Chapter 1

Introduction to transportation engineering

1.1 Overview

Mobility is a basic human need. From the times immemorial, everyone travels either for food or leisure. A closely associated need is the transport of raw materials to a manufacturing unit or finished goods for consumption. Transportation fulfills these basic needs of humanity. Transportation plays a major role in the development of the human civilization. For instance, one could easily observe the strong correlation between the evolution of human settlement and the proximity of transport facilities. Also, there is a strong correlation between the quality of transport facilities and standard of living, because of which society places a great expectation from transportation facilities. In other words, the solution to transportation problems must be analytically based, economically sound, socially credible, environmentally sensitive, practically acceptable and sustainable. Alternatively, the transportation solution should be safe, rapid, comfortable, convenient, economical, and eco-friendly for both men and material.

1.2 Transportation system

In the last couple of decades transportation systems analysis has emerged as a recognized profession. More and more government organizations, universities, researchers, consultants, and private industrial groups around the world are becoming truly multi-modal in their orientation and are opting a systematic approach to transportation problems.

1.2.1 Diverse characteristics

The characteristics of transportation system that makes it diverse and complex are listed below:

1. **Multi-modal:** Covering all modes of transport; air, land, and sea for both passenger and freight.

2. **Multi-sector:** Encompassing the problems and viewpoints of government, private industry, and public.

3. **Multi-problem:** Ranging across a spectrum of issues that includes national and international policy, planning of regional system, the location and design of specific facilities, carrier management issues, regulatory, institutional and financial policies.

4. **Multi-objective:** Aiming at national and regional economic development, urban development, environment quality, and social quality, as well as service to users and financial and economic feasibility.

5. **Multi-disciplinary:** Drawing on the theories and methods of engineering, economics, operations research, political science, psychology, other natural, and social sciences, management and law.

1.2.2 Study context

The context in which transportation system is studied is also very diverse and are mentioned below:

1. **Planning range:** Urban transportation planning, producing long range plans for 5-25 years for multi-modal transportation systems in urban areas as well as short range programs of action for less than five years.

2. **Passenger transport:** Regional passenger transportation, dealing with inter-city passenger transport by air, rail, and highway and possible with new modes.

3. **Freight transport:** Routing and management, choice of different modes of rail and truck.

4. **International transport:** Issues such as containerization, inter-modal co-ordination.

1.2.3 Background: A changing world

The strong interrelationship and the interaction between transportation and the rest of the society especially in a rapidly changing world is significant to a transportation planner. Among them four critical dimensions of change in transportation system can be identified; which form the background to develop a right perspective.

1. **Change in the demand:** When the population, income, and land-use pattern changes, the pattern of demand changes; both in the amount and spatial distribution of that demand.

2. **Changes in the technology:** As an example, earlier, only two alternatives (bus transit and rail transit) were considered for urban transportation. But, now new systems like LRT, MRTS, etc offer a variety of alternatives.

3. **Change in operational policy:** Variety of policy options designed to improve the efficiency, such as incentive for car-pooling, bus fare, road tolls etc.

4. **Change in values of the public:** Earlier all beneficiaries of a system was monolithically considered as users. Now, not one system can be beneficial to all, instead one must identify the target groups like rich, poor, young, work trip, leisure etc.

1.2.4 Role of transportation engineer

In spite of the diversity of problem types, institutional contexts and technical perspectives there is an underlying unity: a body of theory and set of basic principles to be utilized in every analysis of transportation systems. The core of this is the transportation system analysis approach. The focus of this is the interaction between the transportation and activity systems of region. This approach is to *intervene, delicately and deliberately in the complex fabric of society to use transport effectively in coordination with other public and private actions to achieve the goals of that society*. For this the analyst must have substantial understanding of the transportation systems and their interaction with activity systems; which requires understanding of the basic theoretical concepts and available empirical knowledge.

1.2.5 Basic premise of a transportation system

The first step in formulation of a system analysis of transportation system is to examine the scope of analytical work. The basic premise is the explicit treatment of the total transportation system of region and the interrelations between the transportation and socioeconomic context.They can be stated as:

P1 The total transportation system must be viewed as a single multi-modal system.

P2 Considerations of transportation system cannot be separated from considerations of social, economic, and political system of the region.

This follows the following steps for the analysis of transportation system:

- **S1** Consider all modes of transportation
- **S2** Consider all elements of transportation like persons, goods, carriers (vehicles), paths in the network facilities in which vehicles are going, the terminal, etc.
- **S3** Consider all movements of passengers and goods for every O-D pair.
- **S4** Consider the total trip for every flows for every O-D over all modes and facilities.

As an example, consider the study of intra-city passenger transport in metro cities.

- Consider all modes: i.e rail, road, buses, private automobiles, trucks, new modes like LRT, MRTS, etc.
- Consider all elements like direct and indirect links, vehicles that can operate, terminals, transfer points, intra-city transit like taxis, autos, urban transit.
- Consider diverse pattern of O-D of passenger and goods.
- Consider service provided for access, egress, transfer points and mid-block travel etc.

Once all these components are identified, the planner can focus on elements that are of real concern.

1.3 Major disciplines of transportation

Transportation engineering can be broadly consisting of the four major parts:

1. Transportation Planning
2. Geometric Design
3. Pavement Design
4. Traffic Engineering

A brief overview of the topics is given below: Transportation planning deals with the development of a comprehensive set of action plan for the design, construction and operation of transportation facilities.

1.3.1 Transportation planning

Transportation planning essentially involves the development of a transport model which will accurately represent both the current as well as future transportation system.

1.3.2 Geometric design

Geometric design deals with physical proportioning of other transportation facilities, in contrast with the structural design of the facilities. The topics include the cross-sectional features, horizontal alignment, vertical alignment and intersections. Although there are several modes of travel like road, rail, air, etc.. the underlying principles are common to a great extent. Therefore emphasis will be normally given for the geometric design of roads.

1.3.3 Pavement analysis and design

Pavement design deals with the structural design of roads, both (bituminous and concrete), commonly known as (flexible pavements and rigid pavements) respectively. It deals with the design of paving materials, determination of the layer thickness, and construction and maintenance procedures. The design mainly covers structural aspects, functional aspects, drainage. Structural design ensures the pavement has enough strength to withstand the impact of loads, functional design emphasizes on the riding quality, and the drainage design protects the pavement from damage due to water infiltration.

1.3.4 Traffic engineering

Traffic engineering covers a broad range of engineering applications with a focus on the safety of the public, the efficient use of transportation resources, and the mobility of people and goods. Traffic engineering involves a variety of engineering and management skills, including design, operation, and system optimization. In order to address the above requirement, the traffic engineer must first understand the traffic flow behavior and characteristics by extensive collection of traffic flow data and analysis. Based on this analysis, traffic flow is controlled so that the transport infrastructure is used optimally as well as with good service quality. In short, the role of traffic engineer is to protect the environment while providing mobility , to preserve scarce resources while assuring economic activity, and to assure safety and security to people and vehicles, through both acceptable practices and high-tech communications.

1.4 Other important disciplines

In addition to the four major disciplines of transportation, there are several other important disciplines that are being evolved in the past few decades. Although it is difficult to categorize them into separate well defined disciplines because of the significant overlap, it may be worth the effort to highlight the importance given by the transportation community. They can be enumerated as below:

1. **Public transportation:** Public transportation or mass transportation deals with study of the transportation system that meets the travel need of several people by sharing a vehicle. Generally this focuses on the urban travel by bus and rail transit. The major topics include characteristics of various modes; planning, management and operations; and policies for promoting public transportation.

2. **Financial and economic analysis** Transportation facilities require large capital investments. Therefore it is imperative that who ever invests money should get the returns. When government invests in transportation, its objective is not often monetary returns; but social benefits. The economic analysis of transportation project tries to quantify the economic benefit which includes saving in travel time, fuel consumption, etc. This will help the planner in evaluating various projects and to optimally allocate funds. On the contrary, private sector investments require monetary profits from the projects. Financial evaluation tries to quantify the return from a project.

3. **Environmental impact assessment** The depletion of fossil fuels and the degradation of the environment has been a severe concern of the planners in the past few decades. Transportation; in spite of its benefits to the society is a major contributor to the above concern. The environmental impact assessment attempts in quantifying the environmental impacts and tries to evolve strategies for the mitigation and reduction of the impact due to both construction and operation. The primary impacts are fuel consumption, air pollution, and noise pollution.

4. **Accident analysis and reduction** One of the silent killers of humanity is transportation. Several statistics evaluates that more people are killed due to transportation than great wars and natural disasters. This discipline of transportation looks at the causes of accidents, from the perspective of human, road, and vehicle and formulate plans for the reduction.

5. **Intelligent transport system** With advent to computers, communication, and vehicle technology, it is possible in these days to operate transportation system much effectively with significant reduction in the adverse impacts of transportation. Intelligent transportation system offers better mobility, efficiency, and safety with the help of the state-of-the-art-technology.

In addition disciplines specific to various modes are also common. This includes railway engineering, port and harbor engineering, and airport engineering.

1.5 Summary

Transportation engineering is a very diverse and multidisciplinary field, which deals with the planning, design, operation and maintenance of transportation systems. Good transportation is that which provides safe, rapid, comfortable, convenient, economical, and environmentally compatible movement of both goods and people. This profession carries a distinct societal responsibility. Transportation planners and engineers recognize the fact that transportation systems constitute a potent force in shaping the course of regional development. Planning and development of transportation facilities generally raises living standards and enhances the aggregate of community values.

1.6 Problems

1. Which analysis helps in finding the monetary returns from a project?
   - (a) Accident analysis
   - (b) Financial and economic analysis
   - (c) Intelligent transport system

   (d) Environmental impact assessment

2. The study of the transportation system that meets the travel need of several people by sharing a vehicle is
   (a) Mass transportation
   (b) Intelligent transport system
   (c) Passenger transport
   (d) None of the above

1.7 Solutions

1. Which analysis helps in finding the monetary returns from a project?
   (a) Accident analysis
   (b) Financial and economic analysis$\surd$
   (c) Intelligent transport system
   (d) Environmental impact assessment
2. The study of the transportation system that meets the travel need of several people by sharing a vehicle is
   (a) Mass transportation$\surd$
   (b) Intelligent transport system
   (c) Passenger transport
   (d) None of the above

Chapter 2

Introduction to Highway Engineering

2.1 Overview

Road transport is one of the most common mode of transport. Roads in the form of trackways, human pathways etc. were used even from the pre-historic times. Since then many experiments were going on to make the riding safe and comfort. Thus road construction became an inseparable part of many civilizations and empires. In this chapter we will see the different generations of road and their characteristic features. Also we will discuss about the highway planning in India.

2.2 History of highway engineering

The history of highway enginnering gives us an idea about the roads of ancient times. Roads in Rome were constructed in a large scale and it radiated in many directions helping them in military operations. Thus they are considered to be pioneers in road construction. In this section we will see in detail about Ancient roads, Roman roads, British roads, French roads etc.

2.2.1 Ancient Roads

The first mode of transport was by foot. These human pathways would have been developed for specific purposes leading to camp sites, food, streams for drinking water etc. The next major mode of transport was the use of animals for transporting both men and materials. Since these loaded animals required more horizontal and vertical clearances than the walking man, track ways emerged. The invention of wheel in Mesopotamian civilization led to the development of animal drawn vehicles. Then it became necessary that the road surface should be capable of carrying greater loads. Thus roads with harder surfaces emerged. To provide adequate strength to carry the wheels, the new ways tended to follow the sunny drier side of a path. These have led to the development of foot-paths. After the invention of wheel, animal drawn vehicles were developed and the need for hard surface road emerged. Traces of such hard roads were obtained from various ancient civilization dated as old as 3500 BC. The earliest authentic record of road was found from Assyrian empire constructed about 1900 BC.

2.2.2 Roman roads

The earliest large scale road construction is attributed to Romans who constructed an extensive system of roads radiating in many directions from Rome. They were a remarkable achievement and provided travel times across

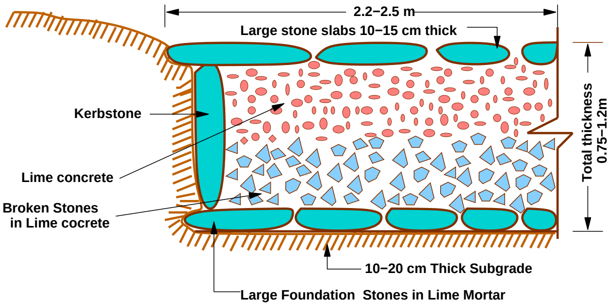


Figure 2:1: Roman roads

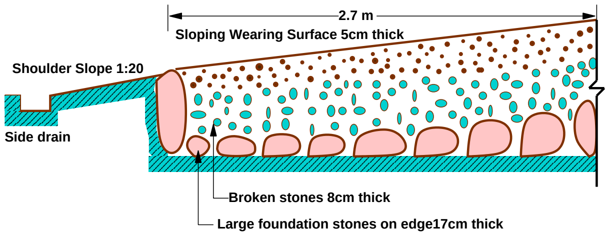


Figure 2:2: French roads

Europe, Asia minor, and north Africa. Romans recognized that the fundamentals of good road construction were to provide good drainage, good material and good workmanship. Their roads were very durable, and some are still existing. Roman roads were always constructed on a firm - formed subgrade strengthened where necessary with wooden piles. The roads were bordered on both sides by longitudinal drains. The next step was the construction of the *agger*. This was a raised formation up to a 1 meter high and 15 m wide and was constructed with materials excavated during the side drain construction. This was then topped with a sand leveling course. The agger contributed greatly to moisture control in the pavement. The pavement structure on the top of the agger varied greatly. In the case of heavy traffic, a surface course of large 250 mm thick hexagonal flag stones were provided. A typical cross section of roman road is given in Figure 2:1 The main features of the Roman roads are that they were built straight regardless of gradient and used heavy foundation stones at the bottom. They mixed lime and volcanic puzzolana to make mortar and they added gravel to this mortar to make concrete. Thus concrete was a major Roman road making innovation.

2.2.3 French roads

The next major development in the road construction occurred during the regime of Napoleon. The significant contributions were given by Tresaguet in 1764 and a typical cross section of this road is given in Figure 2:2. He developed a cheaper method of construction than the lavish and locally unsuccessful revival of Roman practice. The pavement used 200 mm pieces of quarried stone of a more compact form and shaped such that they had at least one flat side which was placed on a compact formation. Smaller pieces of broken stones were then compacted into the spaces between larger stones to provide a level surface. Finally the running layer was made with a layer of 25 mm sized broken stone. All this structure was placed in a trench in order to keep the running surface level with the surrounding country side. This created major drainage problems which were counteracted by making the surface as impervious as possible, cambering the surface and providing deep side ditches. He gave

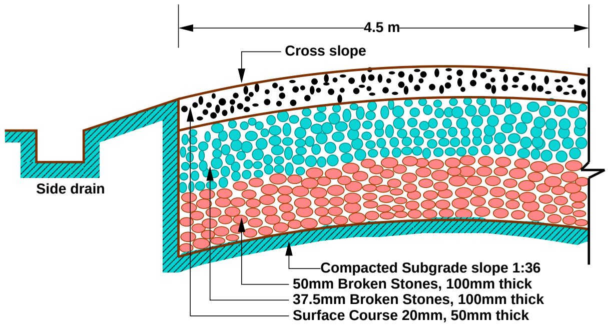


Figure 2:3: British roads

much importance for drainage. He also enunciated the necessity for continuous organized maintenance, instead of intermittent repairs if the roads were to be kept usable all times. For this he divided the roads between villages into sections of such length that an entire road could be covered by maintenance men living nearby.

