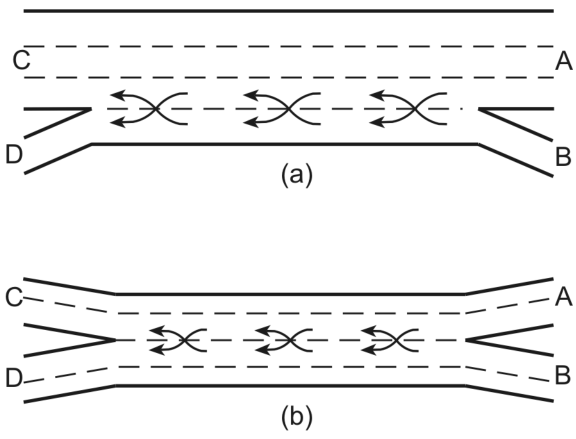


Figure 7.1: Type A weaves: (a) ramp-weave
(b) major weave with crown line

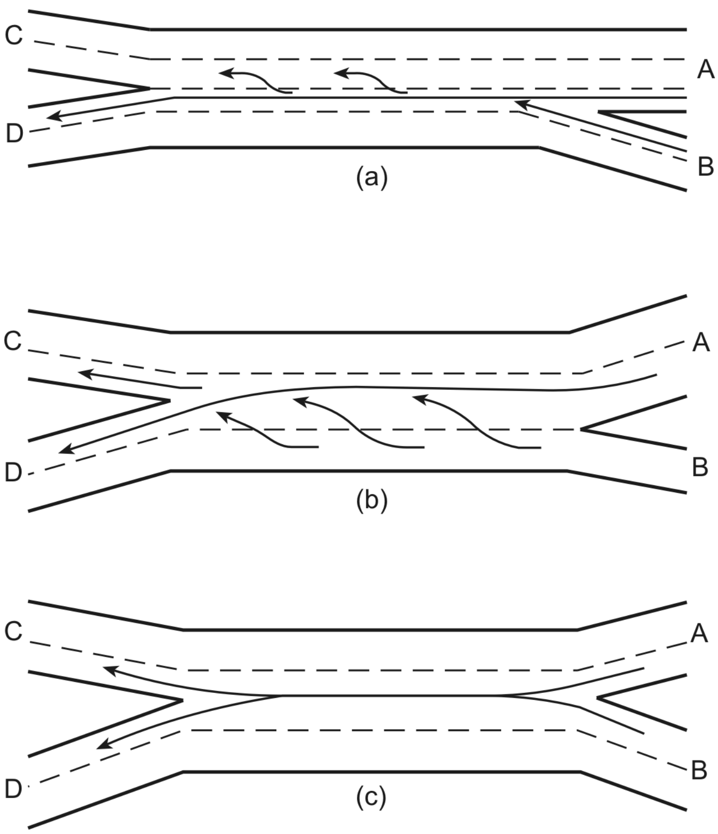


Figure 7.2: Type B weaves:
(a) major weave with lane balance at exit
(b) major weave with merging at entrance
(c) major weave with merging at entrance and lane balance at exit

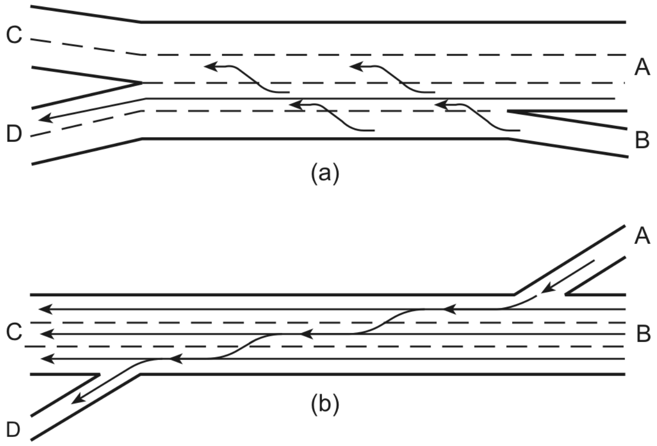


Figure 7.3: Type C weaves: (a) major weave without lane balance
(b) two-sided weave

either up- or downstream of the weaving area and with an additional lane being provided either to the on- or to the off-ramp. It follows that a Type A weaving section can be easily converted into a Type B weave. At any site at which a Type A weave appears, it would thus be prudent to check the operation at the site for both types of weave. Type C weaves rarely occur in South Africa.

7.4 LOCATION AND SPACING OF INTERCHANGES

The location of interchanges is based primarily on service to adjacent land. On rural freeways bypassing small communities, the provision of a single interchange may be adequate, with larger communities requiring more. The precise location of interchanges would depend on the particular needs of the community but, as a general guide, would be on roads recognised as being major components of the local system. Rural interchanges are typically spaced at distances of eight kilometres apart or more. This distance is measured from centreline to centreline of crossing roads.

The generous spacing applied to rural interchanges would not be able to serve intensively developed urban areas adequately. As an illustration of context sensitive design, trip lengths are shorter and speeds lower on urban freeways than on rural freeways. As drivers are accustomed to taking a variety of alternative actions in rapid succession a spacing of closer than eight kilometres can be considered.

At spacings appropriate to the urban environment, reference to a centreline-to-centreline distance is too coarse to be practical. The point at issue is that weaving takes place between interchanges and the available distance is a function of the layout of successive interchanges. For a common centreline-to-centreline spacing, the weaving length available between two diamond interchanges is significantly different from that between a Par-Clo-A followed by a Par-Clo-B. Weaving distance is defined in the Highway Capacity Manual 2000 and other sources as the distance between the point at which the separation between the ramp and the adjacent lane is

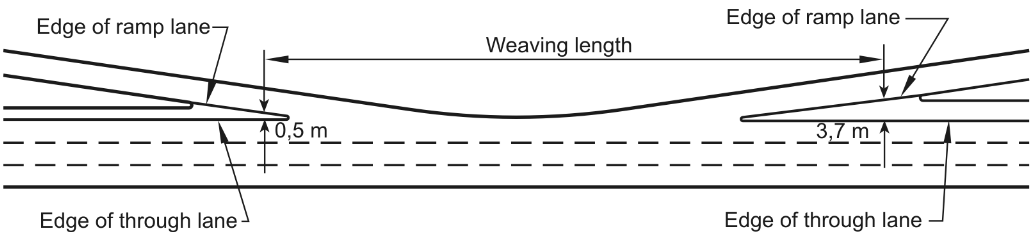


Figure 7.4: Weaving distance

0,5 metres to the point at the following off-ramp at which the distance between ramp and lane is 3,7 m as illustrated in Figure 7.4.

If this definition is adopted, the weaving length becomes a function of the rates of taper applied to the on- and off-ramps. Reference to the Yellow Line Break Point (YLBP) distance is totally unambiguous and is the preferred option.

Table 7.1: Interchange spacing in terms of signage requirements		
Configuration	Urban areas	Rural areas
Access to access	1 300 m	2 200 m
Access to systems	2 100 m	3 300 m
Systems to access	1 400 m	2 200 m
Systems to systems	2 500 m	3 800 m

Three criteria for the spacing of interchanges can be considered. In the first instance, the distance required for adequate signage should ideally dictate spacing of successive interchanges. If it is not possible to achieve these distances, consideration can be given to a relaxation based on achieving Level Of Service (LOS) D conditions on the freeway. The third criterion is that of turbulence, which is applied to the merge-diverge situation.

(1) The distance required by the SADC Road Traffic Signs Manual to provide adequate sign posting which, in turn, influences the safe operation of the freeway, is used to define the minimum distance between ramps. The minimum distances to be used for detailed design purposes, as measured between Yellow Line Break Points, for different areas and interchange types should be not less than the values stated in Table 7.1.

(2) In exceptional cases, where it is not possible to meet the above requirements, relaxation of these may be considered. It is, however, necessary to ensure that densities in the freeway left hand lane are not so high that the flow of traffic breaks down. Densities associated with LOS E would make it difficult, if not actu-

ally impossible, for drivers to be able to change lanes. Formulae according to which densities can be estimated are provided in the Highway Capacity Manual (2000). Drivers need time to locate a gap and then to position themselves correctly in relation to the gap while simultaneously adjusting their speed to that required for the lane change. The actual process of changing lanes also requires time.

(3) In the case of the merge-diverge manoeuvre, turbulence caused on the left lane of the freeway by a close succession of entering and exiting vehicles becomes an issue. According to Roess and Ulerio this turbulence manifests itself over a distance of roughly 450 metres upstream of an off-ramp and downstream of an on-ramp. A spacing of 900 metres would suggest that the entire length of freeway between interchanges would be subject to turbulent flow. The likelihood of breakdown in the traffic flow would thus be high and the designer should ensure that space is available for one area of turbulence to subside before onset of the next.

(4) In off-peak periods, vehicles would be moving between interchanges at the design speed or higher. The geometry of the on- and off-ramps should be such that they can accommodate manoeuvres at these speeds. Increasing the taper rates or reducing the length of the speed change lanes purely to achieve some or other hypothetically acceptable Yellow Line Break Point distance does not constitute good design.

(5) The spacing between successive interchanges will have an impact on traffic operations on the crossing roads and vice versa. If the crossing road can deliver vehicles to the freeway faster than they can carry out the merge, stacking of vehicles will occur on the on-ramp with the queue possibly backing up on to the crossing road itself. Stacking can also occur on an off-ramp if the crossing road ramp terminal cannot accommodate the rate of flow arriving from the freeway. The queue could conceivably back up onto the freeway, which would create an extremely hazardous situation.

It should be realised that relaxations below the distances recommended under (1) above will result in an increase in the driver workload. Failure to accommodate acceptable levels of driver workload in relation to reaction times can be expected to result in higher than average crash rates. Twomey et al demonstrate that, at spacings between noses of greater than 2 500 metres, the crash rate is fairly constant, i.e. the presence of the following interchange is not a factor in the crash rate. At spacings of less than 2500 m between noses, the crash rate increases until, at about 500 m between noses, the crash rate is nearly double that of the 2500 m spacing. This is illustrated in Figure 7.2 below

The question that must be addressed is the benefit that the community can expect to derive in exchange for the cost of the higher accident rate. By virtue of the fact that freeway speeds tend to be high, there is a high probability that many of the accidents would be fatal. It is therefore suggested that the decision to reduce the interchange spacing below those listed in Table 7.1 should not be taken lightly.

It would be necessary to undertake a full-scale engineering analysis of the situation that would include:

- estimation of future traffic volumes at a

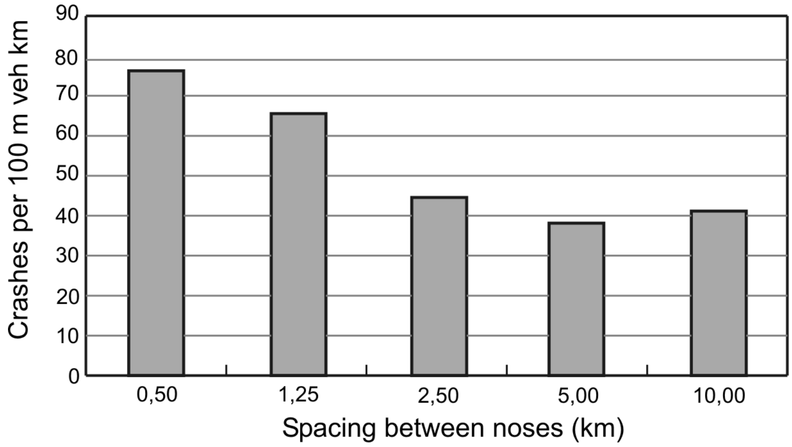


Figure 7.5: Relationship between interchange spacing and accident rate

ten to twenty year time horizon, comprising weaving and through volumes in the design year;

- calculation of traffic densities;
- assessment of the local geometry in terms of sight distances, and horizontal and vertical alignment;
- development of a sign sequence; and
- a form of benefit/cost analysis relating community benefits to the decrease in traffic safety.

Density offers some indication of the level of exposure to risk and, for want of any better measure, it is suggested that a density higher than 22 vehicles/kilometre/lane, corresponding to LOS D, would not result in acceptable design. It would be necessary to pay attention to remedial actions to prevent interchange constraints, such as inadequate ramp capacity, signalling or crossroad volumes, causing back up onto the freeway

In summary: Spacings of interchanges in terms of their YLBP distances should desirably be in accordance with Table 7.1. If these spacings are not achievable and an interchange is absolutely vital for service to the community and adjacent land uses, relaxations may be considered but then, to minimise the risk of crashes, the density calculated according to the above-mentioned formulae should not exceed 22 vehicles/kilometre/lane, which corresponds to LOS D.

7.5 BASIC LANES AND LANE BALANCE

Basic lanes are those that are maintained over an extended length of a route, irrespective of local changes in traffic volumes and requirements for lane balance. Alternatively stated, the basic number of lanes is a constant number of lanes assigned to a route, exclusive of auxiliary lanes.

The number of basic lanes changes only when there is a significant change in the general level of traffic volumes on the route. Short sections of the route may thus have insufficient capacity, which problem can be overcome by the use of auxiliary lanes. In the case of spare capacity, reduction in the number of lanes is not recommended because this area could, at some future time, become a bottleneck. Unusual traffic demands, created by accidents, maintenance or special events, could also result in these areas becoming bottlenecks.

The basic number of lanes is derived from consideration of the design traffic volumes and capacity analyses. To promote the smooth flow of traffic there should be a proper balance of lanes at points where merging or diverging manoeuvres occur. In essence, there should be one lane where the driver has the choice of a change of direction without the need to change lanes.

At merges, the number of lanes downstream of the merge should be one less than the number of lanes upstream of the merge. This is typified by a one-lane ramp merging with a two-lane carriageway that, after the merge, continues as a two-lane carriageway as is the case on a typical Diamond Interchange layout. This rule precludes a two-lane ramp immediately merging with the carriageway without the addition of an auxiliary lane.

At diverges, the number of lanes downstream of the diverge should be one more than the number upstream of the diverge. The only exception to this rule is on short weaving sections, such as at Cloverleaf Interchanges, where a condition of this exception is that there is an auxiliary lane through the weaving section.

When two lanes diverge from the freeway, the above rule indicates that the number of freeway lanes beyond the diverge is reduced by one. This can be used to drop a basic lane to match anticipated flows beyond the diverge. Alternatively, it can be an auxiliary lane that is dropped.

Basic lanes and lane balance are brought into harmony with each other by building on the basic lanes, adding or removing auxiliary lanes as required. The principle of lane balance should always be applied in the use of auxiliary lanes. Operational problems on existing roadways can be directly attributed to a lack of lane balance and failure to maintain route continuity.

The application of lane balance and coordination with basic number of lanes is illustrated in Figure 7.6

7.6 AUXILIARY LANES

As in the case of the two-lane two-way road cross-section with its climbing and passing lanes, and the intersection with its right- and left-turning lanes, the auxiliary lane also has its role to play in the freeway cross-section and the interchange. In a sense, the application of the auxiliary lane in the freeway environment is identical to its application elsewhere. It is added to address a local operational issue and, as soon as the need for the auxiliary lane is past, it is dropped.
Important features to consider in the application and design of the auxiliary lane are thus:

- The need for an auxiliary lane;
- The terminals;
- Driver information

7.6.1 The need for an auxiliary lane

Auxiliary lanes are normally required on free-

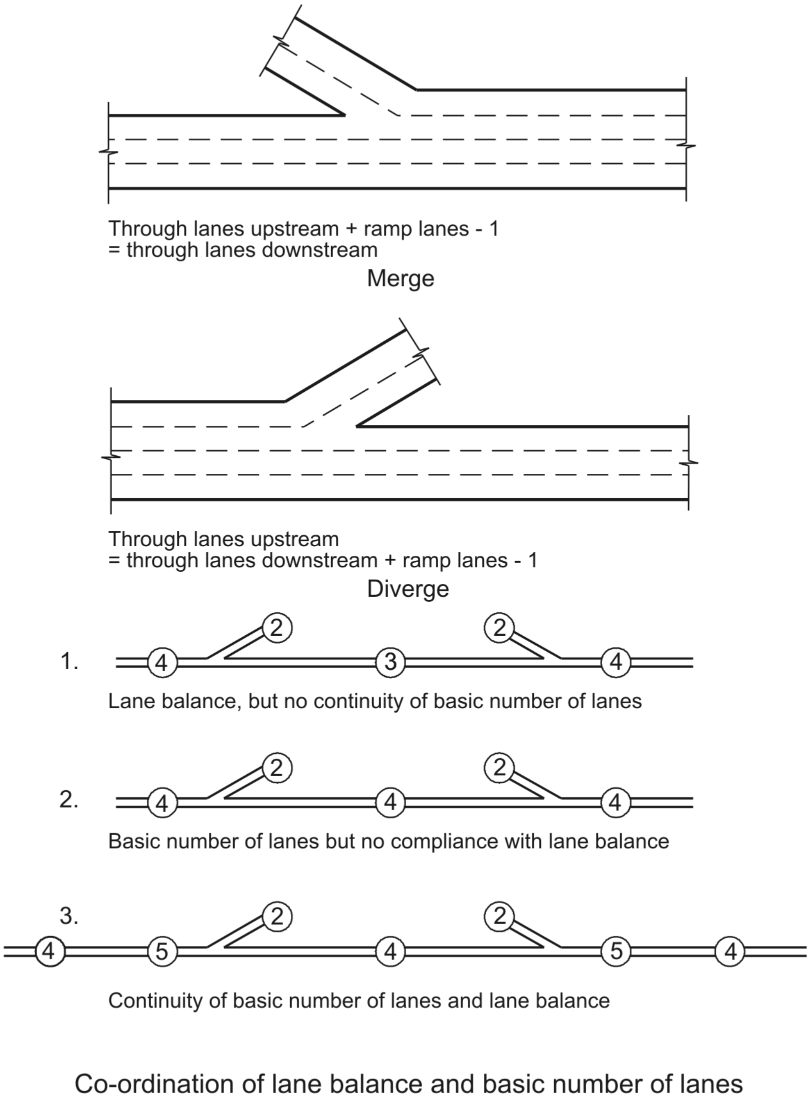


Figure 7.6: Coordination of lane balance with basic number of lanes

ways either as:

- climbing lanes; or
- to support weaving; or
- to support lane balance.

The climbing lane application is similar to that discussed in Chapter 4 in respect of two-lane two-way roads whereas the weaving and lane balance applications are unique to the freeway situation.

Climbing lanes

Ideally, maximum gradients on freeways are in the range of three to four per cent ensuring that most vehicles can maintain a high and fairly constant speed. However, in heavily rolling country it is not always possible to achieve this ideal without incurring excessive costs in terms of earthworks construction. Because of the heavy volumes of traffic that necessitate the provision of a freeway, lane changing to over-take a slow-moving vehicle is not always easy and, under peak flow conditions, may actually be impossible. Speed differentials in the traffic stream are thus not only extremely disruptive but may also be potentially dangerous. Both conditions, i.e. disruption and reduction in safety, require consideration.

If a gradient on a freeway is steeper than four per cent, an operational analysis should be carried out to establish the impact of the gradient on the Level of Service. A drop through one level, e.g from LOS B through LOS C to LOS D, would normally suggest a need for a climbing lane.

As discussed in Chapter 4, crash rates increase exponentially with increasing speed differential. For this reason, international warrants for climbing lanes normally include a speed differential in the range of 15 to 20 km/h. South Africa has adopted a truck speed reduction of 20 km/h as its speed-based warrant for climbing lanes. If, on an existing freeway, the measured truck speed reduction in the outermost lane is thus 20 km/h or higher, the provision of a climbing lane should be considered. In the case of a new design, it will be necessary to construct a speed profile of the truck traffic to evaluate the need for a climbing lane.

Weaving

In the urban environment, interchanges are fairly closely spaced and local drivers are very inclined to use freeways as part of the local circulation system - a form of rat-running in reverse and as undesirable as the normal form of rat-running where the higher order road is bypassed through the use of local residential streets as long-distance urban routes. To ensure that the freeway is not unduly congested because of this practice, an auxiliary lane can be provided between adjacent interchanges resulting in Type A weaving as described in Section 7.3.

If a large number of vehicles are entering at the upstream interchange, it may be necessary to provide a two-lane entrance ramp. Some of these vehicles may exit at the following interchange but those wishing to travel further will have to weave across traffic from still further upstream that intends exiting at the following interchange and then merge with through traffic on the freeway. The auxiliary lane is then extended beyond the downstream interchange to allow a separation between the two manoeuvres. Similarly, a large volume of exiting vehicles may necessitate a two-lane exit, in which case the auxiliary lane should be extended upstream. Type B weaving thus comes into being. The desired length of the extension of the auxiliary lane beyond the two interchanges is normally assessed in terms of the probability of merging vehicles locating an acceptable gap in the opposing traffic flow.

Lane balance

As discussed in Section 7.5, lane balance requires that:

- In the case of an exit, the number of lanes downstream of the diverge should be one more than the number upstream; and
- In the case of an entrance, the number of lanes downstream of the merge should be one less than the number upstream

This is illustrated in Figure 7.6.

Single-lane on- and off-ramps do not require auxiliary lanes to achieve lane balance in terms of the above definition. It should be noted that, unless two-lane on- and off-ramps are provided, the Type A weave is actually a violation of the principles of lane balance.

To achieve lane balance at an exit, three lanes upstream of the diverge should be followed by a two-lane off-ramp in combination with two basic lanes on the freeway. The continuity of basic lanes requires that the outermost of the three upstream lanes should be an auxiliary lane.

If all three upstream lanes are basic lanes, it is possible that traffic volumes beyond the off-ramp may have reduced to the point where three basic lanes are no longer necessary. Provision of a two-lane exit would thus be a convenient device to achieve a lane drop while simultaneously maintaining lane balance. The alternative would be to provide a single-lane off-ramp, carrying the basic lanes through the interchange and dropping the outside lane some distance beyond the on-ramp terminal. In view of the additional construction costs involved, this approach is not recommended.

If a two-lane on-ramp is joining two basic lanes, lane balance will require that there be three lanes beyond the merge. The outside lane of the three could become a new basic lane if the increase in traffic on the freeway merits it. On the other hand, it is possible that the flow on the ramp is essentially a local point of high density and that two basic lanes are all that are required downstream of the on-ramp. In this case, the outside lane could be dropped as soon as convenient.

7.6.2 Auxiliary lane terminals

An auxiliary lane is intended to match a particular situation such as, for example, an unacceptably high speed differential in the traffic stream. It follows that the full width of auxiliary lane must be provided over the entire distance in which the situation prevails. The terminals are thus required to be provided outside the area of need and not as part of the length of the auxiliary lane.

ntering and exiting from auxiliary lanes require a reverse curve path to be followed. It is thus suggested that the taper rates discussed in Chapter 4.4.3 be employed rather than those normally applied to on- and off-ramps. The entrance taper should thus be about 100 metres long and the exit taper about 200 metres long.

7.6.3 Driver information

The informational needs of drivers relate specifically to needs with regard to the exit from the auxiliary lane and include an indication of:

- the presence of a lane drop;
- the location of the lane drop; and
- the appropriate action to be undertaken

These are fully discussed in Chapter 4.4.3 and reference should be made to that chapter for further detail.

7.7 INTERCHANGE TYPES

7.7.1 General

There is a wide variety of types of interchanges that can be employed under the various circumstances that warrant the application of interchanges. The major determinant of the type of interchange to be employed at any particular site is the classification and characteristics of the intersecting road. Intersecting roads are typically freeways or urban arterials but may also be collectors.

In the case of freeways as intersecting roads, reference is made to systems interchanges. Systems interchanges exclusively serve vehicles that are already on the freeway system.

Access to the freeway system from the surrounding area is via interchanges on roads other than freeways, for which reason these interchanges are known as access interchanges. Service areas, providing opportunities to buy fuel, or food or simply to relax for a while are typically accessed via an interchange. In some instances, the services are duplicated on either side of the freeway, in which access is via a left-in/left-out configuration. The requirements in terms of deflection angle, length of ramp and spacing that apply to interchange ramps apply equally to left-in/left-out ramps. In effect, this situation could be described as being an interchange without a crossing road.

The primary difference between systems and access/service interchanges is that the ramps on systems interchanges have free-flowing terminals at both ends, whereas the intersecting road ramp terminals on an access interchange are typically in the form of at-grade intersections.

Interchanges can also be between non-freeway roads, for example between two heavily trafficked arterials. In very rare instances there may even be an application for an interchange between a major and a local road, as suggested above in the case where local topography may force a grade separation between the two roads.

In addition to the classification and nature of the intersecting road, there are a number of controls guiding the selection of the most appropriate interchange form for any particular situation. In the sense of context sensitive design, these include;

- Safety;
- Adjacent land use;
- Design speed of both the freeway and the intersecting road;
- Traffic volumes of the through and turning movements;
- Traffic composition;
- Number of required legs;
- Road reserve and spatial requirements;
- Topography;
- Service to adjacent communities;
- Environmental considerations, and
- Economics.

The relative importance of these controls may vary from interchange to interchange. For any particular site, each of the controls will have to be examined and its relative importance assessed. Only after this process will it be possible to study alternative interchange types and configurations to determine the most suitable in terms of the more important controls.

While the selection of the most appropriate type and configuration of interchange may vary between sites, it is important to provide consistent operating conditions in order to match driver expectations.

7.7.2 Systems interchanges

As stated above, at-grade intersections are inappropriate to systems interchanges and their avoidance is mandatory. For this reason, hybrid interchanges, in which an access interchange is contained within a systems interchange, are to be avoided.

Hybrid interchanges inevitably lead to an unsafe mix of high and low speed traffic. Furthermore, signposting anything up to six possible destinations within a very short distance is, at best, difficult. Selecting the appropriate response generates an enormous workload for the driver so that the probability of error is substantial. Past experience suggests that these interchange configurations are rarely successful.

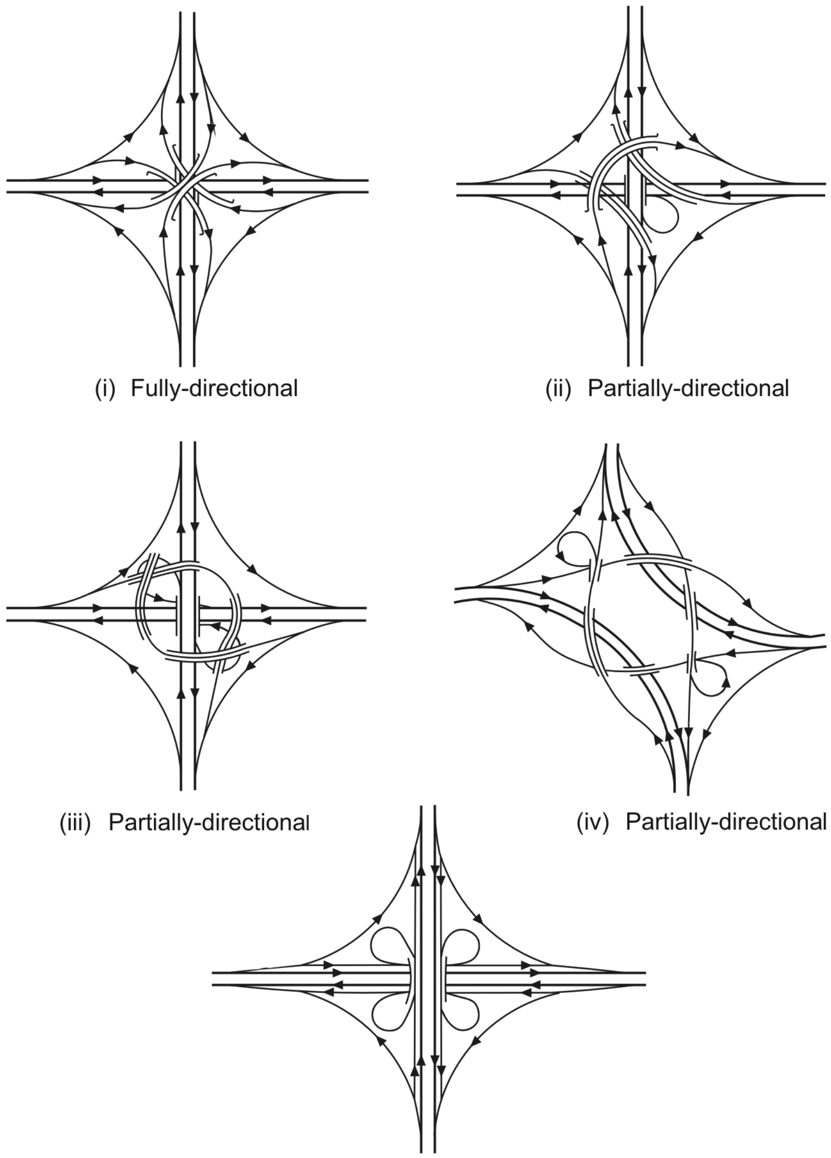
Directional interchanges provide high-speed connections to left and to right provided that the ramp exits and entrances are on the left of the through lanes. Where turning volumes are low or space is limited, provision of loops for right turning traffic can be considered. Directional interchanges that include one or more loops are referred to as being partially-directional. If all right turns are required to take place on loops, the cloverleaf configuration emerges. Various forms of systems interchanges are illustrated below.

Four-legged interchanges

The fully directional interchange illustrated in Figure 7.7 (i) provides single exits from all four directions and directional ramps for all eight turning movements. The through roads and ramps are separated vertically on four levels.

Partially directional interchanges allow the number of levels to be reduced. The Single Loop Partially-directional Interchange, illustrated in Figure 7.7 (ii), and the Two Loop arrangement, illustrated in Figure 7.7 (iii) and (iv), require three levels. The difference between Figures 7.7 (iii) and (iv) is that, in the former case, the freeways cross and, in the latter, route continuity dictates a change in alignment. Loop ramps are normally only used for lighter volumes of right-turning traffic. A three-loop arrangement is, in effect, a cloverleaf configuration, with one of the loops being replaced by a directional ramp and is not likely to occur in practice, largely because of the problem of weaving discussed below.

The principal benefit of the cloverleaf is that it requires only a simple one-level structure, in contrast to the complex and correspondingly costly structures necessary for the directional and partially directional configurations. The major weakness of the cloverleaf is that it requires weaving over very short distances. Provided weaving volumes are not high and sufficient space is available to accommodate the interchange, the cloverleaf can, however, be considered to be an option. If weaving is required to take place on the main carriageways, the turbulence so created has a serious effect on the flow of traffic through the interchange area. The cloverleaf also has the characteristic of confronting the driver with two exits from the freeway in quick succession. Both these problems can be resolved by providing collector-distributor roads adjacent to the through carriageways.



(i) Fully-directional
(ii) Partially-directional
(iii) Partially-directional
(iv) Partially-directional
(v) Full cloverleaf

Figure 7.7: Four-legged systems interchanges

Three-legged interchanges

Various fully-directional and partially-directional three legged interchanges are illustrated in Figure 7.8. In Figure 7.8 (i), one single structure providing a three-level separation is required. Figure 7.8 (ii) also requires three levels of roadway but spread across two structures hence reducing the complexity of the structural design.

It is also possible with this layout to slightly reduce the height through which vehicles have to climb. Figure 7.8 (iii) illustrates a fully-directional interchange that requires only two but widely separated structures. If North is assumed as being at the top of the page, vehicles turning from West to South have a slightly longer path imposed on them so that this should, ideally be the lesser turning volume.

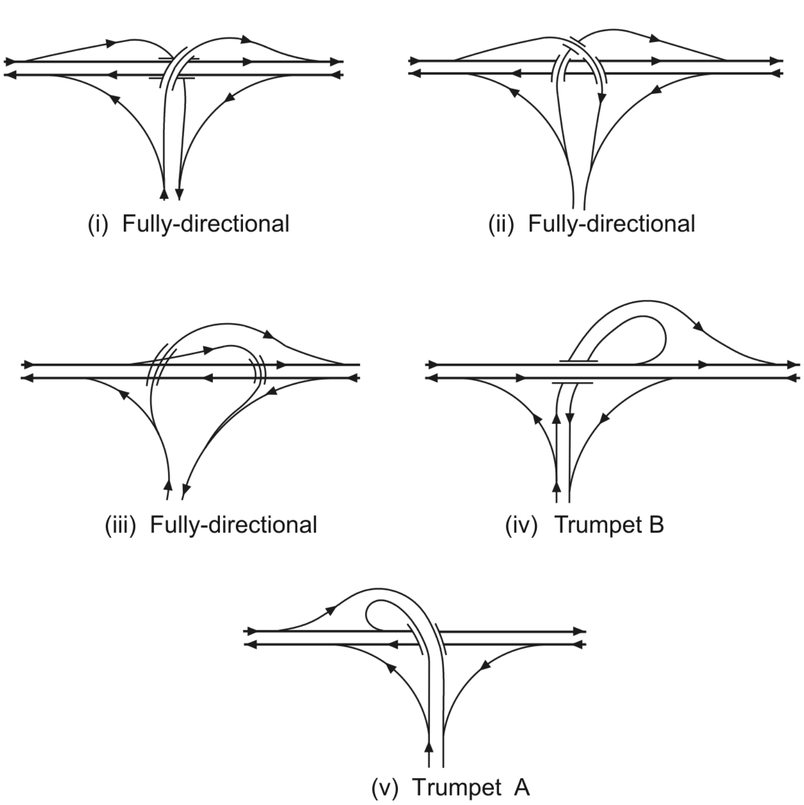


Figure 7.8: Three-legged systems interchanges

Figures 7.8 (iv) and (v) show semi-directional interchanges. Their names stem from the loop ramp located within the directional ramp creating the appearance of the bell of a trumpet. The letters "A" and "B" refer to the loop being in Advance of the structure or Beyond it. The smaller of the turning movements should ideally be on the loop ramp but the availability of space may not always make this possible.

7.7.3 Access and service interchanges

In the case of the systems interchange, all traffic enters the interchange area at freeway speeds. At access and service interchanges, vehicles entering from the crossing road may be doing so from a stopped condition, so that it is necessary to provide acceleration lanes to ensure that they enter the freeway at or near freeway speeds. Similarly, exiting vehicles should be provided with deceleration lanes to accommodate the possibility of a stop at the crossing road.

As previously discussed, there is distinct merit in the crossing road being taken over the freeway as opposed to under it. One of the advantages of the crossing road being over the freeway is that the positive and negative gradients respectively support the required deceleration and acceleration to and from the crossing road. The final decision on the location of the crossing road is, however, also dependent on other controls such as topography and cost.

Access interchanges normally provide for all turning movements. If, for any reason, it is deemed necessary to eliminate some of the turning movements, the return movement, for any movement that is provided, should also be provided. Movements excluded from a particular interchange should, desirably, be provided at the next interchange upstream or downstream as, without this provision, the community served loses amenity.

There are only two basic interchange types that are appropriate to access and service interchanges. These are the Diamond and the Par-Clo interchanges. Each has a variety of possible configurations.

Trumpet interchanges used to be considered suitable in cases where access was to provided to one side only, for example to a bypass of a town or village. In practice, however, once a bypass has been built it does not take long before development starts taking place on the other side of the bypass. The three-legged interchange then has to be converted into a four-legged interchange. Conversion to a Par-Clo can be achieved at relatively low cost. Other than in the case of the Par-Clo AB, one of the major movements is forced onto a loop ramp. The resulting configuration is thus not appropriate to the circumstances. In practice, the interchange should be planned as a Diamond in the first instance, even though the crossing road, at the time of construction, stops immediately beyond the interchange.

Diamond Interchanges

There are three basic forms of Diamond, being:

- The Simple Diamond;
- The Split Diamond, and the
- Single Point Interchange.

The Simple Diamond is easy for the driver to understand and is economical in its use of space. The major problem with this configuration is that the right turn on to the crossing road can cause queuing on the exit ramp. In extreme cases, these queues can extend back onto the freeway, creating a hazardous situation. Where the traffic on the right turn is very heavy, it may be necessary to consider placing it on a loop ramp. This is the reverse of the situation on systems interchanges where it is the lesser volumes that are located on loop ramps. It has the advantage that the right turn is converted into a left-turn at the crossing road ramp terminal. By the provision of auxiliary lanes, this turn can operate continuously without being impeded by traffic signals.
The Simple Diamond can take one of two configurations: the Narrow Diamond and the Wide Diamond.

The Narrow Diamond is the form customarily applied. In this configuration, the crossing road ramp terminals are very close in plan to the freeway shoulders to the extent that, where space is heavily constricted, retaining walls are located just outside the freeway shoulder breakpoints. Apart from the problem of the right turn referred to above, it can also suffer from a lack of intersection sight distance at the crossing road ramp terminals. This problem arises when the crossing road is taken over the freeway and is on a minimum value crest curve on the structure. In addition, the bridge balustrades can also inhibit sight distance. In the case where the crossing road ramp terminal is signalised, this is less of a problem, although a vehicle accidentally or by intent running the red signal could create a dangerous situation.

The Wide Diamond was originally intended as a form of stage construction, leading up to conversion to a full Cloverleaf Interchange. The time span between construction of the Diamond and the intended conversion was, however, usually so great that, by the time the upgrade became necessary, standards had increased to the level whereby the loop ramps could not be accommodated in the available space. The decline in the popularity of the Cloverleaf has led to the Wide Diamond also falling out of favour.

The Wide Diamond has the problem of imposing a long travel distance on right-turning vehicles but is not without its advantages. The crossing road ramp terminals are located at the start of the approach fill to the structure. To achieve this condition, the ramps have to be fairly long so that queues backing up onto the freeway are less likely than on the Narrow Diamond. The crossing road ramp terminals are also at ground level, which is a safer alternative than having the intersections on a high fill. Finally, because the ramp terminals are remote from the structure, intersection sight distance is usually not a problem.

The Split Diamond can also take one of two forms: the conventional Split and the transposed Split. This configuration is normally used when the crossing road takes the form of a one-way pair. The problems of sight distance and queues backing up are not normally experienced on Split Diamonds and the most significant drawback is that right-turning vehicles have to traverse three intersections before being clear of the interchange. It is also necessary to construct frontage roads linking the two one-way streets to provide a clear route for right-turning vehicles.

The transposed Split has the ramps between the two structures. This results in a very short distance between the entrance and succeeding exit ramps, with significant problems of weaving on the freeway. Scissor ramps are the extreme example of the transposed Split. These require either signalisation of the crossing of the two ramps or a grade separation. The transposed Split has little to recommend it and has fallen into disuse, being discussed here only for completeness of the record.

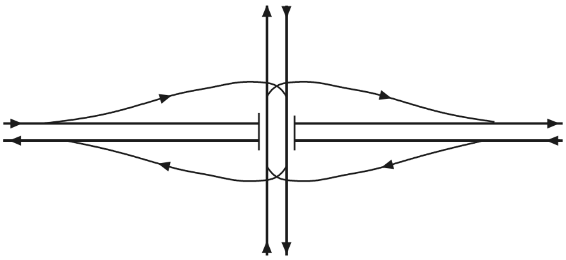
The Single Point Interchange brings the four ramps together at a point over the freeway. This interchange is required where space is at a premium or where the volume of right-turning traffic is very high. The principal operating difference between the Single Point and the Simple Diamond is that, in the former case, the right turns take place outside each other and in the latter they are "hooking" movements. The capacity of the Single Point Interchange is thus higher than that of the Simple Diamond. It does, however, require a three-phase signal plan and also presents pedestrians with wide unprotected crossings.

The various configurations of the Diamond Interchange are illustrated in Figure 7.9

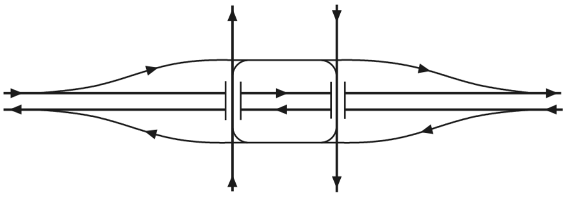
Par-Clo interchanges

Par-Clo interchanges derive their name as a contraction of PARtial CLOverleaf, mainly because of their appearance, but also because they were frequently a first stage development of a Cloverleaf Interchange. In practice, they could perhaps be considered rather as a distorted form of Simple Diamond Interchange.

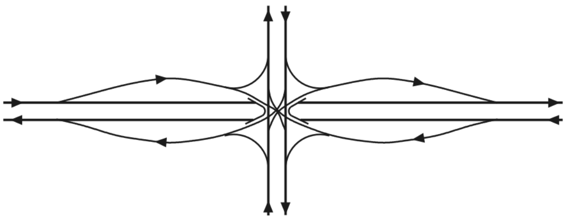
Three configurations of Par-Clo Interchange are possible: the Par-Clo A, the Par-Clo B and the Par-Clo AB. As in the case of the Trumpet Interchange, the letters have the significance of the loops being in advance of or beyond the structure. The Par-Clo AB configurationhas the loop in advance of the structure for the one direction of travel and beyond the structure for the other. In all cases, the loops are on opposite sides of the freeway. Both the Par-Clo A and the Par-Clo B have alternative configurations: the A2 and A4 and the B2 and B4. These



Simple diamond



Split diamond



Single point diamond

Figure 7.9: Diamond interchanges

configurations refer to two quadrants only being occupied or alternatively to all four quadrants having ramps. The four-quadrant layout does not enjoy much, if any, usage in South Africa.

The various layouts are illustrated in Figures 7.10, 7.11 and 7.12.

Internationally, the Par-Clo A4 is generally regarded as being the preferred option for an interchange between a freeway and a heavily trafficked arterial. In the first instance, the loops serve vehicles entering the freeway whereas, in the case of the Par-Clo B, the high-speed vehicles exiting the freeway are confronted by the loop. This tends to surprise many drivers and loops carrying exiting traffic have higher accident rates than the alternative layout. Secondly, the left turn from the crossing road is remote from the intersections on the crossing road and the only conflict is between right-turning vehicles exiting from the freeway and through traffic on the crossing road. This makes two-phase signal control possible.

The Par-Clo AB is particularly useful in the situation where there are property or environmental restrictions in two adjacent quadrants on the same side of the crossing road. Examples include a road running alongside a river or the situation, frequently found in South Africa, of a transportation corridor containing parallel road and rail links in close proximity to each other.

The Rotary Interchange illustrated in Figure 7.12 has the benefit of eliminating intersections on the crossing road, replacing them by short weaving sections. Traffic exiting from the free-

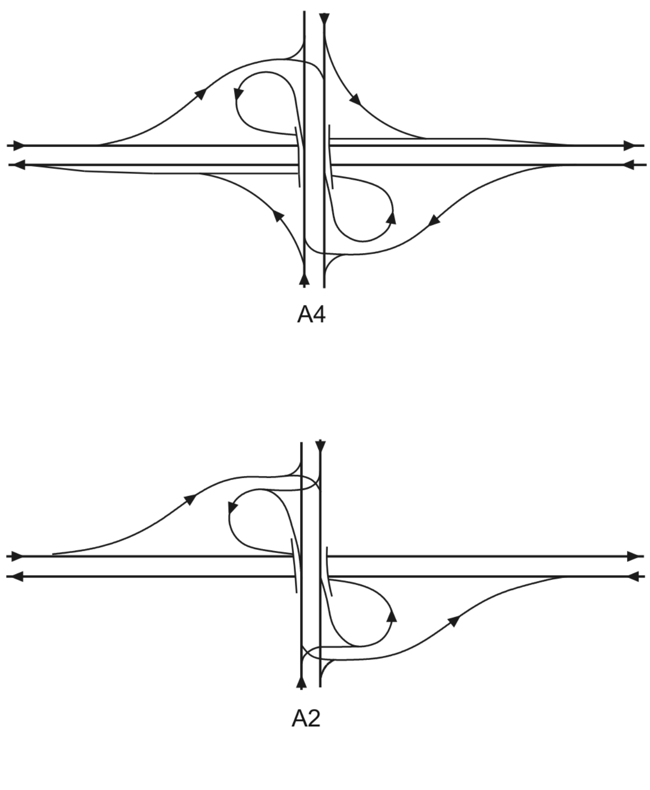


Figure 7.10: Par-Clo A interchanges

way may experience difficulty in adjusting speed and merging with traffic on the rotary.

Rotaries have also been used in the United Kingdom as systems interchanges. In this configuration, a two-level structure is employed. One freeway is located at ground level and the other freeway on the upper level of the structure

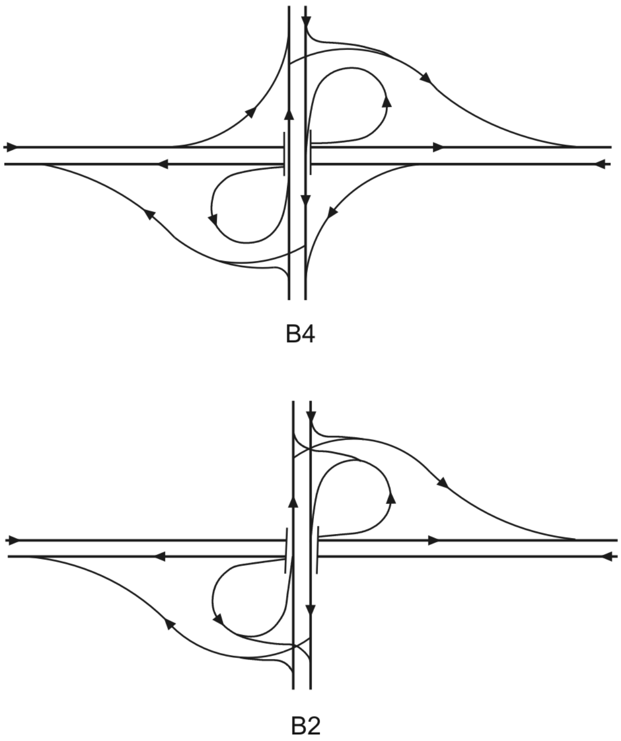


Figure 7.11: Par-Clo B interchanges

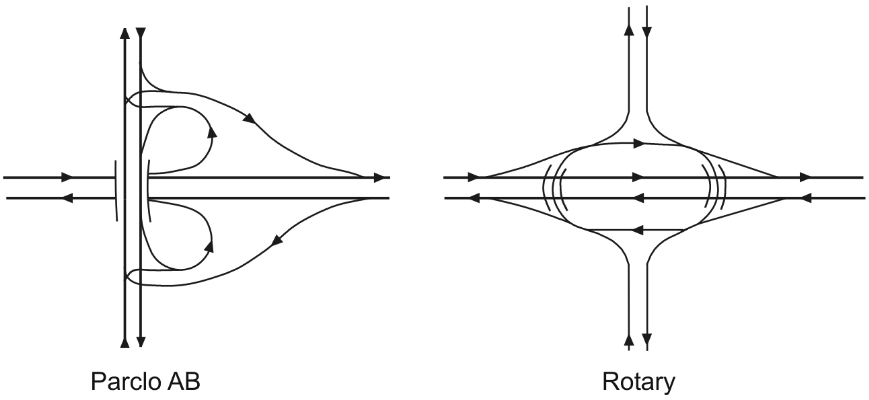


Figure 7.12: Par-Clo AB interchanges and rotaries

with the rotary sandwiched between them. This is the so-called "Island in the Sky" concept. The Rotary is an interchange form that is unknown in South Africa so that its application would compromise consistency of design and thus be contrary to drivers' expectations.

7.7.4 Interchanges on non-freeway roads

Interchanges where a non-freeway is the major route are a rarity in South Africa. This application would arise where traffic flows are so heavy that a signalised intersection cannot provide sufficient capacity. In this case, the crossing road terminals would be provided on the road with the lower traffic volume. As a general rule, a simple and relatively low standard Simple Diamond or a Par-Clo Interchange should suffice.

An intersection with a particularly poor accident history may also require upgrading to an interchange. The accident history would provide some indication of the required type of interchange.

Where the need for the interchange derives purely from topographic restraints, i.e. where traffic volumes are low, a Jug Handle Interchange, illustrated in Figure 7.13, would be adequate. This layout, also known as a Quarter Link, provides a two-lane-two-way connection between the intersecting roads located in whatever quadrant entails the minimum construction and property acquisition cost.

Drivers would not expect to find an interchange on a two-lane two-way road and, in terms of driver expectancy, it may therefore be advisable to introduce a short section of dual carriageway at the site of the interchange

7.8 RAMP DESIGN

7.8.1 General

A ramp is defined as a roadway, usually one-way, connecting two grade-separated through roads. It comprises an entrance terminal, a mid-section and an exit terminal.

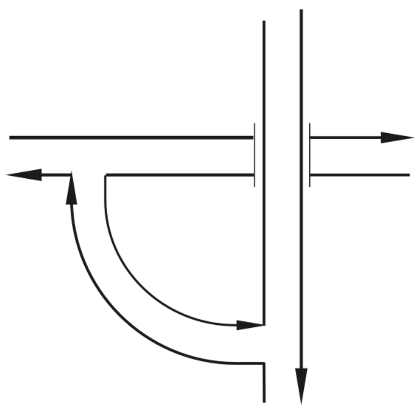


Figure 7.13: Jug Handle interchange

The general configuration of a ramp is determined prior to the interchange type being selected. The specifics of its configuration, being the horizontal and vertical alignment and cross-section, are influenced by a number of considerations such as traffic volume and composition, the geometric and operational characteristics of the roads which it connects, the local topography, traffic control devices and driver expectations.

A variety of ramp configurations can be used. These include:

- The outer connector, which serves the left turn and has free-flowing terminals at either end;
- The diamond ramp, serving both the left- and right-turns with a free-flowing terminal on the freeway and a stop-condition terminal on the crossing-road;
- The Par-Clo ramp, which serves the right turn and has a free-flowing terminal on the freeway and a stop-condition terminal on the crossing road, with a 180° loop between them;
- The loop ramp, serving the right turn and which has free-flowing terminals at both ends and a 270° degree loop between them;
- The directional ramp also serving the right turn, with a curve only slightly in excess of 90° degrees and free-flowing terminals at either end, and
- The collector-distributor road intended to remove the weaving manoeuvre from the freeway.

The express-collector system discussed later is a transfer roadway and is not an interchange ramp.

7.8.2 Design speed

Guideline values for ramp design speeds are given in Table 7.2. Strictly speaking, the design speed of a ramp could vary across its length from that of the freeway to that of the at-grade intersection, with the design speed at any point along the ramp matching the operating speed of the vehicles accelerating to or decelerating from the design speed of the freeway. The design speeds given in the table apply to the controlling curve on the mid-section of the ramp. The ramp design speed is shown as a design domain because of the wide variety of site conditions, terminal types and ramp shapes.

Table 7.2: Ramp design speed

Roadway design speed (km/h)	Ramp Design Speed Domain (km/h)
60	40 – 50
70	40 – 60
80	40 – 70
90	50 – 80
100	50 – 90
110	60 – 100
120	60 – 110
130	70 - 110

In the case of a directional ramp between freeways, vehicles must be able to operate safely at the higher end of the ranges of speeds shown in Table 7.2.

Factors demanding a reduction in design speed include site limitations, ramp configurations and economic factors. Provided the reduction is not excessive, drivers are prepared to reduce speed in negotiating a ramp so that the lower design speed is not in conflict with driver expectations.

For directional ramps and outer connectors, higher values in the speed domain are appropriate and, in general, ramp designs should be based on the upper limit of the domain. The constraints of the site, traffic mix and form of interchange, however, may force a lower design speed.

A Par-Clo or a loop ramp cannot be designed to a high design speed. A ramp design speed of 70 km/h, being the low end of the domain for a freeway design speed of 130 km/h, would require a radius of between 150 metres and 200 metres, depending on the rate of superelevation selected. It is not likely, particularly in an urban area, that the space required to accommodate a loop with this radius would be available. Even if the space were available, the additional travel distance imposed by the greater radius would nullify the advantages of the higher travel speed. The added length of roadway also adds the penalty of higher construction cost. This penalty also applies to the ramp outside the loop because it has to be longer to contain the larger loop. In general, loop ramps are designed for speeds of between 40 and 50 km/h. Because of the substantial difference between the freeway design speed and that of the loop ramp, it is advisable not to have a loop on an exit ramp if this can be avoided.

Safety problems on ramp curves are most likely for vehicles travelling faster than the design speed. This problem is more critical on curves with lower design speeds because drivers are more likely to exceed these design speeds than the higher ranges. Trucks can capsize when travelling at speeds only marginally higher than the design speed. To minimise the possibility of trucks exceeding the design speed it is suggested that the lower limits of design speeds shown in Table 7.2 should not be used for ramps carrying a substantial amount of truck traffic.

If site specific constraints preclude the use of design speeds that more-or-less match anticipated operating speeds, the designer should seek to incorporate effective speed controls, such as advisory speed signing, special pavement treatments, long deceleration lanes and the use of express-collector systems in the design.

7.8.3 Sight distance on ramps

It is necessary for the driver to be able to see the road markings defining the start of the taper on exit ramps and the end of the entrance taper. At the crossing road ramp terminal, lanes are often specifically allocated to the turning movements with these lanes being developed in advance of the terminal. The driver has to position the vehicle in the lane appropriate to the desired turn. It is therefore desirable that decision sight distance be provided on the approaches to the ramp as well as across its length.

Appropriate values of decision sight distance are given in Table 3.7.

7.8.4 Horizontal alignment

Minimum radii of horizontal curvature on ramps are as shown in Table 4.1 for various values of e_{max}. In general, the higher values of e_{max} are used in freeway design and the selected value should also be applied to the ramps.

Achieving the step down of radii from higher to lower design speeds on a loop may require the application of compound curves. In general, the ratio between successive radii should be 1,5 : 1 and, as a further refinement, they could be connected by transition curves. The length of each arc is selected to allow for deceleration to the speed appropriate to the next radius at the entry to that arc. In the case of stepping up through successive radii, the same ratio applies but the design speed for the radius selected should match the desired speed at the far end of each arc.

It is recommended that the designer develops a speed profile for the loop and bases the selection of radii and arc lengths on this speed profile. As a rough rule of thumb, the length of each arc should be approximately a third of its radius.

If a crossover crown line, discussed below, is not used, the crossfall on the exit or entrance taper between the Yellow Line Break Point and the nose is controlled by that on the through lanes. Superelevation development can thus only commence at the nose. The distance required to achieve the appropriate superelevation thus determines the earliest possible location of the first curve on the ramp.

Ramps are relatively short and the radii of curves on ramps often approach the minimum for the selected design speed. Furthermore, if there is more than one curve on a ramp, the distance between the successive curves will be short. Under these restrictive conditions, transition curves should be considered.

Ramps are seldom, if ever, cambered and superelevation typically involves rotation around one of the lane edges. Drivers tend to position their vehicles relative to the inside edge of any curve being traversed, i.e. they steer towards the inside of the curve rather than away from the outside. For aesthetic reasons, the inside edge should thus present a smoothly flowing three-dimensional alignment with the outside edge rising and falling to provide the superelevation. Where a ramp has an S- or reverse curve alignment, it follows that first one edge and then the other will be the centre of rotation, with the changeover taking place at the point of zero crossfall.

In view of the restricted distances within which superelevation has to be developed, the crossover crown line is a useful device towards rapid development. A crossover crown is a line at which an instantaneous change of crossfall takes place and which runs diagonally across the lane. The crossover crown could, for example, be located along the yellow line defining the edge of the left lane of the freeway, thus enabling initiation of superelevation for the first curve on the ramp earlier than would otherwise be the case. The crossover crown should, however, be used with caution as it may pose a problem to the driver, particularly to the driver of a vehicle with a high load. This is because the vehicle will sway as it traverses the crossover crown and, in extreme cases, may prove difficult to control. The algebraic difference in slope across the crossover crown should thus not exceed four to five per cent.

7.8.5 Vertical alignment

Gradients

The profile of a ramp typically comprises a mid-section with an appreciable gradient coupled with terminals where the gradient is controlled by the adjacent road. If the crossing road is over the freeway, the positive gradient on the off-ramps will assist a rapid but comfortable deceleration and the negative gradient on the on-ramp will support acceleration to freeway speeds. In theory, thus, the higher the value of gradient, the better. Values of gradient up to eight per cent can be considered but, for preference, gradients should not exceed six per cent.

Diamond ramps are usually fairly short, possibly having as little as 120 to 360 metres between the nose and the crossing road. The effect of the midsection gradient, while possibly helpful, is thus restricted. However, a steep gradient (8 per cent) in conjunction with a high value of superelevation (10 per cent) would have a resultant of 12,8 per cent at an angle of 53° to the centreline of the ramp. This would not contribute to drivers' sense of safety. In addition, the drivers of slow-moving trucks would have to steer outwards to a marked extent to maintain their path within the limits of the ramp width. This could create some difficulty for them. It is suggested that designers seek a combination of superelevation and gradient such that the gradient of the resultant is less than ten per cent. Table 7.3 provides an indication of the gradients of the resultants of combinations of superelevation and longitudinal gradient.

Table 7.3: Maximum resultant gradients

Superelevation (%)	Longitudinal gradient (%)			
	2	4	6	8
4	4,47	5,65	7,21	8,94
6	6,31	7,20	8,49	10,00
8	8,25	8,94	10,00	11,31
10	10,19	10,73	11,63	12,80

The combinations of gradient and superelevation shown shaded in Table 7.3 should be avoided.

Curves

As suggested above, decision sight distance should be available at critical points on ramps. The K-values of crest curves required to meet this requirement are given in Table 7.4.

These values of K are based on the required sight distance being contained within the length of the vertical curve. It is, however, unlikely that, within the confined length of a ramp, it would be possible to accommodate the length of vertical curve required for this condition to materialise.

On sag curves, achieving decision sight distance is not a problem so that the K-values shown in Table 4.14 can be applied. If the interchange area is illuminated, as would typically be the case in an urban area, the K-values for comfort can be used. Unilluminated, i.e. rural, interchanges would require the application of K-values appropriate to headlight sight distance.

Table 7.4: K-Values of crest curvature for decision sight distance

Design speed (km/h)	K-Value
40	118
50	161
60	210
70	266
80	328
90	473
100	643
110	738
120	840
130	948

The designer should therefore have recourse to Equation 4.21 to calculate the K-value for the condition of the curve being shorter than the required sight distance. This equation is repeated below for convenience.

$$K = \frac{2S}{A} - \frac{200(h_1^{0,5} + h_2^{0,5})^2}{A^2} \qquad 4.21$$

where K = Distance required for a 1% change of gradient (m)

S = Stopping sight distance for selected design speed (m)

h1 = Driver eye height (m)

h2 = Object height (m)

A = Algebraic difference in gradient between the approaching and departing grades (%)

7.8.6 Cross-section

Horizontal radii on ramps are typically short and hence are usually in need of curve widening. The width of the ramps normally adopted is thus 4 metres. Because of the inconvenience of changing the lane width of comparatively short sections, this width is applied across the entire length of the ramp.

In the case of a loop ramp, the radius could be as low as fifty metres. At this radius, a semi trailer would require a lane width of 5,07 metres. It is necessary for the designer to consider the type of vehicle selected for design purposes and to check whether the four metre nominal width is adequate. If semi trailers are infrequent users of the ramp, encroachment on the shoulders could be considered.

The calculation of required lane width is described in detail in Section 4.2.6. As suggested above, regardless of the width required, it should be applied across the entire length of the ramp.

Ramp shoulders typically have a width of the order of 2 metres, with this width applying both to the inner and to the outer shoulder. In conjunction with the nominal lane width of 4 metres, the total roadway width is thus 8 metres. This width would allow comfortably for a truck to pass a broken-down truck. In addition, it would provide drivers with some sense of security in the cases where the ramp is on a high fill.

7.8.7 Terminals

The crossing road ramp terminals may be free-flowing, in which case their design is as discussed below. Crossing road ramp terminals that are at-grade intersections should be designed according to the recommendations contained in Chapter 6. It is worth noting that, from an operational point of view, what appears to be a four-legged intersection, e.g. the crossing road ramp terminals on a Diamond Interchange, is, in fact, two three-legged intersections back-to-back. Ideally, the crossing road ramp terminal should be channelised to reduce the possibility of wrong-way driving.

Vehicles entering or exiting from a freeway should be able to do so at approximately the operating speed of the freeway. Given the fact that the crossing road terminal is invariably a signalised or stop control at-grade intersection, the change in speed across the ramp is substantial. Provision should thus be made for acceleration and deceleration to take place clear of the freeway so as to minimise interference with the through traffic and reduce the potential for crashes. The auxiliary lanes provided to accommodate this are referred to as speed-change lanes or acceleration or deceleration lanes. These terms describe the area adjacent to the travelled way of the freeway, including that portion of the ramp taper where a merging vehicle is still clear of the through lane, and do not imply a definite lane of uniform width.

The speed change lane should have sufficient length to allow the necessary adjustment in speed to be made in a comfortable manner and, in the case of an acceleration lane, there should also be sufficient length for the driver to find and manoeuvre into a gap in the through traffic stream before reaching the end of the acceleration lane. The length of the speed change lane is based on:

- The design speed on the through lane, i.e. the speed at which vehicles enter or exit from the through lanes;
- The control speed of the ramp midsection, i.e. the design speed of the smallest radius curve on the ramp, and
- The tempo of acceleration or deceleration applied on the speed change lane.

In Chapter 3, reference was made to a deceleration rate of 3 m/s^2 as the basis for the calculation of stopping sight distance. Research has shown that deceleration rates applied to off-ramps are a function of the freeway design speed and the ramp control speed. As both speeds increase, so does the deceleration rate, which varies between 1,0 m/s² and 2,0 m/s². For convenience, the deceleration rate used to develop Table 7.5 has been set at 2,0 m/s².

Deceleration should only commence once the exiting vehicle is clear of the through lane. Assuming that a vehicle with a width of 2,5 metres is correctly positioned relative to the ramp edge line to be centrally located within an ultimate ramp width of 4 metres, its tail end will clear the edge of the through lane at a distance, L, from the Yellow Line Break Point where

$$L = 3{,}2 / T$$

with T = Taper rate (as listed in Table 7.7)

The maximum legal length of any vehicle on South African roads is 22 metres and this should be added to the distance, L, to establish the distance from the Yellow Line Break Point, at which the deceleration can commence so that

$$L_T = 3{,}2/T + 22$$

Values of L_T, are listed in Table 7.5. From this table it follows that, in the case of a diamond ramp without curves, the distance from the Yellow Line Break Point to the crossing road ramp terminal should be not less than 437 metres.

The acceleration rate can, according to American literature, be taken as 0,7 m/s^2. The length of the acceleration lane is thus as shown in Table 7.6. An important feature of the acceleration lane is the gap acceptance length, which should be a minimum of 100 metres to 150 metres, depending on the nose width. The length of the entering vehicle is not relevant in determination of the taper length.

The lengths of the deceleration and acceleration lanes shown in Tables 7.5 and 7.6 apply to gradients of between - 3 per cent and + 3 per cent. Acceleration lanes will have to be longer on upgrades and may be made shorter on downgrades, with the reverse applying to deceleration lanes.

The actual entrance or exit points between the freeway and the ramp can take the form of a taper or a parallel lane. The parallel lane has the problem of forcing a reverse curve path, possibly followed after the nose by a further curve to the left, whereas the taper involves only a single change of direction. As such, drivers prefer the taper. In cases where an extended acceleration or deceleration distance is required, for example where the crossing road passes under the freeway resulting in the on-ramp being on an upgrade and the off-ramp on a downgrade, a straight taper would result in an inordinately long ramp. The parallel lane allows the speed change to take place in an auxiliary lane immediately adjacent to the through lanes thus reducing the spatial demands of the interchange.
Furthermore, in the case of the on-ramp, the parallel configuration provides an extended distance to find a gap in the adjacent traffic stream. Typical configurations of on- and off-ramp terminals are illustrated in Figures 7.14 to 7.17.

Taper

Two different criteria apply to the selection of the taper rate, depending on whether the ramp is an

Table 7.5: Length of deceleration lanes (m)									
Freeway design speed (km/h)	L_T (m)	Ramp control speed (km/h)							
		0	40	50	60	70	80	90	100
60	67	70	60	21					
70	76	95	85	45	25				
80	76	125	115	75	55	30			
90	86	155	150	110	85	60	30		
100	92	190	185	145	125	100	70	35	
110	102	235	225	185	165	140	110	75	40
120	108	280	270	230	210	180	155	120	85
130	112	325	320	275	255	230	200	170	135

Table 7.6: Length of acceleration lanes (m)									
Freeway design speed (km/h)	L_T (m)	Ramp controlling speed (km/h)							
		0	40	50	60	70	80	90	100
60	45	200	110	60					
70	54	270	180	130	70				
80	54	350	265	215	155	80			
90	64	450	360	310	250	175	95		
100	70	550	460	415	350	280	200	105	
110	80	670	580	530	470	395	315	220	115
120	86	790	700	655	595	525	440	350	240
130	90	930	840	795	735	660	580	485	380

exit from or entrance to the freeway.