2.2.4 British roads

The British government also gave importance to road construction. The British engineer John Macadam introduced what can be considered as the first scientific road construction method. Stone size was an important element of Macadam recipe. By empirical observation of many roads,he came to realize that 250 mm layers of well compacted broken angular stone would provide the same strength and stiffness and a better running surface than an expensive pavement founded on large stone blocks. Thus he introduced an economical method of road construction.

The mechanical interlock between the individual stone pieces provided strength and stiffness to the course. But the inter particle friction abraded the sharp interlocking faces and partly destroy the effectiveness of the course. This effect was overcome by introducing good quality interstitial finer material to produce a well-graded mix. Such mixes also proved less permeable and easier to compact. A typical cross section of British roads is given in Figure 2:3.

2.2.5 Modern roads

The modern roads by and large follow Macadam's construction method. Use of bituminous concrete and cement concrete are the most important developments. Various advanced and cost-effective construction technologies are used. Development of new equipments help in the faster construction of roads. Many easily and locally available materials are tested in the laboratories and then implemented on roads for making economical and durable pavements.

Scope of transportation system has developed very largely. Population of the country is increasing day by day. The life style of people began to change. The need for travel to various places at faster speeds also increased. This increasing demand led to the emergence of other modes of transportation like railways and travel by air. While the above development in public transport sector was taking place,the development in private transport was at a much faster rate mainly because of its advantages like accessibility, privacy, flexibility, convenience and comfort. This led to the increase in vehicular traffic especially in private transport network. Thus road space available was becoming insufficient to meet the growing demand of traffic and congestion started. In addition, chances for accidents also increased. This has led to the increased attention towards control of vehicles so that the transport infrastructure was optimally used. Various control measures like traffic signals, providing roundabouts and medians, limiting the speed of vehicle at specific zones etc. were implemented.

With the advancement of better roads and efficient control, more and more investments were made in the road sector especially after the World wars. These were large projects requiring large investment. For optimal utilization of funds, one should know the travel pattern and travel behavior. This has led to the emergence of transportation planning and demand management.

2.3 Highway planning in India

Excavations in the sites of Indus valley, Mohenjo-dero and Harappan civilizations revealed the existence of planned roads in India as old as 2500-3500 BC. The Mauryan kings also built very good roads. Ancient books like *Arthashastra* written by Kautilya, a great administrator of the Mauryan times, contained rules for regulating traffic, depths of roads for various purposes, and punishments for obstructing traffic.

During the time of Mughal period, roads in India were greatly improved. Roads linking North-West and the Eastern areas through gangetic plains were built during this time.

After the fall of the Mughals and at the beginning of British rule, many existing roads were improved. The construction of Grand-Trunk road connecting North and South is a major contribution of the British. However, the focus was later shifted to railways, except for feeder roads to important stations.

2.3.1 Modern developments

The first World war period and that immediately following it found a rapid growth in motor transport. So need for better roads became a necessity. For that, the Government of India appointed a committee called Road development Committee with Mr.M.R. Jayakar as the chairman. This committee came to be known as Jayakar committee.

Jayakar Committee

In 1927 Jayakar committee for Indian road development was appointed. The major recommendations and the resulting implementations were:

- Committee found that the road development of the country has become beyond the capacity of local governments and suggested that Central government should take the proper charge considering it as a matter of national interest.
- They gave more stress on long term planning programme, for a period of 20 years (hence called twenty year plan) that is to formulate plans and implement those plans with in the next 20 years.
- One of the recommendations was the holding of periodic road conferences to discuss about road construction and development. This paved the way for the establishment of a semi-official technical body called Indian Road Congress (IRC) in 1934
- The committee suggested imposition of additional taxation on motor transport which includes duty on motor spirit, vehicle taxation, license fees for vehicles plying for hire. This led to the introduction of a development fund called Central road fund in 1929. This fund was intended for road development.

- A dedicated research organization should be constituted to carry out research and development work. This resulted in the formation of Central Road Research Institute (CRRI) in 1950.

Nagpur road congress 1943

The second World War saw a rapid growth in road traffic and this led to the deterioration in the condition of roads. To discuss about improving the condition of roads, the government convened a conference of chief engineers of provinces at Nagpur in 1943. The result of the conference is famous as the Nagpur plan.

- A twenty year development programme for the period (1943-1963) was finalized. It was the first attempt to prepare a co-ordinated road development programme in a planned manner.
- The roads were divided into four classes:
  - National highways which would pass through states, and places having national importance for strategic, administrative and other purposes.
  - State highways which would be the other main roads of a state.
  - District roads which would take traffic from the main roads to the interior of the district . According to the importance, some are considered as *major district roads* and the remaining as *other district roads*.
  - Village roads which would link the villages to the road system.
- The committee planned to construct 2 lakh kms of road across the country within 20 years.
- They recommended the construction of star and grid pattern of roads throughout the country.
- One of the objective was that the road length should be increased so as to give a road density of 16kms per 100 sq.km

Bombay road congress 1961

The length of roads envisaged under the Nagpur plan was achieved by the end of it, but the road system was deficient in many respects. The changed economic, industrial and agricultural conditions in the country warranted a review of the Nagpur plan. Accordingly a 20-year plan was drafted by the Roads wing of Government of India, which is popularly known as the Bombay plan. The highlights of the plan were:

- It was the second 20 year road plan (1961-1981)
- The total road length targeted to construct was about 10 lakhs.
- Rural roads were given specific attention. Scientific methods of construction was proposed for the rural roads. The necessary technical advice to the Panchayaths should be given by State PWD's.
- They suggested that the length of the road should be increased so as to give a road density of 32kms/100 sq.km
- The construction of 1600 km of expressways was also then included in the plan.

Lucknow road congress 1984

This plan has been prepared keeping in view the growth pattern envisaged in various fields by the turn of the century. Some of the salient features of this plan are as given below:

- This was the third 20 year road plan (1981-2001). It is also called *Lucknow road plan.*
- It aimed at constructing a road length of 12 lakh kilometres by the year 1981 resulting in a road density of 82kms/100 sq.km
- The plan has set the target length of NH to be completed by the end of seventh, eighth and ninth five year plan periods.
- It aims at improving the transportation facilities in villages, towns etc. such that no part of country is farther than 50 km from NH.
- One of the goals contained in the plan was that expressways should be constructed on major traffic corridors to provide speedy travel.
- Energy conservation, environmental quality of roads and road safety measures were also given due importance in this plan.

2.4 Summary

This lecture cover a brief history of highway engineering, highlighting the developments of road construction. Significant among them are Roman, French, and British roads. British road construction practice developed by Macadam is the most scientific and the present day roads follows this pattern. The highway development and classification of Indian roads are also discussed. The major classes of roads include National Highway, State highway, District roads, and Village roads. Finally, issues in highway alignment are discussed.

2.5 Problems

1. Approximate length of National highway in India is:
   (a) 1000 km
   (b) 5000 km
   (c) 10000 km
   (d) 50000 km
   (e) 100000 km
2. The most accessible road is
   (a) National highway
   (b) State highway
   (c) Major District road
   (d) Other District road
   (e) Village road

2.6 Solutions

1. Approximate length of National highway in India is:
   (a) 1000 km
   (b) 5000 km
   (c) 10000 km
   (d) 50000 km $\surd$
   (e) 100000 km
2. The most accessible road is
   (a) National highway
   (b) State highway
   (c) Major District road
   (d) Other District road
   (e) Village road $\surd$

Chapter 3

Role of transportation in society

3.1 Overview

Transportation is a non separable part of any society. It exhibits a very close relation to the style of life, the range and location of activities and the goods and services which will be available for consumption. Advances in transportation has made possible changes in the way of living and the way in which societies are organized and therefore have a great influence in the development of civilizations. This chapter conveys an understanding of the importance of transportation in the modern society by presenting selected characteristics of existing transportation systems, their use and relationships to other human activities.

Transportation is responsible for the development of civilizations from very old times by meeting travel requirement of people and transport requirement of goods. Such movement has changed the way people live and travel. In developed and developing nations, a large fraction of people travel daily for work,shopping and social reasons. But transport also consumes a lot of resources like time,fuel, materials and land.

3.2 Economic role of transportation

Economics involves production, distribution and consumption of goods and services. People depend upon the natural resources to satisfy the needs of life but due to non uniform surface of earth and due to difference in local resources, there is a lot of difference in standard of living in different societies. So there is an immense requirement of transport of resources from one particular society to other. These resources can range from material things to knowledge and skills like movement of doctors and technicians to the places where there is need of them.

3.2.1 The place, time, quality and utility of goods

An example is given to evaluate the relationship between place, time and cost of a particular commodity. If a commodity is produced at point A and wanted by people of another community at any point B distant x from A, then the price of the commodity is dependent on the distance between two centers and the system of transportation between two points. With improved system the commodity will be made less costly at B.

3.2.2 Changes in location of activities

The reduction of cost of transport does not have same effect on all locations. Let at any point B the commodity is to be consumed. This product is supplied by two stations A and K which are at two different distances from B. Let at present the commodity is supplied by A since it is at a lesser distance but after wards due to improvement in road network between B and K,the point K becomes the supply point of product.

3.2.3 Conclusions

- Transport extends the range of sources of supply of goods to be consumed in an area, making it possible for user to get resources at cheap price and high quality.
- The use of more efficient systems of supply results in an increase in the total amount of goods available for consumption.
- Since the supply of goods is no longer dependent on the type of mode, items can be supplied by some alternative resources if usual source cannot supply what is needed.

3.3 Social role of transportation

Transportation has always played an important role in influencing the formation of urban societies. Although other facilities like availability of food and water, played a major role, the contribution of transportation can be seen clearly from the formation, size and pattern, and the development of societies, especially urban centers.

3.3.1 Formation of settlements

From the beginning of civilization, the man is living in settlements which existed near banks of major river junctions, a port, or an intersection of trade routes. Cities like New York, Mumbai and Moscow are good examples.

3.3.2 Size and pattern of settlements

The initial settlements were relatively small developments but with due course of time, they grew in population and developed into big cities and major trade centers. The size of settlements is not only limited by the size of the area by which the settlement can obtain food and other necessities, but also by considerations of personal travels especially the journey to and from work. The increased speed of transport and reduction in the cost of transport have resulted in variety of spatial patterns.

3.3.3 Growth of urban centers

When the cities grow beyond normal walking distance, then transportation technology plays a role in the formation of the city. For example, many cities in the plains developed as a circular city with radial routes, where as the cities beside a river developed linearly. The development of automobiles, and other factors like increase in personal income, and construction of paved road network, the settlements were transformed into urban centers of intense travel activity.

3.4 Political role of transportation

The world is divided into numerous political units which are formed for mutual protection, economic advantages and development of common culture. Transportation plays an important role in the functioning of such political units.

3.4.1 Administration of an area

The government of an area must be able to send/get information to/about its people. It may include laws to be followed, security and other needful information needed to generate awareness. An efficient administration of a country largely depends on how effectively government could communicate these information to all the country. However, with the advent of communications, its importance is slightly reduced.

3.4.2 Political choices in transport

These choices may be classified as communication, military movement, travel of persons and movement of freight. The primary function of transportation is the transfer of messages and information. It is also needed for rapid movement of troops in case of emergency and finally movement of persons and goods. The political decision of construction and maintenance of roads has resulted in the development of transportation system.

3.5 Environmental role of transportation

The negative effects of transportation is more dominating than its useful aspects as far as transportation is concerned. There are numerous categories into which the environmental effects have been categorized. They are explained in the following sections.

3.5.1 Safety

Growth of transportation has a very unfortunate impact on the society in terms of accidents. Worldwide death and injuries from road accidents have reached epidemic proportions. -killed and about 15 million injured on the road accidents annually. Increased variation in the speeds and vehicle density resulted in a high exposure to accidents. Accidents result in loss of life and permanent disability, injury, and damage to property. Accidents also causes numerous non-quantifiable impacts like loss of time, grief to the near ones of the victim, and inconvenience to the public. The loss of life and damage from natural disasters, industrial accidents, or epidemic often receive significant attention from both government and public. This is because their occurrence is concentrated but sparse. On the other hand, accidents from transport sector are widespread and occurs with high frequency.

For instance, a study has predicted that death and disabilities resulting from road accidents in comparison with other diseases will rise from ninth to third rank between 1990 and 2020. Road accidents as cause to death and disability could rank below heart disease and clinical depression, and ahead of stroke and all infectious diseases. Significant reduction to accident rate is achieved in the developing countries by improved road designed maintenance, improved vehicle design, driver education, and law enforcements. However in the developing nations, the rapid growth of personalized vehicles and poor infrastructure, road design, and law enforcement has resulted in growing accident rate.

3.5.2 Air Pollution

All transport modes consume energy and the most common source of energy is from the burning of fossil fuels like coal, petrol, diesel, etc. The relation between air pollution and respiratory disease have been demonstrated by various studies and the detrimental effects on the planet earth is widely recognized recently. The combustion of the fuels releases several contaminants into the atmosphere, including carbon monoxide, hydrocarbons, oxides of nitrogen, and other particulate matter. Hydrocarbons are the result of incomplete combustion of fuels. Particulate matters are minute solid or liquid particles that are suspended in the atmosphere. They include aerosols, smoke, and dust particles. These air pollutants once emitted into the atmosphere , undergo mixing and disperse into the surroundings.

3.5.3 Noise pollution

Sound is acoustical energy released into atmosphere by vibrating or moving bodies where as noise is unwanted sound produced. Transportation is a major contributor of noise pollution, especially in urban areas. Noise is generated during both construction and operation. During construction, operation of large equipments causes considerable noise to the neighborhood. During the operation, noise is generated by the engine and exhaust systems of vehicle, aerodynamic friction, and the interaction between the vehicle and the support system (road-tire, rail-wheel). Extended exposure to excessive sound has been shown to produce physical and psychological damage. Further, because of its annoyance and disturbance, noise adds to mental stress and fatigue.

3.5.4 Energy consumption

The spectacular growth in industrial and economic growth during the past century have been closely related to an abundant supply of inexpensive energy from fossil fuels. Transportation sector is unbelieved to consume more than half of the petroleum products. The compact of the shortage of fuel was experienced during major wars when strict rationing was imposed in many countries. The impact of this had cascading effects on many factors of society, especially in the price escalation of essential commodities. However, this has few positive impacts; a shift to public transport system, a search for energy efficient engines, and alternate fuels. During the time of fuel shortage, people shifted to cheaper public transport system. Policy makers and planners, thereafter gave much emphasis to the public transit which consume less energy per person. The second impact was in the development of fuel-efficient engines and devices and operational and maintenance practices. A fast depleting fossil fuel has accelerated the search for energy efficient and environment friendly alternate energy source. The research is active in the development of bio-fuels, hydrogen fuels and solar energy.