In the first case, the vehicle is merely required to achieve a change of direction from the freeway to the ramp. The taper rate should thus be sufficiently flat to ensure that the vehicle path can be accommodated within the lane width. In Table 7.7, the radii of curvature corresponding to a superelevation of 2,0 per cent for the various design speeds are listed as are the taper rates corresponding to the condition of the wheel path having an offset of 150 millimetres inside the Yellow Line Break Point.

These taper rates are higher than the rate of 1 : 15 currently applied on South African freeways. This rate derives from the application of an "operating speed" previously assumed to be 85 per cent of the design speed and the further assumption that the wheel path could be

Table 7.7: Taper rates for exit ramps			
Design speed (km/h)	Radius (m) for 2% superelevation	Taper rate 1:	Taper length, L_T (m)
60	1 000	14	67
70	1 500	17	76
80	1 500	17	76
90	2 000	20	86
100	2 500	22	92
110	3 000	25	102
120	3 500	27	108
130	4 000	28	112

allowed to pass over the Yellow Line Break Point. In practice, a 1: 15 rate requires that drivers accept a higher level of side friction in negotiating the change of direction. Passenger cars can negotiate tapers of this magnitude with ease so that, in the case of existing interchanges, it is not always necessary to incur the expenditure of upgrading to a flatter taper. Trucks do, however, sometimes roll over at the start of the ramp taper.

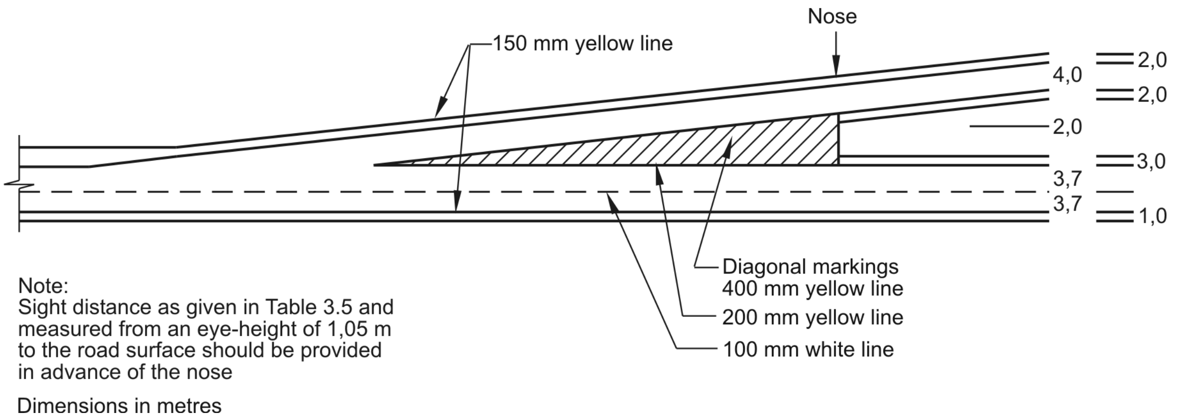


Figure 7.14: Single lane exit

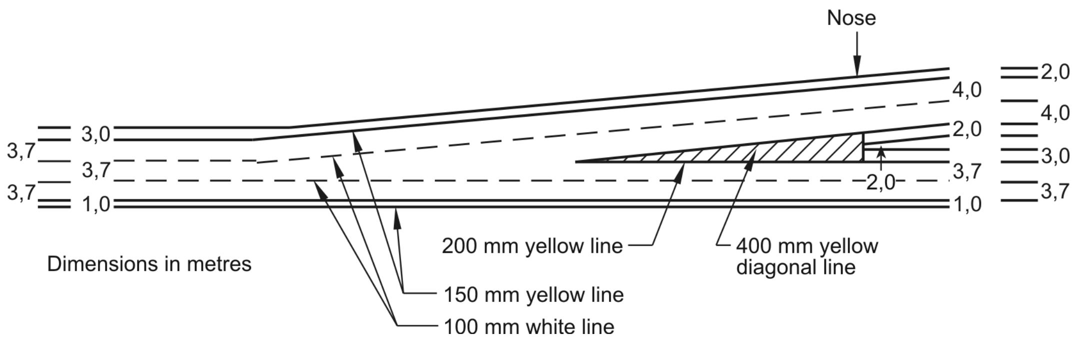
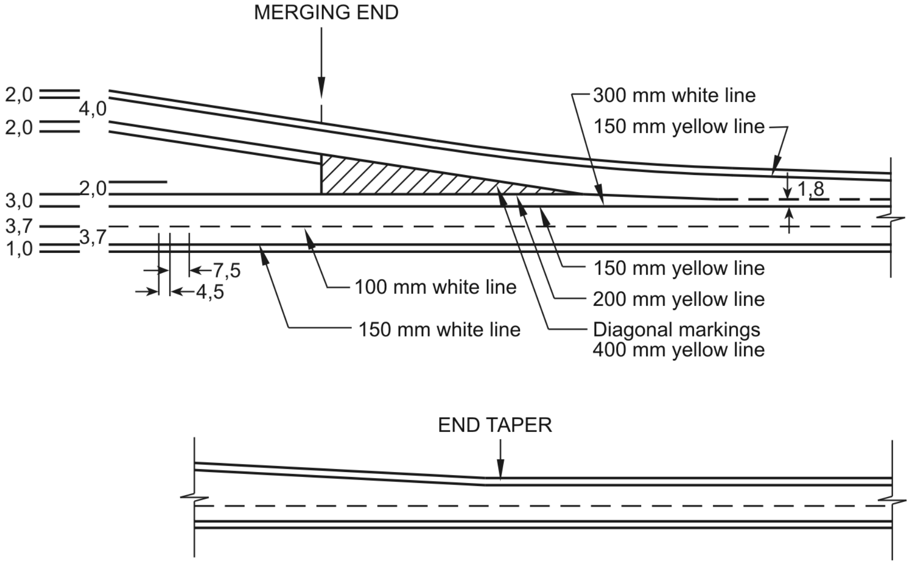


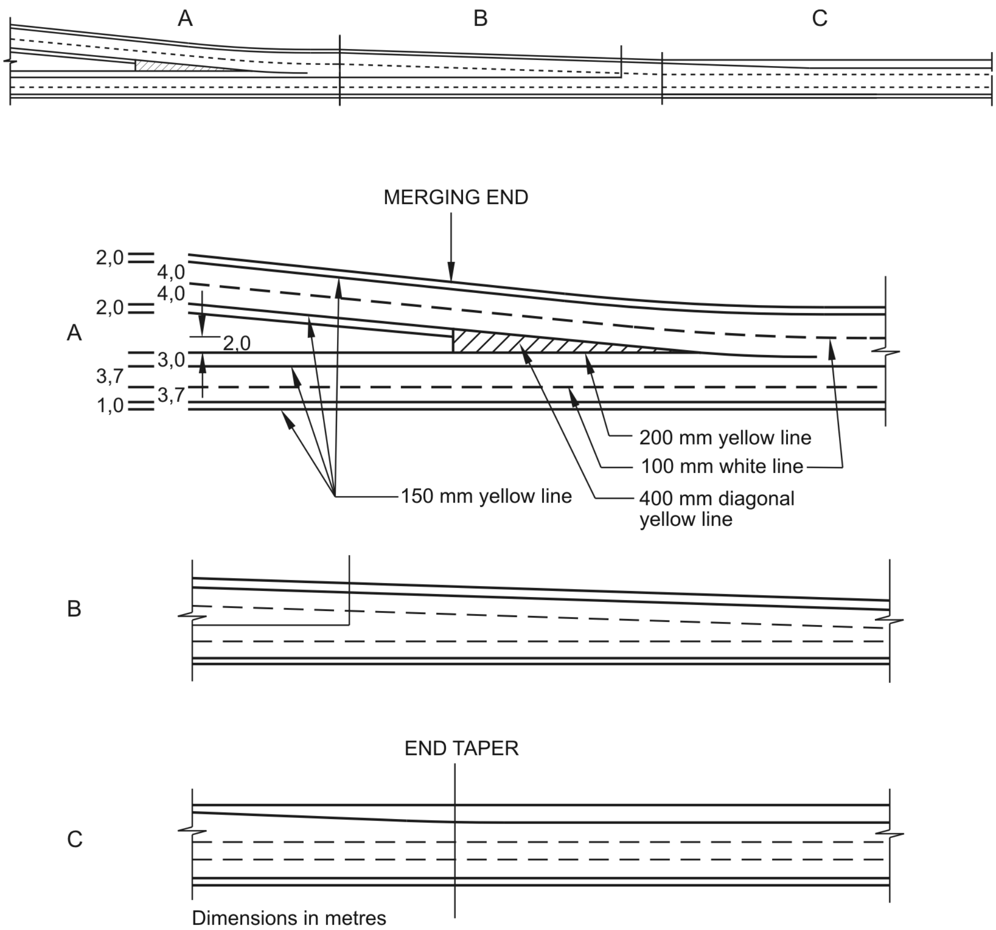
Figure 7.15: Two-lane exit

Note:
For 50 m advance of the merging end, the ramp should not be more than 0,75 m lower than the freeway



Dimensions in metres

Figure 7.16: One lane entrance



Dimensions in metres

Note:
For 50 m in advance of the merging end, the ramp should not be more than 0,75 m lower than the freeway

Figure 7:17: Two-lane entrance

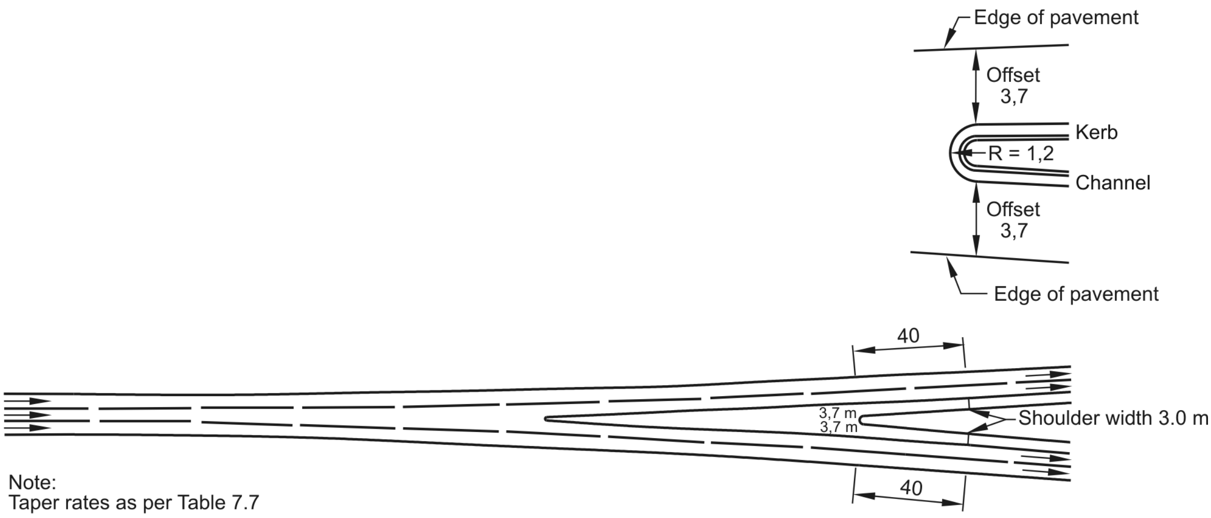


Figure 7.18: Major fork

If there is a high percentage of truck traffic and if the incidence of roll-overs is unacceptably high, it may be necessary to consider upgrading to the tapers suggested in Table 7.7.

In the case of the entrance ramp, the driver of the merging vehicle should be afforded sufficient opportunity to locate a gap in the opposing traffic and position the vehicle correctly to merge into this gap at the speed of the through traffic. Taper rates flatter than those proposed in Table 7.7 should thus be applied to on-ramps. A taper of 1:50 provides a travel time of about 13 seconds between the nose and the point at which the wheel paths of merging and through vehicles would intersect. This has been found in practice to provide sufficient opportunity to prepare to merge into gaps in the through flow because drivers typically begin to locate usable gaps in the freeway flow prior to passing the nose.

Parallel

The parallel lane configuration is also initiated by means of a taper but, in this case, the taper rate does not have to be as flat as suggested above. As in the case of the climbing lane, a taper length of 100 metres would be adequate.

Parallel entrances and exits have a major advantage over straight tapers in respect of the selection of curves on the ramps. In the case of a design speed of 120 km/h, a straight taper provides a distance of just short of 300 metres between the Yellow Line Break Point and the nose. A portion of this distance should ideally be traversed at the design speed of the freeway with deceleration commencing only after the vehicle is clear of the outside freeway lane. About 220 metres is available for deceleration and this suggests that a vehicle exiting at 120 km/h is likely still to be travelling at a speed of about 95 km/h when passing the nose. The first curve on the ramp could, allowing for superelevation development to 10 per cent, be located 100 metres beyond the nose so that, at the start of the curve, the vehicle speed would be in the range of 80 to 90 km/h. The first curve should thus have a radius of not less than 250 metres. In the case of the parallel ramp, the radius could obviously be significantly shorter, depending on the length of the deceleration lane preceding the nose.

7.9 COLLECTOR - DISTRIBUTOR ROADS

Collector-distributor roads are typically applied to the situation where weaving manoeuvres would be disruptive if allowed to occur on the freeway. Their most common application, therefore, is at Cloverleaf interchanges. The exit and entrance tapers are identical to those applied to any other ramps. The major difference between Cloverleaf interchange C-D roads and other ramps is that they involve two exits and two entrances in quick succession. The two exits are, firstly, from the freeway and, secondly, the split between vehicles turning to the left and those intending to turn to the right. The two entrances are, firstly, the merge between the two turning movements towards the freeway and, secondly, the merge with the freeway through traffic.

The distance between the successive exits should be based on signing requirements so as to afford drivers adequate time to establish whether they have to turn to the right or to the left to reach their destination. Nine seconds is generally considered adequate for this purpose and seeing that vehicles may be travelling at the design speed of the freeway as they pass the nose of the first exit, the distance between the noses should desirably be based on this speed.

The distance between successive entrances is based on the length required for the acceleration lane length quoted in Table 7.6.

7.10 OTHER INTERCHANGE DESIGN FEATURES

7.10.1 Ramp metering

Ramp metering consists of traffic signals installed on entrance ramps in advance of the entrance terminal to control the number of vehicles entering the freeway. The traffic signals may be pretimed or traffic-actuated to release the entering vehicles individually or in platoons. It is applied to restrict the number of vehicles that are allowed to enter a freeway in order to ensure an acceptable level of service on the freeway or to ensure that the capacity of the freeway is not exceeded. The need for ramp metering may arise owing to factors such as:

- Recurring congestion because traffic demand exceeds the provision of road infrastructure in an area;
- Sporadic congestion on isolated sections of a freeway because of short term traffic loads from special events, normally of a recreational nature;
- As part of an incident management system to assist in situations where a accident downstream of the entrance ramp causes a temporary drop in the capacity of the freeway; and
- Optimising traffic flow on freeways

Ramp metering also supports local transportation management objectives such as:

- Priority treatments with higher levels of service for High Occupancy Vehicles; and
- Redistribution of access demand to other on-ramps.

It is important to realise that ramp metering should be considered a last resort rather than as a first option in securing an adequate level of service on the freeway. Prior to its implementation, all alternate means of improving the capacity of the freeway or its operating characteristics or reducing the traffic demand on the freeway should be explored. The application of ramp metering should be preceded by an engineering analysis of the physical and traffic conditions on the freeway facilities likely to be affected. These facilities include the ramps, the ramp terminals and the local streets likely to be affected by metering as well as the freeway section involved.

The stopline should be placed sufficiently in advance of the point at which ramp traffic will enter the freeway to allow vehicles to accelerate to approximately the operating speed of the freeway, as would normally be required for the design of ramps. It will also be necessary to ensure that the ramp has sufficient storage to accommodate the vehicles queuing upstream of the traffic signal.

The above requirement will almost certainly lead to a need for reconstruction of any ramp that is to be metered. The length of on-ramps is typically determined by the distance required to enable a vehicle to accelerate to freeway speeds. Without reconstruction, this could result in the ramp metering actually being installed at the crossing-road ramp terminal.

7.10.2 Express-collector systems

Express-collector systems are used where traffic volumes dictate a freeway width greater than four lanes in each direction. The purpose of the express-collector is to eliminate weaving on the mainline lanes by limiting the number of entrance and exit points while satisfying the demand for access to the freeway system.

An express-collector system could, for example, be started upstream of one interchange and run through it and the following, possibly closely spaced, interchange, terminating downstream of the second. The terminals at either end of the express-collector system would have the same standards as applied to conventional on- and off-ramps. The interchange ramps are connected to the express-collector system and not directly to the freeway mainline lanes.

Traffic volumes and speeds on the express-collector roads are typically much lower than those found on the mainline lanes, allowing for lower standards being applied to the ramp geometry of the intervening interchanges.

The minimum configuration for an express-collector system is to have a two-lane C-D road on either side of a freeway with two lanes in each direction. The usual configuration has more than two mainline lanes in each direction

A similar configuration is known as a dual-divided freeway, sometimes referred to as a dual-dual freeway. In this case, the C-D roads are taken over a considerable distance and may have more than two lanes. Furthermore, connections are provided at intervals between the outer and the core lanes along the length of the freeway. These, unfortunately, create the effect of right-side entrances and exits to and from the outer lanes and are thus contrary to drivers' expectations. The geometry of the situation is otherwise similar to that of the express-collector system.

Back to Menu

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 8

ROADSIDE SAFETY

8.1 INTRODUCTION

8.1.1 General

Road crashes, to varying degrees, are caused by defects attributable to the vehicle, the driver or the road or by combination of these defects. In addition, a significant percentage of deaths on roads in South Africa occur as a result of pedestrians being on the road. The road, or more fully, the road environment, has been estimated to contribute to 28 per cent of all road crashes in South Africa. Various studies have indicated that up to 40 per cent or more of crash reduction, which could reasonably be expected on the road system, could accrue from the provision of safer roads. The cost of road crashes to society in South Africa exceeds the annual expenditure on roads, thus the expenditure of considerable sums of money can be justified in reducing the crash rates on roads through improved design appropriate and standards and by catering for the presence of pedestrians on roads. Crashes resulting from simply leaving the roadway regardless of the underlying cause, represent a substantial portion of the total road crash problem i.e. "run-off-the-road" (ROR) accidents account for 25 per cent of all road crashes in South Africa. They occur on both straight and curved sections of road and generally involve either rollover of the vehicle or collision with fixed objects, such as trees, roadside structures etc.

It is thus obvious that the roadside environment and its design have a vital role to play in improving road safety and that, in the design of new roads or the upgrading of existing ones, particular attention should be given to safety as a prime design criterion.

A safe road should:

- Warn and inform road users of changes in the approaching road environment;
- Guide and control road users safely through the road environment;
- Provide a forgiving roadside environment;
- Provide a controlled release of information;
- Provide an aesthetically pleasing landscape;
- Maintain road user interest and concentration;
- Not surprise road users;
- Give consistent messages to road users; and
- Provide good visibility for all road users.

Numerous research projects have established relationships between crashes and geometric design elements (as well as operating speeds and traffic volumes).

The various geometric design features of a road, shown in Table 8.1, affect safety by:

- Influencing the ability of the driver to maintain vehicle control and identify hazards. Significant features include:
  - Lane and shoulder width;
  - Horizontal and vertical alignment;
  - Sight distance;
  - Superelevation; and

- Pavement surface and drainage.
- Influencing the number and types of opportunities that exist for conflict between vehicles. Significant features include:
  - Access control;
  - Intersection design;
  - Number of lanes;and
  - Medians.
- Affecting the consequences of an out-of-control vehicle leaving the travel lanes. Significant features include:
  - Shoulder width and type;
  - Edge drop;
  - Roadside conditions;
  - Side slopes; and
  - Traffic barriers.

In addition to geometric features, a variety of other factors affect road safety, including other elements of the overall road environment, such as:

- Pavement condition;
- Weather;
- Lighting;
- Traffic flows;
- Traffic regulation;
- Presence of pedestrians;
  - Intoxication and
  - Age;
- Vehicle characteristics, such as:
  - Size;
  - Mass; and
  - Braking capability.

The effect of road design is somewhat obscured by the presence of these extraneous factors and most accidents result from a combination of factors interacting in ways that prevent a single factor being identified as the cause of a crash. However, even when a vehicle leaves the road owing to driver error or mechanical failure, good roadside design can mitigate the severity of a crash. This interaction between road, driver and vehicle characteristics unfortunately complicates attempts to estimate the accident reduction potential of a particular safety improvement.

The construction of a road is typically a trade-off between standards and the cost of providing them. High design standards might be expensive to provide. However, the cost to society of road crashes and deaths often exceeds the total annual expenditure on roads. Reducing initial construction (or capital) costs of road projects can result in increased life cycle costs if the cost of accidents, injuries and deaths is included in the economic calculations. It is the design engineer's responsibility to inform the client of the consequences of inadequate expenditure on safety.

It is often extremely difficult, if not impossible, to correct safety defects at a later stage without major reconstruction. For this reason, designing for safety should occur at the outset, or be provided for in stage construction drawings. Road safety audits on the design, carried out by an independent person or team should take place at various stages in the project, as provided for in Volume 4 of the South African Road Safety Manual. Although this design guide focuses on road design features, the psychological aspects of driver behaviour are always present. An error in perception or judgement or a faulty action on the part of the driver can easily lead to a crash.

Roads should be designed in such a manner that only one decision at a time is required from a driver, ensuring that he/she is never surprised by an unexpected situation and that adequate

Table 8.1 Design elements that influence road safety	
Category	**Design Element**
Horizontal Alignment	Curve radius Curve length Superelevation Runoff length/transition
Vertical Alignment	Grade Critical length of grade Vertical curves (sag and crest)
Cross-section	Number of lanes Lane width Shoulder type Shoulder width Median type Median width
Roadside	Side slopes Horizontal clearance to obstruction (clear zone) Ditch design Traffic barriers (roadside) Median barriers
Intersection	Sight triangle
Interchange	Ramp terminal sight triangle Taper rate/length Successive exits/entrances

time is provided to make the decision. Research has shown that the number of accidents increases as the number of decisions required by the driver increases. This matter is discussed in more detail in chapter 3 of this document, as well as in Volume 1 of the South African Road Safety Manual.

Standardisation in road design features and traffic control devices plays an important role in reducing the number of required decisions, as the driver becomes aware of what to expect on a certain type of road. However it should be noted that standardization alone does not necessarily ensure a safe facility, hence the requirement for a safety audit of the design.

8.1.2 Safety objectives

Broadly, there are three avenues of action possible for mitigating road crashes and their consequences;

- By reducing the possibility that run-off will occur;
- Once the run-off does occur, by providing opportunities for the driver of the vehicle to recover and return to the road without incident, and
- If a crash does occur, by providing design elements to reduce the severity of that collision.

The first of these areas of endeavour is dealt with by incorporating features in the overall design of the road which will reduce the possibility that run-off will occur. These features are influenced to a large extent by the selected design speed.

This chapter deals with the latter two of these three possibilities. It provides the designer with guidance on the design of roadside environment that includes elements to allow for recovery on the part of the driver, as well as features which are intended to reduce the severity of such accidents as do occur.

More specifically, it is recommended that the following safety objectives be adopted when a road is designed:

- Separate potential conflict points and reduce potential conflict areas;
- Control the relative speeds of the conflicting vehicles;
- Guide the driver through unusual sections;
- Ensure that the needs of pedestrians and cyclists (if relevant) are also considered;
- Provide a roadside environment that forgives a driver's errant or inappropriate behaviour, by attention to details such as the safe placement of roadside furniture and by the location and selection of types of traffic barrier.

8.1.3 The "Forgiving Roadside" approach

The forgiving roadside concept, coined in the 1960's, relates to the approach of making provision for errant vehicles leaving the roadway by incorporating design elements that reduce the consequences of such departure. This concept is an integral part of modern road design philosophy and approaches. Transportation Research Circular 435 states that:-

"Basically a forgiving roadside is one free of obstacles which could cause serious injuries to occupants of an errant vehicle. To the extent possible, a relatively flat, unobstructed roadside recovery area is desirable, and when these conditions cannot be provided, hazardous features in the recovery area should be made breakaway or shielded with an appropriate barrier".

When a vehicle leaves the traffic lane, the path of the vehicle and any object in or near that path become contributing factors to the degree of severity of the crash. By designing a forgiving roadside the severity of crashes can be reduced. The concept of designing a forgiving roadside should not be regarded as a by-product of the application of safety criteria to each design element, but as an integral part of the total engineering for the road.

The need for a forgiving roadside is paramount on the outside of horizontal curves with radii of less than 1 000 metres, where the possibility of an errant vehicle running off the road is greatest. However, this statement does not imply that horizontal curves with radii in excess of 1 000 metres are always safe since the phenomena of "risk adaptation", whereby drivers concentrate less and drive at higher speeds on sections of road they consider to be safe, should be taken into account.

8.1.4 Design Focus

The focus of design measures outlined in this chapter is primarily one of improving road safety through roadside hazard management by the design and provision of appropriate recovery and protection measures. The effectiveness of road safety features depends greatly on five aspects;

- Knowledge of the safety characteristics and limitations of roadside features by the designer (and the maintenance personnel);
- The correct choice of appropriate treatment;
- The correct installation of the roadside safety features;
- The maintenance of the roadside safety features and roadside environment; and
- Regular monitoring of installations to ensure they perform adequately.

There are two needs that are the key to effective attention to safety in the roadside design process.

1. The need for explicit evaluation of design trade-offs with an impact on road safety. In the traditional design process, attention to safety has usually been implicit, not explicit. The common myth among designers is that if current "standards" are met, then the road is safe. The reality is that road design "standards" are often no more than a limit : one should not provide less than the standard stipulates but, within limits, to provide more is often better. Furthermore, just meeting the standard does not mean that an appropriate amount of safety has been provided.

2. The need to recognize that the design of the roadside environment is a highly complex and probabilistic process. There are many levels of interaction between different roadside design components, between roadside elements and other aspects of the facility design, and between the road itself, the driver and the vehicle. As a result, information touching on road design issues necessarily is available from many sources. Designers should not rely on this Guide as the sole source of information on roadside design issues, particularly when dealing with unusual or local conditions that depart from generally accepted situational norms.

Particular attention should be paid to the South Africa Road Safety Manual produced by the South African Committee of Land Transport Officials (COLTO).

8.1.5 Roadside safety analysis

The design of the roadside environment is a complex problem. Evaluation of alternative designs and choosing between them are difficult tasks, which involve

- degrees of uncertainty with respect to the occurrence of crashes;
- the outcome of crashes in terms of severity; and
- the real costs of the property damage, injuries, and fatalities which can result.

Nonetheless, such analysis which provides an explicit framework for considering design trade-offs is a much more desirable approach to roadside safety design than meeting arbitrary "standards" whose underpinnings may or may not be appropriate to a given situation. Figure 8.1 illustrates an algorithm for conducting a roadside safety analysis.

The process is generally based on two fundamental models.

- Predictive models that provide a way of

estimating collision frequencies and severities under a wide variety of conditions; and
- Cost-effectiveness models that provide a way of quantifying the life-cycle costs (and benefits) associated with any given set of safety measures.

Predictive models have been developed and deployed by a number of agencies in North America. Although the latest AASHTO Roadside Design Guide probably represents the most current and widely accepted effort in this regard, designers should be aware that the state of the art in this area is continually developing and should be monitored regularly for new models and techniques which may have application to their design challenges.

The techniques of cost-effectiveness analysis are well established and are applied for a variety of purposes in transportation and highway design agencies. A number of alternative approaches are available but, most commonly, the tools used by transportation agencies are built on life-cycle costing models and use present worth or annualised cost techniques as their underlying analysis methodology. All these approaches are built on fundamental assumptions regarding parameters such as discount rates and unit crash costs. In order to enforce consistent and comparable results across the road authority, these basic assumptions are usually set as a matter of policy and represent a "given" for designers to use in their analyses.

8.1.6 Road safety audits

First developed in the United Kingdom, Australia and New Zealand, this process, which specifically investigates the roadside safety of a particular project, has a proven potential to improve the safety of both proposed and existing facilities.

Road safety audits, especially during the design stage, create the opportunity to eliminate, as far as possible, road safety problems in the provision of new road projects. They should be seen, however, as part of the broader scope of the philosophy of roadside hazard management.

A road safety audit is a formal examination of any road project which interacts with road users, in which a qualified and independent examiner reports on the projects accident potential and safety performance. The audit may be conducted at the project's:
- Feasibility stage;
- Draft design stage;
- Detailed design stage;
- Pre-opening stage; and
- On existing roads.

The earlier a road is audited within the design and development process, the better.

The subject is dealt with more fully in Chapter 2.5 of this document, and particularly in Volume 6 of the South African Road Safety Manual.

8.2 ROADSIDE HAZARDS AND CLEAR ZONE CONCEPT

8.2.1 Overview

Research has shown that in 50 per cent of all run-off-the-road accidents the vehicle leaves the road in a skidding manner. Roadside hazards

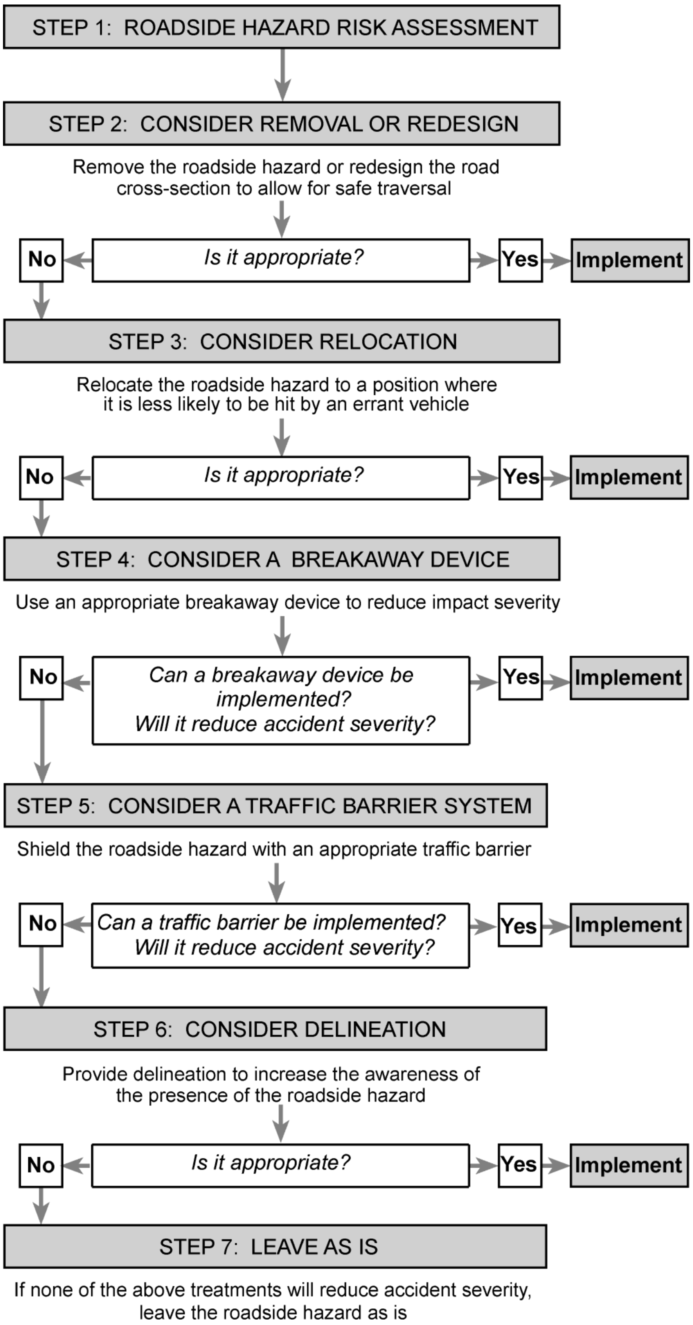


Figure 8.1: Roadside safety analysis

can significantly increase the severity of crashes and it is necessary to manage roadside hazards in such a manner as to decrease the severity of these crashes. Figure 8.1 illustrates this process.

Common existing roadside hazards include:

- Trees;
- Supports and poles (for lighting, utilities, signage;
- Drainage structures such as culverts, drains, drop inlets;
- Bridge abutments/piers;
- Side slopes such as embankments;
- Ends of traffic barriers, bridge railings;
- Incorrectly positioned traffic barriers, i.e. <3 metres off the roadway;

- Obsolete roadside furniture;
- SOS call boxes; and
- Fire hydrants

Accidents involving roadside objects are significant for both the urban and rural road environments. In South Africa, approximately 25 per cent of all accidents involve vehicles running off the road. In 1996 alone, the accident costs related to fixed object accidents amounted to R3,5 billion (1997 Rand).

It is not feasible to provide sufficient width adjacent to the carriageway that will allow all errant vehicles to recover. Therefore it is necessary to reach a compromise or level of risk management. The most widely accepted form of risk management for roadside hazards is the 'clear zone concept'. The clear zone is the horizontal width (measured from the edge of the traffic lane) that is kept free from hazards to allow an errant vehicle to recover. The clear zone is a compromise between the recovery area for every errant vehicle, the cost of providing that area and the probability of an errant vehicle encountering a hazard. The clear zone should be kept free from non-frangible hazards where economically possible; alternatively, hazards within the clear zone should be protected. The clear zone width is dependent on:

- Speed;
- Traffic volumes;
- Side slopes; and
- Horizontal geometry.

It should be noted that the clear zone width is not a magical number and, where possible, hazards beyond the desirable clear zone should be minimized.

Clear zone widths vary throughout the world depending on land availability and design policy. The concept originated in the United States in the early 1960's and has progressively been refined and updated. The clear zone width varies between 4.0 and 10 metres with the upper end of the scale being more appropriate for high-speed National Roads. More recent studies have found that the first 4,0 - 5,0 metres provide most of the potential benefit from clear zones.

8.2.2 Elements of the clear zone

The clear zone falls within an area called the recovery zone. The recovery zone is the total unobstructed traversable area available along the edge of the road and, by convention, it is measured from the edge of the closest travel lane. The recovery zone may have recoverable slopes, non-recoverable slopes and a clear run-out area.

Figure 8.2 illustrates the clear zone concept in the context of the roadside recovery zone.

Recoverable slopes are those on which a driver may, to a greater or lesser extent, retain or regain control of a vehicle. A non-recoverable slope may be traversable, but a vehicle will continue to the bottom. A clear run-out area is located at the toe of a non-recoverable slope, and is available for safe use by an errant vehicle. There is also provision for a smooth transition between slopes to allow for the safe passage of vehicles.

The clear zone is the total, fixed-object-free area available to the errant vehicle. The design domain for the clear zone width has been found to depend on traffic volume and speed, road

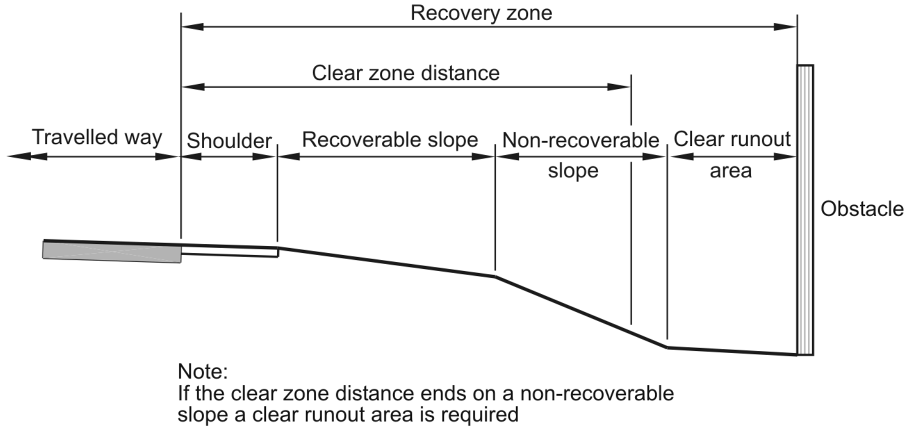


Figure 8.2: Roadside recovery zone

geometry, embankment height, side slope and environmental conditions such as rain, snow, ice, and fog. The wider the clear zone, the less the frequency and severity of collisions with fixed objects. However, there is a point beyond which any further expenditure to move or protect the fixed objects is not warranted because the marginal risk reduction is too small.

8.2.3 Factors influencing the clear zone design domain

When originally introduced, the clear zone concept dictated a single value of 9 metres and was based on limited research. The concept was formally introduced in the 1974 version of the AASHTO report entitled Highway Design and Operational Practices Related to Highway Safety where the authors noted:

"…for adequate safety, it is desirable to provide an unencumbered roadside recovery area that is as wide as practical on a specific highway section. Studies have indicated that on high-speed highways, a width of 9 metres or more from the edge of the travelled way permits about 80 per cent of the vehicles leaving a roadway out of control to recover….."

The last portion of this statement requires emphasis. Provision of the recommended clear zone does not guarantee that vehicles will not encroach further than the recommended clear zone distance. Quite the contrary, the clear zone principle embodies the explicit fact that some portion of the vehicles that encroach will go beyond the clear zone itself.

Steeper embankment slopes tend to increase vehicle encroachment distances. Conversely, on low-volume or low-speed facilities, the 9,0 m distance was found to be excessive and could seldom be justified. As a result, as the concept evolved, design practice moved to a variable clear zone distance definition and a better understanding of the wide range of factors that influence the limits of its design domain was gained.

The approach set out in paragraph 8.2.4 below, and borrowed from Canadian practice reflects the influence of:

- Design speed;
- Traffic volumes;
- The presence of cut or fill slopes;
- The steepness of slopes; and
- Horizontal curve adjustments.

Designers should, however, recognize the limitations of the figures presented below. AASHTO provides a caution to designers on the issue: ".....the numbers obtained from these curves represent a reasonable measure of the degree of safety suggested for a particular roadside; but they are neither absolute nor precise. In some cases, it is reasonable to leave a fixed object within the clear zone; in other instances, an object beyond the clear zone distance may require removal or shielding. Use of an appropriate clear zone distance amounts to a compromise between safety and construction costs."

8.2.4 Determining width of clear zone

Table 8.2 provides an indication of the appropriate width of a clear zone on a straight section of road, measured in metres from the edge of the lane, according to design speed, traffic volumes and cut or fill slope values. The values in Table 8.2 are taken from the 1996 AASHTO Roadside Design Guide, and suggest only the approximate centre of a range to be considered and not a precise distance, since, in making their choice, designers should also consider specific site conditions.

Where side slopes are steeper than 1 : 4 (i.e. non-trafficable) designers should give consideration to the provision of a protection barrier.

For sections of road with horizontal curvature, these distances should be increased on the outside of curves by a factor that depends on the operating speed and the radius of the curve. Figure 8.3 provides guidelines on adjustment factors for clear zones on the outside of curves.

8.2.5 Best practices in respect of roadside vegetation

Single-vehicle collisions with trees account for a sizeable proportion of all fixed object collisions. Unlike typical roadside hardware, with the exception of landscaping, trees are not a design element over which the designers have direct control. While policies and approaches vary by agency, a number of best practices are presented here to assist the designer in dealing with this complex and important issue.

Depending on their size, trees within the clear zone constitute a serious hazard. Generally, a tree with a trunk diameter greater than 150 mm is considered a fixed object.

When trees or shrubs with multiple trunks, or groups of small trees are close together, because of their combined cross-sectional area, they may be considered as having the effect of a single tree

Typically, large trees should be removed from within the selected clear zone for new construction and reconstruction projects. Segments of a highway can be analysed to identify groups of trees or individual trees that are candidates for removal or shielding.

Table 8.2 Clear zone distances (metres)					
Design Speed (km/h)	Design ADT	Fill slopes		Cut slopes	
		1:4 to 1:5	1:6 or flatter	1:4 to 1:5	1:6 or flatter
<60	<750	2,0 – 3,0	2,0 – 3,0	2,0 – 3,0	2,0 --3,0
	750 – 1500	3,5 – 4,5	3,0 – 3,5	3,0 – 3,5	3,0 – 3,5
	1500 – 6000	4,5 – 5,0	3,5 – 4,5	3,5 – 4,5	3,5 – 4,5
	>6000	5,0 – 5,5	4,5 – 5,0	4,5 – 5,0	4,5 – 5,0
70 – 80	<750	3,5 – 4,5	3,0 – 3,5	2,5 – 3,0	3,0 – 3,5
	750 – 1500	5,0 – 6,0	4,5 – 5,0	3,5 – 4,5	4,5 – 5,0
	1500 – 6000	60 – 8,0	5,0 – 5,5	4,5 – 5,0	5,0 – 5,5
	>6000	7,5 – 8,5	6,0 – 6,5	5,5 – 6,0	6,0 – 6,5
90	<750	4,5 – 5,5	3,5 – 4,5	3,0 – 3,5	3,0 – 3,5
	750 – 1500	6,0 – 7,5	5,0 – 5,5	4,5 – 5,0	5,0 – 5,5
	1500 – 6000	7,5 – 9,0	6,0 – 6,5	5,0 – 5,5	6,0 – 6,5
	>6000	8,0 – 10,0	6,5 – 7,5	6,0 – 6,5	6,5 – 7,5
100	<750	6,0 – 7,5	5,0 – 5,5	3,5 – 4,5	4,5 – 5,0
	750 – 1500	8,0 – 10,0	6,0 – 7,5	5,0 – 5,5	6,0 – 6,5
	1500 – 6000	10,0 – 12,0	8,0 – 9,0	5,5 – 6,5	7,5 – 8,0
	>6000	11,0 – 13,5	9,0 – 10,0	7,5 – 8,0	8,0 – 8,5
>110	<750	6,0 – 8,0	5,5 – 6.0	4,5 – 5,0	4,5 – 4,9
	750 – 1500	8,5 – 11,0	7,5 – 8.0	5,5 – 6,0	6,0 – 6,5
	1500 – 4000	10,5 – 13,0	8,5 – 10.0	6,5 – 7,5	8,0 – 8,5
	>6000	11,5 – 14,0	9,0 – 10.5	8,0 – 9,0	8,5 – 9,0

While tree removal generally generates some public resistance, it will reduce the severity of any crashes.

Tree removal often has adverse environmental impacts. It is important that this measure only be used when it is the only solution. For example, slopes of 1:3 or flatter may be traversable but a vehicle on a 1:3 slope will usually reach the bottom. If there are numerous trees at the toe of the slope, the removal of isolated trees on the slope will not significantly reduce the vehicle/tree collision risk although some isolated trees may be candidates for removal if they are noticeably close to the roadway. If a tree or group of trees is in a vulnerable location but cannot be removed, traffic barriers can be used to shield them.

Maintenance of the roadside plays an important role in helping to control vegetation and tree problems by mowing within the clear zone and eliminating seedlings before they create a hazard.

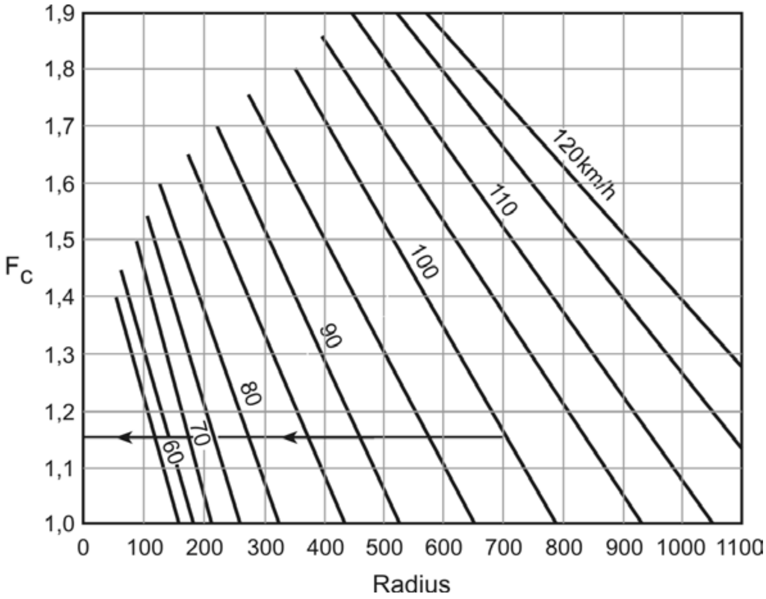


Figure 8.3: Adjustment for clear zones on curves

8.3 SIGN AND OTHER SUPPORTS

8.3.1 Basis for Design

Although the objective of roadside design is to provide an adequate clear zone to allow errant vehicles to recover without a crash, this is not always possible. For various reasons, including traffic operation, certain obstacles may have to remain within the clear zone.

These obstacles include:

- Traffic signposts;
- Utility poles;
- Roadway illumination features; and
- Structures, including headwalls of drainage structures.

Collisions with sign and lighting supports constitute a significant portion of all vehicle crashes and thus merit serious attention. The roadside hazard danger associated with utility and signpost poles increases with an increase in traffic flow, pole density (poles per km of road) and the offset from the edge of the road. The hazard of poles located on the outside of horizontal curves and adjacent to pavements with low skid resistance pavements is greater than at other sites.

These poles can be treated in a number of ways, namely:

- By relocating them to a safer location (this can include moving a lighting pole to the inside of a horizontal curve rather than on the outside)
- Removal of some of the poles by:
    - Increasing the pole spacing;
    - The combining of a number of utilities or signs per pole; or
    - Installing underground cables;
- Shielding the utility poles with an appropriate traffic barrier system and provision of a proper end-treatment;
- If appropriate, installation of a break away device;
- Providing a high skid resistance surfacing on curves; and
- Attaching delineators to the device to increase its visibility if no other measure can be implemented.

Figure 8.1 illustrates the proposed methodology.

8.3.2 Breakaway supports

Definition

The term "breakaway support", developed in the late 1960's, refers to all types of signs, luminaires and traffic signal supports that are safely displaced under vehicle impact, whether the release mechanism is a slip plane, plastic hinge, fracture element or a combination of these.

The AASHTO Standard Specifications for Structural Supports for Highway Signs, Luminaires and Traffic Signals specifies that all supports located in the clear zone widths of high-speed facilities should be equipped with a breakaway device unless they are protected with a suitable traffic barrier system. In the case of urban or low speed facilities, the use of breakaway devices is not advisable. The AASHTO Specifications seem to be inadequate as an occupant can sustain serious injuries when striking the car interior during a vehicle impact at 40 km/h with a non-yielding object. The reason for this guideline, however, is that there is a probability of cyclists and pedestrians being injured by the yielding support(s). It is therefore important that the designer should consider the relative risks related to each location before a design is selected.

A breakaway support is designed for loading in shear and normally for impact at bumper height (typically 500 mm above ground level). It is critical that the support be properly installed as to ensure that loading takes place at the correct height. Loading above the design height may cause the breakaway device to fail to activate because the bending moment in the breakaway support may be sufficient to keep the support in place. Incorrect loading can also take place when the support is installed close to ditches or steep slopes, causing a vehicle to become airborne and hit the support at the wrong position.

The soil type of the breakaway system is important as it may also affect the activation of the mechanism. In the case of fracture-type supports such as high carbon U-channel posts, telescoping tubes and wood supports, the supports may slip through saturated or loose soil during impact, absorbing energy and changing the breakaway mechanism.

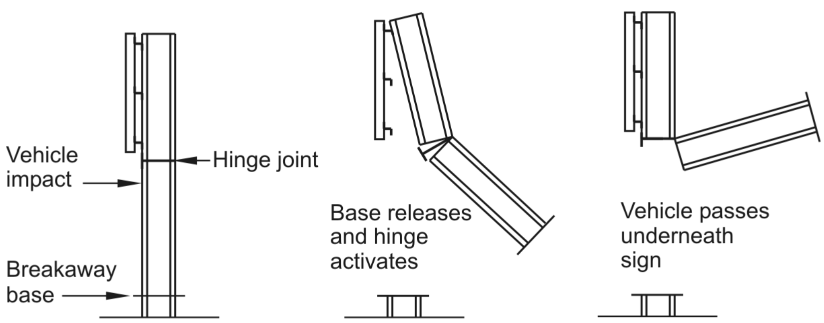


Figure 8.4: Breakaway supports

If a sign support is installed at a depth less than 1 metre, it will pull out of the soil during impact. Installations with anchor plates or those installed deeper than 1 m are particularly sensitive to the foundation conditions. For small sign supports using base-bending or yielding mechanisms, the performance of the supports in strong soils is more critical.

The maintenance requirements are critical in the selection of a particular breakaway device. The following maintenance requirements should be considered:

(a) The availability of breakaway devices will influence the costs associated with installation and maintenance or replacement after impact. An installation that can be reused can be more cost-effective than mechanisms that have to be replaced.

(b) The durability of a support is important as it will determine the life span of a support that is not struck as compared to that of a non-breakaway support.

(c) A breakaway device yields when hit if it is properly installed and maintained. The mechanism should then be replaced or repaired. Consequently, the availability of material, maintenance personnel and availability of personnel after an impact for each breakaway design should influence the selection thereof.

Acceptance criteria for breakaway supports

The AASHTO guide, Standard Specification for Structural Support for Highway Signs, Luminaires and Traffic Signals, should be used to determine whether a sign support conforms to the criteria for a breakaway support.

The broad criteria which breakaway supports should meet include:

- Dynamic performance criteria, i.e. implicit velocity breakaway thresholds;
- Maximum remaining stub height of 100mm;
- The need for the vehicle to remain upright during and after the collision; and
- No significant deformation of the vehicle or intrusion into the passenger compartment during or after impact.

Non- breakaway sign supports

The first requirement for sign supports is the need to structurally support the devices that are mounted upon them. Signs and other devices should be carefully placed in order to minimize the hazard that they can represent to motorists. The following practices should be borne in mind by designers when developing signing plans for their projects:

(a) Sign supports should not be placed in drainage ditches, where erosion might affect the proper operation of breakaway supports.

(b) Wherever possible, signs should be placed behind existing roadside barriers (beyond the deflection distance), on existing structures, or in non-accessible areas. If this cannot be achieved, then breakaway supports should be used.

(c) Only when the use of breakaway supports is not practicable should a traffic barrier or crash cushion be used to shield sign supports.

8.3.3 Design and Location Criteria for Sign Supports

Roadway signs fall into three primary classes: overhead signs, large roadside signs, and small roadside signs.

Overhead Signs

Since overhead signs, including cantilevered signs, require massive support systems that cannot be made breakaway, they should be installed on or relocated to nearby overpasses or other structures, where possible.

All overhead sign supports located within the clear zone should be shielded with a crashworthy barrier. In such instances, the sign gantry should be located beyond the design deflection distance of the barrier.

Large Roadside Signs

Large roadside signs are generally greater than 5,0 square metres in area. Typically, they have two or more supports that are breakaway.

The hinge for breakaway supports on large roadside signs should be at least 2100 mm above ground, so that the likelihood of the sign or upper section of the support penetrating the windshield of an impacting vehicle is minimized. The required impact performance is shown in Figure 8.4.

No supplementary signs should be attached below the hinges if their placement is likely to interfere with the breakaway action of the support post or if the supplementary sign is likely to strike the windscreen of an impacting vehicle.

The design requirements for breakaway support systems for roadside signs are documented in a number of publications. The South African Road Safety Manual should be used as the basis for the design of these.

Small road side signs

Small roadside signs are supported on one or more posts and have a sign panel area of less than 5,0 square metres. Although not perceived as significant obstacles, small signs can cause serious damage to impacting automobiles, and wooden posts should be used as far as possible.

The bottom of the sign panel should be a minimum 2100 mm above ground and the top of the panel should be a minimum 2700 mm above ground to minimize the possibility of the sign panel and post rotating on impact and striking the windshield of a vehicle.

The requirements for breakaway support systems for roadside signs are documented in a number of publications. The South African Road Safety Manual should be used as the basis for design.

Consideration to various factors should be given when selecting, designing and locating breakaway and other supports. These include:

- Road environment : urban or rural;
- Terrain where device is installed;
- Proximity to drainage ditches or structures;
- Soil type used as a base for the breakaway support;
- Maintenance requirements of the support (i.e., the simplicity of maintenance, availability of material and the durability of the support); and
- Expected impact frequency.

8.3.4 Design approach for lighting supports

Lighting supports should be of the frangible base, slip base or frangible coupling type. They are designed to release in shear when hit at a typical bumper height of about 500 mm.

As long as the side slopes between the roadway and the luminaire support are 6:1 or flatter, vehicles should strike the support appropriately, and breakaway action can be assured.

Superelevation, side slope, rounding and vehicle departure angle and speed will influence the striking height of a typical bumper. Designers should consider this fact when developing illumination plans for their projects.

As a general rule, a lighting support will fall near the line of the path of an impacting vehicle. Designers should be aware that these falling poles represent a threat to bystanders such as pedestrians, bicyclists and uninvolved motorists.

Poles with breakaway features should not exceed 17 m in height - the current maximum height of accepted hardware.

The mass of a breakaway lighting support should not exceed 450 kg.

Foundations for lighting supports should be designed with consideration being given to the surrounding soil conditions that could influence the effectiveness of the breakaway mechanism.

When a lighting support is located near a traffic barrier and if it is within the design deflection distance of the barrier, it should be either a breakaway design, or the barrier should be strengthened locally to minimize its deflection.

Higher mounting heights can reduce the number of lights needed on a facility. High mast lighting - which requires far fewer supports located much further from the roadway - can be beneficial. While consideration of this approach is recommended to designers, the massive nature of these high mast structures requires analysis and planning in the design and placement of the high mast supports.

8.4 TRAFFIC SAFETY BARRIERS

8.4.1 Overview

Traffic safety barriers are systems utilized to shield road users from potential hazards alongside the travelled way and should be able to redirect or contain:

- An errant vehicle without imposing intolerable vehicle occupant forces;
- Vehicles in range of sizes, weights and designs; or
- An errant vehicle over a range of impact speeds and impact angles.

Traffic barriers are obstacles on the roadside and vehicles striking barriers can cause occupant injury and/or vehicle damage. A traffic barrier should be installed only if it is likely to reduce the severity of potential collisions. It is therefore of the utmost importance that, in selection of the traffic barrier, due cognisance be taken of the characteristics of the particular barrier system. Barrier systems differ not only in purpose but also in terms of deflection and redirecting properties.

Traffic barriers are either classified as being impact attenuation devices or longitudinal barriers.

The purpose of an impact attenuation device is to cause a vehicle to decelerate and come to a halt. A longitudinal traffic barrier redirects a vehicle parallel to the roadway.

Figure 8.5 shows a functional classification of traffic barriers.

8.4.2 Determining Need for Safety Barriers

Barriers are installed on the basis of warrant analysis. Traditionally, these warrants have been based on a subjective analysis of certain roadside elements or conditions within the clear zone. If the consequences of a vehicle running off the road and striking a barrier are believed to be less serious than the consequences if no barrier existed, the barrier is considered warranted.

Although this approach can be used, there are often instances where the distinction between the two conditions is not immediately obvious. In addition, this approach does not allow for consideration of the cost-effectiveness of treatment or non-treatment.

In recent years, techniques have been developed which allow warrants for barrier installation to be established on the basis of a benefit cost analysis in which such factors as design speed, traffic volume, installation and maintenance costs, and collision costs are taken into consideration.

Typically, such an approach is used to evaluate three options:

- The removal or alteration of the area of concern so that it no longer requires shielding;
- The installation of an appropriate barrier; or
- Leaving the area of concern unshielded (usually only considered on low-volume and/or low-speed facilities).

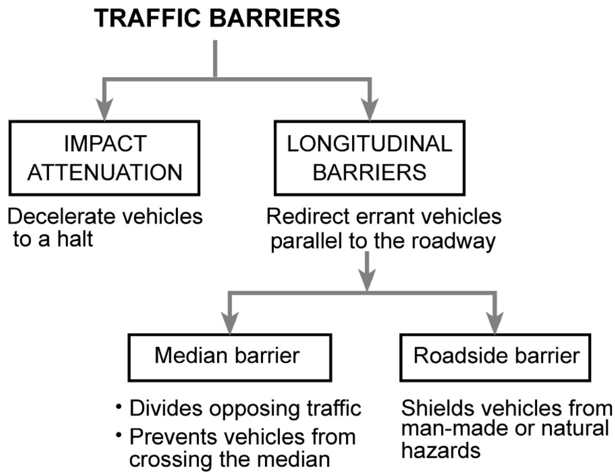


Figure 8.5: Classification of traffic barriers

Once a barrier is found to be necessary for a embankment, it should be provided over the entire length of the embankment and not simply terminated when the embankment height becomes less than the warranted height.

Barrier warrants for roadside obstacles are based on their location within the clear zone and are a function of the nature of the obstacle, its distance from the travelled portion of the roadway and the likelihood that it will be hit by an errant vehicle.

Conventional criteria used for embankments and roadside hazards are not usually applicable to the pedestrian/bicyclist case, and these are usually resolved through a careful individual evaluation of each potential project.

As with roadside barriers, warrants for median barriers have been established on the basis that a barrier should be installed only if the consequences of striking the barrier are less severe than the consequences that would result if no barrier existed. The primary purpose of a median barrier is to prevent an errant vehicle from crossing a median on a divided highway and encountering oncoming traffic. As such, the development of median barrier warrants has been based on an evaluation of median crossover collisions and related research studies. In determining the need for barriers on medians, median width and average daily traffic volumes are the basic factors generally used in the analysis. However the incidence of illegal cross-median movements may also justify the use of median barriers.

Warrants for implementing impact attenuation divides (crash cushions) are based on shielding a fixed object within the clear zone that is considered to be a hazard and cannot be removed, relocated, made breakaway, or adequately shielded by a longitudinal barrier.

In considering the use of traffic barriers, designers should note that, even when these are properly designed and constructed, they might not protect errant vehicles and their occupants completely. After installation of these, the severity of collisions generally decreases but, as the number of installations increases, the frequency of minor collisions may also increase. For this reason, where cost-effective, the designer should make every effort to design without traffic barriers. This can be done by clearing the roadside of obstacles, flattening embankment slopes and introducing greater median separation where possible. It should be noted, however, that, whilst a particular barrier system is chosen based on the containment level required, regular monitoring is essential to allow the system to be replaced by a more adequate one if experience indicates the need for this.

8.4.3 Longitudinal roadside barriers

Classification and performance characteristics

Figure 8.6 shows the classification of longitudinal barriers based on their deflection characteristics. It should be noted that the deflection characteristics of a barrier system are not an indication of its effectiveness or safety.

Misconceptions exist regarding the advantages of the different longitudinal barrier types. Some engineers firmly believe that one system is better than another based on its deflection charac-

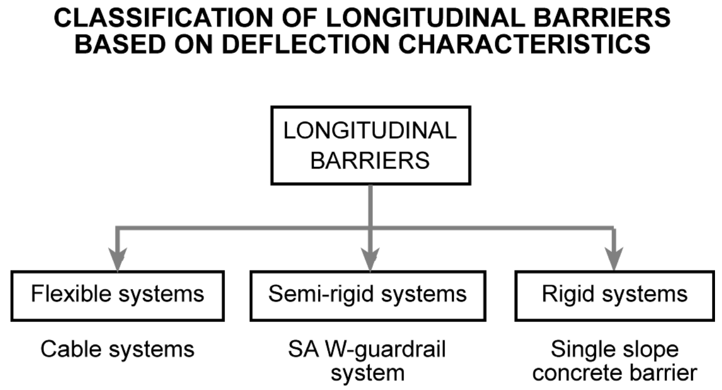


Figure 8.6: Classification of longitudinal barriers

teristics. The deflection characteristics of a particular system are not a measure of its effectiveness. The mechanisms by which a vehicle is restrained after impacting a traffic barrier differ completely depending on the type of barrier selected. The reaction of a vehicle on impact with different types of barriers is thus also different.

In accomplishing their task of guiding and redirecting impinging vehicles, a longitudinal barrier should balance the need to prevent penetration of the barrier with the need to protect the occupants of the vehicle. Various barrier technologies achieve this in various ways and can be grouped into three distinct types:

1) Flexible systems, resulting in large lateral barrier deflections, but the lowest vehicle deceleration rates. Such systems have application in places where a substantial area behind the barrier is free of obstructions and/or other hazards within the zone of anticipated lateral deflection. These barriers usually consist of a weak post-and-beam system, and their design deflections are typically in the range of 3,2 metres to 3,7 metres.but can be as low as1,7 metres.

2) Semi-rigid systems, providing reduced lateral barrier deflections, but higher vehicle deceleration rates. These barrier systems have application in areas where lateral restrictions exist and where anticipated deflections have to be limited. They usually consist of a strong post-and-beam system and have design deflections ranging from 0,5 to 1,7 metres.

3) Rigid systems, usually taking the form of a continuous concrete barrier. These technologies result in no lateral deflection, but impose the highest vehicle deceleration rates. They are usually applied in areas where there is very little room for deflection or where the penalty for penetrating the barrier is very high. Numerous shapes are available, including a high version for use where there is a high percentage of trucks.

Designers should familiarize themselves with, and design to, the specific performance characteristics of their selected or candidate technologies.

Selection guidelines

Roadside barriers may be subjected to a wide range of impacts by errant vehicles and provide a wide range of protection to the occupants of such vehicles. It is therefore necessary to determine the level of protection that they will provide.

The procedures and criteria for assessing the safety performance of traffic barriers and other features have been standardized in the USA through the publication of various National Co-operative Highway Research Program (NCHRP) Reports. The current test battery is described in NCHRP Report 350 "Recommended Procedures for the Safety Evaluation Performance of Highway Features" (and includes various test levels defined by the size of the test (design) vehicle, impact speed and angle). This provides the designer with the opportunity to match test conditions with the anticipated operational conditions on the road.

In selecting an appropriate traffic barrier it is essential that designers have a good understanding of the protection level expected from the barrier and they should note that, if they choose a particular system that is inadequate, they might make themselves, or their agency, liable for damages.

The following factors should therefore be seriously considered before a particular barrier is selected:

- Performance capability;
- Site conditions;
    - Compatibility;
    - Life cycle costs;
    - Maintenance;
    - Aesthetics; and
    - Field experience.

The requirements of NHCRP Report 350 should be regarded as the minimum.

Performance capability

The "design conditions" for a particular barrier need to be assessed carefully because areas with poor geometrics, high traffic volumes, high speeds and a large proportion of heavy vehicles might not be consistent with the "conditions" assumed when the barrier had been tested. Such sites might require barriers with a higher than normal performance level.

Site conditions

Site conditions play a major role in the selection of appropriate barriers. The slope approaching a flexible barrier should, for example, not exceed 10 per cent and rigid barriers should not be used where the expected impact angle is large. Narrow fill sections could result in conditions where post spacing and post support might be inadequate to allow them to perform as intended.

A number of site-specific aspects will have a major influence on the selection of a particular type of barrier to meet the performance requirements at that location. These aspects include:

- Compatibility.
  All barriers are subject to damage and require intermittent maintenance. Keeping the number of different barrier types to a minimum therefore simplifies maintenance. Special barrier designs should only be considered when site or operational conditions cannot be satisfied with the standard barrier.

- Life cycle costs.

  It is prudent to realize that any barrier system accrues costs throughout its life. High initial costs could mean low maintenance costs, whilst low initial costs could mean higher maintenance costs. In addition, expected accident costs should be considered in the calculation of life cycle costs.

- Maintenance.

  Most systems require very little routine maintenance. When the barrier has been involved in a crash, the subsequent rehabilitation costs may be significant, to the point of being excessive in the case of a high accident location. It should be noted that only material specifically designed for that particular system should be used for maintenance, and the tendency to "mix and match" should be avoided.

- Aesthetics.

  It is important to realize that all traffic barriers are visual obstructions. Should this become a particular concern it is necessary to ensure that alternative systems that may be considered are able to meet the performance requirements. Aesthetics should, under no circumstances, be given preference before safety considerations.

- Field experience.

  Site personnel's' experience of the performance, cost and maintenance requirements of installed systems as well as the traffic police services' experience of the performance of particular barrier systems under impact conditions, should not be under-estimated by the designer. Early identification of potential problems can ensure that future installations operate effectively.

Warrants for use

Roadside hazards that warrant shielding by barriers include embankments and roadside obstacles. Warrants for the use of barriers on embankments generally use embankment height and side slope as the parameters in the analysis and essentially compare the collision severity of hitting a barrier with the severity of going down the embankment. Figure 8.7, adapted from Australian guidelines, provides guidance for the installation of such barriers on embankments.