3.5.5 Other impacts

Transportation directly or indirectly affects many other areas of society and few of then are listed below:

Almost all cities uses 20-30 percent of its land in transport facilities. Increased travel requirement also require additional land for transport facilities. A good transportation system takes considerable amount of land from the society.

Aesthetics of a region is also affected by transportation. Road networks in quite country side is visual intrusion. Similarly, the transportation facilities like fly-overs are again visual intrusion in urban context.

The social life and social pattern of a community is severely affected after the introduction of some transportation facilities. Construction of new transportation facilities often require substantial relocation of residents and employment opportunities.

3.6 Summary

The roles of transportation in society can be classified according to economic, social, political and environmental roles. The social role of transport has caused people to live in permanent settlements and has given chances of sustainable developments. Regarding political role, large areas can now be very easily governed with the help of good transportation system. The environmental effects are usually viewed negatively.

3.7 Problems

1. Safety criteria of transportation is viewed under
   (a) Political role of transportation
   (b) Environmental role of transportation
   (c) Social role of transportation
   (d) None of these
2. Which of the following is not a negative impact of transportation?
   (a) Safety
   (b) Aesthetics
   (c) Mobility
   (d) Pollution

3.8 Solutions

1. Safety criteria of transportation is viewed under
   (a) Political role of transportation
   (b) Environmental role of transportation$\surd$
   (c) Social role of transportation
   (d) None of these
2. Which of the following is not a negative impact of transportation?
   (a) Safety
   (b) Aesthetics
   (c) Mobility$\surd$
   (d) Pollution

Chapter 4

Factors affecting transportation

4.1 Overview

The success of transportation engineering depends upon the co-ordination between the three primary elements, namely the vehicles, the roadways, and the road users. Their characteristics affect the performance of the transportation system and the transportation engineer should have fairly good understanding about them. This chapter elaborated salient human, vehicle, and road factors affecting transportation.

4.2 Human factors affecting transportation

Road users can be defined as drivers, passengers, pedestrians etc. who use the streets and highways. Together, they form the most complex element of the traffic system - the human element - which differentiates Transportation Engineering from all other engineering fields. It is said to be the most complex factor as the human performances varies from individual to individual. Thus, the transportation engineer should deal with a variety of road user characteristics. For example, a traffic signal timed to permit an average pedestrian to cross the street safely may cause a severe hazard to an elderly person. Thus, the design considerations should safely and efficiently accommodate the elderly persons, the children, the handicapped, the slow and speedy, and the good and bad drivers.

4.2.1 Variability

The most complex problem while dealing human characteristics is its variability. The human characteristics like ability to react to a situation, vision and hearing, and other physical and psychological factors vary from person to person and depends on age, fatigue, nature of stimuli, presence of drugs/alcohol etc. The influence of all these factors and the corresponding variability cannot be accounted when a facility is designed. So a standardized value is often used as the design value. The 85^{th} percentile value of different characteristics is taken as a standard. It represents a characteristic that 85 per percent of the population can meet or exceed. For example. if we say that the 85^{th} percentile value of walking speed is about 2 m/s, it means that 85 per cent of people has walking speed faster than 2 m/s. The variability is thus fixed by selecting proper 85th percentile values of the characteristics.

4.2.2 Critical characteristics

The road user characteristics can be of two main types, some of them are quantifiable like reaction time, visual acuity etc. while some others are less quantifiable like the psychological factors, physical strength, fatigue, and dexterity.

4.2.3 Reaction time

The road user is subjected to a series of stimuli both expected and unexpected. The time taken to perform an action according to the stimulus involves a series of stages like:

- **Perception**: Perception is the process of perceiving the sensations received through the sense organs, nerves and brains. It is actually the recognitions that a stimulus on which a reaction is to happen exists.
- **Intellection**: Intellection involves the identification and understanding of stimuli.
- **Emotion**: This stage involves the judgment of the appropriate response to be made on the stimuli like to stop, pass, move laterally etc.
- **Volition**: Volition is the execution of the decision which is the result of a physical actions of the driver.

For example., if a driver approaches an intersection where the signal is red, the driver first sees the signal (perception), he recognizes that is is a red/STOP signal, he decides to stop and finally applies the brake(volition). This sequence is called the PIEV time or perception-reaction time. But apart from the above time, the vehicle itself traveling at initial speed would require some more time to stop. That is, the vehicle traveling with initial speed u will travel for a distance, $d = vt$ where, t is the above said PIEV time. Again, the vehicle would travel some distance after the brake is applied.

4.2.4 Visual acuity and driving

The perception-reaction time depends greatly on the effectiveness of drivers vision in perceiving the objects and traffic control measures. The PIEV time will be decreased if the vision is clear and accurate. Visual acuity relates to the field of clearest vision. The most acute vision is within a cone of 3 to 5 degrees, fairly clear vision within 10 to 12 degrees and the peripheral vision will be within 120 to 180 degrees. This is important when traffic signs and signals are placed, but other factors like dynamic visual acuity, depth perception etc. should also be considered for accurate design. Glare vision and color vision are also equally important. Glare vision is greatly affected by age. Glare recovery time is the time required to recover from the effect of glare after the light source is passed, and will be higher for elderly persons. Color vision is important as it can come into picture in case of sign and signal recognition.

4.2.5 Walking

Transportation planning and design will not be complete if the discussion is limited to drivers and vehicular passengers. The most prevalent of the road users are the pedestrians. Pedestrian traffic along footpaths, sidewalks, crosswalks, safety zones, islands, and over and under passes should be considered. On an average, the pedestrian walking speed can be taken between 1.5 m/sec to 2 m/sec. But the influence of physical, mental, and emotional factors need to be considered. Parking spaces and facilities like signals, bus stops, and over and under passes are to be located and designed according to the maximum distance to which a user will be willing to walk. It was seen that in small towns 90 per cent park within 185 m of their destinations while only 66 per cent park so close in large city.

4.2.6 Other Characteristics

Hearing is required for detecting sounds, but lack of hearing acuity can be compensated by usage of hearing aids. Lot of experiments were carried out to test the drive vigilance which is the ability of a drive to discern environmental signs over a prolonged period. The results showed that the drivers who did not undergo any type of fatiguing conditions performed significantly better than those who were subjected to fatiguing conditions. But the mental fatigue is more dangerous than skill fatigue. The variability of attitude of drivers with respect to age, sex, knowledge and skill in driving etc. are also important.

Two of the important constituents of transportation system are drivers and users/passengers. Understanding of certain human characteristics like perception - reaction time and visual acuity and their variability are to be considered by Traffic Engineer. Because of the variability in characteristics, the 85^{Th} percentile values of the human characteristics are fixed as standards for design of traffic facilities.

4.3 Vehicle factors

It is important to know about the vehicle characteristics because we can design road for any vehicle but not for an indefinite one. The road should be such that it should cater to the needs of existing and anticipated vehicles. Some of the vehicle factors that affect transportation is discussed below.

4.3.1 Design vehicles

Highway systems accommodate a wide variety of sizes and types of vehicles, from smallest compact passenger cars to the largest double and triple tractor-trailer combinations. According to the different geometric features of highways like the lane width, lane widening on curves, minimum curb and corner radius, clearance heights etc some standard physical dimensions for the vehicles has been recommended. Road authorities are forced to impose limits on vehicular characteristics mainly:

- to provide practical limits for road designers to work to,
- to see that the road space and geometry is available to normal vehicles,
- to implement traffic control effectively and efficiently,
- take care of other road users also.

Taking the above points into consideration, in general, the vehicles can be grouped into motorized two wheeler's, motorized three wheeler's, passenger car, bus, single axle trucks, multi axle trucks, truck trailer combinations, and slow non motorized vehicles.

4.3.2 Vehicle dimensions

The vehicular dimensions which can affect the road and traffic design are mainly: width, height, length, rear overhang, and ground clearance. The width of vehicle affects the width of lanes, shoulders and parking facility. The capacity of the road will also decrease if the width exceeds the design values. The height of the vehicle affects the clearance height of structures like over-bridges, under-bridges and electric and other service lines and also placing of signs and signals. Another important factor is the length of the vehicle which affects the extra width of pavement, minimum turning radius, safe overtaking distance, capacity and the parking facility. The rear overhang control is mainly important when the vehicle takes a right/left turn from a stationary point. The ground clearance of vehicle comes into picture while designing ramps and property access and as bottoming out on a crest can stop a vehicle from moving under its own pulling power.

4.3.3 Weight, axle configuration etc.

The weight of the vehicle is a major consideration during the design of pavements both flexible and rigid. The weight of the vehicle is transferred to the pavement through the axles and so the design parameters are fixed on the basis of the number of axles. The power to weight ratio is a measure of the ease with which a vehicle can move. It determines the operating efficiency of vehicles on the road. The ratio is more important for heavy vehicles. The power to weight ratio is the major criteria which determines the length to which a positive gradient can be permitted taking into consideration the case of heavy vehicles.

4.3.4 Turning radius and turning path

The minimum turning radius is dependent on the design and class of the vehicle. The effective width of the vehicle is increased on a turning. This is also important at an intersection, round about, terminals, and parking areas.

4.3.5 Visibility

The visibility of the driver is influenced by the vehicular dimensions. As far as forward visibility is concerned, the dimension of the vehicle and the slope and curvature of wind screens, windscreen wipers, door pillars, etc should be such that:

- visibility is clear even in bad weather conditions like fog, ice, and rain;
- it should not mask the pedestrians, cyclists or other vehicles;
- during intersection maneuvers.

Equally important is the side and rear visibility when maneuvering especially at intersections when the driver adjusts his speed in order to merge or cross a traffic stream. Rear vision efficiency can be achieved by properly positioning the internal or external mirrors.

4.4 Acceleration characteristics

The acceleration capacity of vehicle is dependent on its mass, the resistance to motion and available power. In general, the acceleration rates are highest at low speeds, decreases as speed increases. Heavier vehicles have lower rates of acceleration than passenger cars. The difference in acceleration rates becomes significant in mixed traffic streams. For example, heavy vehicles like trucks will delay all passengers at an intersection. Again, the gaps formed can be occupied by other smaller vehicles only if they are given the opportunity to pass. The presence of upgrades make the problem more severe. Trucks are forced to decelerate on grades because their power is not sufficient to maintain their desired speed. As trucks slow down on grades, long gaps will be formed in the traffic stream which cannot be efficiently filled by normal passing maneuvers.

4.5 Braking performance

As far as highway safety is concerned, the braking performance and deceleration characteristics of vehicles are of prime importance. The time and distance taken to stop the vehicle is very important as far as the design of various traffic facilities are concerned. The factors on which the braking distance depend are the type of the road and its condition, the type and condition of tire and type of the braking system. The distance to decelerate from one speed to another is given by:

$$d = \frac{v^2 - u^2}{f + g} \tag{4.1}$$

where d is the braking distance, v and u are the initial and final speed of the vehicle, f is the coefficient of forward rolling and skidding friction and g is the grade in decimals. The main characteristics of a traffic system influenced by braking and deceleration performance are:

- Safe stopping sight distance: The minimum stopping sight distance includes both the reaction time and the distance covered in stopping. Thus, the driver should see the obstruction in time to react to the situation and stop the vehicle.
- Clearance and change interval: The Clearance and change intervals are again related to safe stopping distance. All vehicles at a distance further away than one stopping sight distance from the signal when the Yellow is flashed is assumed to be able to stop safely. Such a vehicle which is at a distance equal or greater than the stopping sight distance will have to travel a distance equal to the stopping sight distance plus the width of the street, plus the length of the vehicle. Thus the yellow and all red times should be calculated to accommodate the safe clearance of those vehicles.
- Sign placement: The placement of signs again depends upon the stopping sight distance and reaction time of drivers. The driver should see the sign board from a distance at least equal to or greater than the stopping sight distance.

From the examples discussed above, it is clear that the braking and reaction distance computations are very important as far as a transportation system is concerned. Stopping sight distance is a product of the characteristics of the driver, the vehicle and the roadway. and so this can vary with drivers and vehicles. Here the concept of design vehicles gains importance as they assist in general design of traffic facilities thereby enhancing the safety and performance of roadways.

4.6 Road factors

4.6.1 Road surface

The type of pavement is determined by the volume and composition of traffic, the availability of materials, and available funds. Some of the factors relating to road surface like road roughness, tire wear, tractive resistance, noise, light reflection, electrostatic properties etc. should be given special attention in the design, construction and maintenance of highways for their safe and economical operation. Unfortunately, it is impossible to build road surface which will provide the best possible performance for all these conditions. For heavy traffic volumes, a smooth riding surface with good all-weather anti skid properties is desirable. The surface should be chosen to retain these qualities so that maintenance cost and interference to traffic operations are kept to a minimum.

4.6.2 Lighting

Illumination is used to illuminate the physical features of the road way and to aid in the driving task. A luminaire is a complete lighting device that distributes light into patterns much as a garden hose nozzle distributes water. Proper distribution of the light flux from luminaires is one of the essential factors in efficient roadway lighting. It is important that roadway lighting be planned on the basis of many traffic information such as night vehicular traffic, pedestrian volumes and accident experience.

4.6.3 Roughness

This is one of the main factors that an engineer should give importance during the design, construction, and maintenance of a highway system. Drivers tend to seek smoother surface when given a choice. On four-lane highways where the texture of the surface of the inner-lane is rougher than that of the outside lane, passing vehicles tend to return to the outside lane after execution of the passing maneuver. Shoulders or even speed-change lanes may be deliberately roughened as a means of delineation.

4.6.4 Pavement colors

When the pavements are light colored(for example, cement concrete pavements) there is better visibility during day time whereas during night dark colored pavements like bituminous pavements provide more visibility. Contrasting pavements may be used to indicate preferential use of traffic lanes. A driver tends to follow the same pavement color having driven some distance on a light or dark surface, he expects to remain on a surface of that same color until he arrives a major junction point.

4.6.5 Night visibility

Since most accidents occur at night because of reduced visibility, the traffic designer must strive to improve nighttime visibility in every way he can. An important factor is the amount of light which is reflected by the road surface to the drivers' eyes. Glare caused by the reflection of oncoming vehicles is negligible on a dry pavement but is an important factor when the pavement is wet.

4.6.6 Geometric aspects

The roadway elements such as pavement slope, gradient, right of way etc affect transportation in various ways. Central portion of the pavement is slightly raised and is sloped to either sides so as to prevent the ponding of water on the road surface. This will deteriorate the riding quality since the pavement will be subjected to many failures like potholes etc. Minimum lane width should be provided to reduce the chances of accidents. Also the speed of the vehicles will be reduced and time consumed to reach the destination will also be more. Right of way width should be properly provided. If the right of way width becomes less, future expansion will become difficult and the development of that area will be adversely affected. One important other road element is the gradient. It reduces the tractive effort of large vehicles. Again the fuel consumption of the vehicles climbing a gradient is more. The other road element that cannot be avoided are curves. Near curves, chances of accidents are more. Speed of the vehicles is also affected.

4.7 Summary

The performance, design and operation of a transportation system is affected by several factors such as human factors, vehicle factors, acceleration characteristics, braking performance etc. These factors greatly influence the geometric design as well as design of control facilities. Variant nature of the driver, vehicle, and roadway characteristics should be given importance for the smooth, safe, and efficient performance of traffic in the road.

4.8 Problems

1. What is the standard percentile value taken for fixing the variability of human characteristics?
   - (a) 85^{th} percentile
   - (b) 90^{th} percentile
   - (c) 95^{th} percentile
   - (d) 80^{th} percentile
2. The range of fairly clear vision is within
   - (a) $7°$ to $8°$
   - (b) $3°$ to $5°$
   - (c) $10°$ to $12°$
   - (d) $25°$ to $45°$

4.9 Solutions

1. What is the standard percentile value taken for fixing the variability of human characteristics?
   - (a) 85^{th} percentile$\surd$
   - (b) 90^{th} percentile
   - (c) 95^{th} percentile
   - (d) 80^{th} percentile
2. The range of fairly clear vision is within
   - (a) $7°$ to $8°$
   - (b) $3°$ to $5°$
   - (c) $10°$ to $12°$$\surd$
   - (d) $25°$ to $45°$

Chapter 5

Travel demand modeling

5.1 Overview

This chapter provides an introduction to travel demand modeling, the most important aspect of transportation planning. First we will discuss about what is modeling, the concept of transport demand and supply, the concept of equilibrium, and the traditional four step demand modeling. We may also point to advance trends in demand modeling.

5.2 Transport modeling

Modeling is an important part of any large scale decision making process in any system. There are large number of factors that affect the performance of the system. It is not possible for the human brain to keep track of all the player in system and their interactions and interrelationships. Therefore we resort to models which are some simplified, at the same time complex enough to reproduce key relationships of the reality. Modeling could be either physical, symbolic, or mathematical In physical model one would make physical representation of the reality. For example, model aircrafts used in wind tunnel is an example of physical models. In symbolic model, with the complex relations could be represented with the help of symbols. Drawing time-space diagram of vehicle movement is a good example of symbolic models. Mathematical model is the most common type when with the help of variables, parameters, and equations one could represent highly complex relations. Newton's equations of motion or Einstein's equation $E = mc^2$, can be considered as examples of mathematical model. No model is a perfect representation of the reality. The important objective is that models seek to isolate key relationships, and not to replicate the entire structure. Transport modeling is the study of the behavior of individuals in making decisions regarding the provision and use of transport. Therefore, unlike other engineering models, transport modeling tools have evolved from many disciplines like economics, psychology, geography, sociology, and statistics.

5.3 Transport demand and supply

The concept of demand and supply are fundamental to economic theory and is widely applied in the field to transport economics. In the area of travel demand and the associated supply of transport infrastructure, the notions of demand and supply could be applied. However, we must be aware of the fact that the transport demand is a *derived* demand, and not a need in itself. That is, people travel not for the sake of travel, but to practice in activities in different locations

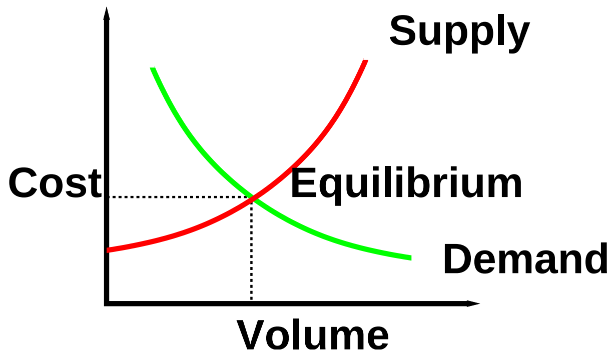


Figure 5:1: Demand supply equilibrium

The concept of equilibrium is central to the supply-demand analysis. It is a normal practice to plot the supply and demand curve as a function of cost and the intersection is then plotted in the equilibrium point as shown in Figure 5:1 The demand for travel T is a function of cost C is easy to conceive. The classical approach defines the supply function as giving the quantity T which would be produced, given a market price C. Since transport demand is a derived demand, and the benefit of transportation on the non-monetary terms(time in particular), the supply function takes the form in which C is the unit cost associated with meeting a demand T. Thus, the supply function encapsulates response of the transport system to a given level of demand. In other words, supply function will answer the question what will be the level of service of the system, if the estimated demand is loaded to the system. The most common supply function is the link travel time function which relates the link volume and travel time.

5.4 Travel demand modeling

Travel demand modeling aims to establish the spatial distribution of travel explicitly by means of an appropriate system of zones. Modeling of *demand* thus implies a procedure for predicting what travel decisions people would like to make given the generalized travel cost of each alternatives. The base decisions include the choice of destination, the choice of the mode, and the choice of the route. Although various modeling approaches are adopted, we will discuss only the classical transport model popularly known as four-stage model(FSM).

The general form of the four stage model is given in Figure 5:2. The classic model is presented as a sequence of four sub models: trip generation, trip distribution, modal split, trip assignment. The models starts with defining the study area and dividing them into a number of zones and considering all the transport network in the system. The database also include the current (base year) levels of population, economic activity like employment, shopping space, educational, and leisure facilities of each zone. Then the *trip generation* model is evolved which uses the above data to estimate the total number of trips generated and attracted by each zone. The next step is the allocation of these trips from each zone to various other destination zones in the study area using *trip distribution* models. The output of the above model is a trip matrix which denote the trips from each zone to every other zones. In the succeeding step the trips are allocated to different modes based on the modal attributes using the *modal split* models. This is essentially slicing the trip matrix for various modes generated to a mode specific trip matrix. Finally, each trip matrix is assigned to the route network of that particular mode using the *trip assignment models.* The step will give the loading on each link of the network.

The classical model would also be viewed as answering a series of questions (decisions) namely how many

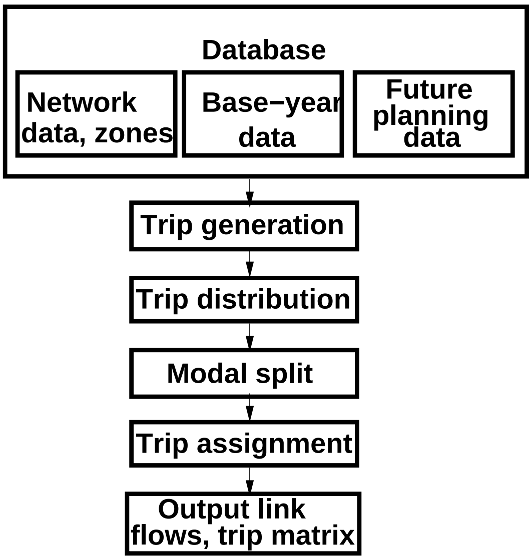


Figure 5:2: General form of the four stage modeling

trips are generated, where they are going, on what mode they are going, and finally which route they are adopting. The current approach is to model these decisions using discrete choice theory, which allows the lower level choices to be made conditional on higher choices. For example, route choice is conditional on the mode choice. This hierarchical choices of trip is shown in Figure 5:3 The highest level to find all the trips T_i originating from a zone is calculated based on the data and aggregate cost term C_i***. Based on the aggregate travel cost C_{ij}** from zone i to the destination zone j, the probability $p_{m|ij}$ of trips going to zone j is computed and subsequently the trips T_{ij}** from zone i to zone j by all modes and all routes are computed. Next, the mode choice model compute the probability $p_{m|ij}$ of choosing mode m based on the travel cost C_{jm}* from zone i to zone j, by mode m is determined. Similarly, the route choice gives the trips T_{ijmr} from zone i to zone j by mode m through route r can be computed. Finally the travel demand is loaded to the supply model, as stated earlier, will produce a performance level. The purpose of the network is usually measured in travel time which could be converted to travel cost. Although not practiced ideally, one could feed this back into the higher levels to achieve real equilibrium of the supply and demand.

5.5 Summary

In a nutshell, travel demand modeling aims at explaining where the trips come from and where they go,and what modes and which routes are used. It provides a zone wise analysis of the trips followed by distribution of the trips, split the trips mode wise based on the choice of the travelers and finally assigns the trips to the network. This process helps to understand the effects of future developments in the transport networks on the trips as well as the influence of the choices of the public on the flows in the network.

5.6 Problems

1. Link travel time function relates travel time and

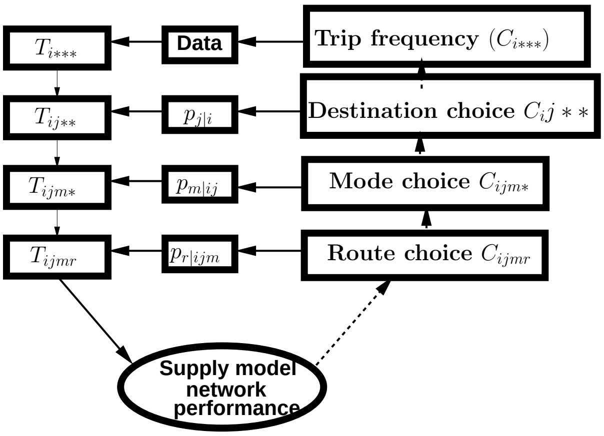


Figure 5:3: Demand supply equilibrium

(a) link volume

(b) link cost

(c) level of service

(d) none of the above

2. What is the first stage of four-stage travel demand modeling?

(a) Trip generation

(b) Trip distribution

(c) Modal split

(d) Traffic assignment

5.7 Solutions

1. Link travel time function relates travel time and

(a) link volume√

(b) link cost

(c) level of service

(d) none of the above

2. What is the first stage of four-stage travel demand modeling?

(a) Trip distribution

(b) Trip generation√

(c) Modal split

(d) Traffic assignment

Chapter 6

Data Collection

6.1 Overview

The four-stage modeling, an important tool for forecasting future demand and performance of a transportation system, was developed for evaluating large-scale infrastructure projects. Therefore, the four-stage modeling is less suitable for the management and control of existing software. Since these models are applied to large systems, they require information about travelers of the area influenced by the system. Here the data requirement is very high, and may take years for the data collection, data analysis, and model development. In addition, meticulous planning and systematic approach are needed for accurate data collection and processing. This chapter covers three important aspects of data collection, namely, survey design, household data collection, and data analysis. Finally, a brief discussion of other important surveys is also presented.

6.2 Survey design

Designing the data collection survey for the transportation projects is not easy. It requires considerable experience, skill, and a sound understanding of the study area. It is also important to know the purpose of the study and details of the modeling approaches, since data requirement is influenced by these. Further, many practical considerations like availability of time and money also has a strong bearing on the survey design. In this section, we will discuss the basic information required from a data collection, defining the study area, dividing the area into zones, and transport network characteristics.

6.2.1 Information needed

Typical information required from the data collection can be grouped into four categories, enumerated as below.

1. **Socio-economic data:** Information regarding the socio-economic characteristics of the study area. Important ones include income, vehicle ownership, family size, etc. This information is essential in building trip generation and modal split models.

2. **Travel surveys:** Origin-destination travel survey at households and traffic data from cordon lines and screen lines (defined later). Former data include the number of trips made by each member of the household, the direction of travel, destination, the cost of the travel, etc. The latter include the traffic flow, speed, and travel time measurements. These data will be used primarily for the calibration of the models, especially the trip distribution models.

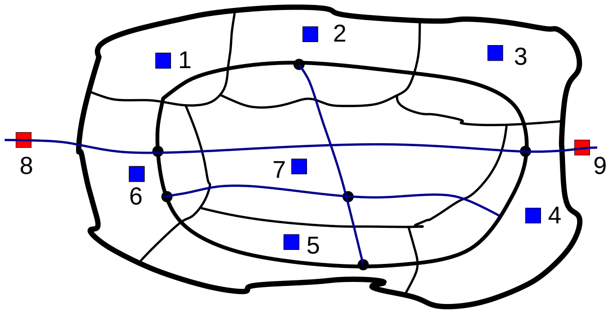


Figure 6:1: zoning of a study area

3. **Land use inventory:** This includes data on the housing density at residential zones, establishments at commercial and industrial zones. This data is especially useful for trip generation models.

4. **Network data:** This includes data on the transport network and existing inventories. Transport network data includes road network, traffic signals, junctions etc. The service inventories include data on public and private transport networks. These particulars are useful for the model calibration, especially for the assignment models.

6.2.2 Study area

Once the nature of the study is identified, the study area can be defined to encompass the area of expected policy impact. The study area need not be confirmed by political boundaries, but bounded by the area influenced by the transportation systems. The boundary of the study area is defined by what is called as *external cordon* or simply the cordon line. A sample of the zoning of a study area is shown in figure 6:1 Interactions with the area outside the cordon are defined via external stations which effectively serve as doorways to trips, into, out of, and through the study area. In short, study area should be defined such that majority of trips have their origin and destination in the study area and should be bigger than the area-of-interest covering the transportation project.

6.2.3 Zoning

Once the study area is defined, it is then divided into a number of small units called traffic analysis zones (TAZ) or simply *zones*. The zone with in the study area are called *internal zones*.

Zones are modeled as if all their attributes and properties were concentrated in a single point called the zone*centroid*. The centroids are connected to the nearest road junction or rail station by centroid connectors. Both centroid and centroid connectors are notional and it is assumed that all people have same travel cost from the centroid to the nearest transport facility which is the average for a zone. The intersection from outside world is normally represented through *external zones*. The external zones are defined by the catchment area of the major transport links feeding to the study area. Although the list is not complete, few guidelines are given below for selecting zones.

1. zones should match other administrative divisions, particularly census zones.

2. zones should have homogeneous characteristics, especially in land use, population etc.

3. zone boundaries should match cordon and screen lines, but should not match major roads.
4. zones should be as smaller in size as possible so that the error in aggregation caused by the assumption that all activities are concentrated at the zone centroids is minimum.

6.2.4 Network

Transport network consists of roads,junctions, bus stops, rails, railway station etc. Normally road network and rail network are represented separately. Road network is considered as directed graph of nodes and links. Each node and links have their own properties. Road link is normally represented with attributes like starting node, ending node, road length, free flow speed, capacity, number of lanes or road width, type of road like divided or undivided etc. Road junctions or nodes are represented with attributes like node number, starting nodes of all links joining the current node, type of intersection (uncontrolled, round about, signalized, etc.). Similarly public transport network like bus transit network and rail network are represented, but with attributes relevant to them. These may include frequency of service, fare of travel, line capacity, station capacity etc. This completes the inventory of base-year transportation facility.

6.3 Household data

To understand the behavior and factors affecting the travel, one has got the origin of travel when the decision for travel is made. It is where people live as family which is the household. Therefore household data is considered to be the most basic and authentic information about the travel pattern of a city.