However such warrant procedures are regarded as less than adequate because they do not take into account the probability of a crash occurring against the barrier or the cost of installing a barrier versus leaving the slope unprotected. The development of cost analysis techniques provides the designer with an approach to analysing the need for roadside embankment protection barriers. In South Africa, however, there is a lack of reliable data to carry out such analyses and it is necessary for the designer to make site-specific analyses, using Figure 8.7 as a guide.

The significance of this figure is that it provides a range of values of fill slope for which, at certain heights of fill, a barrier may be more or less hazardous than the embankment it protects. For example, at a fill height of 6 metres, a fill slope steeper than 1:3 would warrant the use of a barrier while a fill slope flatter than 1:4 would not require protection. On the intervening slopes, the designer should use his or her discretion in determining the need for a barrier.

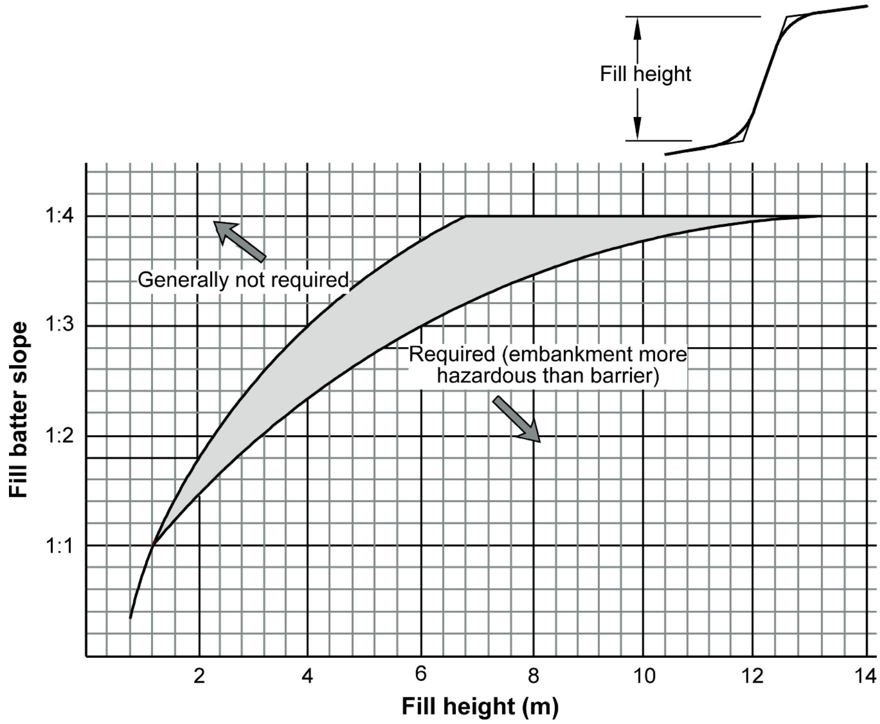


Figure 8.7: Warrants for use of roadside barriers

In respect of roadside obstacles that need protection from errant vehicles (or vice versa) warrants for shielding or otherwise can be developed using a quantitative cost-effective analysis, which takes the characteristics of the obstacle and its likelihood of being hit into account. However, once again the designer must examine each site specifically to determine the necessity or otherwise for shielding. Table 8.3 provides an overview of the types of non-traversable terrain and fixed objects that are normally considered for shielding.

In some situations, a measure of physical protection may be required for pedestrians or bicyclists using, or in close proximity to, a major street or highway. Examples of such cases could include;

- A barrier adjacent to a school boundary or property to minimize potential vehicle contact;
- Shielding businesses or residences near the right of way in locations where there is a history of run-off-the road crashes; or
- Separating pedestrians and/or cyclists from vehicle flows in circumstances where high-speed vehicle intrusions onto boulevards or sidewalk areas might occur.

In all these cases, conventional criteria will not serve to provide warrants for barriers, and the designer should be aware of the needs and circumstances of the individual situation when deciding on appropriate action.

Longitudinal barrier placement

A typical longitudinal roadside barrier installa-

Table 8.3: Roadside obstacles normally considered for shielding	
Terrain or Obstacle	**Comment**
Bridge piers, abutments, railing ends	Shielding analysis required
Boulders	Judgement: nature of object: likelihood of impact
Culverts, pipes (smooth)	Judgement: based on size, shape, location
Cut slopes (smooth)	Shielding analysis not generally required
Cut slopes (rough)	Judgement: based on likelihood or impact
Ditches (parallel)	Analysis generally required
Embankments	Judgement: based on fill height and slope
Retaining walls	Judgement: based on wall smoothness and angle of impact
Sign and luminaire supports	Shielding analysis for isolated signals in the clear zone on high speed (80 km/h or greater) facility
SOS telephones	Shielding analysis required
Traffic signal supports	Shielding analysis for isolated signals in the clear zone on high speed (80 km/h or greater) facility
Trees	Judgement: site specific
Utility poles	Judgement: case by case basis
Permanent bodies of water	Judgement: depth of water, likelihood of encroachment

tion, with its associated elements for a two-lane, two-way road, is illustrated in Figure 8.8. The length of need as illustrated in this figure is illustrated in more detail in Figure 8.10.

The factors to be considered in barrier installation are the following:

- Offset of the barrier from the travelled way;
- Rail deflection distance;
- Terrain effects;
- Flare rate; and
- Length of need.

Barriers should ideally be set as far away from the travelled way as possible. This ensures that:

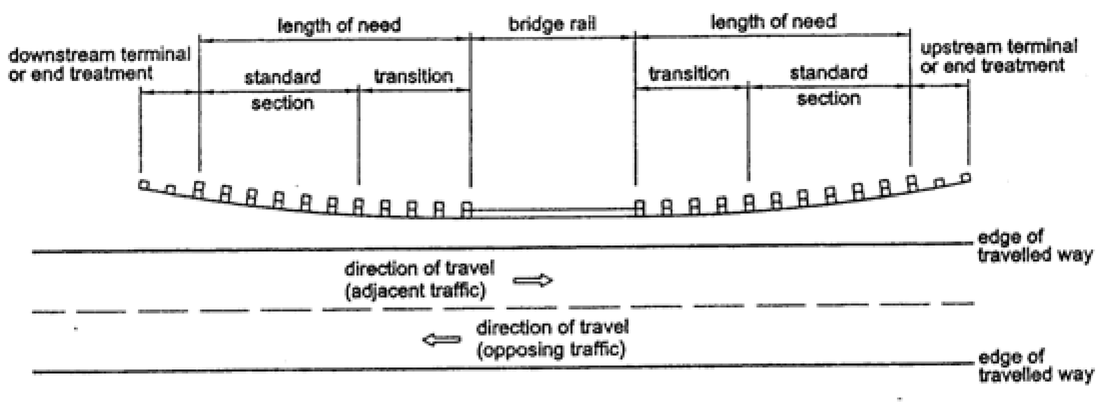


Figure 8.8: Roadside barrier elements

Barriers should ideally be set as far away from the travelled way as possible. This ensures that:

- There is more recovery area to regain control of the vehicle;
- There is better sight distance;
- Less barrier is required to shield the hazard; and
- Adverse driver reaction to the barrier is reduced.

However placing the barrier away from the roadway and closer to the hazard may have disadvantages. These are:

- The possible impact angle increases, leading to higher risk of the vehicle penetrating the rail as well as increased collision severity. The South African Road Safety Manual provides a detailed discussion on this issue.
- The roadside area in front of the barrier has to be traversable and as flat as possible.

Barrier deflection

The expected deflection of a barrier should not exceed the available space between rail and the object being shielded. If the available space between the rail and the obstacle is not adequate for non-rigid barrier systems then the barrier can be stiffened in advance of, and alongside, the fixed object. This can be achieved through reducing the post spacing, increasing post sizes or increasing the rail stiffness by nesting rail elements. However care should be exercised when considering this step since the total system characteristics might be altered.

Other areas of concern include the possibility of rolling over when vehicles with a high centre of gravity impact a barrier or of vehicles dropping over the edge when a barrier, positioned too close to the edge, deflects on impact.

Table 8 4. Recommended minimum offset distances

Design Speed (km/h)	Offset Distance (m) measured from the edge of the travelled way
50	1,1
60	1,4
80	2,0
100	2,4
120	3,2
130	3,7

Recommended offset distances measured from the edge of the travelled way are shown in Table 8.4. Barriers are typically placed at a distance of 0,3 metres beyond the edge of the usable shoulder so that the greater of the distance in Table 8.4 or the width of the shoulder plus 0,3 metres should be used.

A minimum distance of 600 mm behind flexible guardrails to the edge of an embankment would generally provide enough resistance to lateral movement of the posts to resist the rail tension.

Terrain effects

Roadside features such as kerbs and drainage inlets affect the bumper height and suspension and may cause errant vehicles to snag or vault the barrier.

Kerbs should preferably be sited behind the guardrail face. Barrier offsets less than 230 mm behind the kerb would still be acceptable. The height of the rail should be carefully considered to limit the possibility of the bumper or a wheel under-riding the rail. This may be achieved by setting the rail height relative to the road surface in front of the kerb.

Slopes

Roadside barriers perform best when installed on slopes of 1:10 or flatter. Slope changes may cause vehicles to impact higher on the barrier than normal, increasing the possibility of vaulting. Should barriers be installed beyond a slope change, they should be set back at least 3,5 metres from the slope break line to allow the vehicle trajectory to stabilize. Installation of guardrails on slopes steeper than 1:6 is not recommended because inadequate lateral support for the guardrail posts would result. If this location is unavoidable, consideration should be given to deeper postholes.

Flare rate

A barrier flare may be used to increase the barrier offset from the edge of the roadway. This is normally used to position the barrier terminal further from the roadway, to adjust the existing roadside features, to reduce the total length of rail and to reduce driver reaction to the close proximity of the barrier rail next to the road.

Flared barriers can, however, also lead to increased impact angles causing higher impact severity, as well as to larger rebound angles causing greater conflicts with other vehicles.

The maximum recommended flare rates are shown in Table 8.5. Flatter rates may be used particularly where extensive grading would be

Table 8.5 Recommended maximum flare rates for barrier design

Design Speed (km/h)	Barrier inside the min. offset line	Barrier beyond the min. offset line	
		Rigid Barriers	Non-rigid barriers
50	1:13	1:8	1:7
60	1:16	1:10	1:8
80	1:21	1:14	1:11
100	1:26	1:18	1:14
110	1:30	1:20	1:15
120	1:40	1:25	1:18
130	1:50	1:30	1:20

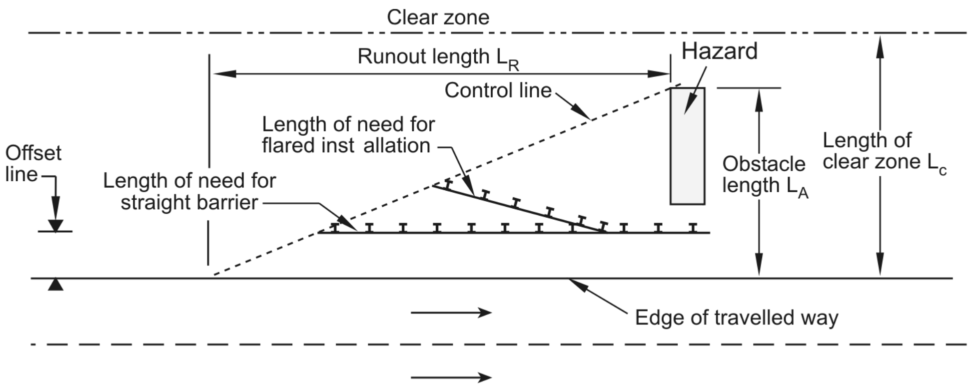


Figure 8.9: Length of need for adjacent traffic

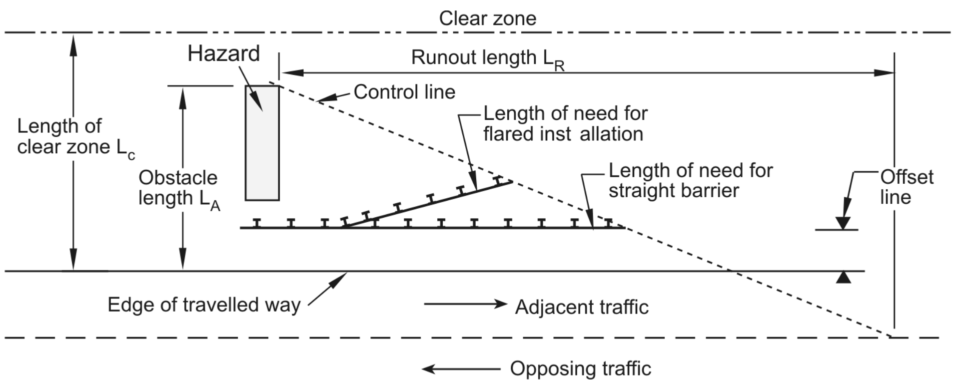


Figure 8.10: Length of need for opposing traffic

required to provide a 1:10 approach slope to the barrier.

Length of need

The variables to be considered in the design process of barriers are shown in Figure 8.9 for the approach side towards a hazard and in Figure 8.10 for the trailing side beyond the hazard, providing for the shielding of the hazard for opposing traffic.

Run-out length is the theoretical distance required for a vehicle leaving the roadway to come to a stop prior to impacting a hazard. The design of a traffic barrier requires provision to be made for sufficient length to restrict such a vehicle from reaching the hazard. The recommended run-out lengths are shown in Table 8.6.

The run-out length is measured along the edge of the road. A control line is established

Table 8.6 Recommended run-out lengths for barrier design

Design speed (km/h)	Run-out length (m)			
	ADT <800	800<ADT<2000	2000<ADT<6000	ADT>6000
50	40	45	50	50
60	50	55	60	70
80	75	80	90	100
100	100	105	120	130
110	110	120	135	145
120 or higher	125	135	150	160

between the end of the run-out length and the far side of the hazard to be shielded. The length of need for a standard barrier would then be the length between the near side of the hazard and the position where the barrier intersects the control line. If the barrier is designed for a continuous hazard such as a river or a critical fill embankment, then the control line would be between the end of the run-out length and the end of the desirable clear zone. The same principle is adopted to determine the length of need for opposing traffic.

The standard guardrail ends at the end of the length of need. An acceptable end-treatment should be added to this length to determine the total length of installation.

An application of the length of need principles to practical design problems is given in Volume 6 of the South Africa Road Safety Manual.

8.4.4 Median barriers

Most of the principles with respect to longitudinal barriers also apply to median barriers. Regarding warrants for their use, median barriers should only be installed if the consequences that would result if they did not exist are more severe than the consequences of striking them. However, excessive incidence of illegal cross-median movements might justify the use of median barriers.

For median widths of 15 metres or greater, median barriers are generally not required, whilst, for median widths of 10 metres and less with ADT's in excess of 30 000 vpd, and 8 metres and less with ADTs below 30 000 vpd, median barriers are generally justified. These figures presuppose that the particular section of roadway under consideration does not suffer from an adverse cross-median collision history and that unauthorized cross-median U-turns do not take place.

Once the need for a median barrier is established, the designer should consider several factors in developing the barrier layout. These include:

- Terrain effects;
- Flare rate of the barrier;
- Treatment of rigid objects in the median; and
- Openings in the median as a result of underpasses.

Terrain effects

For a median barrier to be effective, it is essential that, at the time of impact, the vehicle has all its wheels on the ground and that its suspension system is neither compressed nor extended. Kerbs and sloped medians are of particular concern, since a vehicle, which traverses one of these features prior to impact, may go over or under the barrier or snag on its support posts.

Kerbs offer no safety benefits on high-speed roads and are not recommended where median barriers are present.

Medians should be relatively flat (slopes of 1:10 or less) and free of rigid objects. Where this is not the case, carefully considered placement of the median barrier is needed. AASHTO notes three conditions where specific guidelines for median barrier placement should be followed:

- In depressed medians or medians with a ditch section the slopes and ditch section should first be checked to determine whether a roadside barrier is warranted. If both slopes require shielding, a roadside barrier should be placed near the shoulder on each side of the median. If only one slope requires shielding, a median barrier should be placed near the shoulder of the adjacent travelled way.

- If neither slope requires shielding but both are steeper than 1:10, a median barrier should be placed on the side with the steeper slope, when warranted.

- If both slopes are relatively flat, then a median barrier may be placed at or near the centre of the median if vehicle override is not likely.

For stepped medians that separate travelled ways with significant differences in elevation, a median barrier should be placed near the shoulder adjacent to each travelled way if the embankment slope is steeper than 1:10. If the cross-slope is flatter than 1:10, a barrier could be placed at or near the centre of the median.

Placement criteria are not clearly defined for raised medians or median berms. Research suggests that the cross section of a median berm itself, if high and wide enough, can redirect vehicles impacting at relatively shallow angles.

As a general rule, if the cross section is inadequate for redirecting errant vehicles, a semi-rigid barrier should be placed at the apex of the cross-section. If the slopes are not traversable, roadside barriers should be used near the shoulder adjacent to each of the travelled ways.

Flare rates

If a median barrier has to be flared at a rigid object in the median, the flare rates for roadside barriers should be used for the median barrier flare as well.

Rigid objects

A special case may result in circumstances where a median barrier is not warranted but where a rigid object warrants shielding. Typical examples are bridge piers, overhead sign support structures, and high mast lighting installations. If shielding is necessary for one direction of travel only, or if the object is in a depressed median and shielding from either or both directions of travel is necessary, the criteria for roadside barriers should be used.

If shielding for both directions of travel is necessary and if the median side slopes are steeper than 1:10 the designer may investigate the possibility of a crash cushion (or an earth berm) to shield the object. A second possibility involves the use of semi-rigid barriers with crash cushions or earth berms to shield the barrier ends.

Median openings as a result of underpasses

In certain instances, the cost implications of providing underpasses have the result that an opening in the median occurs. In such instances the use of transverse barriers (or concrete balustrades) shielded by impact attenuation devices should be considered.

End treatments

Traffic barriers (both roadside and median types) themselves represent fixed objects. Impact with their untreated terminal sections can have severe consequences, primarily because of the very high deceleration rates experienced by vehicle occupant under such circumstances, but also often because penetration of the passenger compartment by the barrier itself is a distinct possibility. There are a number of different end treatments available for the various types of barriers.

A proper end terminal has two functions:

- In any non-rigid barrier system, the end terminal should act as an anchor to allow the full tensile strength of the system to be developed during downstream angled impacts on the barrier.
- Regardless of the type of barrier, the end terminal should be crashworthy, i.e. it must keep the vehicle stable and it must keep the vehicle occupants away from rigid points creating high deceleration resulting in serious injuries or death during impact.

Experience has shown that metal beam systems often result in penetration of the passenger compartment, and that high-speed impacts with concrete barriers result in intolerable deceleration forces. In designing crashworthy end treatments, designers must create treatments that provide vehicle deceleration rates that are within recommended limits for survivability.

A number of principles relevant to barrier end treatments are offered:

1. Crashworthy end treatments are essential if a barrier terminates within the clear zone. Such a terminal must not spear, vault, or roll a vehicle in either head-on or angled hits.
2. Barrier end treatments should gradually stop or redirect an impacting vehicle when a barrier is hit end on. The end treatment should also be capable of redirecting a vehicle impacting the side of the terminal.
3. The end treatment should have the same redirectional characteristics as the barrier to which it is attached for impacts at or near the end of the terminal and within the length of need. The end should be properly anchored and capable of developing the full tensile strength of the barrier elements.
4. Where space is available, a barrier can sometimes be introduced far enough from approaching traffic so that the end can be considered non-hazardous and no additional end treatment is required. Flare rates, in this case, should be in accordance with those mentioned above. Positive end anchorage is required in semi-flexible systems in order to preclude penetration of the barrier within the length of need. Care should be taken, however, to ensure that this flaring back does not create a hazard for traffic in the opposing direction.
5. End treatments involving turned down terminals parallel to the direction of travel may cause impacting vehicles to vault and roll over or ride up the terminal and hit the object the barrier is intended to protect. Consequently, turned down terminals should not be used on the approach ends of roadside or median barriers on high-speed, high-volume roads unless they are also flared.
6. Termination of a barrier in a back slope eliminates the danger of an untreated barrier end and reduces the opportunity for errant vehicles to penetrate the end of the barrier.

7. A number of end treatments have been developed for metal beam barriers that utilize a combination of a breakaway mechanism and a cable with a flared configuration to address the spearing and roll-over potential and to develop the full tensile strength of the rail for downstream impacts.

8. Where an end treatment is designed as a "gating" device, i.e., to allow for controlled penetration of a vehicle when impacted, through a breakaway mechanism, care should be taken to provide an adequate run-out area behind the end treatment.

9. The concrete safety shape barrier can be terminated by tapering the end. However, this treatment should only be used where speeds are low (60 km/h or less) and space is limited. Flaring the barrier beyond the clear zone should be considered on higher speed facilities where space is available.

10. Proprietary mechanical end treatments are often suitable only for limited types of barrier applications. When adopting such technologies, designers should ensure not only the efficacy of the technology of their choice but also its compatibility with the barrier technology being used. In addition to information generally available from the manufacturers and suppliers of these treatments, road agencies and others compile and provide appropriate guidance in respect of crash testing results and system compatibility recommendations.

11. All systems should be installed with a level surface leading to the treatment. The use of kerb and gutter is discouraged, but if they are needed, only the mountable type should be specified.

The principles noted above provide a rule of thumb approach. Road designers should still investigate physical site restrictions such as longitudinal space, hazard width, slopes and surface types. At locations with a high likelihood of collisions, the costs of accidents and repair should be factored into the decision matrix in addition to the initial installation costs.

Designers should note that new technologies are continually being developed and tested. Nothing in this Guide relieves the designer of the responsibility of keeping abreast of these new technologies and their potential application to the roadside barrier end treatment problem.

8.5 IMPACT ATTENUATION DEVICES

8.5.1 Function

Impact attenuators, sometimes called crash cushions, are best suited for use in places where fixed objects cannot be removed, relocated or made breakaway, and cannot be adequately shielded by a longitudinal barrier. They have proven to be an effective and safe means of shielding particular types of roadside obstacles, and accomplish their task by absorbing energy at a controlled rate, thereby causing the vehicle to decelerate so as to reduce the potential for serious injury to its occupants. Most operational impact attenuation devices have been designed and tested by their manufacturers and acceptable units can usually be selected directly from design charts.

Typical objects and areas that can benefit from the use of impact attenuators include:

- A freeway exit ramp gore area in an elevated or depressed structure where a bridge rail end or a pier requires shielding;

- The ends of roadside or median barriers;
- Rigid objects like cantilever sign gantries within the clear zone;
- Construction work zones; and
- Toll booths.

It is difficult to develop an easy selection process for determining the most appropriate impact attenuator for a specific situation. This is owing to the large number of factors influencing the choice. The choice could therefore be narrowed down to the use of impact attenuators that have been installed in South Africa and for which a track record (however small at this stage) is being built up. Another major consideration is the ease with which these impact attenuators can be routinely maintained or reinstated after an impact. Certain impact attenuators are marketed specifically as low maintenance attenuators.

the preliminary design of the impact attenuator:

- Hazard characteristics - type width and height;
- Site geometry - including space available for installation;
- Traffic pattern - bi-directional or uni-directional traffic;
- Slopes - preferably on flat surface, but with a slope of no more than 1:50 over the length of the attenuator;
- Design speed;
- Kerb and roadway elevation; preferably no kerb within 16 metres of attenuator;
- Probable angle of impact;
- Base type and base features;
- Site features - are there any unique site features;
- Orientation - an attenuator should be oriented to maximize likelihood of head-on impact, though a maximum angle of up to 10 degrees between roadway centre line and attenuation device is acceptable; and
- Placement area.

Table 8.7 : Impact attenuator and end terminal application

Appropriate Device Type	Distance		Probable Impact Angle
	Hazard to traffic (m)	Opposing traffic to hazard (m)	
Re-directive			
Bi-directional	<3	<9	$>5^{\circ}$
Uni-directional	<3	>9	$>5^{\circ}$
Non-redirective	>3	>9	$<5^{\circ}$

8.5.2 Design/selection of impact attenuators

The detail design of impact attenuators should be done in conjunction with the manufacturer of a specific attenuator and will be dependent on the actual attenuator chosen for installation.
The following factors should be considered for

8.5.3 Functional considerations

Attenuators as well as barrier end-treatments can be installed as bi-directional or uni-directional as well as with redirective or non-redirective capabilities. As a general rule the chosen system should be able to redirect an errant vehicle if the hazard being shielded is less than 3 m

Table 8.8 : Space requirements for plastic drum attenuators						
Design Speed (km/h)	Space requirements (m)					
	Minimum			Preferred		
	N*	L*	F*	N*	L*	F*
70	1.8	4.3	0.6	3.6	8.5	1.2
80	1.8	5.2	0.6	3.6	10.0	1.2
90	1.8	6.2	0.6	3.6	12.0	1.2
100	1.8	7.2	0.6	3.6	14.0	1.2
110	1.8	8.3	0.6	3.6	16.0	1.2

* N, F and L are defined in Figure 8.11 below

from the edge of the travelled way. Typical functional considerations for attenuators and barrier end terminals are given in Table 8.7.

The designer should allow for enough space to install an attenuator in the most effective way and to ensure that its performance will not be compromised by insufficient placement areas. The particular system's requirements in terms of installation should also be met.

Figure 8.11 and Table 8.8 show the space to be reserved for sand-filled plastic drum attenuators under different design speed conditions.

8.5.4 Sand-filled plastic barrel impact attenuators

Sand-filled plastic barrel impact attenuators work on the principle of conservation of momentum and therefore do not need a backdrop in front of the hazard being shielded.

Recommendation pertaining to these devices are as follows:

- Single rows of barrels should not be allowed for permanent installation;
- Barrels should be spaced some 150 mm apart and stop 300 mm to 600 mm short of the hazard being shielded;
- Barrels should be positioned in such a way that rigid corners of the hazard are overlapped by barrels by some 760 mm (300 mm minimum) to reduce the severity of angled impacts near to the rear of the attenuator. Where such attenuators are subject to bi-directional traffic flow, the array of barrels should be flush with the edge of the hazard so as to ensure that reverse direction traffic does not inadvertently impact the rear end of the barrel arrangement.

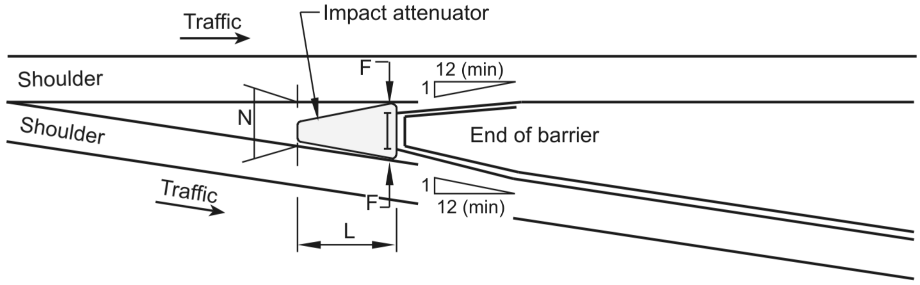


Figure 8.11: Space requirement for plastic drum attenuators

- If speeds higher than 95 km/h are anticipated, barriers can be lengthened. Since most serious accidents occur at excessive speeds, an "over-design" is acceptable where space permits.

A number of typical arrangements of sand-filled barrel attenuators are shown in Figure 8.12. The legend illustrates the mass of sand contained in each barrel.

Wall up to 0,9 m wide
Attenuator: length 7,3 m,
width 1,8 m
12 modules rated for 90 km/h

Wall up to 1,8 m wide
Attenuator: length 8,2 m,
width 2,7 m
17 modules rated for 100 km/h

Wide wall or bridge rail
Up to 3,7 m wide
Attenuator: length 7,3 m
width 4,6 m
32 modules rated for 90 km/h

Gore with overpass piers and guardrail
Attenuator: length 7,3 m,
width 2,7 m
16 modules rated for 90 km/h

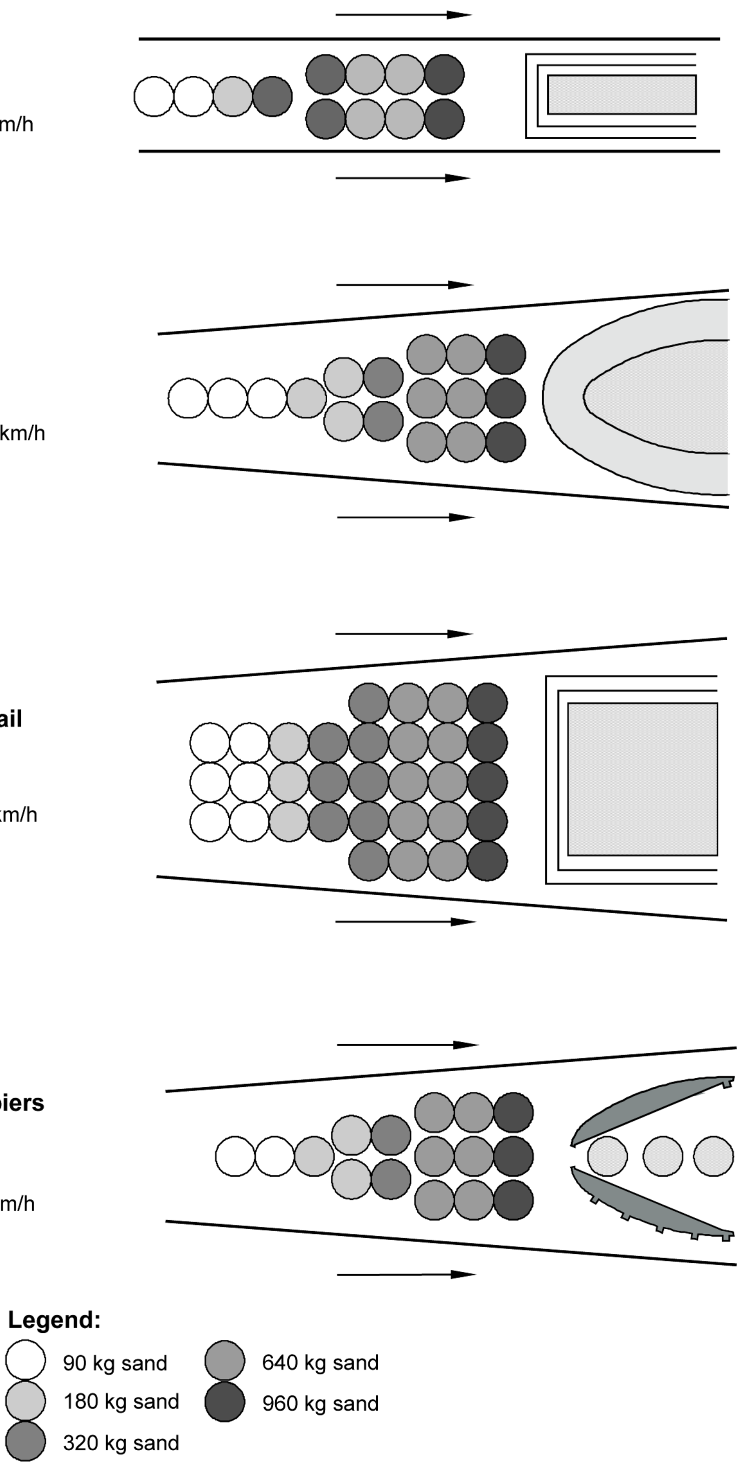


Figure 8.12: Typical arrangement of sand-filled barrel attenuators

8.6 RUNAWAY VEHICLE FACILITIES

8.6.1 Introduction

Runaway-vehicle escape ramps and associated arrestor beds are specifically designed to reduce the safety hazard associated with out-of-control heavy vehicles where long steep grades occur.

The following factors are generally associated with such incidents:

- Gradient;
- Driver error such as failure to downshift gears;
- Equipment failure such as defective brakes;
- Inexperience with the vehicle;
- Unfamiliarity with the specific location;
- Driver impairment due to fatigue or alcohol; and
- Inadequate signing of downgrade.

The purpose of escape ramps and arrestor beds is to stop, without serious injury or serious damage to vehicles, to adjacent property or to other road users, those vehicles that become out-of-control on long downhill gradients. Deceleration rates of between 5 m/s^2 and 6 m/s^2 are obtained by a full width level arrestor bed without the use of vehicle brakes. (A 10 per cent down gradient on the bed surface can reduce the deceleration by about 1 m/s^2). It should be noted that, under high deceleration, inadequately restrained vehicle occupants, or insecurely attached cargo may not be safely contained.