Ideally one should take the details of all the people in the study to get complete travel details. However, this is not feasible due to large requirement of time and resources needed. In addition this will cause difficulties in handling these large data in modeling stage. Therefore, same sample households are randomly selected and survey is conducted to get the household data. Higher sample size is required fro large population size, and vice-versa. Normally minimum ten percent samples are required for population less than 50,000. But for a population more than one million require only one percent for the same accuracy.

6.3.1 Questionnaire design

The next step in the survey is the questionnaire design. A good design will ensure better response from the respondent and will significantly improve the quality of data. Design of questionnaire is more of an art than a science. However few guiding principles can be laid out. The questionnaire should be simple, direct, should take minimum time, and should cause minimum burden to the respondent. Traditional household survey has three major sections; household characteristics, personal characteristics, and trip details.

Household characteristics This section includes a set of questions designed to obtain socioeconomic information about the household. Relevant questions are:number of members in the house, no.of employed people, number of unemployed people, age and sex of the members in the house etc., number of two-wheelers in the house, number of cycles, number of cars in the house etc., house ownership and family income.

Personal characteristics This part includes questions designed to classify the household members(older than 5) according to the following aspects:relation to the head of the household (e.g. wife, son), sex, age, possession of a driving license, educational level, and activity.

Trip data This part of the survey aims at detecting and characterizing all trips made by the household members identified in the first part. A trip is normally defined as any movement greater than 300 meters from an origin to a destination with a given purpose. Trips are characterized on the basis of variables such as: origin and destination, trip purpose, trip start and ending times, mode used, walking distance, public-transport line and transfer station or bus stop (if applicable).

6.3.2 Survey administration

Once the questionnaire is ready, the next step is to conduct the actual survey with the help of enumerators. Enumerators has to be trained first by briefing them about the details of the survey and how to conduct the survey. They will be given random household addresses and the questionnaire set. They have to first get permission to be surveyed from the household. They may select a typical working day for the survey and ask the members of the household about the details required in the questionnaire. They may take care that each member of the household should answer about their own travel details, except for children below 12 years. Trip details of children below 5 years are normally ignored. Since the actual survey may take place any time during the day, the respondents are required to answer the question about the travel details of the previous day.

There are many methods of the administration of the survey and some of them are discussed below:

1. **Telephonic:** The enumerator may use telephone to fix an appointment and then conduct detailed telephonic interview. This is very popular in western countries where phone penetration is very high.

2. **Mail back:** The enumerator drops the questionnaire to the respondent and asks them to fill the details and mail them back with required information. Care should be taken to design the questionnaire so that it is self explanatory.

3. **Face-to-face** In this method, the enumerator visits the home of the respondent and asks the questions and fills up the questionnaire by himself. This is not a very socially acceptable method in the developed countries, as these are treated as intrusion to privacy. However, in many developed countries, especially with less educated people, this is the most effective method.

6.4 Data preparation

The raw data collected in the survey need to be processed before direct application in the model. This is necessary, because of various errors, except in the survey both in the selection of sample houses as well as error in filling details. In this section, we will discuss three aspects of data preparation; data correction, data expansion, and data validation.

6.4.1 Data correction

Various studies have identified few important errors that need to be corrected, and are listed below.

1. **Household size correction** It may be possible that while choosing the random samples, one may choose either larger or smaller than the average size of the population as observed in the census data and correction should be made accordingly.

2. **Socio-demographic corrections** It is possible that there may be differences between the distribution of the variables sex, age, etc. between the survey, and the population as observed from the census data. This correction is done after the household size correction.

3. **Non-response correction** It is possible that there may not be a response from many respondents, possible because they are on travel everyday. Corrections should be made to accommodate this, after the previous two corrections.

4. **Non-reported trip correction** In many surveys people underestimate the non-mandatory trips and the actual trips will be much higher than the reported ones. Appropriate correction need to be applied for this.

6.4.2 Sample expansion

The second step in the data preparation is to amplify the survey data in order to represent the total population of the zone. This is done with the help of expansion factor which is defined as the ratio of the total number of household addressed in the population to that of the surveyed. A simple expansion factor F_i for the zone i could be of the following form.

$$F_i = \frac{a}{b-d} \tag{6.1}$$

where a is the total number of household in the original population list, b is the total number of addresses selected as the original sample, and d is the number of samples where no response was obtained.

6.4.3 Validation of results

In order to have confidence on the data collected from a sample population, three validation tests are adopted usually. The first simply considers the consistency of the data by a field visit normally done after data entry stage. The second validation is done by choosing a computational check of the variables. For example, if age of a person is shown some high unrealistic values like 150 years. The last is a logical check done for the internal consistency of the data. For example, if the age of a person is less than 18 years, then he cannot have a driving license. Once these corrections are done, the data is ready to be used in modeling.

6.5 Other surveys

In addition to the household surveys, these other surveys are needed for complete modeling involving four stage models. Their primary use is for the calibration and validation of the models, or act as complementary to the household survey. These include O-D surveys, road side interviews, and cordon and screen line counts.

6.5.1 O-D survey

Sometime four small studies, or to get a feel of the O-D pattern without doing elaborate survey, work space interviews are conducted to find the origin-destination of employers in a location. Although they are biased in terms of the destination, they are random in terms of the mode of travel.

6.5.2 Road side interviews

These provide trips not registered in a household survey, especially external-internal trips. This involves asking questions to a sample of drivers and passengers of vehicles crossing a particular location. Unlike household survey, the respondent will be asked with few questions like origin, destination, and trip purpose. Other information like age, sex, and income can also be added, but it should be noted that at road-side, drivers will not be willing to spend much time for survey.

6.5.3 Cordon and screen-line survey

These provide useful information about trips from and to external zones. For large study area, internal cordon-line can be defined and surveying can be conducted. The objective of the survey is primarily to collect the origin and destination zones and for this many suitable methods can be adopted. It could be either recording the license plate number at all the external cordon points or by post-card method.

Screen lines divide the study area into large natural zones, like either sides of a river, with few crossing points between them. The procedure for both cordon-line and screen-line survey are similar to road-side interview. However, these counts are primarily used for calibration and validation of the models.

6.6 Summary

Data collection is one of the most important steps in modeling. Only if accurate data is available, modeling becomes successful. Survey design is discussed in detail. Household data gives important information required for data collection. Questionnaire should be simple, less time consuming and should be designed such that the required information is obtained with less burden on the respondent. Data collected should be prepared well before application. Various corrections should be made in data collection before they are used in modeling. Finally, other types of surveys are also discussed.

6.7 Problems

1. The data that is useful for developing trip generation models is
   (a) Travel survey data
   (b) Land-use inventory data
   (c) Network data
   (d) None of these
2. Which of the following is not a criterion for zoning?
   (a) zones should match other administrative divisions, particularly census zones.
   (b) zones should have homogeneous characteristics, especially in land use, population etc.
   (c) zone boundaries should match cordon and screen lines, but should not match major roads.
   (d) zones should have regular geometric shape.

6.8 Solutions

1. The data that is useful for developing trip generation models is
   (a) Travel survey data

(b) Land-use inventory data$\surd$

(c) Network data

(d) None of these

2. Which of the following is not a criterion for zoning?

   (a) zones should match other administrative divisions, particularly census zones.

   (b) zones should have homogeneous characteristics, especially in land use, population etc.

   (c) zone boundaries should match cordon and screen lines, but should not match major roads.

   (d) zones should have regular geometric shape$\surd$

Chapter 7

Trip generation

7.1 Overview

Trip generation is the first stage of the classical first generation aggregate demand models. The trip generation aims at predicting the total number of trips generated and attracted to each zone of the study area. In other words this stage answers the questions to "how many trips" originate at each zone, from the data on household and socioeconomic attributes. In this section basic definitions, factors affecting trip generation, and the two main modeling approaches; namely growth factor modeling and regression modeling are discussed.

7.1.1 Types of trip

Some basic definitions are appropriate before we address the classification of trips in detail. We will attempt to clarify the meaning of journey, home based trip, non home based trip, trip production, trip attraction and trip generation.

Journey is an out way movement from a point of origin to a point of destination, where as the word "trip" denotes an outward and return journey. If either origin or destination of a trip is the home of the trip maker then such trips are called home based trips and the rest of the trips are called non home based trips. Trip production is defined as all the trips of home based or as the origin of the non home based trips. See figure 7:1

Trips can be classified by trip purpose, trip time of the day, and by person type. Trip generation models are found to be accurate if separate models are used based on trip purpose. The trips can be classified based on the purpose of the journey as trips for work, trips for education, trips for shopping, trips for recreation and

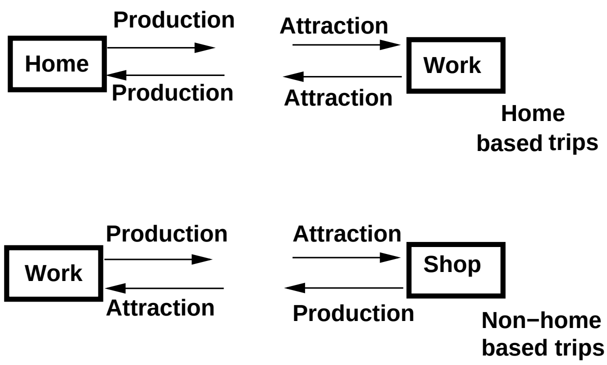


Figure 7:1: trip types

other trips. Among these the work and education trips are often referred as mandatory trips and the rest as discretionary trips. All the above trips are normally home based trips and constitute about 80 to 85 percent of trips. The rest of the trips namely non home based trips, being a small proportion are not normally treated separately. The second way of classification is based on the time of the day when the trips are made. The broad classification is into peak trips and off peak trips. The third way of classification is based on the type of the individual who makes the trips. This is important since the travel behavior is highly influenced by the socio economic attribute of the traveler and are normally categorized based on the income level, vehicle ownership and house hold size.

7.1.2 Factors affecting trip generation

The main factors affecting personal trip production include income, vehicle ownership, house hold structure and family size. In addition factors like value of land, residential density and accessibility are also considered for modeling at zonal levels. The personal trip attraction, on the other hand, is influenced by factors such as roofed space available for industrial, commercial and other services. At the zonal level zonal employment and accessibility are also used. In trip generation modeling in addition to personal trips, freight trips are also of interest. Although the latter comprises about 20 percent of trips, their contribution to the congestion is significant. Freight trips are influenced by number of employees, number of sales and area of commercial firms.

7.2 Growth factor modeling

Growth factor modes tries to predict the number of trips produced or attracted by a house hold or zone as a linear function of explanatory variables. The models have the following basic equation:

$$T_i = f_i t_i \tag{7.1}$$

where T_i is the number of future trips in the zone and t_i is the number of current trips in that zone and f_i is a growth factor. The growth factor f_i depends on the explanatory variable such as population (P) of the zone , average house hold income (I) , average vehicle ownership (V). The simplest form of f_i is represented as follows

$$f_i = \frac{P_i^d \times I_i^d \times V_i^d}{P_i^c \times I_i^c \times V_i^c} \tag{7.2}$$

where the subscript " d" denotes the design year and the subscript "c" denotes the current year.

Example

Given that a zone has 275 household with car and 275 household without car and the average trip generation rates for each groups is respectively 5.0 and 2.5 trips per day. Assuming that in the future, all household will have a car, find the growth factor and future trips from that zone, assuming that the population and income remains constant.

Solution

Current trip rate $t_i = 275 \times 2.5 + 275 \times 5.0 = 2062.5$ trips / day.

Growth factor $F_i = \frac{V_i^d}{V_i^c} = \frac{550}{275} = 2.0$

Therefore, no. of future trips $T_i = F_i t_i = 2.0 \times 2062.5 = 4125$ trips / day.
The above example also shows the limitation of growth factor method. If we think intuitively, the trip rate will remain same in the future.

Therefore the number of trips in the future will be 550 house holds $\times$ 5 trips per day = 2750 trips per day .

It may be noted from the above example that the actual trips generated is much lower than the growth factor method. Therefore growth factor models are normally used in the prediction of external trips where no other methods are available. But for internal trips , regression methods are more suitable and will be discussed in the following section.

7.3 Regression methods

The general form of a trip generation model is

$$T_i = f(x_1, x_2, x_3, ...x_i, ...x_k) \tag{7.3}$$

Where xi's are prediction factor or explanatory variable. The most common form of trip generation model is a linear function of the form

$$T_i = a_0 + a_1 x_1 + a_2 x_2 + ...a_i x_i ... + a_k x_k \tag{7.4}$$

where a_i 's are the coefficient of the regression equation and can be obtained by doing regression analysis. The above equations are called multiple linear regression equation, and the solutions are tedious to obtain manually. However for the purpose of illustration, an example with one variable is given.

Example

Let the trip rate of a zone is explained by the household size done from the field survey. It was found that the household size are 1, 2, 3 and 4. The trip rates of the corresponding household is as shown in the table below. Fit a linear equation relating trip rate and household size.

Household size(x)				
	1	2	3	4
Trips	1	2	4	6
per	2	4	5	7
day(y)	2	3	3	4
Σy	5	9	12	17

Solution The linear equation will have the form $y = bx + a$ where y is the trip rate, and x is the household size, a and b are the coefficients. For a best fit, b is given by

$$\begin{aligned} b &= \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} \\ a &= \bar{y} - b\bar{x} \\ \Sigma x &= 3 \times 1 + 3 \times 2 + 3 \times 3 + 3 \times 4 = 30 \\ \Sigma x^2 &= 3 \times (1^2) + 3 \times (2^2) + 3 \times (3^2) + 3 \times (4^2) = 90 \end{aligned}$$

$$\begin{aligned}
\Sigma y &= 5+9+12+17 = 43 \\
\Sigma xy &= 1\times 1 + 1\times 2 + 1\times 2 \\
&+ 2\times 2 + 2\times 4 + 2\times 3 \\
&+ 3\times 4 + 3\times 5 + 3\times 3 \\
&+ 4\times 6 + 4\times 7 + 4\times 4 \\
&= 127 \\
\bar{y} &= 43/12 = 3.58 \\
\bar{x} &= 30/12 = 2.5 \\
b &= \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} \\
&= \frac{((12\times 127) - (30\times 43))}{((12\times 90) - (30)^2)} = 1.3 \\
a &= \bar{y} - b\bar{x} = 3.58 - 1.3\times 2.5 = +0.33 \\
\bar{y} &= 1.3x - 0.33
\end{aligned}$$

7.4 Summary

Trip generation forms the first step of four-stage travel modeling. It gives an idea about the total number of trips generated to and attracted from different zones in the study area. Growth factor modeling and regression methods can be used to predict the trips. They are discussed in detail in this chapter.

7.5 Problems

1. The trip rate (y) and the corresponding household sizes (x) from a sample are shown in table below. Compute the trip rate if the average household size is 3.25 (Hint: use regression method).