8.6.2 Types of escape ramps

There are six different types of truck escape ramp:

- Sand pile;

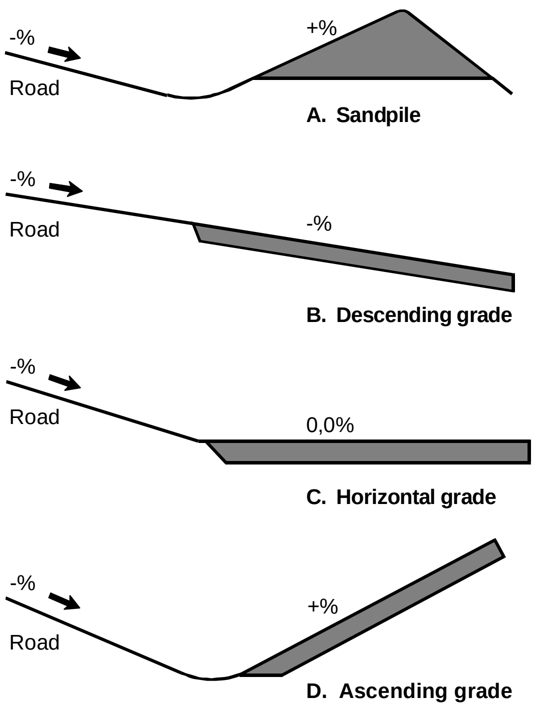


Figure 8.13: Typical arrestor beds

- Ascending grade arrestor bed;
- Horizontal grade arrestor bed;
- Descending grade arrestor bed; and
- Roadside arrestor bed.

Figure 8.13 illustrates four types in general use.

The sand pile types are composed of loose, dry sand and are usually no more than 130 metres in length. The influence of gravity is dependent on the slope of the surface of the sand pile. The increase in rolling resistance to reduce overall lengths is supplied by the loose sand. The deceleration characteristics of the sand pile are severe and the sand can be affected by weather. Because of these characteristics, sand piles are less desirable than arrestor beds. They may be suitable where space is limited and the compacting properties of the sand pile are appropriate.

Descending grade ramps are constructed parallel and adjacent to the through lanes of the highway. They require the use of loose aggregate in an arrestor bed to increase rolling resistance and thus slow the vehicle. The descending-grade ramps can be rather lengthy because the gravitational effect is not acting to help reduce the speed of the vehicle.

In a horizontal-grade ramp, the effect of the force of gravity is zero and the increase in rolling resistance has to be supplied by an arrestor bed composed of loose aggregate. This type of ramp will be longer than those using gravitational forces to help stop the vehicle.

An ascending-grade ramp uses both the arresting bed and the effect of gravity, in general reducing the length of ramp necessary to stop the vehicle. The loose material in the arresting bed increases the rolling resistance, as in the other types of ramps, while the force of gravity acts in the opposite direction to the movement of the vehicle. The loose bedding material also serves to hold the vehicle in place on the ramp grade after it has come to a safe stop. Ascending grade ramps without an arrestor bed are not encouraged in areas with moderate to high commercial vehicle usage as heavy vehicles may roll back and jack-knife upon coming to rest.

Each one of the ramp types is applicable to a particular situation where an emergency escape ramp is desirable and should be compatible with the location and topography. The most effective escape ramp is an ascending ramp with an arrestor bed. On low volume roads with less than approximately 1000 vehicles per day, clear run-off areas without arrestor beds are acceptable.

8.6.3 Criteria for provision of escape ramps

On hills where there is a history of accidents involving runaway vehicles, consideration should always be given to the provision of escape ramps with arrestor beds. A measure of effectiveness can be assessed by an analysis of personal injury and "damage only" accidents plus the incidence of damage to property, based on local records.

On new construction, where long steep gradients are unavoidable, and where the probability of damage caused by runaway vehicles is greater than normal, the provision of arrestor beds should be considered as an integral part of the project design. As a guide for provision, where gradients are over 5 per cent, an arrestor bed should be considered if the gradient (in per cent) squared, multiplied by the approach length from the crest, in kilometres, is over 60.

On long, straight stretches of down grade where a sufficiently steep or long up-grade occurs before any bend is met, observations of heavy vehicle driver choice indicate that arrestor beds are unlikely to be used.

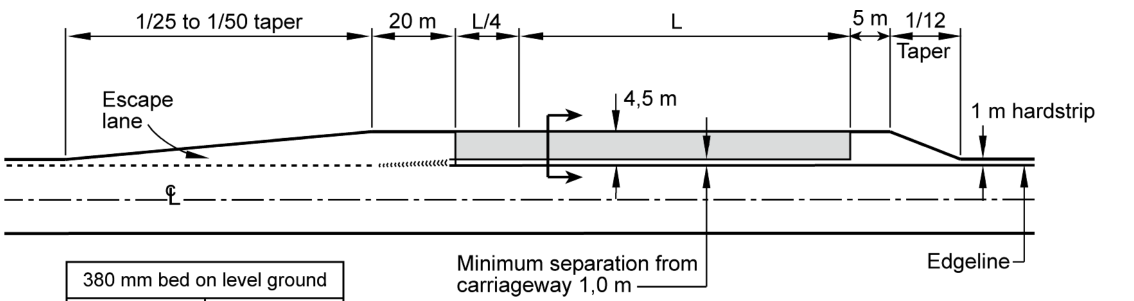
8.6.4 Location of runaway-vehicle facilities

On both new and existing roads, engineering judgement is required to determine the location of arrestor beds. Relevant factors to be considered include:

- The location of previous accidents;
- The length of downgrade;
- The conditions at the bottom of the grade;
- The percentage of heavy vehicles;
- Horizontal alignment;
- Topography (i.e. effect on cost of earth works); and
- Toll plazas at the bottom of steep grades

Runaway-vehicle facilities should not be constructed where an out-of-control vehicle would need to cross oncoming traffic.

On divided carriageways, safety ramps may also be located in the median if sufficient space is available. This would be in conflict with driver expectations and prominent advance warning signs prior to the safety ramp exit would have to be provided. Because of the conflict with driver expectations, providing arrestor beds in the



380 mm bed on level ground	
Expected max. entry speed km/hr	Suggested length L (metres)
50	23
60	32
70	44
85	66
100	90
120	130
130	155
	Bed length to be 25% greater

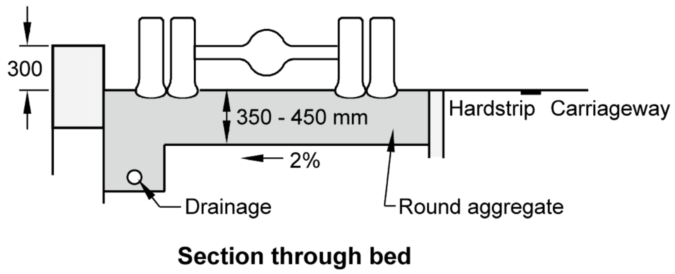


Figure 8.14: Layout of arrestor bed adjacent to carriageway

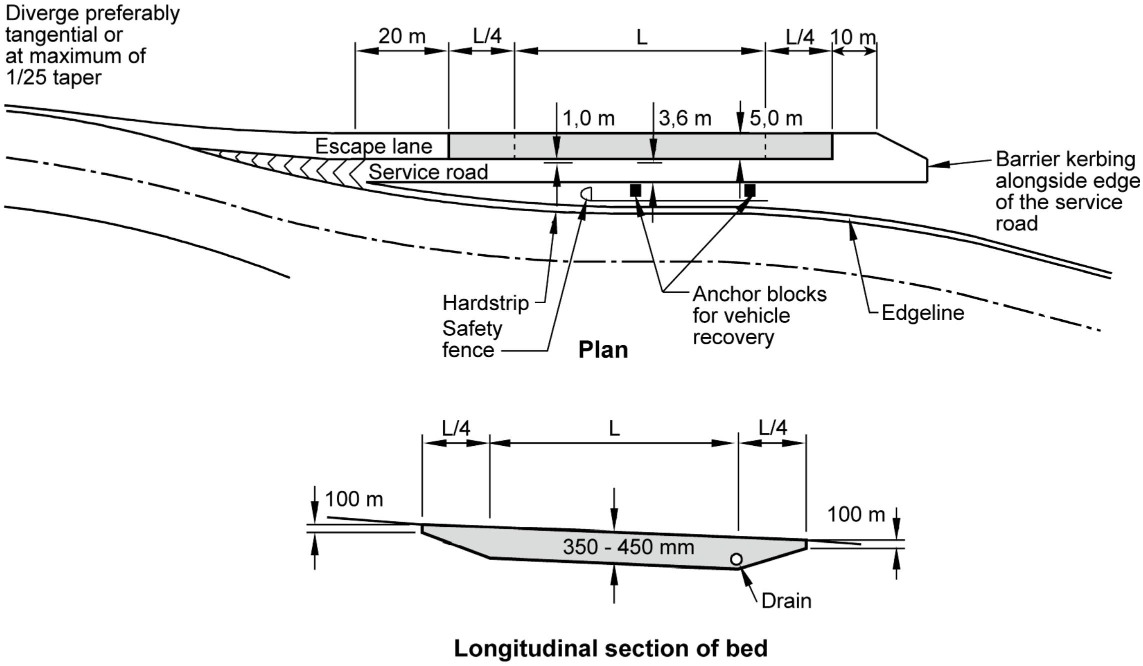


Figure 8.15: Layout of arrestor bed remote from carriageway

median is not a recommended practice but, in the case of a sharp curve to the left at the bottom of a steep downgrade this location may prove to be necessary.

For safety ramps to be effective, their location is critical. They should be located prior to, or at the start of, the smaller radius curves along the alignment. For example, an escape ramp after the tightest curve will be of little benefit if trucks are unable to negotiate the curves leading up to it. As vehicle brake temperature is a function of the length of the grade, escape ramps are generally located within the bottom half of the steeper section of the alignment.

Where possible, arrestor beds should not be located in verges on the outside of right-hand curves, or at any other location where there is a likelihood of vehicles mistakenly entering the bed during the hours of darkness. Where arrestor beds are constructed in built-up areas, local schools should be informed that these are dangerous locations so that children are discouraged from playing in them.

Lack of suitable sites for the installation of ascending type ramps may necessitate the installation of horizontal or descending arrestor beds. Suitable sites for horizontal or descending arrestor beds can also be limited, particularly if the downward direction is on the outside or fill side of the roadway formation. The entrance to the bed should be clearly signed for drivers of runaway vehicles and clearway restrictions should be applied so that the entrance is kept freely accessible. Adequate visibility to the entrance must be available so as to give drivers time to manoeuvre their vehicles into the bed and, as a minimum requirement, visibility equal to the desirable stopping sight distance appropriate to the anticipated maximum speed of runaway vehicles should be provided.

8.6.5 Arrestor bed design features

Suggested layouts for an arrestor bed immediately adjacent to the carriageway, are shown in Figure 8.14 and, in Figure 8.15, for a bed separated from the road. Where space permits, the arrangement separated from the road should be used.

Length of bed

The length of bed required to halt runaway vehicles is dependent on the predominant vehicle type, the likely speed of entry into the bed, the type and depth of aggregate used and the slope of the arrestor bed. The bed length for all-purpose roads should cope with the critical design vehicle, generally a large articulated vehicle with multiple axle groups. This vehicle is likely to have the highest entry velocity and the lowest average deceleration rate.

Table 8.9 gives suggested lengths for horizontal grade arrestor beds (excluding the initial depth transition zone). Where the bed surface is aligned on a downgrade its length should be increased by 3 per cent for each one degree of slope.

Width of bed

As a general guide a constant bed width of 4,0 - 5,0 metres is adequate. Barrier kerbing with a 300 mm up stand should be installed at the side of the bed remote from the carriageway to assist in restricting sideways movement. The use of safety fencing may also be desirable. The bed should be separated from the main carriageway, where possible, by at least 2,0 metres, and flush kerbing may be required locally where the road is provided with 1,0 metre wide hard strips.

Depth of bed material

Beds should have depths between 300 mm and 450 mm with the depth gradually increasing over an initial length in order to provide for smooth vehicle entry. Where entry velocities are less than 75 km/h, vehicle deceleration is significantly higher for beds which contain greater depths of bed material, whereas, at speeds above 75 km/h, decelerations tend to be

Table 8.9: Length of Arrestor Bed (Over Uniform Gravel Depth)

Expected maximum entry speed (km/hr)	Suggested Length (L*) (m)
70 and less	45
80	60
100	90
120	130
130	155

* L as shown on Figure 8.13 (bed length to be 25 per cent greater)

independent of bed depth. The greater (450 mm) depth gives around 50 per cent greater stopping ability than the minimum (350 mm) provision and should be considered where bed length is restricted.

Type of material

To achieve a high deceleration rate it is necessary that vehicle tyres sink into the bed material. Rounded uncrushed gravel and single size cubic aggregate or similar artificial lightweight aggregate has performed satisfactorily in tests and should be used in preference to angular gravel (i.e. crushed rock) or sand, which tend to restrict wheel penetration and compact with time and usage.

Arrestor bed material should be free draining and adequate drainage should be provided so that in freezing or saturated conditions it still retains its function of wheel penetration, thereby bringing vehicles to a standstill. A suitable specification for the bed material is given below as a guide. This may need to be modified to allow locally available suitable materials to be used, in the light of further experience and testing.

Typical specification for arrestor bed material

The material should be clean, uncrushed, hard durable natural gravel consisting primarily of smooth round particles. Alternatively, an appropriate artificial lightweight aggregate may be used. The following particle size distribution is suitable:

BS Sieve Size	Percentage by mass passing
10 mm	100%
5 mm	0%

8.7 BRAKE CHECK AND BRAKE REST AREAS

A brake check area or a compulsory truck stop is an area set aside before the steep descent as distinct from a brake rest area which is an area set aside for commercial vehicles part way down or at the bottom of the descent.

These facilities should be provided on routes that have long steep downgrades and commercial vehicle numbers of around 500 per day, especially on National Roads and principal traffic routes. These areas, when used, will ensure that drivers begin the descent at zero speed and in a low gear that may make the difference between controlled and out-of-control operation on the downgrade. They would also provide an opportunity to display information about the grade ahead, escape ramp locations and maximum safe recommended descent speeds.

These areas may need to be large enough to store several semi-trailers, the actual numbers depending on volume and predicted arrival rate.

Their location would need good visibility with acceleration and deceleration tapers provided. Adequate signage should be provided to advise drivers in advance of the facilities. Special signs, specific to the site, may need to be designed for these areas.

Back to Menu

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 9

ROAD BETTERMENT

9.1 Introduction

3R projects are essentially aimed at extending the life of a road through maintenance of the pavement at the levels of:

- Resurfacing, which involves the application of a new wearing course to the existing surface, with this wearing course being anything from a fog spray to an asphalt overlay;
- Restoration, typically consisting of patching of potholes and repair of failed sections that may extend for several metres;
- Rehabilitation, where the entire base course layer - and perhaps the lower layers as well - are scarified, reshaped and recompacted.

Reference is sometimes made to 4R, with the fourth R referring to reconstruction.

In the case of rehabilitation or reconstruction projects, the possibility of modifying the geometry of the road where it is currently sub-standard also merits consideration. Unfortunately, many opportunities for low-cost safety improvements are lost because of the overriding focus on pavement improvements. Safety needs are often not addressed until little time remains to accommodate the design time for geometric improvements and the available funding is allocated totally to the pavement. Furthermore, there is a tendency to focus on widening of lanes or shoulders rather than on the reconstruction of sharp curves or replacement of narrow culverts and bridges. Very often, improvements at these locations may be more safety cost-effective than routine cross-section improvements.

Modification may simply entail a modest upgrade of the cross-section over the length of the section to be rehabilitated. It may also include changes to selected portions of the horizontal or vertical alignments.

9.1.1 Balanced context-sensitive design of 3R projects

The objective of geometric modification is to enhance the safety of the road. For this reason, a decision to simply increase the widths of the lanes or shoulders or both may not be appropriate. Realignment of a markedly substandard horizontal curve may have an impact on safety that is greater than that of merely adding 300 mm to the width of the lanes. The designer should thus study the entire length of the road section in depth and develop a package of improvements that will have the maximum impact on safety in terms of the funds available.

A road with a design speed of 120 km/h and a cross-section comprising 3,7 metre lanes and 2,5 metre shoulders may conceivably be the ideal in a given situation. It may also be unaffordable. The road authority then has to consider the relative merits of improving safety over only a short section of the road, leaving the rest as is, or of making more modest safety improvements over the whole length of the road. The latter is generally the preferred option in terms of consistency of design and matching driver expectations because the high-standard road section may encourage speeds that are inappropriate to the balance of the road.

An important feature of geometric improvements carried out as part of a 3R project is that, after upgrading, there should still be balance between the various elements of the road. A high-standard cross-section generally creates the impression in the driver's mind of a high design speed, which may be at variance with the reality of the situation as defined by the horizontal alignment.

9.1.2 The myth of the sub-standard road

In reality, there is no such thing as a sub-standard road. Two possible situations can, however, give rise to the belief that a road is not to standard. These are:

- A mismatch between the various dimensions, for example a low design speed vertical alignment superimposed on a high design speed horizontal alignment. Similarly, a narrow cross-section applied to a road that, in all other respects, supports high-speed travel.
- A low design speed of the road in relation to the environment being traversed.

With regard to the first point, if there is no mismatch, i.e. if the various dimensions of the road are in balance with one another, a design speed can be established that will match the geometry of the road. The second point suggests that this design speed may be inappropriate to the circumstances. An example would be where property boundaries impose an alignment comprising short tangents and low-radius curves on a road traversing an essentially flat landscape.

Where, for either of the reasons listed above, a road is described as being sub-standard, it may be argued that any improvement is better than none even if the result is still modestly sub-standard. The counter argument could be that drivers will proceed cautiously on a road which is "obviously dangerous" and the accident rate will be correspondingly low. If the design speed of the road is, however, improved so as to be closer to the speed drivers wish to select - but without actually achieving it - they may be unaware of the shortfall and thus tend to overdrive the alignment with a consequent increase in the accident rate. The designer should be sensitive to the merits of both arguments and, within the flexibility offered by the adoption of guidelines as opposed to rigid standards, achieve the best possible compromise within the context of the financial and physical restraints attached to the specific project.

It is suggested that, as a general rule, improvements to the alignment and cross-section of a road be in balance to a single design speed and that, if this speed is low for the circumstances, advisory speeds be posted.

9.1.3 Possible methodologies for determination of road improvements

The most obvious reason for improving a road is a mismatch in dimensions. This is because a design speed that is too low may require full reconstruction at almost every horizontal and vertical curve. This may be outside the budget typically available for betterment works. The designer should thus establish precisely where dimensional mismatches occur along the road and the extent of the mismatch.

The mismatch referred to above is that between the horizontal and the vertical alignments and the cross-section. A further mismatch is described in detail in Section 4.2.2, being differences between successive elements of the horizontal alignment. These differences are between the design and operating speeds on individual curves and between operating speeds on successive curves. Design speeds and operating speeds along the route should thus be calculated and compared.

In terms of road safety, differences between design and operating speeds along the road are possibly more likely to lead to crashes than would mismatches between road dimensions. The designer should, therefore, in the first instance, tabulate the design and operating speeds along the road, highlighting differences greater than those allowable in terms of Section 4.2.2. The magnitude of these differences offers a yardstick for prioritising the required improvements. In addition, the tabulation would suggest a range of curve radii that could be considered at each point along the road.

After evaluation of the horizontal alignment, the vertical alignment can be checked, specifically for the K-values of the crest curves, as these dictate the availability of stopping sight distance. This design speed should be equal to, or preferably greater than, the design speed of the horizontal alignment. The speed of passenger cars is dictated by the horizontal rather than the vertical alignment so that a low design speed on a crest curve would tend to be overdriven, i.e. drivers would have less stopping sight distance than they need for safety. The speed of trucks, on the other hand, is dictated largely by gradients and it is unlikely that they would exceed the design speed of a crest vertical curve.

The coordination of the vertical and horizontal alignments as described in Chapter 5 should also be checked, although financial constraints may make it difficult to devote funding purely to improvement of the aesthetics of the road. However, if it is necessary to rebuild a horizontal or a vertical curve for whatever reason, the aesthetics of the situation should be considered at the same time.

Having achieved an acceptable level of internal coherence of the horizontal alignment and an appropriate design speed for the vertical alignment, the designer can then consider the cross-section.

With time, "desirable" lane widths have increased from 2,7 metres to 3, 7 metres (9 feet to 12 feet) and shoulder widths from 0,3 metres to 3,0 metres, with the latter including a 0,5 metre allowance for rounding. The upper limit of design speeds has increased from 100 km/h to 130 km/h. There is thus a distinct likelihood that a road that is at or beyond the end of its design life may have a horizontal and vertical alignment matching an acceptable design speed, but still be narrower than is considered desirable. Under these circumstances, improving the cross-section over the full extent of the betterment project should be considered.

The principle of balanced design should not however be ignored. Even if the funds for improvements were unlimited, it would not be wise to develop a cross-section appropriate to a high design speed when the alignment is, in fact, appropriate to a lower design speed. As stated above, a mismatch between the dimensions of the road constitutes poor design.

When a table defining and prioritising all the geometric improvements that should be made as part of the 3R project has been prepared, the final step in the process would be to locate these improvements relative to the identified pavement improvements. It is likely that the best return on investment would be achieved if the investment were located at a point where both the pavement and the geometry have to be improved.

9.2 COST-EFFECTIVENESS OF GEOMETRIC IMPROVEMENTS

The construction costs involved in geometric improvements cannot be quantified on a countrywide basis. The availability of materials, the remoteness of sites, the presence of site-specific complications and the variation in skills levels all contribute to variations in cost that would tend to make a national average a very rough guide at best. Costing should thus be on a project-by-project basis.

Central to any system whereby geometric improvements can be prioritised, is knowledge of the value of the benefits that are likely to accrue from any improvement. Benefits take the form of:

- reduction in the number and/or severity of crashes; and
- reduction in travel time across the improved road section.

The cost of crashes in 2001 Rands is listed in Table 9.1. For convenience, a weighted average of R 30 k has been adopted.

Crash prediction models are used in conjunction with average crash costs to derive the annual savings derived from a given geometric improvement over the design life of the improvement. These models are normally structured to predict the number of crashes that are likely to occur given a number of preconditions such as lane width, shoulder width and radius of curvature.

Where road improvements are contemplated, the focus is on the marginal reduction in crashes that is likely to be achieved by a deliberate change in these preconditions. The lanes or the shoulders may be widened or the radius of a

Table 9.1: Cost of crashes by severity

Severity	Cost (R)
Fatal	470 634
Serious injury	108 202
Slight injury	28 223
Property damage only	18 525
Weighted average	29 642

horizontal curve increased. A vertical curve may be flattened or an intersection relocated. It is difficult to assign a reduction in crash rate to any one specific geometric element and even more difficult to evaluate the consequences of more than one variation in the road geometry. For example, it is reasonable to assume that an increase in lane width would result in a reduction in the crash rate and an increase in shoulder width would also reduce the crash rate. The question then arises whether the reduction in crashes because of a combination of lane and shoulder width increases would equal the sum of the reductions caused by the individual increases. Similarly, an unchanged width between shoulder breakpoints would result in a reduction in shoulder width if the lane width were increased. Would there then be a net increase or a net decrease in the crash rate? In the sections that follow, these matters are discussed in more depth.

9.3 CRASH PREDICTION MODELS

In North America and Europe, various road agencies have developed and maintain crash recording systems to monitor the safety of the road networks for which they are responsible. Useful though these are as a point of departure, they provide historical or retrospective information whereas effective management requires a prospective viewpoint. It is necessary to know what the current safety performance of a road is and what it is likely to become if certain remedial actions are undertaken.

Estimates of the future safety performance of a road were based on one of four approaches:

- Averages from historical accident data;
- Predictions from statistical models based on regression analysis;
- Results of before-and-after studies; or
- Expert judgments by experienced engi neers.

When used alone, each of these approaches has weaknesses.

9.3.1 Averages from historical data

South Africa does not have useful crash databases. In countries where these do exist, it has been found that historical data are highly variable. Consequently, estimates of the expected accident rate based on a short-duration data set are likely to be unreliable.

A form of historical data that has been and still is used in South Africa refers to "red spot location". A location with a high frequency of crashes is identified on the basis of its experiencing more than a specified number of crashes of varying severity within a period typically of one to three years. However, statistical theory suggests that, because of the random nature of crashes, a location with a high short-term crash rate is likely to experience fewer accidents in the future. This phenomenon, known as regression to the mean, has been verified experimentally through the conducting of before-and-after studies with *no intervening remedial action being undertaken.* The after studies often showed an improvement in the safety of the location.

Historical crash data can thus lead to an incorrect - specifically a pesimistic - interpretation of the situation.

9.3.2 Regression analysis

Statistical models are developed on the basis of a database of crash and roadway characteristics. An appropriate functional form is selected for the model and its parameters are then determined by means of multiple regression analysis. Regression models are accurate in their prediction of the total crash experience for a location or class of locations but are less effective in isolating the effects of individual geometric or control features.

Regression models assume a statistical correlation between roadway features and crashes, although these correlations could be spurious or not representative of a cause-and-effect relationship. For example, the fact that crash rates tend to decrease with time could suggest that the passage of time is sufficient to cause the reduction. Furthermore, a strong correlation between independent variables would make it impossible to isolate their individual effects.

9.3.3 Before-and-after studies

The principal weakness of before-and-after studies has already been discussed. Because of regression to the mean, the user of such a study cannot be sure that it represents the true effect of the improvement on the safety of the crash location being studied. These studies tend to provide an over-optimistic interpretation of the value of the improvement.

However, safety experts are of the opinion that, if the bias resulting from regression to the mean could be eliminated, the before-and-after study would be a powerful tool in the assessment of the safety effects of geometric and traffic control features.

9.3.4 Expert judgment

Expert judgment developed after many years of experience in the highway safety field plays an important role in the development of reliable safety estimates. Although experts have difficulty in providing quantitative i.e. absolute estimates, they are very capable of developing relative or judgments of the form: A is likely to be more or less than B or C would not be more than 20 per cent of D. They thus need a frame of reference based on historical data, statistical models or before-and-after studies to make useful judgments.

9.4 A NEW APPROACH

Having discussed the weaknesses of the four basic forms of assessment of safety, it is suggested that combining all four into a single form of assessment would lead to a more reliable estimate of safety than could be achieved if each one were used individually.

The Interactive Highway Safety Design Model (IHSDM), briefly introduced in Chapter 2 of this Guideline, includes a crash prediction module based on a combination of all four forms of assessment. The algorithms in this module have been developed in the United States for rural two-lane highways. Separate algorithms have been developed for roadway segments and at-grade intersections. The two algorithms can be used together to predict crash experience for an entire highway section or improvement project. The roadway segment algorithm predicts all non-intersection-related crashes. For example, a ran-off-road crash within 15 metres of an intersection could be considered by the investigating officer to be unrelated to the intersection. For modeling purposes, crashes occurring within 76 metres (250 feet) of an intersection because of the presence of the intersection are considered to be intersection-related crashes.

Each of these algorithms is composed of two components, being the base model and crash modification factors (CMF). They take the form;

$$N_{rs} = N_{br}\left(CMF_{1r}\ CMF_{2r} \cdots CMF_{nr}\right) \quad 9.1$$

where N_{rs} = predicted number of total road way segment crashes per year after application of accident modification factors

N_{br}= predicted number of total road way segment crashes per year for nominal or base condition

CMF= crash modification factors for roadway segments

The effect of average daily traffic (ADT) on predicted crash frequency is incorporated as part of the base condition and the effects of geometric design and traffic control measures are incorporated through the CMFs.

South African fatality and crash rates are higher than those of the United States for which the relationships were derived. The differences between South African and American road safety indicators are illustrated in Figures 9.1, 9.2 and 9.3.

Although Figures 9.1 and 9.2 display an encouraging trend in the South African rates, the ratio between South African and American crash rates for all types of crashes has stayed fairly constant at a factor of 2,5 over the period shown in the graphs.

In the American application of IHSDM, allowance is made for a calibration procedure to modify the national relationship to the accident history of the individual States or local areas

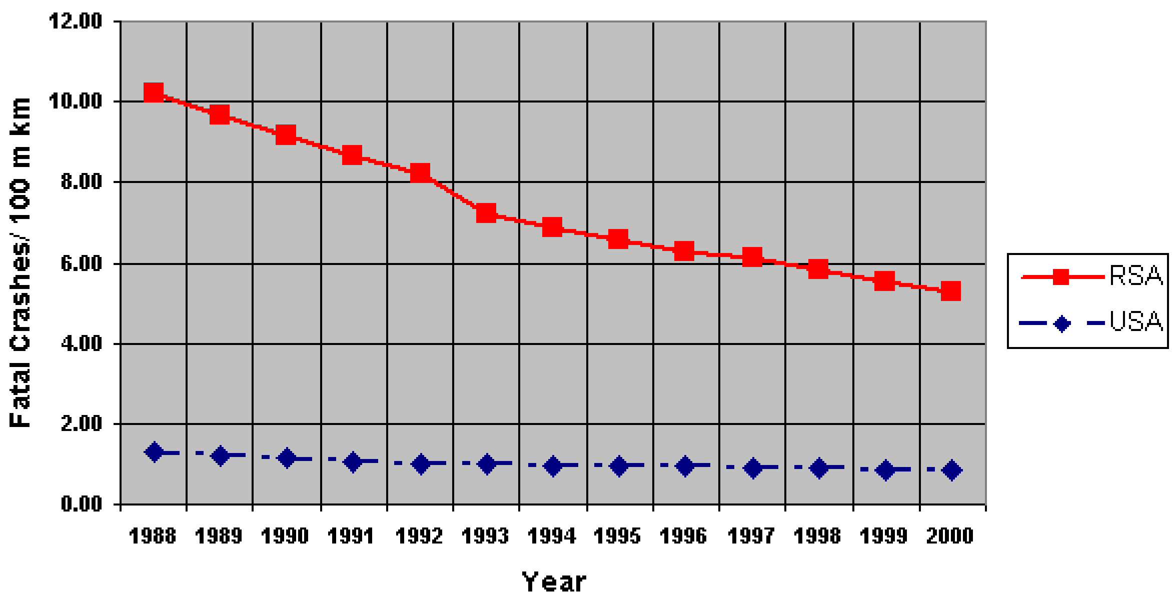


Figure 9.1: Fatal Crash rates

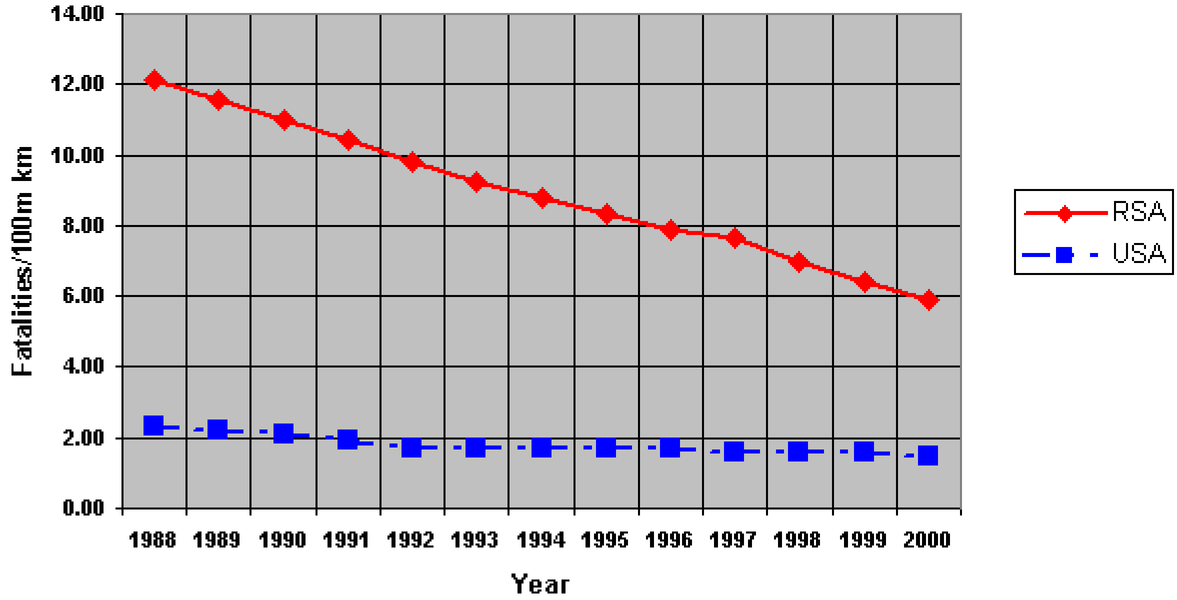


Figure 9.2: Fatility rates

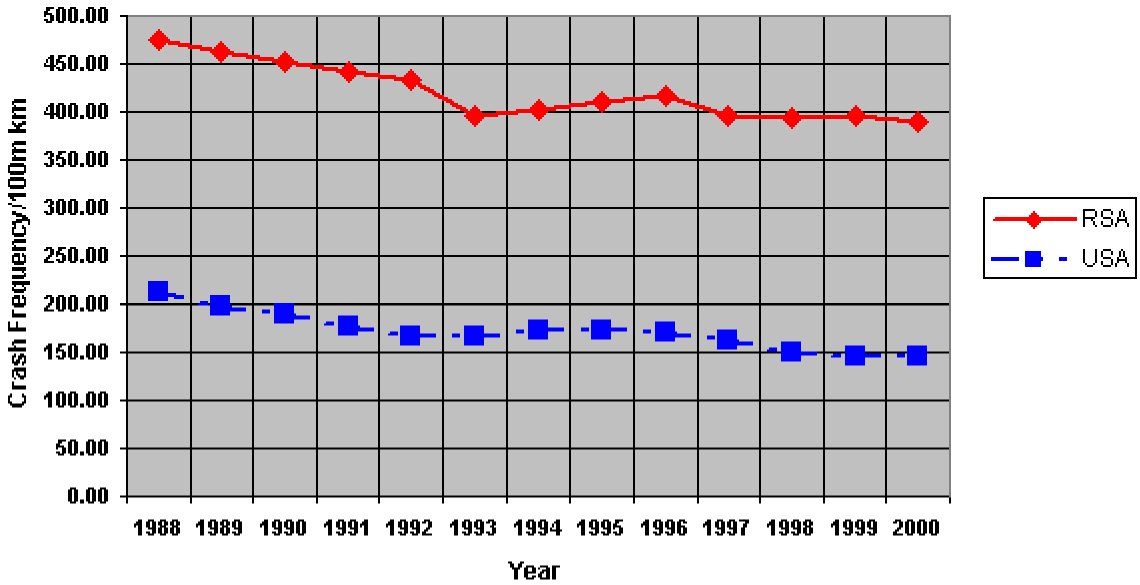


Figure 9.3: Crash rates for all types of crashes

within each State. A calibration factor of 2,5 should, on the basis of Figure 9.3, be applied to Equation 9.1 to match South African conditions.

The individual States of the United States generally have crash and road information superior to that available in South Africa so that their calibration procedure can be more precise than that proposed above. The calibration procedure involves developing an inventory of the road network stratified in terms of ADT in order to extract:

- Number of kilometres of tangent road way;
- Number of kilometres of roadway on horizontal curves;
- Average degree of curvature for hori zontal curves;
- Number of kilometres of level roadway;

- Number of kilometres of roadway on grade; and
- Average percentage of gradient for roadway on grade.

Values for the various geometric parameters, i.e.:

- Average lane and shoulder width;
- Shoulder type (paved, gravel or turf);
- Driveway density (number per kilometre); and
- Average roadside hazard rating
- For horizontal curves;
  - o No spiral transition present;
  - o Super elevation not deficient

also have to be derived.

The above information is input into the crash prediction module and the predicted number of crashes calculated. The actual number of crashes recorded in the database, divided by the predicted number of crashes, provides the calibration factor.

9.4.1 Base model for roadway segments

The predicted number of crashes for the base condition is given as

The roadside hazard rating is that devised by Zegeer et al and has the following structure:

Rating = **1**

- Wide clear slopes not less than 9 metres from the pavement edgeline;
- Side slopes flatter than 1:4;
- Recoverable.

Rating = **2**

- Clear zone for 6 to 7,5 metres from pavement edgeline;
- Side slope ofabout 1:4;
- Recoverable.

Rating = **3**

- Clear zone for 3 metres from pavement edgeline;
- Side slope of about 1: 3 to 1:4;
- Marginally recoverable.

Rating = **4**

- Clear zone for 1,5 to 3 metres from pavement edgeline;
- Side slope of about 1:3 to 1:4;
- May have guardrail about 1,5 to 2 metres from pavement edgeline;
- May have exposed trees, poles or other objects about 3 metres from pavement edgeline;
- Marginally forgiving, but increased chance of a reportable roadside crash.

$$N_{br} = 4.EXPO.e^{0,6409-0,277LW-0,194SW+0,0668RHR+0.005DD}.W \qquad 9.2$$

where Nbr = Predicted number of crashes for base condition

EXPO = Exposure (10^6 veh km/year)

= ADT x 365 x L/10^6

LW = Lane width (metres)

SW = Shoulder width (metres)

RHR = Roadside hazard rating (integer value between 1 and 7)

DD = Driveway density (number /km)

W = Weighting factor for specific road segment

Rating = 5

- Clear zone of 1,5 to 3 metres from pavement edgeline;
- Side slope of about 1:3;
- May have guardrail 0 to 1,5 metres from pavement edgeline;
- May have rigid obstacles or embankment within 2 to 3 metres of pavement edgeline;
- Virtually non-recoverable.

Rating = 6

- Clear zone of 1,5 metres or less from pavement edgeline;
- Side slope of about 1:2;
- No guardrail;
- Exposed rigid obstacles within 0 to 2 metres of the pavement edgeline;
- Non-recoverable.

Rating = 7

- Clear zone of 1,5 metres or less from pavement edgeline;
- Side slope of 1:2 or steeper;
- Cliff or vertical rock cut;
- No guardrail;
- Non-recoverable with high likelihood of severe injuries from roadside collision.

9.4.2 Base model for intersections

Base models have been developed for:

- Three-legged STOP-controlled inter sections;
- Four-legged STOP-controlled intersec tions; and
- Four-legged signalised intersections.

These predict crash frequency per year for intersection-related crashes that occurred within 76 metres of an intersection. They are limited to intersections of two-lane two-way roads without auxiliary lanes at the intersections.

For three-legged intersections:

$$N_{bi} = \exp(-10{,}79 + 0{,}79 \ln ADT_1 + 0{,}49 \ln ADT_2) \qquad 9.4$$

For four-legged intersections:

$$N_{bi} = \exp(-9{,}34 + 0{,}60 \ln ADT_1 + 0{,}61 \ln ADT_2) \qquad 9.5$$

For four-legged signalised intersections:

$$N_{bi} = \exp(-5{,}73 + 0{,}60 \ln ADT_1 + 0{,}20 \ln ADT_2) \qquad 9.6$$

Where ADT_1 and ADT_2 refer to the average daily traffic on the major and minor roads respectively and N_{bi} is the predicted number of

The weighting applied in Equation 9.2 is calculated as

$$W = \sum W_{H_i} e^{257{,}8/R_i} . \sum W_{V_j} e^{0{,}465 A_j} . \sum W_{G_k} e^{0{,}105 G_k} \qquad 9.3$$

where
W_{Hi} = weight factor for the i^{th} horizontal curve
= proportion of segment length contained within Curve i
R_i = radius of Curve i (metres)
W_{Vj} = weight factor for the j^{th} vertical curve
= proportion of segment length contained within Curve j
A_j = algebraic difference in gradient across the j^{th} vertical curve (per cent)
W_{Gk} = weight factor for k^{th} grade
= proportion of segment length contained in k^{th} grade
G_k = absolute value of gradient for the k^{th} grade (per cent).

crashes at intersections for the nominal or base condition.

9.4.3 Crash modification factors

- o Intersection sight distance

The CMF's in respect of lane widths and shoulder types and widths are shown in Tables 9.2 and 9.3.

Table 9.2: Crash modification factors for various lane widths				
ADT (veh/day)	Crash Modification Factor for Lane widths of			
	2,7 m	3,0 m	3,3 m	3,7 m
1	1,02	1,01	1,00	1,0
500	1,02	1,01	1,00	1,0
1000	1,07	1,04	1,01	1,0
1500	1,12	1,07	1,01	1,0
2000	1,18	1,11	1,02	1,0
2500	1,18	1,11	1,02	1,0

As shown in Equation 9.1, the Crash Modification Factors (CMF) are multipliers applied to the base condition. The CMFs developed for the American IHSDM model were based on a panel's best judgment of the relative merits of available research findings with the credibility of the model being supported by a sensitivity analysis. The CMFs incorporated in the model include:

- Roadway segments:
  - o Lane width;
  - o Shoulder type and width;
- Horizontal curves;
  - o Length;
  - o Radius;
  - o Presence or absence of transition curves;
  - o Superelevation;
- Gradients;
- Driveway density;
- Passing lanes/short four-lane sections;
- Roadside design;
- Intersections:
  - o Angle of skew
  - o Traffic control form
  - o Exclusive right-turn lanes

The CMF in respect of horizontal curvature is shown in Equation 9.7.

$$CMF = \frac{2{,}48L_C + \frac{24{,}46}{R} - 0{,}012S}{2{,}48L_C} \qquad 9.7$$

where L_C = Curve length (km)

R = Radius of curve (m)

S = 1 if transition curve present

= 0 if no transition curve is present

If transition curves are present, the length variable, L_c, represents the length of the circular portion of the curve.

The CMF for superelevation deficiency is expressed as

$$CMF = 1{,}001.e_d^{0{,}0781} \qquad 9.8$$

= 1,0 for e_d < 1%

where e_d = superelevation deficiency (per cent)

It should be noted that, as discussed in Chapter 4, a superelevation deficiency of four per cent or more constitutes poor design.

Table 9.3: Crash modification factors for various shoulder types and widths						
Shoulder type and total width (m)	ADT (veh/day)					
	1	500	1000	1500	2000	2500
Paved						
0,0	1,04	1,04	1,08	1,13	1,18	1,18
0,6	1,02	1,02	1,05	1,08	1,11	1,11
1,2	1,01	1,01	1,02	1,04	1,05	1,05
1,8	1,00	1,00	1,00	1,00	1,00	1,00
2,5	0,99	0,99	0,98	0,97	0,95	0,95
Gravel						
0,0	1,04	1,04	1,08	1,13	1,18	1,18
0,6	1,03	1,03	1,06	1,08	1,11	1,11
1,2	1,01	1,01	1,03	1,04	1,06	1,06
1,8	1,01	1,01	1,01	1.01	1,01	1,01
2,5	1,00	1,00	0,99	0,99	0,96	0,96
Composite (50% paved)						
0,0	1,04	1,04	1,08	1,13	1,18	1,18
0,6	1,03	1,03	1,06	1,09	1,11	1,11
1,2	1.02	1.02	1,03	1,05	1,06	1,06
1,8	1,01	1,01	1,01	1,01	1,01	1,01
2,5	1,01	1,01	1,00	0,99	0,97	0,97
Turf						
0,0	1,04	1,04	1,08	1,13	1,18	1,18
0,6	1,04	1,04	1,06	1,09	1,12	1,12
1,2	1,02	1,02	1,04	1,06	1,07	1,07
1,8	1,03	1,03	1,03	1,03	1,03	1,03
2,5	1,03	1,03	1,02	1,00	0,99	0,99

The CMF for gradient is given in Equation 9.9. The equation is based on the absolute value of gradient simply because an upgrade in one direction is a downgrade in the opposite direction. The gradient factor is applied to the entire grade, i.e. from one Vertical Point of Intersection (VPI) to the next.

$$CMF = 1 + 0{,}035\,G \qquad 9.9$$

where G = Gradient (%)

The crash rate is strongly correlated with the number of accesses or driveways along the road. Expressed in terms of a driveway density in driveways per kilometre, the CMF is shown in

Equation 9.10 as

$$CMF = (0{,}424 + 0{,}04DD)\left(\frac{ADT}{500}\right)^{0{,}0911 - 0{,}015.DD} \qquad 9.10$$

where DD = Driveway density (Driveways/km)

ADT = Average daily traffic

The CMF for passing lanes or climbing lanes is taken as 0,75 for total crashes in both directions of travel over the length of the passing or climbing lane. Where auxiliary lanes are provided on both sides of the road over a short length of the road segment, the CFM improves to 0,65 for the length of the auxiliary lanes.

The quality of roadside design is represented by the roadside hazard rating described in Section 9.4.1. The calculated CMF applies to the total roadway segment over which the roadside hazard rating applies and is thus independent of the length of the segment. The CMF is calculated as shown in Equation 9.11.

$$CMF = 0{,}7915 + 0{,}0718.RHR \qquad 9.11$$

Where RHR = Roadside hazard rating

The number of legs in an intersection has a significant effect on the crash rate. This is attributable to the difference in the number of conflict points in four- as opposed to three-legged intersections. These differences are accommodated in the base models described in Section 9.4.2 and thus do not generate a CMF. It is, however, widely accepted that departures from an angle of skew of 90^{O}, whether positive or negative, is detrimental to safety. This impact is captured in Equations 9.12(a) and 9.12(b) for three-legged and four-legged STOP-controlled intersections respectively.

$$CMF = \exp(0{,}0040\ \theta) \qquad 9.12(a)$$

$$CMF = \exp(0{,}0054\ \theta) \qquad 9.12(b)$$

Where θ = intersection skew angle expressed as the absolute value of the difference between 90^{O} and the actual intersection angle.

Signal control separates the movements from conflicting approaches so that the CMF for skew at signalized four-legged intersections is 1,00 for all angles.

The safety differences between STOP-controlled and signalized intersections are accounted for by separate base models rather than by a CMF. The nominal case for STOP-control has STOP-control on the minor approaches only. Minor-road YIELD-control is treated identically to STOP-control. A CMF of 0,53 is applied to the CMF for minor-road STOP control to provide for conversion to all-way STOP control. This suggests that all-way STOP control has a crash rate that is 47 per cent lower than that for minor-leg STOP control. This should not, however be interpreted as an argument in favour of arbitrarily replacing minor-leg control by all-way control. The CMF applies only when all-way STOP control is, in fact, warranted.

The nominal or base condition for provision of right-turn lanes at intersections is the absence of turning lanes. No data are available to quantify the effect of right-turn lanes on the minor legs of intersections. The CMFs for right-turn lanes on the major legs are presented in Table 9.4. The CMFs for installation of right-turn lanes on both major legs of a four-legged intersection are simply the square of the value for a right-turn lane on a single approach. The CMFs apply to total intersection-related crashes.

The nominal or base condition for intersection sight distance is the availability of adequate sight distance in all quadrants of the intersection. Where, for reasons of road alignment and terrain, sight distance is less than that specified for a design speed of 20 km/h less than the design speed of the major road, it is considered to be limited. Restrictions of sight distance by specific obstructions such as trees and bushes do not qualify for consideration of CMFs, it being assumed that these obstructions would be removed.

The CMFs for restricted sight distance are:

- 1,05 for sight distance restriction in one quadrant of the intersection;
- 1,10 for sight distance restriction in two quadrants of the intersection;
- 1,15 for sight distance restriction in three quadrants of the intersection; and
- 1,20 for sight distance restriction in four quadrants of the intersection.

In the case of all-way STOP and signal control, a CMF of 1 applies.

9.5 ECONOMIC ANALYSIS OF GEOMETRIC IMPROVEMENTS

9.5.1 Crash reduction

In Section 9.4, a method for predicting the number of crashes on a specified stretch of roadway was offered. This methodology supports the derivation of the probable difference in accident rate resulting from a specific geometric improvement. The procedure involved would be to calculate the number of crashes that could be expected to occur:

- if the geometry of the road was unaltered; and
- if the geometry was upgraded.

Upgrading could be an increase in the lane and/or shoulder width across the length of the road or improvement of a specific horizontal or vertical curve or any combination of these options.

Benefits would accrue from these upgrades throughout the design life of the road. As ADT is an important component of the accident prediction model, an academically correct

Table 9.4: Crash modification factors for right-turn lanes

Intersection type	Intersection traffic control	Number of major approaches on which right turns are installed	
		One approach	Both approaches
Three-legged	STOP sign	0,78	-
	Traffic signal	0,85	-
Four-legged	STOP sign	0,76	0,58
	Traffic signal	0,82	0,67

approach would involve calculating the predicted number of crashes with and without the proposed upgrades for each successive year on the basis of an annually increasing ADT, attaching a cash value to the reduction in crashes, thereafter discounting benefits to the current year and summing them for comparison with the anticipated construction cost.

In view of the reliance placed on an external model and the coarseness of the calibration factor, this level of refinement is deemed to be unnecessary. It is suggested that the Net Present Worth of an annual series with a duration equal to the design life of the proposed upgrade would provide an adequate indication of the merits or otherwise of the upgrade. In effect, the assumption is made that the ADT will remain constant for the life of the upgrade.

The Net Present Worth is expressed as

$$P = A_s \left[\frac{(1+i)^n - 1}{i(1+i)^n} \right] \qquad 9.7$$

where	P	=	Net present worth
	A_S	=	Annual saving arising from improvements in safety
	i	=	Interest rate
	n	=	Design life of upgrade (years)

In Section 9.2, reference was made to the average cost of a crash as being of the order of R 30 000. The annual saving to be inserted in Equation 9.7 is thus the product of the reduction in the predicted number of crashes annually and R 30 000.

9.5.2 Time savings

Improvements in geometry can lead to increases in travel speed on a road and hence to reductions in travel time. The travel speed will vary along the length of the road because of changes in the geometry as well as changes in traffic-related variables such as flow, directional split and traffic composition, i.e. the percentage of various vehicle types in the traffic stream. The traffic-related variables have to be eliminated from the comparison between the two conditions (with and without geometric upgrades) and this is done by determining the speed profiles of vehicles with headways typically longer than 10 seconds.

Calculation of the benefit, being the value of the reduction in travel time arising from improvements to the geometry of the road segment, requires that a speed profile be employed to determine the journey times between the terminals of the road segment under consideration. The development of the speed profile is discussed in the following section.

For one year, the benefit is expressed as:

$$A_T = ADT\,(T_1 - T_2)V \qquad 9.8$$

where	A_T	=	Annual saving arising from reduction in travel time
	ADT	=	Average daily traffic
	T_1	=	Travel time with upgrades
	T_2	=	Travel time without upgrades
	V	=	Cash value of one hour saved

The benefit, AT, is inserted into Equation 9.7 to derive the Net Present Worth of the time savings accruing over the design life of the geometric upgrades.

9.5.3 The speed profile

Development of the speed profile requires an ability to predict speed on the basis of the geometric elements that present themselves. Much research on speed prediction has been carried out worldwide and some of this work is discussed in Chapter 4. In general it has been found that lane and shoulder widths are not statistically significant as descriptors of speed and they do not thus appear in the prediction models.

For convenience, Equation 4.1 is repeated below

$$V_{85} = 105{,}31 + 1{,}62 \times 10^{-5} \times B^2 - 0{,}064 \times B \qquad (4.1)$$

where V_{85} = 85^{th} percentile speed (km/h)

B = Bendiness = 57 300/R (degrees/km) = 57 300 θ/L

R = Radius (metres)

θ = Deviation angle (radians)

L = Total length of curve (metres).

9.5.4 Evaluation of geometric improvements

Assuming that improvements could be made at several points along the road, it is highly unlikely that sufficient funding would be available to build all of them. Summing the benefits resulting from safety improvements and time saving improvements is thus pointless. A further argument against summing all benefits is that the prime objective of a 3R project is to extend the life of the pavement. Savings arising from geometric improvements should be seen as a bonus and not as a prime objective.

The designer should create a strip map listing all the planned pavement enhancement activities by stake value along the road segment under consideration. Brief descriptions of the nature of proposed geometric improvements and their Net Present Worth should then be added to the strip map. With this strategic overview of the 3R project, it should be possible to select the geometric improvements that could form part of the project.

It is not possible to lay down hard and fast rules concerning the further definition of the project. A geometric improvement may have such a high Net Present Worth that it changes the prioritization of the pavement remedial works. On the other hand, an improvement, even with a high Net Present Worth, may fall in an area where no remedial works are intended and, because of this, is abandoned. Clearly, a high level of engineering judgment would have to brought to bear on a 3R project to determine the best possible combination of pavement and geometric improvements.

Back to Menu

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 10

GRADE SEPARATION STRUCTURES

10.1 DESIGN CONSISTENCY

The provisions contained in this chapter should be used in conjunction with the SANRAL Code of Procedure for the Planning and Design of Structures. In the event of inconsistencies, the Code of Procedure for the Planning and Design of Structures shall be the governing document.

This chapter relates primarily to clearances for structures for geometric design purposes and the clearances given should be regarded as minimum requirements. Any variations to these clearances should be agreed with the client prior to the incorporation into the design. Nevertheless, the onus rests on the designer to ascertain their applicability to given conditions and the designer should ensure that the clearances provided do not impede the required sight distance. Due cognisance should be taken of possible future carriageway or road widening in accommodating the specific clearances and no part of the structure should encroach on the envelope defining the clearance requirements. However, where the specific horizontal clearance cannot be achieved, a longitudinal protective barrier should be provided.

Where structures either over or under railway lines are required, the regulations of the Code of Procedure issued by the relevant railway authority should be applied to the geometric design. In cases where these regulations differ from the suggestions contained in these guidelines, the matter should be discussed with the Roads Agency's Regional Manager.

In this guideline, the words “underpass” and “overpass” refer to the position of the minor road relative to the major road.

10.2 FUTURE CAPACITY REQUIREMENTS

For capacity analysis in respect of structures, a design life of 50 years should be used, and any structure over or under a national road should provide for the eventual development of the national road to its full standard during this lifetime. At the time of construction of a structure over a single-carriageway national road on permanent alignment, the matter of the construction of either one or both spans of the eventual two-span structure should be discussed with the Road Agency's Regional Manager.

In the case where a double carriageway national road crosses over a secondary road, the decision on whether or not to fill in the gap in the median caused by not "decking" the two separate sections of the underpass under the National road carriageway, should be carefully considered with adequate weight being given to safety as a result of the creation of a hazardous situation for vehicles on the national road.

In general on divided carriageways a bridge median with a width of 10 metres or less should be decked.

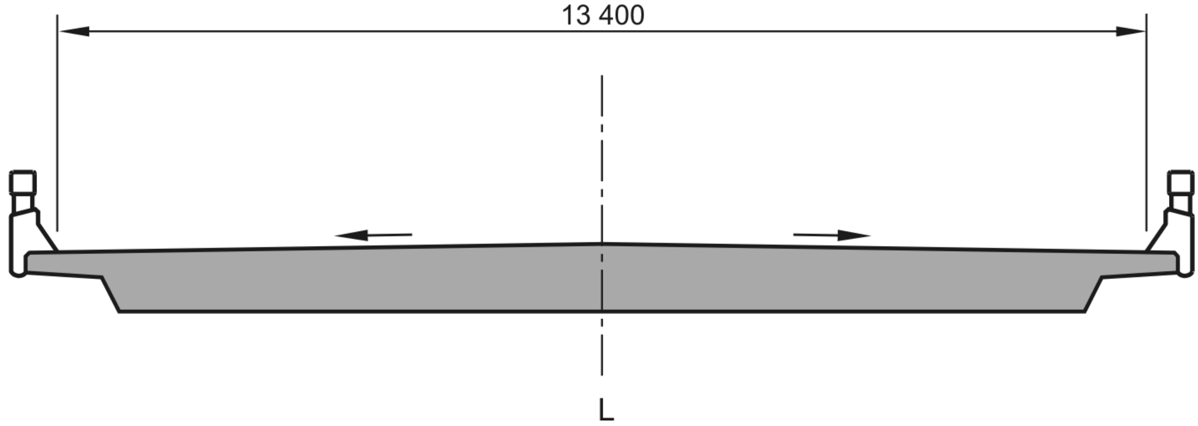


Figure 10.1: National Road underpasses - Two-lane single carriageway bridges

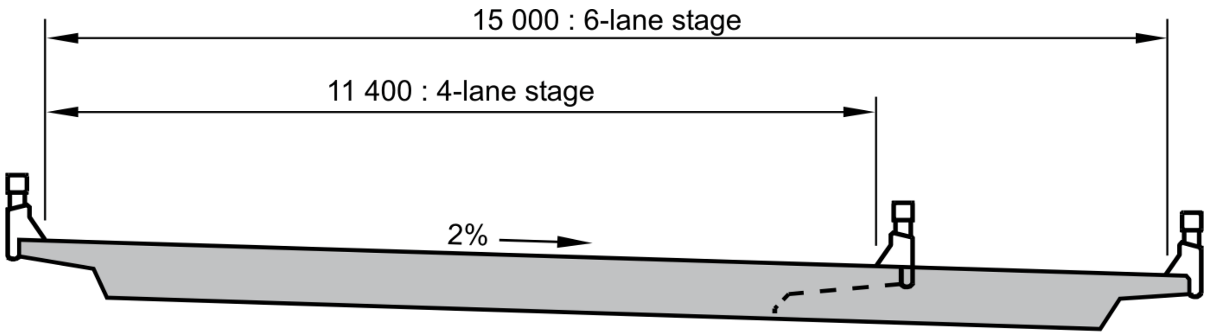


Figure 10.2: National Road underpasses- Dual carriageways

10.3 UNDERPASSES

10.3.1 National road structures

Where a National Road is carried over a travelled way, the full width of the National Road carriageway and shoulders should be provided between kerbs on all structures as shown in Figures 10.1 and 10.2.

10.3.2 Cross roads

Table 10.1 gives standard underpass widths in metres.

The normal vertical clearance of an underpass is 5,2 metres. However, the requirements of the Provincial or other authority should also be taken into account.

Table 10.1: Standard underpass widths

Separate right turn provision	Basic Lanes on Cross Road		
	2	4	6
Nil	13,0	25,4	32,8
Median ‘Slots”	16,7	25,4	32,8
Auxiliary lane for each direction	-	32,8	-

Note: For four- and six-lane underpass cross roads, a five metre median is employed.

Bridge designers will normally make use of the median to accommodate a centre pier. The median would normally accommodate an adequate length of right turn slot 3,7 metres wide, except where the first intersection on the cross-road is very close to the structure.

10.3.3 Agricultural underpasses

The normal horizontal clearance between abutments for an agricultural underpass is 4,0 metres. If local conditions allow, this dimension may be reduced. Proposals for wider clearances should be motivated and referred to the National Road Agency's Regional Manager. If the agricultural underpass provides the only access to the farm, it should have a minimum vertical clearance of 4,3 metres. Any other agricultural underpasses provided would normally have a vertical clearance of 4 metres, or 2,5 metres if this is acceptable to the farmer.

10.3.4 Cattle and equestrian underpasses

The vertical and horizontal clearance for cattle and equestrian underpasses should normally be 3 metres x 3 metres. Skewed crossings should be avoided and crossings should be straight so that the entire length can be seen from each end.

10.3.5 Pedestrian underpasses

Pedestrian underpasses should only be considered in cases where an overhead structure is impractical. The minimum vertical and horizontal clearances for pedestrian underpasses will be 2,5 metres x 2,5 metres.

In urban areas, lighting should be provided for pedestrian underpasses, as well as for pedestrian ramps.

The maximum longitudinal slope on ramps should not exceed 9 per cent.

10.4 OVERPASSES

10.4.1 Cross roads

Table 10.2 gives the standard widths in metres between guardrails on overpasses for roads crossing the national road. These overpasses are dimensioned in more detail in Figures 10.3 to 10.6.

Table 10.2: Standard widths on overpasses			
Separate right turn provision	**Basic Lanes on Cross Road**		
	2	**4**	**6**
Nil	12,4	2 x 12,4 [2]	2x 16,1 [2]
Median turning lanes	16,1 [1]	2 x 14,6 [1]	2 x18,3 [2]
Auxiliary lane for each direction	-	2 x 16,1 [2]	2 x 19.8 [2]

Notes

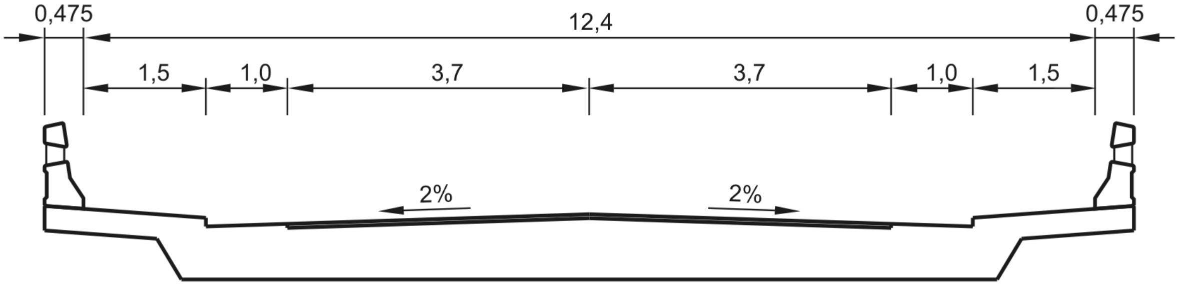
1. In this type of road there is no median but an extra lane is added for back-to-back right hand turn storage.

2. All roads having 4 or more lanes shall be constructed as two separate carriage ways with, in the interests of safety, a median of adequate width.

The vertical clearance above any point of the road surface which is under a structure should, for new bridges, be not less than 5,2 metres. This allows for future resurfacing under the bridge. The critical point is not necessarily on the centreline of the road passing under the structure. The geometry of each individual case must be evaluated in terms of the local gradient and superelevation, if any, as well as the shape of the structure, which may also be carrying a road with its own gradient and superelevation.

Due to the vulnerability of the individual beams of precast prestressed beam and slab decks to impact, a minimum vertical clearance of 5,6 metres should be used for this type of deck. Alternatively, the beam bottom flanges should be joined together to form a continuous solid soffit.

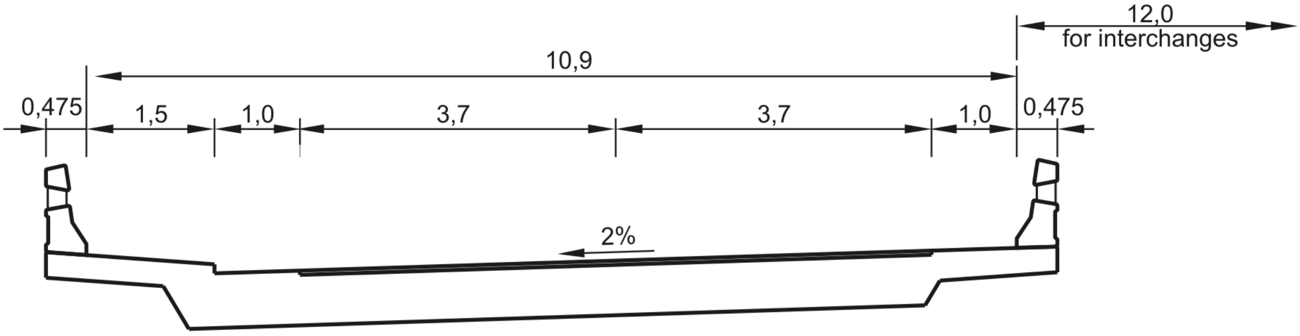
In the case of existing bridges, a relaxation to not less than 4,9 metres can be considered to make provision for rehabilitation of the pave-



Note: Sidewalks are not obligatory. If not provided, 1,5 metres to be added to shoulder. Sidewalks to allow for 300 mm beyond lane edge.

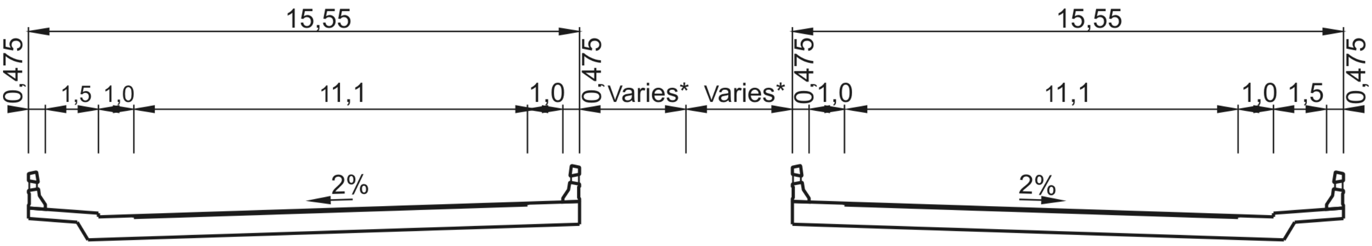
Dimensions in metres.