	Householdsize(x)			
	1	2	3	4
Trips	1	3	4	5
per	3	4	5	8
day(y)	3	5	7	8

Solution Fit the regression equation as below.

$$\begin{aligned}
\Sigma x &= 3\times 1 + 3\times 2 + 3\times 3 + 3\times 4 = 30 \\
\Sigma x^2 &= 3\times(1^2) + 3\times(2^2) + 3\times(3^2) + 3\times(4^2) = 90 \\
\Sigma y &= 7+12+16+21 = 56 \\
\Sigma xy &= 1\times 1 + 1\times 3 + 1\times 3 \\
&+ 2\times 3 + 2\times 4 + 2\times 5 \\
&+ 3\times 4 + 3\times 5 + 3\times 7
\end{aligned}$$

$$
\begin{aligned}
&+ \quad 4 \times 5 + 4 \times 8 + 4 \times 8 \\
&= \quad 163 \\
\bar{y} &= \quad 56/12 = 4.67 \\
\bar{x} &= \quad 30/12 = 2.5 \\
b &= \quad \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} \\
&= \quad \frac{((12 \times 163) - (30 \times 56))}{((12 \times 90) - (30)^2)} = 1.533 \\
a &= \quad \bar{y} - b\bar{x} = 4.67 - 1.533 \times 2.5 = 0.837 \\
y &= \quad 0.837 + 1.533x
\end{aligned}
$$

When average household size =3.25, number of trips becomes,
$y = 0.837 + 1.533 \times 3.25 = 5.819$

Chapter 8

Trip distribution

8.1 Overview

The decision to travel for a given purpose is called trip generation. These generated trips from each zone is then distributed to all other zones based on the choice of destination. This is called trip distribution which forms the second stage of travel demand modeling. There are a number of methods to distribute trips among destinations; and two such methods are growth factor model and gravity model. Growth factor model is a method which respond only to relative growth rates at origins and destinations and this is suitable for short-term trend extrapolation. In gravity model, we start from assumptions about trip making behavior and the way it is influenced by external factors. An important aspect of the use of gravity models is their calibration, that is the task of fixing their parameters so that the base year travel pattern is well represented by the model.

8.2 Definitions and notations

8.2.1 Trip matrix

The trip pattern in a study area can be represented by means of a trip matrix or origin-destination (O-D)matrix. This is a two dimensional array of cells where rows and columns represent each of the zones in the study area. The notation of the trip matrix is given in figure 8:1.

The cells of each row i contain the trips originating in that zone which have as destinations the zones in the corresponding columns. T_{ij} is the number of trips between origin i and destination j. O_i is the total number

Zones	1	2	...	j	...	n	O_i
1	T_{11}	T_{12}	...	T_{1j}	...	T_{1n}	O_1
2	T_{21}	T_{22}	...	T_{2j}	...	T_{2n}	O_2
$\vdots$	...	...	...	...	...	...	$\vdots$
	T_{i1}	T_{i2}	...	T_{ij}	...	T_{in}	O_i
$\vdots$	...	...	...	...	...	...	$\vdots$
n	T_{ni}	T_{n2}	...	T_{nj}	...	T_{nn}	O_n
D_j	D_1	D_2	...	D_j	...	D_n	T

where $D_j = \Sigma_i T_{ij}$, $O_i = \Sigma_j T_{ij}$, and $T = \Sigma_{ij} T_{ij}$.

Figure 8:1: Notation of a trip matrix

of trips between originating in zone i and D_j is the total number of trips attracted to zone j. The sum of the trips in a row should be equal to the total number of trips emanating from that zone. The sum of the trips in a column is the number of trips attracted to that zone. These two constraints can be represented as: $\Sigma_j T_{ij} = O_i$ $\Sigma_i T_{ij} = D_j$ If reliable information is available to estimate both O_i and D_j, the model is said to be doubly constrained. In some cases, there will be information about only one of these constraints, the model is called singly constrained.

8.2.2 Generalized cost

One of the factors that influences trip distribution is the relative travel cost between two zones. This cost element may be considered in terms of distance, time or money units. It is often convenient to use a measure combining all the main attributes related to the dis-utility of a journey and this is normally referred to as the generalized cost of travel. This can be represented as

$$c_{ij} = a_1 t^v_{ij} + a_2 t^w_{ij} + a_3 t^t_{ij} + a_4 t_{nij} + a_5 F_{ij} + a_6 \phi_j + \delta \tag{8.1}$$

where t^v_{ij} is the in-vehicle travel time between i and j, t^w_{ij} is the walking time to and from stops, t^t_{ij} is the waiting time at stops, F_{ij} is the fare charged to travel between i and j, ϕ_j is the parking cost at the destination, and δ is a parameter representing comfort and convenience, and a_1,a_2,.... are the weights attached to each element of cost function.

8.3 Growth factor methods

8.3.1 Uniform growth factor

If the only information available is about a general growth rate for the whole of the study area, then we can only assume that it will apply to each cell in the matrix, that is a uniform growth rate. The equation can be written as:

$$T_{ij} = f \times t_{ij} \tag{8.2}$$

where f is the uniform growth factor t_{ij} is the previous total number of trips, T_{ij} is the expanded total number of trips. Advantages are that they are simple to understand, and they are useful for short-term planning. Limitation is that the same growth factor is assumed for all zones as well as attractions.

8.3.2 Example

Trips originating from zone 1,2,3 of a study area are 78,92 and 82 respectively and those terminating at zones 1,2,3 are given as 88,96 and 78 respectively. If the growth factor is 1.3 and the cost matrix is as shown below, find the expanded origin-constrained growth trip table.

	1	2	3	o_i
1	20	30	28	78
2	36	32	24	92
3	22	34	26	82
d_j	88	96	78	252

Solution Given growth factor = 1.3, Therefore, multiplying the growth factor with each of the cells in the matrix gives the solution as shown below.

	1	2	3	O_i
1	26	39	36.4	101.4
2	46.8	41.6	31.2	119.6
3	28.6	44.2	33.8	106.2
D_j	101.4	124.8	101.4	327.6

8.3.3 Doubly constrained growth factor model

When information is available on the growth in the number of trips originating and terminating in each zone, we know that there will be different growth rates for trips in and out of each zone and consequently having two sets of growth factors for each zone. This implies that there are two constraints for that model and such a model is called doubly constrained growth factor model. One of the methods of solving such a model is given by Furness who introduced *balancing factors* a_i and b_j as follows:

$$T_{ij} = t_{ij} \times a_i \times b_j \tag{8.3}$$

In such cases, a set of intermediate correction coefficients are calculated which are then appropriately applied to cell entries in each row or column. After applying these corrections to say each row, totals for each column are calculated and compared with the target values. If the differences are significant, correction coefficients are calculated and applied as necessary. The procedure is given below:

1. Set $b_j = 1$
2. With b_j solve for a_i to satisfy trip generation constraint.
3. With a_i solve for b_j to satisfy trip attraction constraint.
4. Update matrix and check for errors.
5. Repeat steps 2 and 3 till convergence.

Here the error is calculated as: $E = \Sigma|O_i - O_i^1| + \Sigma|D_j - D_j^1|$ where O_i corresponds to the actual productions from zone i and O_i^1 is the calculated productions from that zone. Similarly D_j are the actual attractions from the zone j and D_j^1 are the calculated attractions from that zone.

8.3.4 Advantages and limitations of growth factor model

The advantages of this method are:

1. Simple to understand.
2. Preserve observed trip pattern.
3. useful in short term-planning.

The limitations are:

1. Depends heavily on the observed trip pattern.

2. It cannot explain unobserved trips.

3. Do not consider changes in travel cost.

4. Not suitable for policy studies like introduction of a mode.

Example

The base year trip matrix for a study area consisting of three zones is given below.

	1	2	3	o_i
1	20	30	28	78
2	36	32	24	92
3	22	34	26	82
d_j	88	96	78	252

The productions from the zone 1,2 and 3 for the horizon year is expected to grow to 98, 106, and 122 respectively. The attractions from these zones are expected to increase to 102, 118, 106 respectively. Compute the trip matrix for the horizon year using doubly constrained growth factor model using Furness method.

Solution The sum of the attractions in the horizon year, i.e. $\Sigma O_i = 98{+}106{+}122 = 326$. The sum of the productions in the horizon year, i.e. $\Sigma D_j = 102{+}118{+}106 = 326$. They both are found to be equal. Therefore we can proceed. The first step is to fix $b_j = 1$, and find balancing factor a_i. $a_i = O_i/o_i$, then find $T_{ij} = a_i \times t_{ij}$

So $a_1 = 98/78 = 1.26$

$a_2 = 106/92 = 1.15$

$a_3 = 122/82 = 1.49$ Further $T_{11} = t_{11} \times a_1 = 20 \times 1.26 = 25.2$. Similarly $T_{12} = t_{12} \times a_2 = 36 \times 1.15 = 41.4$. etc. Multiplying a_1 with the first row of the matrix, a_2 with the second row and so on, matrix obtained is as shown below.

	1	2	3	o_i
1	25.2	37.8	35.28	98
2	41.4	36.8	27.6	106
3	32.78	50.66	38.74	122
d_j^1	99.38	125.26	101.62	
D_j	102	118	106	

Also $d_j^1 = 25.2 + 41.4 + 32.78 = 99.38$

In the second step, find $b_j = D_j/d_j^1$ and $T_{ij} = t_{ij} \times b_j$. For example $b_1 = 102/99.38 = 1.03$, $b_2 = 118/125.26 = 0.94$ etc.,$T_{11} = t_1 1 \times b_1 = 25.2 \times 1.03 = 25.96$ etc. Also $O_i^1 = 25.96 + 35.53 + 36.69 = 98.18$. The matrix is as shown below:

	1	2	3	o_i	O_i
1	25.96	35.53	36.69	98.18	98
2	42.64	34.59	28.70	105.93	106
3	33.76	47.62	40.29	121.67	122
b_j	1.03	0.94	1.04		
D_j	102	118	106		

	1	2	3	O_i^1	O_i
1	25.96	35.53	36.69	98.18	98
2	42.64	34.59	28.70	105.93	106
3	33.76	47.62	40.29	121.67	122
d_j	102.36	117.74	105.68	325.78	
D_j	102	118	106	326	

Therefore error can be computed as ; $Error = \Sigma|O_i - O_i^1| + \Sigma|D_j - d_j|$

Error $= |98.18-98| + |105.93-106| + |121.67-122| + |102.36-102| + |117.74-118| + |105.68-106| = 1.32$

8.4 Gravity model

This model originally generated from an analogy with Newton's gravitational law. Newton's gravitational law says, $F = GM_1M_2/d_2$ Analogous to this, $T_{ij} = CO_iD_j/c_{ij^n}$ Introducing some balancing factors, $T_{ij} = A_iO_iB_jD_jf(c_{ij})$ where A_i and B_j are the balancing factors, $f(c_{ij})$ is the generalized function of the travel cost. This function is called *deterrence function* because it represents the disincentive to travel as distance (time) or cost increases. Some of the versions of this function are:

$$f(c_{ij}) = e^{-\beta c_{ij}} \tag{8.4}$$

$$f(c_{OJ}) = c_{ij}^{-n} \tag{8.5}$$

$$f(c_{ij}) = c_{ij}^{-n} \times e^{-\beta c_{ij}} \tag{8.6}$$

The first equation is called the exponential function, second one is called power function where as the third one is a combination of exponential and power function. The general form of these functions for different values of their parameters is as shown in figure.

As in the growth factor model, here also we have singly and doubly constrained models. The expression $T_{ij} = A_iO_iB_jD_jf(c_{ij})$ is the classical version of the doubly constrained model. Singly constrained versions can be produced by making one set of balancing factors A_i or B_j equal to one. Therefore we can treat singly constrained model as a special case which can be derived from doubly constrained models. Hence we will limit our discussion to doubly constrained models.

As seen earlier, the model has the functional form, $T_{ij} = A_iO_iB_jD_jf(c_{ij})$

$$\Sigma_i T_{ij} = \Sigma_i A_iO_iB_jD_jf(c_{ij}) \tag{8.7}$$

But

$$\Sigma_i T_{ij} = D_j \tag{8.8}$$

Therefore,

$$D_j = B_jD_j\Sigma_i A_iO_if(c_{ij}) \tag{8.9}$$

From this we can find the balancing factor B_j as

$$B_j = 1/\Sigma_i A_iO_if(c_{ij}) \tag{8.10}$$

B_j depends on A_i which can be found out by the following equation:

$$A_i = 1/\Sigma_j B_jD_jf(c_{ij}) \tag{8.11}$$

We can see that both A_i and B_j are interdependent. Therefore, through some iteration procedure similar to that of Furness method, the problem can be solved. The procedure is discussed below:

1. Set $B_j = 1$, find A_i using equation 8.11
2. Find B_j using equation 8.10
3. Compute the error as $E = \Sigma|O_i - O_i^1| + \Sigma|D_j - D_j^1|$ where O_i corresponds to the actual productions from zone i and O_i^1 is the calculated productions from that zone. Similarly D_j are the actual attractions from the zone j and D_j^1 are the calculated attractions from that zone.
4. Again set $B_j = 1$ and find A_i, also find B_j. Repeat these steps until the convergence is achieved.

Example

The productions from zone 1, 2 and 3 are 98, 106, 122 and attractions to zone 1,2 and 3 are 102, 118, 106. The function $f(c_{ij})$ is defined as $f(c_{ij}) = 1/c_{ij}^2$ The cost matrix is as shown below

$$\begin{bmatrix} 1.0 & 1.2 & 1.8 \\ 1.2 & 1.0 & 1.5 \\ 1.8 & 1.5 & 1.0 \end{bmatrix} \tag{8.12}$$

Solution The first step is given in Table 8:1 The second step is to find B_j. This can be found out as

Table 8:1: Step1: Computation of parameter A_i

i	j	B_j	D_J	$f(c_{ij})$	$B_jD_jf(c_{ij})$	$\Sigma B_jD_jf(c_{ij})$	$A_i = \frac{1}{\Sigma B_jD_jf(c_{ij})}$
	1	1.0	102	1.0	102.00		
1	2	1.0	118	0.69	81.42	216.28	0.00462
	3	1.0	106	0.31	32.86		
	1	1.0	102	0.69	70.38		
2	2	1.0	118	1.0	118	235.02	0.00425
	3	1.0	106	0.44	46.64		
	1	1.0	102	0.31	31.62		
3	2	1.0	118	0.44	51.92	189.54	0.00527
	3	1.0	106	1.00	106		

$B_j = 1/\Sigma A_iO_if(c_{ij})$, where A_i is obtained from the previous step. The detailed computation is given in Table 8:2. The function $f(c_{ij})$ can be written in the matrix form as:

$$\begin{bmatrix} 1.0 & 0.69 & 0.31 \\ 0.69 & 1.0 & 0.44 \\ 0.31 & 0.44 & 1.0 \end{bmatrix} \tag{8.13}$$

Then T_{ij} can be computed using the formula

$$T_{ij} = A_iO_iB_jD_jf(c_{ij}) \tag{8.14}$$

Table 8:2: Step2: Computation of parameter B_j

j	i	A_i	O_i	$f(c_{ij})$	$A_iO_if(c_{ij})$	$\Sigma A_iO_if(c_{ij})$	$B_j = 1/\Sigma A_iO_if(c_{ij})$
	1	0.00462	98	1.0	0.4523		
1	2	0.00425	106	0.694	0.3117	0.9618	1.0397
	3	0.00527	122	0.308	0.1978		
	1	0.00462	98	0.69	0.3124		
2	2	0.00425	106	1.0	0.4505	1.0458	0.9562
	3	0.00527	122	0.44	0.2829		
	1	0.00462	98	0.31	0.1404		
3	2	0.00425	106	0.44	0.1982	0.9815	1.0188
	3	0.00527	122	1.00	0.6429		

Table 8:3: Step3: Final Table

	1	2	3	A_i	O_i	O_i^1
1	48.01	35.24	15.157	0.00462	98	98.407
2	32.96	50.83	21.40	0.00425	106	105.19
3	21.14	31.919	69.43	0.00527	122	122.489
B_j	1.0397	0.9562	1.0188			
D_j	102	118	106			
D_j^1	102.11	117.989	105.987			

For eg, $T_{11} = 102 \times 1.0397 \times 0.00462 \times 98 \times 1 = 48.01$. O_i is the actual productions from the zone and O_i^1 is the computed ones. Similar is the case with attractions also. The results are shown in table 8:3. O_i is the actual productions from the zone and O_i^1 is the computed ones. Similar is the case with attractions also.