Figure 10.3: Two-lane cross road



Note: Sidewalk is not obligatory.

Figure 10.4: Four-lane cross roads



* Minumum of 3 metres when associated with an interchange

May be used for (A) four-lane crossroad with provision for median turning lanes
or (B) six-lane crossroad

Note: Sidewalks are not obligatory
Dimensions in metres.

Figure 10.5: Four- and six-lane cross roads

ment without having to raise the bridge deck. It should be noted that the maximum permissible height of a double-decker bus is 4,6 metres and, for any other vehicle, 4,3 metres.

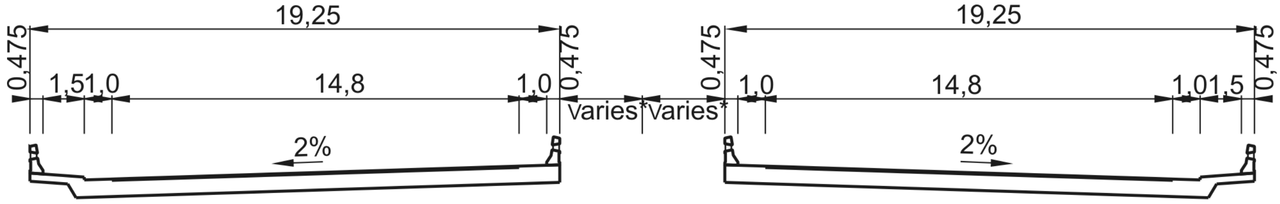
On routes that serve as Superload routes, additional clearances may be necessary. The designer should check whether a route is intended for super loads and, if so, what requirements exist for it..

General Notes

a. In each case, an allowance of 2,5 metre between the edge of the carriageway and the handrail has been made. In the case of interchanges, this is normally made up of 1,0 metre of shoulder and a 1,5 metre sidewalk or a 2,5 metre shoulder and no sidewalk. Where there is no interchange, the controlling local authority may use the 2,5 metres as it sees fit.

b. In the case of minor roads, the clear width between guardrails or handrails may be reduced to 9,4 metres.

c. The requirements of the provincial or other authority should also be taken into account when the dimensions of overpass structures are determined.



* Minumum of 3 metres when associated with an interchange

May be used for (A) sir-lane crossroad with provision for median turning lanes

or (B) six-lane crossroad with auxiliary lanes

Note: 1,5 metre sidewalks are not obligatory

Dimensions in metres.

Figure 10.6: Six-lane cross-roads

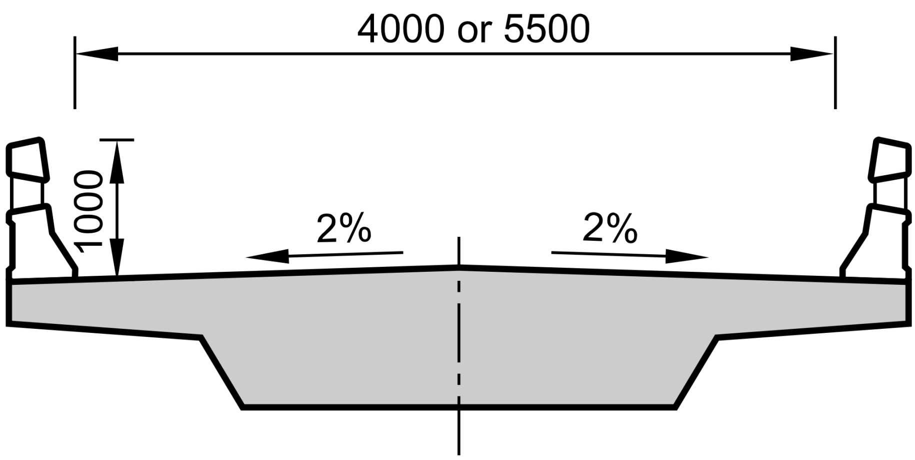


Figure 10.7: Bridge width for agricultural overpass

10.4.2 Agricultural overpasses

Shoulders and foot walks are not necessary on agricultural overpasses. The clearance between guardrails should be either 4,0 metres, as shown in Figure 10.7 where adequate stopping sight distance is obtainable, or 5,5 metres where stopping sight distance is not obtainable with a structure serving more than one property. Any proposed deviation from this standard should be referred to the Regional Manager of the NRA.

10.4.3 Footbridges

Where possible, overhead crossings for pedestrians should have a minimum width of 2,0 metres, and should have handrails normally 1,10 metres in height. The handrails should be designed to be vandal-proof and durable.

Because of the relative lightness of pedestrian footbridges, a vehicle impacting the structure is likely to cause considerable damage, including the possibility of the deck dropping onto the roadway below. A vertical clearance of 5,9 metres is thus normally required for footbridges.

Ramp approaches should be designed to encourage pedestrian usage by following established routes. Where appropriate, these may be provided with steps on the steeper slopes near the ends of the bridge to encourage usage of the bridge.

Vandalism includes deliberately dropping objects onto vehicles passing under footbridges. Where this is likely to occur or has been known to occur, the footbridge should be enclosed by a cage over its full length.

10.4.4 Services

In cases where permission is granted by the NRA for the carrying of services across the national road reserve, these services should be situated in such a way that they will not be visible to traffic on either the National Road or the cross road. The services should be located in the foot walks (where these exist) or under the balustrade.

10.4.5 Foot walks

Where provided on structures carrying cross roads, foot walks should be have a minimum width of 1,25 metres. Pedestrians on the foot walks should be protected from traffic by means of a barrier kerb on the side of the foot walks. For obvious reasons, no foot walks are to be provided on bridges carrying National Road freeways.

10.4.6 Balustrades and parapets

Details for balustrades and parapets should be as specified in the Code of Procedure for the Planning of Design Structures.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 11

TOLL PLAZAS

11.1 INTRODUCTION

Considerable experience in the planning, design and operation of toll plazas has been gained in South Africa on our National Highways. This section summarises the important factors as they impact on the planning and design of major roads. The purpose is to familiarize designers with the concepts and constraints of toll plaza design and provide information for use in layout design and basic planning.

11.2 PLANNING

11.2.1 Positioning of the Plaza

A network analysis of travel volumes, origins and destination and the relative costs of using the tolled road is a prerequisite. Once the financial viability is established i.e. the income stream is in balance with the overall project costs, it is then necessary to position the plaza so as to maximize the defined catchment. This would normally require the evaluation of a number of alternatives as they impact on;

- Road user safety
- Land required
- Cost
- Operational efficiency
- Security

There are a number of measures of effectiveness that can be applied against these broad objectives.

11.2.2 Road User Safety

The road geometry through the plaza and on its approaches is the major determinant of road user safety. The plaza should preferably be located on a tangent and meet the following measures of sight distance.

Stopping sight distance must be continuously available along the road. At the approach to the plaza court, vehicles, and heavy vehicles in particular, require sufficient distance to stop at the rear of a 5-vehicle queue. Because of speed restrictions in the plaza area, an approach speed of 100km/hr can be assumed for design. See Sections 3.5.4 to 3.5.6.

Decision sight distance must be available to the start of the approach tapers as discussed in Section 3.5.8. The column "Interchanges: Sight distance to nose" in Table 3.7 should be used.

11.2.3 Land required

As a general guide, the road reserve for a main line plaza with 8 lanes would require widening on one side by 15 metres and, on the side containing the control building, by approximately 30 metres. The control building itself could occupy a fenced site of approximately 60 metres by 25 metres. These dimensions are average and are heavily influenced by the landform and the surrounding land use.

11.2.4 Cost

The construction cost and operational cost determine the economy of the toll plaza location. The measure of effectiveness is the present value of future expenditure. The following are the important items that should be considered.

- Land costs
- Mass earthworks
- Geotechnical conditions
- Cost of services
- Cost of relocating services
- Stormwater drainage
- Pavement widening
- Plaza construction and running costs
- Accommodation of traffic
- Maintenance costs.

11.2.5 Operational Efficiency

The geometry of the road at the approaches and within the plaza court influences the operational efficiency. Tapers should be such that vehicles can spread comfortably and merge without undue constraint. Approach tapers of 1:8 and merge tapers of 1:5 are acceptable to and from a point 50 metres from the centre line across the plaza.

A maximum longitudinal grade of 1,5% should be maintained within the 60 metre queuing length on either side of the plaza. A grade of 3% should not be exceeded at the approach to the taper areas.

11.2.6 Security

Location factors that influence toll plaza security and which should be evaluated are;

- Land use in the vicinity
- Other access routes to the toll plaza
- Vegetation

11.3 Design Norms and Dimensions

The toll plaza layout will ultimately be governed by the number of lanes required. The standard toll booth module is 5,0 m wide. The lane width is 3,0 metres and the toll island is 2,0 metres wide. The general layout of a plaza, lane arrangement and island details are shown in Figures 11.1 to 11.4. An extra 3,0 m shoulder is recommended at the toll lanes on the periphery to accommodate abnormal vehicles.

For planning purposes, the average queuing and processing time per vehicle can be taken as 30 seconds and the maximum queue length should not exceed 5 vehicles.

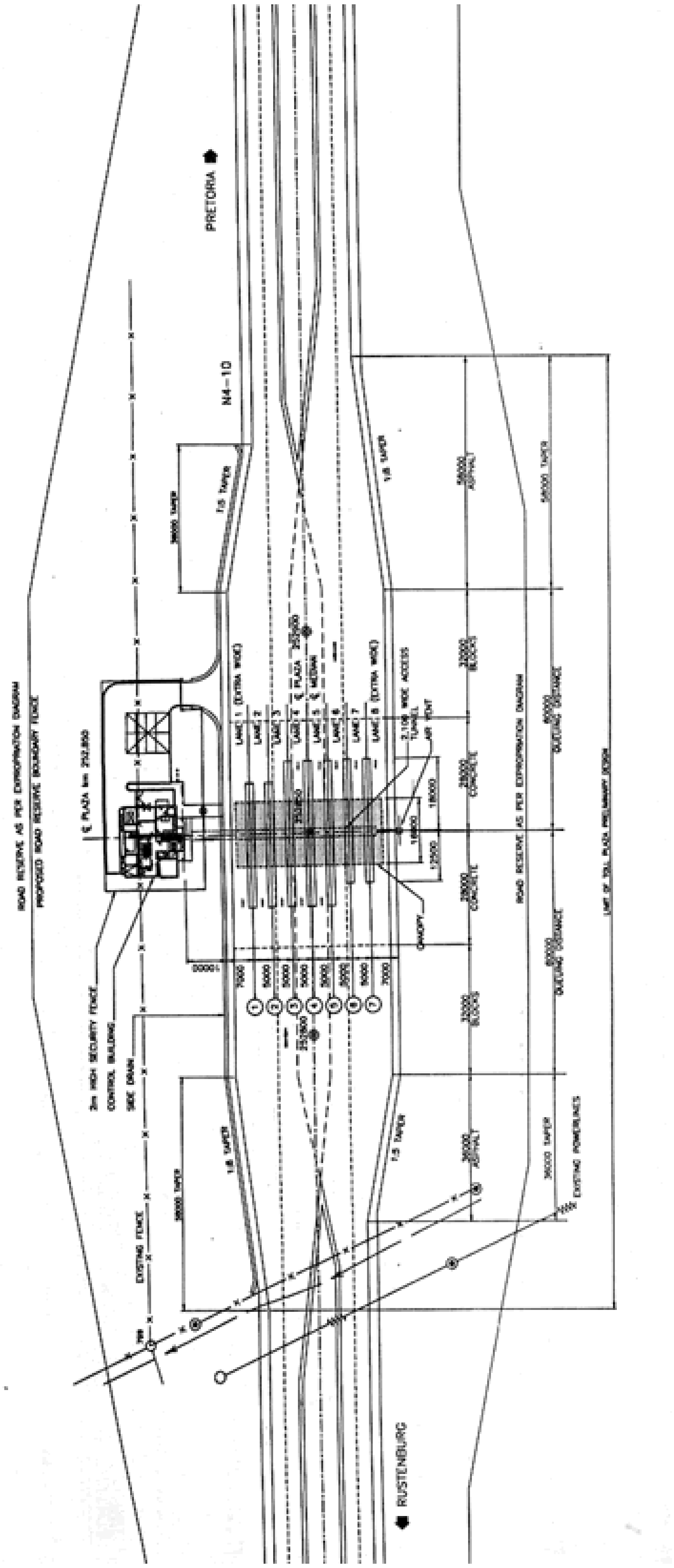


Figure 11.1: Typical layout of toll plaza

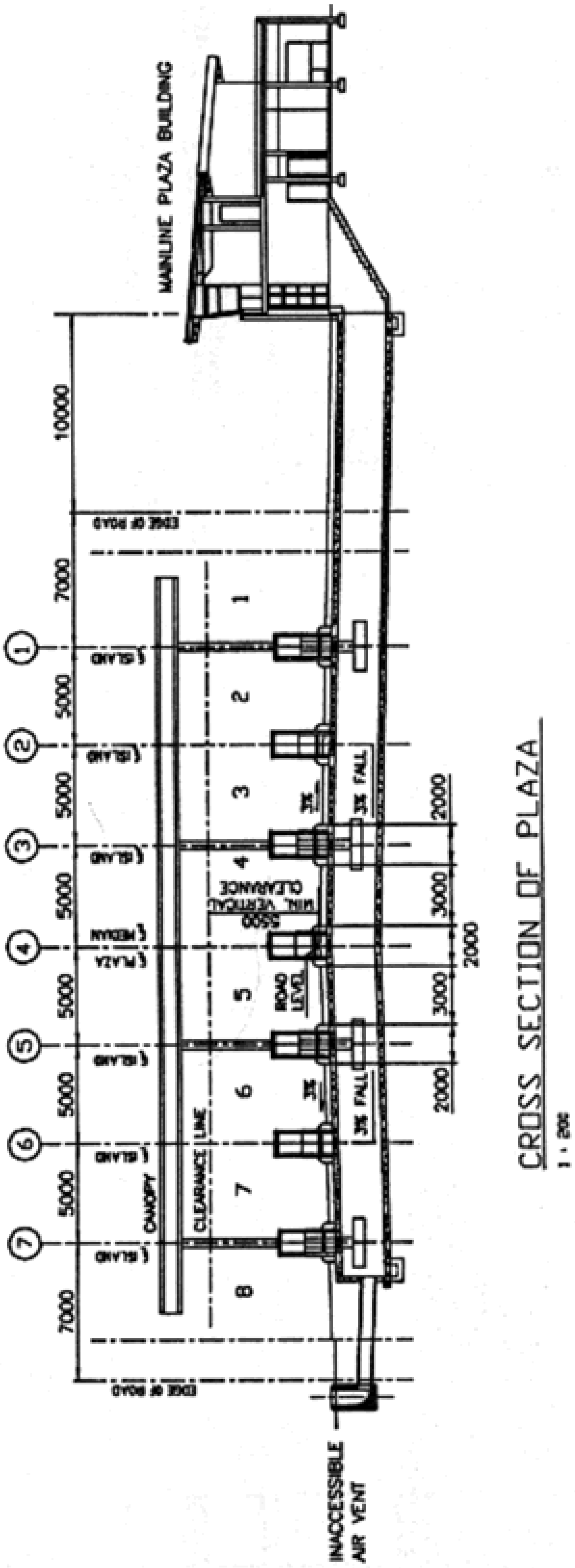


Figure 11.2: Cross-section of toll plaza

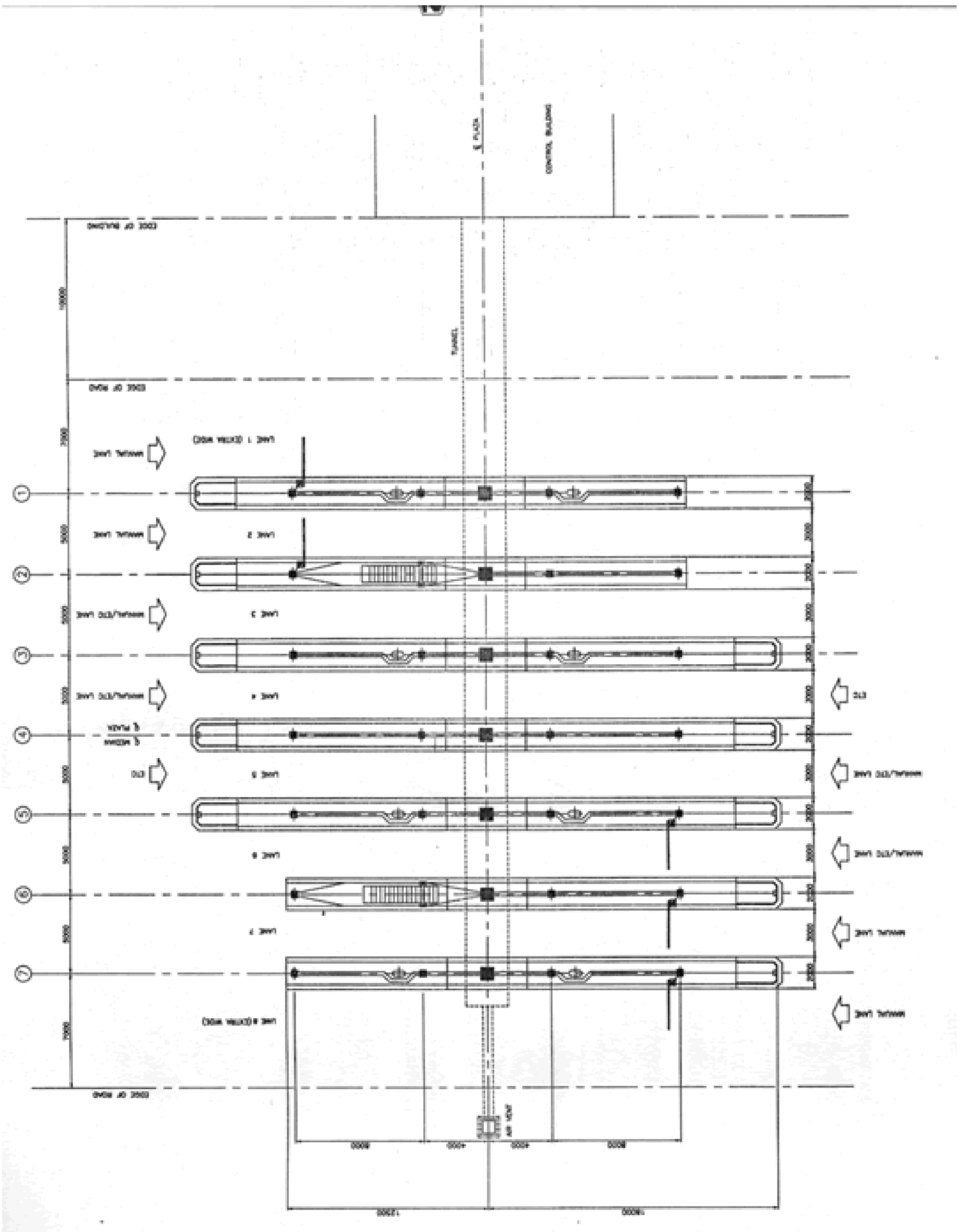


Figure 11.3: Island layout

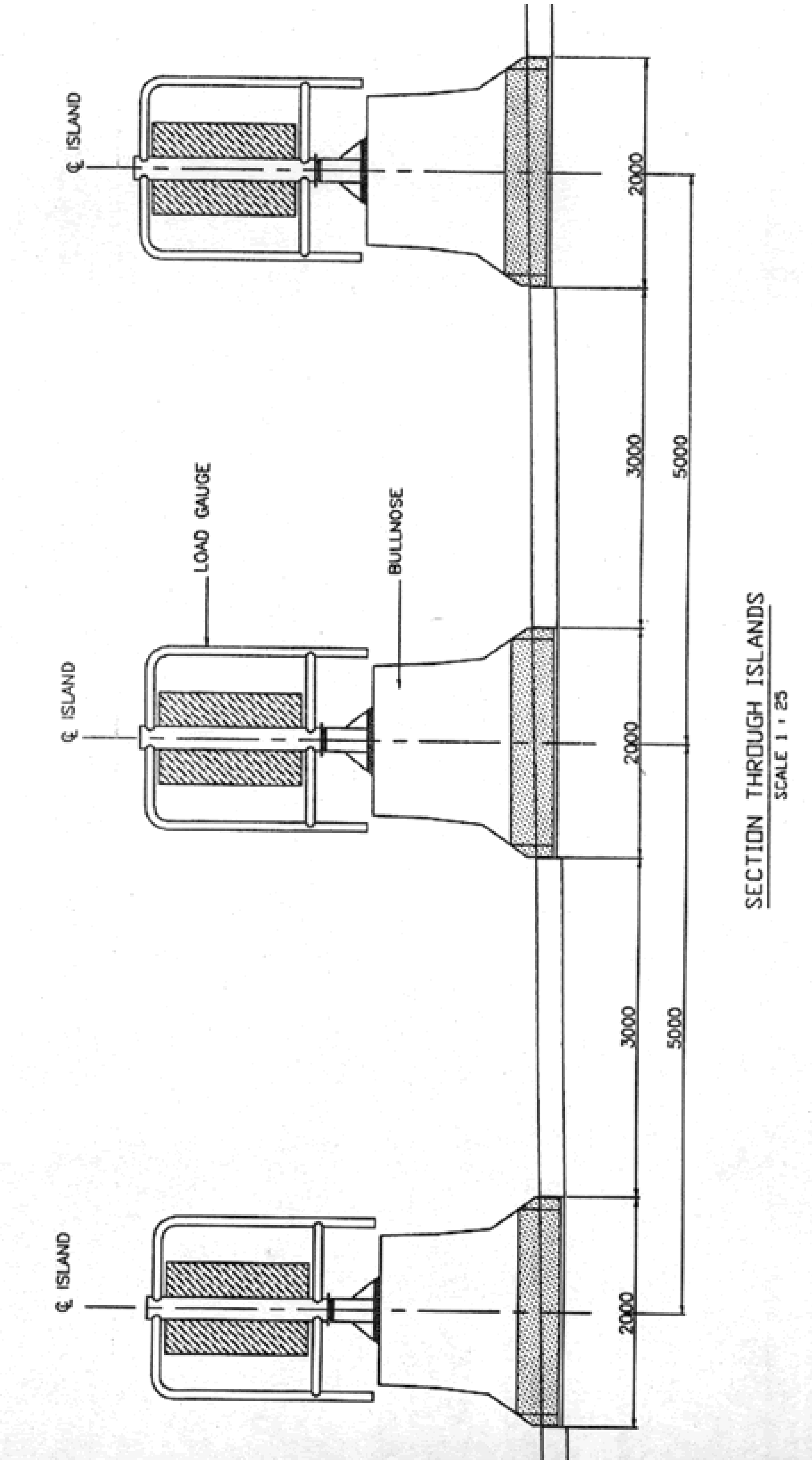


Figure 11.4: Section through islands

BIBLIOGRAPHY

1. Alberta Infrastructure. **Highway geometric design guide**. 1999.
2. American Association of State Highway and Transportation Officials. **Policy on the geometric design of highways and streets.** Washington, 2001
3 Austroads **Draft Guide to the Geometric Design of Rural Roads.** Sydney,Australia 2000.
4 Bared JG, Prosser W and Esse CT. **State-of-the-art design of roundabouts.** Transportation Research Board Record 1579 pp 1-12, Washington 1997.
5. Committee of Land Transport Officials. **South African road safety manual.** Pretoria, 1999
6. Committee of Land Transport Officials. **National guidelines for road access management in South Africa.** Pretoria 2001
7. Committee of State Road Authorities. **TRH 17 : Geometric Design of Rural Roads.** Pretoria, South Africa. 1988.
8. de Beer E, van der Walt J. **Annual traffic safety audit.** Automobile Association of South Africa Road Traffic Safety Foundation. Johannesburg 2000
9. Department of Transport, Chief Directorate of National Roads. **Warrants for the Lighting of Major Roads.** SARB Document RR 89/093. Pretoria, November 1992.
10. Douglass RD, McClelland K and Fitzgerald W. **Context sensitive design.** Proceedings 2nd International Symposium on Highway Geometric Design pp 664-672, Mainz 2000.
11. Durth W. **Implementation of intermediate cross-sections.** International Symposium on Highway Geometric Design Practices, Boston, 1995
12. Federal Highway Administration. **Flexibility in highway design.** Report FHWA-PD-97-062, Washington 1997.
13. Fitzpatrick K et al. **Evaluation of design consistency methods for two-lane rural highways.** Federal Highway Administration Report FHWA-RD-99-173, Washington 1999.
14. Fitzpatrick K, Lienau T and Fambro D. **Driver eye and vehicle heights for use in geometric design.** Transportation Research Board Record 1612 pp 1-9, Washington 1998.
15. Gattis JL and Low ST. **Intersection angle geometry and the driver's field of view.** Transportation Research Board Record 1612 pp 10-16, Washington 1998.
16. Harwood DW, Mason JM and Brydia RE. **Intersection sight distance.** Transportation Research Board NCHRP Report 383. Washington,1996.
17. Harwood DW, Council FM, Hauer E, Hughes WE and A Vogt. **Prediction of the expected safety performance of rural two-lane roads.** Federal Highway Administration Report FHWA-RD-99-207. Washington, 2000.
18. Harwood DW, Pietrucha MT, Wooldridge MD, Brydia RE and Fitzpatrick K. **Median intersection design.** Transportation Research Board NCHRP Report 375, Washington, 1995
19. Hauer E. **Safety in geometric design standards.** Proceedings 2nd International Symposium on Highway Geometric Design pp.11-35, Mainz 2000.

20. Henue KE. **Harmonising transportation and community values**. Institute ofTransportation Engineers Journal Volume 68 Issue 11 pp 32-35, Washington 1998
21. The Highway Agency. **Design Manual for Roads and Bridges.** London.1997
22. Ingham DJ and Burnett SL. **Interchange spacing in Gauteng.** Proceedings 2nd International Symposium on Highway Geometric Design pp 534-546, Mainz 2000.
23. Jordaan PW. **Peaking characteristics of rural road traffic.** Doctoral dissertation, University of Pretoria, Pretoria, 1985.
24. Kahl KB and Fambro DB. **Investigation of object-related accidents affecting stopping sight distance.** Transportation Research Board Record 1500 pp25-30, Washington, 1995.
25. King MR, Williams TP and Ewing R. **States flexing main street design: A report on efforts by various states to 'flex' their highway standards towards better main street design.** Rutgers Transport Policy Institute. New Brunswick, 2000.
26. Lamm R, Wolhuter KM and Ruscher T. **Introduction of a new approach to geometric design and road safety.** 20th Annual South African Transport Conference, Pretoria 2001.
27. Leisch JP. **Freeway and interchange design: A historical perspective.** Transportation Research Board Record 1385, pp 60-68, Washington,1993.
28. Leisch JP. **Operational considerations for systems of interchanges.** Transportation Research Board Record 1385, pp 106-111, Washington 1993.
29. Lunenfeld H. **Human factors associated with interchange design features.** Transportation Research Board Record 1385, pp 84-89, Washington 1993.
30. Mollett C. **A comparative study of road safety between the USA, UK and South Africa.** Conference Proceedings - Road safety on three continents. Pretoria September 2000.
31. National Highway Traffic Safety Administration. **Traffic safety Facts 2000: A compilation of motor vehicle crash data from the fatality analysis reporting system and the general estimates system.** US Department of Transportation, Washington 2001.
32. National Road Traffic Act (Act 93 of 1996).
33. Neumann TR. **Intersection channelisation design guide.** Transportation Research Board NCHRP Report 279, Washington, 1985.
34. Ogden KW, Taylor SY And Veith GJ. **Performance and design of roundabouts.** Traffic Engineering and Management pp 481-506, 1996.
35. Ogden, K.W. Safer Roads - **A Guide to Road Safety Engineering,** Ashgate Publishing Co., Aldershot, UK. 1997.
36. O'Flaherty C.A. **Highways and Traffic.** Edward Arnold, London 1974.
37. Oglesby, C.H. and Hewes, L.I. **Highway Engineering**, John Wiley and Sons, Inc., New York. London 1954.
38. Oglesby, C.H. and Hicks, R.G. **Highway Engineering - Fourth Edition.** John Wiley and Son. New York 1982.
39. Pline J (Editor). **Traffic Engineering Handbook.** Institute of Transportation Engineers, Washington, 1999.

40. Polus A, Craus J and Grinberg I. **Applying the level of service concept to climbing lanes.** Transportation Research Board Record 806, Washington, 1981.

41. Purchase PG. **The geometric design of conventional single-lane traffic circles.** M Thesis, University of Pretoria, Pretoria, 2001.

42. Ribbens H and de Beer EJH. **Pedestrian Facility Guidelines: Manual to Plan, Design, and Maintain Safe Pedestrian Facilities.** Department of Transport, Pretoria,1993.

43. Sampson J. **Implementation of signal spacing standards.** South African Transportation Conference, Pretoria, 2002.

44. Schermers G. **Roundabouts - an alternative to intersection control.** National Institute for Transport and Road Research Technical Report RT/83, CSIR, Pretoria, 1987.

45. Schutte IC. **Format for expressing road collision costs and guidelines for accommodating inflation.** CSIR/Transportek Technical Report TR-2000/13. Pretoria 2000.

46. Slater R.V.J. **A Simplified Approach to Assessing Warrants for the Lighting of Major Roads.** Department of Transport, Pretoria, 1997.

47. Slavik MM, Bester CJ and Sik JR. **Truck climbing ability on the Highveld.** National Institute of Transport and Road Research Technical Note TT/43/81, Pretoria 1981.

48. Southern African Development Community. **Road Traffic Signs Manual.** Department of Transport, Pretoria 1997.

49. Transportation Association of Canada. **Geometric Design Guide for Canadian Roads.** Ottawa,1999.

50. Transportation Research Board. **Highway Capacity Manual Special Report 209.** , National Research Council, Washington, DC. 2000.

51. Transportation Research Board. **NHCRP Report 400. Determination of Stopping Sight Distances.** National Research Council, Washington. DC. 1997.

52. Transportation Research Board. **Record. 1208: Highway Sight Distance Design Issues.** National Research Council, Washington, D.C. 1989.

53. Truck Watch. **The Working Guide for the Transport Industry in South Africa.** Fleetwatch January 2001. Published by Newslink CC. Honeydew. South Africa.

54. Tunnard C and Pushkarev B. **Manmade America: Chaos or control?**. Harmony Books, New York, 1981.

55. Twomey JM, Heckman ML, Hayward JC and Zuk RJ. **Accidents and safety associated with interchanges.** Transportation Research Board Record 1385, Washington, 1993

56. Walker RJ. **Coordination of basic intersection design elements: An overview.** Transportation Research Board Record 1385, Washington, 1993.

57. Wolhuter KM. **Warrants for climbing lanes.** M Thesis, University of Pretoria, Pretoria, 1990.

58. Wolhuter KM. Geometric Design : **An Overview of Research 1988 - 1995.** Department of Transport, Pretoria, South Africa. 1996.

59. Wolhuter KM. **SATTC Code of Practice for the geometric design of trunk roads.** CSIR/-Transportek Report CR-97/049, Pretoria, 1997.

60. Zegeer CV, Reinfurt DW, Hummer J, Herf L, and Hunter W. **Safety effects of cross-section design for two-lane roads.** Transportation Research Board Record 1195, Washington 1988.

61. Zegeer CV, Stewart, R and Council F. **Roadway widths for low-traffic-volume roads.** Transportation Research Board Report 362, Washington 1994.

geometric design guidelines

South African
National Roads Agency Limited

Historically Geometric Design was predicated on the capabilities of the design vehicle. Comparatively recently, Geometric Design has undergone a paradigm shift of note. It is now accepted that a road designed to "standards" is not necessarily safe and, furthermore, that human factors play a greater role in the determination of geometric design standards than do the limitations of the various design vehicles.

This Guideline document, compiled by CSIR on behalf of the South African National Road Agency, replaces the previous G2 Manual and is based on the new design philosophy.

NRA geometric design guidelines

The Guideline discusses the:

- Design philosophy and techniques
- Design controls
- Road elements being:
- The horizontal and vertical alignments; and
- The cross-section
- Aesthetics
- Intersections
- Interchanges and grade separation structures
- Safety features
- Road betterment
- Toll plazas