Therefore error can be computed as ; $Error = \Sigma|O_i - O_i^1| + \Sigma|D_j - D_j^1|$ $Error = |98 - 98.407| + |106 - 105.19| + |122 - 122.489| + ||102 - 102.11| + |118 - 117.989| + |106 - 105.987| = 2.03$

8.5 Summary

The second stage of travel demand modeling is the trip distribution. Trip matrix can be used to represent the trip pattern of a study area. Growth factor methods and gravity model are used for computing the trip matrix. Singly constrained models and doubly constrained growth factor models are discussed. In gravity model, considering singly constrained model as a special case of doubly constrained model, doubly constrained model is explained in detail.

8.6 Problems

The trip productions from zones 1, 2 and 3 are 110, 122 and 114 respectively and the trip attractions to these zones are 120,108, and 118 respectively. The cost matrix is given below. The function $f(c_{ij}) = \frac{1}{c_{ij}}$

$$\begin{bmatrix} 1.0 & 1.2 & 1.8 \\ 1.2 & 1.0 & 1.5 \\ 1.8 & 1.5 & 1.0 \end{bmatrix}$$

Compute the trip matrix using doubly constrained gravity model. Provide one complete iteration.

Solution The first step is given in Table 8:4 The second step is to find B_j. This can be found out as

Table 8:4: Step1: Computation of parameter A_i

i	j	B_j	D_J	$f(c_{ij})$	$B_jD_jf(c_{ij})$	$\Sigma B_jD_jf(c_{ij})$	$A_i = \frac{1}{\Sigma B_jD_jf(c_{ij})}$
	1	1.0	120	1.0	120.00		
1	2	1.0	108	0.833	89.964	275.454	0.00363
	3	1.0	118	0.555	65.49		
	1	1.0	120	0.833	99.96		
2	2	1.0	108	1.0	108	286.66	0.00348
	3	1.0	118	0.667	78.706		
	1	1.0	120	0.555	66.60		
3	2	1.0	108	0.667	72.036	256.636	0.00389
	3	1.0	118	1.00	118		

$B_j = 1/\Sigma A_iO_if(c_{ij})$, where A_i is obtained from the previous step. The function $f(c_{ij})$ can be written in the

Table 8:5: Step2: Computation of parameter B_j

j	i	A_i	O_i	$f(c_{ij})$	$A_iO_if(c_{ij})$	$\Sigma A_iO_if(c_{ij})$	$B_j = 1/\Sigma A_iO_if(c_{ij})$
	1	0.00363	110	1.0	0.3993		
1	2	0.00348	122	0.833	0.3536	0.9994	1.048
	3	0.00389	114	0.555	0.2465		
	1	0.00363	110	0.833	0.3326		
2	2	0.00348	122	1.0	0.4245	1.05	0.9494
	3	0.00389	114	0.667	0.2962		
	1	0.00363	110	0555	0.2216		
3	2	0.00348	122	0.667	0.2832	0.9483	1.054
	3	0.00389	114	1.00	0.44346		

Table 8:6: Step 3: Final Table

	1	2	3	A_i	O_i	O_i^1
1	48.01	34.10	27.56	0.00363	110	109.57
2	42.43	43.53	35.21	0.00348	122	121.17
3	29.53	30.32	55.15	0.00389	114	115
B_j	1.048	0.9494	1.054			
D_j	120	108	118			
D_j^1	119.876	107.95	117.92			

matrix form as:

$$\begin{bmatrix} 1.0 & 0.833 & 0.555 \\ 0.833 & 1.0 & 0.667 \\ 0.555 & 0.667 & 1.0 \end{bmatrix} \tag{8.15}$$

Then T_{ij} can be computed using the formula

$$T_{ij} = A_i O_i B_j D_j f(c_{ij}) \tag{8.16}$$

For eg, $T_{11} = 102 \times 1.0397 \times 0.00462 \times 98 \times 1 = 48.01$. O_i is the actual productions from the zone and O_i^1 is the computed ones. Similar is the case with attractions also. This step is given in Table 8:6 O_i is the actual productions from the zone and O_i^1 is the computed ones. Similar is the case with attractions also.

Therefore error can be computed as ; $Error = \Sigma|O_i - O_i^1| + \Sigma|D_j - D_j^1|$ $Error = |110 - 109.57| + |122 - 121.17| + |114 - 115| + |120 - 119.876 + |108 - 107.95| + |118 - 117.92| = 2.515$

Chapter 9

Modal split

9.1 Overview

The third stage in travel demand modeling is modal split. The trip matrix or O-D matrix obtained from the trip distribution is sliced into number of matrices representing each mode. First the significance and factors affecting mode choice problem will be discussed. Then a brief discussion on the classification of mode choice will be made. Two types of mode choice models will be discussed in detail. ie binary mode choice and multinomial mode choice. The chapter ends with some discussion on future topics in mode choice problem.

9.2 Mode choice

The choice of transport mode is probably one of the most important classic models in transport planning. This is because of the key role played by public transport in policy making. Public transport modes make use of road space more efficiently than private transport. Also they have more social benefits like if more people begin to use public transport , there will be less congestion on the roads and the accidents will be less. Again in public transport, we can travel with low cost. In addition, the fuel is used more efficiently. Main characteristics of public transport is that they will have some particular schedule, frequency etc.

On the other hand, private transport is highly flexible. It provides more comfortable and convenient travel. It has better accessibility also. The issue of mode choice, therefore, is probably the single most important element in transport planning and policy making. It affects the general efficiency with which we can travel in urban areas. It is important then to develop and use models which are sensitive to those travel attributes that influence individual choices of mode.

9.3 Factors influencing the choice of mode

The factors may be listed under three groups:

1. **Characteristics of the trip maker** : The following features are found to be important:
   - (a) car availability and/or ownership;
   - (b) possession of a driving license;
   - (c) household structure (young couple, couple with children, retired people etc.);
   - (d) income;

(e) decisions made elsewhere, for example the need to use a car at work, take children to school, etc;

(f) residential density.

2. **Characteristics of the journey**: Mode choice is strongly influenced by:

   (a) The trip purpose; for example, the journey to work is normally easier to undertake by public transport than other journeys because of its regularity and the adjustment possible in the long run;

   (b) Time of the day when the journey is undertaken.

   (c) Late trips are more difficult to accommodate by public transport.

3. **Characteristics of the transport facility**: There are two types of factors.One is quantitative and the other is qualitative. Quantitative factors are:

   (a) relative travel time: in-vehicle, waiting and walking times by each mode;

   (b) relative monetary costs (fares, fuel and direct costs);

   (c) availability and cost of parking

   Qualitative factors which are less easy to measure are:

   (a) comfort and convenience

   (b) reliability and regularity

   (c) protection, security

A good mode choice should include the most important of these factors.

9.4 Types of modal split models

9.4.1 Trip-end modal split models

Traditionally, the objective of transportation planning was to forecast the growth in demand for car trips so that investment could be planned to meet the demand. When personal characteristics were thought to be the most important determinants of mode choice, attempts were made to apply modal-split models immediately after trip generation. Such a model is called trip-end modal split model. In this way different characteristics of the person could be preserved and used to estimate modal split. The modal split models of this time related the choice of mode only to features like income, residential density and car ownership.

The advantage is that these models could be very accurate in the short run, if public transport is available and there is little congestion. Limitation is that they are insensitive to policy decisions example: Improving public transport, restricting parking etc. would have no effect on modal split according to these trip-end models.

9.4.2 Trip-interchange modal split models

This is the post-distribution model; that is modal split is applied after the distribution stage. This has the advantage that it is possible to include the characteristics of the journey and that of the alternative modes available to undertake them. It is also possible to include policy decisions. This is beneficial for long term modeling.

9.4.3 Aggregate and disaggregate models

Mode choice could be *aggregate* if they are based on zonal and inter-zonal information. They can be called *disaggregate* if they are based on household or individual data.

9.5 Binary logit model

Binary logit model is the simplest form of mode choice, where the travel choice between two modes is made. The traveler will associate some value for the utility of each mode. if the utility of one mode is higher than the other, then that mode is chosen. But in transportation, we have disutility also. The disutility here is the travel cost. This can be represented as

$$c_{ij} = a_1 t^v_{ij} + a_2 t^w_{ij} + a_3 t^t_{ij} + a_4 t_{nij} + a_5 F_{ij} + a_6 \phi_j + \delta \tag{9.1}$$

where t^v_{ij} is the in-vehicle travel time between i and j, t^w_{ij} is the walking time to and from stops, t^t_{ij} is the waiting time at stops, F_{ij} is the fare charged to travel between i and j, ϕ_j is the parking cost, and δ is a parameter representing comfort and convenience. If the travel cost is low, then that mode has more probability of being chosen. Let there be two modes (m=1,2) then the proportion of trips by mode 1 from zone i to zone j is(P^1_{ij}) Let c^1_{ij} be the cost of traveling from zone i to zone$j$ using the mode 1, and c^2_{ij} be the cost of traveling from zone$i$ to zone j by mode 2,there are three cases:

1. if c^2_{ij} - c^1_{ij} is positive, then mode 1 is chosen.
2. if c^2_{ij} - c^1_{ij} is negative, then mode 2 is chosen.
3. if c^2_{ij} - $c^1_{ij} = 0$, then both modes have equal probability.

This relationship is normally expressed by a logit curve as shown in figure 9:1 Therefore the proportion of trips by mode 1 is given by

$$P^1_{ij} = T^1_{ij}/T_{ij} = \frac{e^{-\beta c^1_{ij}}}{e^{-\beta c^1_{ij}} + e^{-\beta c^2_{ij}}} \tag{9.2}$$

This functional form is called logit, where c_{ij} is called the generalized cost and β is the parameter for calibration. The graph in figure 9:1 shows the proportion of trips by mode 1 (T^1_{ij}/T_{ij}) against cost difference.

Example

Let the number of trips from zone i to zone j is 5000, and two modes are available which has the characteristics given in Table 9:1. Compute the trips made by mode bus, and the fare that is collected from the mode bus. If the fare of the bus is reduced to 6, then find the fare collected.

Solution The base case is given below.

Cost of travel by car (Equation)=$c_{car} = 0.03 \times 20 + 18 \times 0.06 + 4 \times 0.1 = 2.08$

Cost of travel by bus (Equation)=$c_{bus} = 0.03 \times 30 + 0.04 \times 5 + 0.06 \times 3 + 0.1 \times 9 = 2.18$

Probability of choosing mode car (Equation)= $p^{car}_{ij} = \frac{e^{-2.08}}{e^{-2.08}+e^{-2.18}} = 0.52$

Probability of choosing mode bus (Equation)= $p^{bus}_{ij} = \frac{e^{-2.18}}{e^{-2.08}+e^{-2.18}} = 0.475$

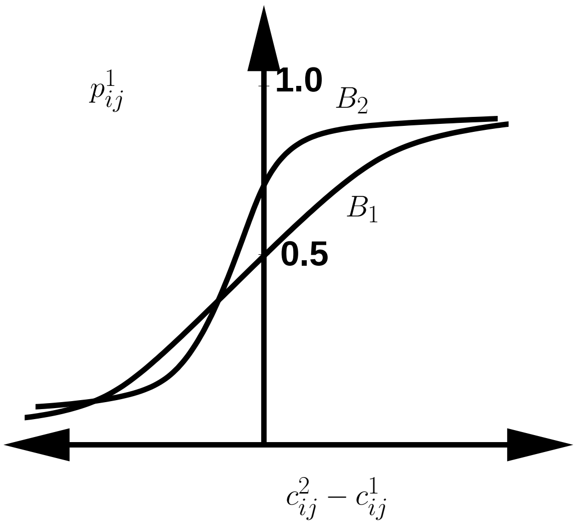


Figure 9:1: logit function

Table 9:1: Trip characterisitcs

	t^u_{ij}	t^w_{ij}	t^t_{ij}	f_{ij}	ϕ_j
car	20	-	18	4	
bus	30	5	3	9	
a_i	0.03	0.04	0.06	0.1	0.1

Table 9:2: Binary logit model example: solution

	t^u_{ij}	t^w_{ij}	t^t_{ij}	f_{ij}	ϕ_j	c_{ij}	p_{ij}	T_{ij}
ar	20	-	18	4		2.08	.52	2600
bus	30	5	3	9		2.18	.475	2400
a_i	.03	.04	.06	.1	.1			

Table 9:3: Trip characteristics

	t^v_{ij}	t^{walk}_{ij}	t^t_{ij}	F_{ij}	ϕ_{ij}
coefficient	0.03	0.04	0.06	0.1	0.1
car	20	-	-	18	4
bus	30	5	3	6	-
train	12	10	2	4	-

Proportion of trips by car = T^{car}_{ij} = 5000×0.52 = 2600

Proportion of trips by bus = T^{bus}_{ij} = 5000×0.475 = 2400

Fare collected from bus = $T^{bus}_{ij} \times F_{ij}$ = 2400×9 = 21600

When the fare of bus gets reduced to 6,

Cost function for bus= $c_{bus} = 0.03 \times 30 + 0.04 \times 5 + 0.06 \times 3 + 0.1 \times 6 = 1.88$

Probability of choosing mode bus (Equation)=$p^{bus}_{ij} = \frac{e^{-1.88}}{e^{-2.08}+e^{-1.88}} = 0.55$

Proportion of trips by bus = T^{bus}_{ij} = 5000×0.55 = 2750

Fare collected from the bus $T^{bus}_{ij} \times F_{ij}$= 2750×6 = 16500

The results are tabulated in table

9.6 Multinomial logit model

The binary model can easily be extended to multiple modes. The equation for such a model can be written as:

$$P^1_{ij} = \frac{e^{-\beta c^1_{ij}}}{\Sigma e^{-\beta c^m_{ij}}} \tag{9.3}$$

9.6.1 Example

Let the number of trips from i to j is 5000, and three modes are available which has the characteristics given in Table 9:3: Compute the trips made by the three modes and the fare required to travel by each mode.

Solution

Cost of travel by car (Equation)=$c_{car} = 0.03 \times 20 + 18 \times 0.1 + 4 \times 0.1 = 2.8$

Cost of travel by bus (Equation)=$c_{bus} = 0.03 \times 30 + 0.04 \times 5 + 0.06 \times 3 + 0.1 \times 6 = 1.88$

Cost of travel by train (Equation)=$c_{train} = 0.03 \times 12 + 0.04 \times 10 + 0.06 \times 2 + 0.1 \times 4 = 1.28$

Probability of choosing mode car (Equation)$p^{car}_{ij} = \frac{e^{-2.8}}{e^{-2.8}+e^{-1.88}+e^{-1.28}} = 0.1237$

Probability of choosing mode bus (Equation$p^{bus}_{ij} = \frac{e^{-1.88}}{e^{-2.8}+e^{-1.88}+e^{-1.28}} = 0.3105$

Table 9:4: Multinomial logit model problem: solution

	t^v_{ij}	t^{walk}_{ij}	t^t_{ij}	F_{ij}	ϕ_{ij}	C	e^C	p_{ij}	T_{ij}
coeff	0.03	0.04	0.06	0.1	0.1	-	-	-	-
car	20	-	-	18	4	2.8	0.06	0.1237	618.5
bus	30	5	3	6	-	1.88	0.15	0.3105	1552.5
train	12	10	2	4	-	1.28	0.28	0.5657	2828.5

Table 9:5: Trip characteristics

	t^v_{ij}	t^{walk}_{ij}	t^t_{ij}	F_{ij}	ϕ_{ij}
coefficient	0.05	0.04	0.07	0.2	0.2
car	25	-	-	22	6
bus	35	8	6	8	-
train	17	14	5	6	-

Probability of choosing mode train (Equation)= $p^{train}_{ij} = \frac{e^{-1.28}}{e^{-2.8}+e^{-1.88}+e^{-1.28}} = 0.5657$

Proportion of trips by car,$T^{car}_{ij} = 5000\times0.1237 = 618.5$

Proportion of trips by bus, $T^{bus}_{ij} = 5000\times0.3105 = 1552.5$

Similarly, proportion of trips by train,$T^{train}_{ij} = 5000\times0.5657 = 2828.5$ We can put all this in the form of a table as shown below 9:4:

- Fare collected from the mode bus = $T^{bus}_{ij} \times F_{ij} = 1552.5\times6 = 9315$
- Fare collected from mode train = $T^{train}_{ij} \times F_{ij} = 2828.5\times4 = 11314$

9.7 Summary

Modal split is the third stage of travel demand modeling. The choice of mode is influenced by various factors. Different types of modal split models are there. Binary logit model and multinomial logit model are dealt in detail in this chapter.

9.8 Problems

1. The total number of trips from zone i to zone j is 4200. Currently all trips are made by car. Government has two alternatives- to introduce a train or a bus. The travel characteristics and respective coefficients are given in table 9:5. Decide the best alternative in terms of trips carried.

   Solution

First, use binary logit model to find the trips when there is only car and bus. Then, again use binary logit model to find the trips when there is only car and train. Finally compare both and see which alternative carry maximum trips.

Cost of travel by car=$c_{car} = 0.05 \times 25 + 0.2 \times 22 + 0.2 \times 6 = 6.85$

Cost of travel by bus=$c_{bus} = 0.05 \times 35 + 0.04 \times 8 + 0.07 \times 6 + 0.2 \times 8 = 4.09$

Cost of travel by trai=$c_{train} = 0.05 \times 17 + 0.04 \times 14 + 0.07 \times 5 + 0.2 \times 6 = 2.96$

Case 1: Considering introduction of bus, Probability of choosing car, $p_{ij}^{car} = \frac{e^{-6.85}}{e^{-6.85}+e^{-4.09}} = 0.059$

Probability of choosing bus, $p_{ij}^{bus} = \frac{e^{-4.09}}{e^{-6.85}+e^{-4.09}} = 0.9403$

Case 2: Considering introduction of train, Probability of choosing car $p_{ij}^{car} = \frac{e^{-6.85}}{e^{-6.85}+e^{-2.96}} = 0.02003$

Probability of choosing train $p_{ij}^{train} = \frac{e^{-2.96}}{e^{-6.85}+e^{-2.96}} = 0.979$

Trips carried by each mode

Case 1: $T_{ij}^{car} = 4200 \times 0.0596 = 250.32$ $T_{ij}^{bus} = 4200 \times 0.9403 = 3949.546$

Case 2: $T_{ij}^{car} = 4200 \times 0.02 = 84.00$ $T_{ij}^{train} = 4200 \times 0.979 = 4115.8$

Hence *train* will attract more trips, if it is introduced.

Chapter 10

Traffic Assignment

10.1 Overview

The process of allocating given set of trip interchanges to the specified transportation system is usually referred to as traffic assignment. The fundamental aim of the traffic assignment process is to reproduce on the transportation system, the pattern of vehicular movements which would be observed when the travel demand represented by the trip matrix, or matrices, to be assigned is satisfied. The major aims of traffic assignment procedures are:

1. To estimate the volume of traffic on the links of the network and obtain aggregate network measures.
2. To estimate inter zonal travel cost.
3. To analyze the travel pattern of each origin to destination(O-D) pair.
4. To identify congested links and to collect traffic data useful for the design of future junctions.

10.2 Link cost function

As the flow increases towards the capacity of the stream, the average stream speed reduces from the free flow speed to the speed corresponding to the maximum flow. This can be seen in the graph shown below.

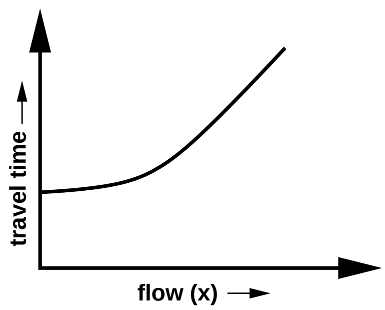


Figure 10:1: Two Link Problem with constant travel time function

That means traffic conditions worsen and congestion starts developing. The inter zonal flows are assigned to the minimum paths computed on the basis of free-flow link impedances (usually travel time). But if the link flows were at the levels dictated by the assignment, the link speeds would be lower and the link travel time would be higher than those corresponding to the free flow conditions. So the minimum path computed prior to the trip assignment will not be the minimum after the trips ae assigned. A number of iterative procedures are done to converge this difference. The relation between the link flow and link impedance is called the link cost function and is given by the equation as shown below: $t = t_0[1 + \alpha(\frac{x}{k})^{\beta}]$

where t and x is the travel time and flow, respectively on the link, t_0 is the free flow travel time, and k is the practical capacity. α and β are the model parameters, for which the value of $\alpha = 0.15$ minimum and $\beta =$ 4.0 are typically used.

The types of traffic assignment models are all-or-nothing assignment (AON), incremental assignment, capacity restraint assignment, user equilibrium assignment (UE), stochastic user equilibrium assignment (SUE), system optimum assignment (SO), etc. The frequently used models all-or-nothing, user equilibrium, and system optimum will be discussed in detail here.

10.3 All-or-nothing assignment

In this method the trips from any origin zone to any destination zone are loaded onto a single, minimum cost, path between them. This model is unrealistic as only one path between every O-D pair is utilized even if there is another path with the same or nearly same travel cost. Also, traffic on links is assigned without consideration of whether or not there is adequate capacity or heavy congestion; travel time is a fixed input and does not vary depending on the congestion on a link. However, this model may be reasonable in sparse and uncongested networks where there are few alternative routes and they have a large difference in travel cost. This model may also be used to identify the *desired path* : the path which the drivers would like to travel in the absence of congestion. In fact, this model's most important practical application is that it acts as a building block for other types of assignment techniques.It has a limitation that it ignores the fact that link travel time is a function of link volume and when there is congestion or that multiple paths are used to carry traffic.

Example

To demonstrate how this assignment works, an example network is considered. This network has two nodes having two paths as links. Let us suppose a case where travel time is not a function of flow as shown in other words it is constant as shown in the figure below.

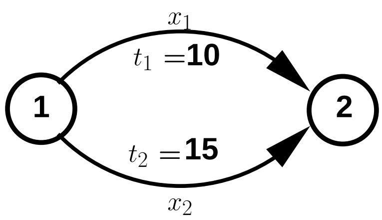


Figure 10:2: Two Link Problem with constant travel time function

Solution The travel time functions for both the links is given by:

$t_1 = 10$

$t_2 = 15$

and total flows from 1 to 2 is given by. $q_{12} = 12$
Since the shortest path is Link 1 all flows are assigned to it making x_1 =12 and $x_2 = 0$.

10.4 User equilibrium assignment (UE)

The user equilibrium assignment is based on Wardrop's first principle, which states that *no driver can unilaterally reduce his/her travel costs by shifting to another route.* User Equilibrium (UE) conditions can be written for a given O-D pair as:

$$f_k(c_k - u) = 0 : \forall k \tag{10.1}$$

$$c_k - u >= 0 : \forall k \tag{10.2}$$

where f_k is the flow on path k, c_k is the travel cost on path k, and u is the minimum cost.

Equation labelqueue2 can have two states.

1. If $c_k - u = 0$, from equation 10.1 $f_k \geq 0$. This means that all used paths will have same travel time.

2. If $c_k - u \geq 0$, then from equation 10.1 $f_k = 0$.

This means that all unused paths will have travel time greater than the minimum cost path. where f_k is the flow on path k, c_k is the travel cost on path k, and u is the minimum cost.

Assumptions in User Equilibrium Assignment

1. The user has perfect knowledge of the path cost.

2. Travel time on a given link is a function of the flow on that link only.

3. Travel time functions are positive and increasing.

The solution to the above equilibrium conditions given by the solution of an equivalent nonlinear mathematical optimization program,

$$\text{Minimize } Z = \sum_a \int_0^{x_a} t_a(x_a)dx, \tag{10.3}$$

$$\begin{aligned} \text{subject to } \sum_k f_k^{rs} &= q_{rs} \ : \forall r, s \\ x_a = \sum_r \sum_s \sum_k \delta_{a,k}^{rs} f_k^{rs} &\ : \forall a \\ f_k^{rs} \geq 0 &\ : \forall \ k, r, s \\ x_a \geq 0 &\ : a \in A \end{aligned}$$

where k is the path, x_a equilibrium flows in link a, t_a travel time on link a, f_k^{rs} flow on path k connecting O-D pair r-s, q_{rs} trip rate between r and sand $\delta_{a,k}^{rs}$ is a definitional constraint and is given by

$$\delta_{a,k}^{r,s} = \begin{cases} 1 & \text{if link a belongs to path k,} \\ 0 & \text{otherwise} \end{cases} \tag{10.4}$$

The equations above are simply flow conservation equations and non negativity constraints, respectively. These constraints naturally hold the point that minimizes the objective function. These equations state user equilibrium principle.The path connecting O-D pair can be divided into two categories : those carrying the flow and those not carrying the flow on which the travel time is greater than (or equal to)the minimum O-D travel time. If the flow pattern satisfies these equations no motorist can better off by unilaterally changing routes. All other routes have either equal or heavy travel times. The user equilibrium criteria is thus met for every O-D pair. The UE problem is convex because the link travel time functions are monotonically increasing function, and the link travel time a particular link is independent of the flow and other links of the networks. To solve such convex problem Frank Wolfe algorithm is useful.

Example

Let us suppose a case where travel time is not a function of flow as shown in other words it is constant as shown in the figure below.

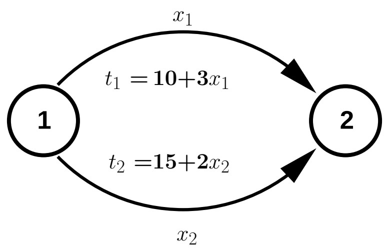


Figure 10:3: Two Link Problem with constant travel time function

Solution Substituting the travel time in equation yield to

$$\begin{aligned} min\ Z(x) &= \int_0^{x_1} 10 + 3x_1 dx_1 + \int_0^{x_2} 15 + 2x_2 dx_2 \\ &= 10x_1 + \frac{3x_1^2}{2} + 15x_2 + \frac{2x_2^2}{2} \end{aligned}$$

subject to

$$x_1 + x_2 = 12$$

Substituting $x_2 = 12 - x_1$, in the above formulation will yield the unconstrained formulation as below :

$$min\ Z(x) = 10x_1 + \frac{3x_1^2}{2} + 15(12 - x_1) + \frac{2(12 - x_1)^2}{2}$$

Differentiate the above equation x_1 and equate to zero, and solving for x_1 and then x_2 leads to the solution x_1 = 5.8, x_2 = 6.2.

10.5 System Optimum Assignment (SO)

The system optimum assignment is based on Wardrop's second principle, which states that *drivers cooperate with one another in order to minimize total system travel time.* This assignment can be thought of as a model in which congestion is minimized when drivers are told which routes to use. Obviously, this is not a behaviorally realistic model, but it can be useful to transport planners and engineers, trying to manage the traffic to minimize travel costs and therefore achieve an optimum social equilibrium.

$$\text{Minimize } Z = \sum_a x_a t_a(x_a) \tag{10.5}$$

subject to

$$\sum_k f_k^{rs} = q_{rs} \;\; : \forall r, s \tag{10.6}$$

$$x_a = \sum_r \sum_s \sum_k \delta_{a,k}^{rs} f_k^{rs} \;\; : \forall a \tag{10.7}$$

$$f_k^{rs} \geq 0 \;\; : \forall \; k, r, s \tag{10.8}$$

$$x_a \geq 0 \;\; : a \in A \tag{10.9}$$

x_a equilibrium flows in link a, t_a travel time on link a, f_k^{rs} flow on path k connecting O-D pair r-s, q_{rs} trip rate between r and s.

Example

To demonstrate how the assignment works, an example network is considered. This network has two nodes having two paths as links. Let us suppose a case where travel time is not a function of flow or in other words it is constant as shown in the figure below.

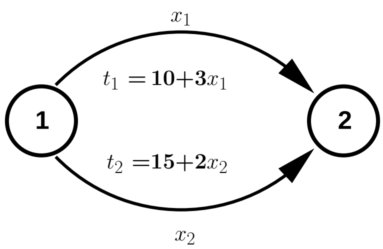


Figure 10:4: Two Link Problem with constant travel time function

Solution Substituting the travel time in equation , we get the following:

$$min \; Z(x) = x_1 * (10 + 3x_1) + x_2 * (15 + 2x_2) \tag{10.10}$$

$$= 10x_1 + 3{x_1}^2 + 15x_2 + 2{x_2}^2 \tag{10.11}$$

Substituting $x_2 = x_1 - 12$

$$min \; Z(x) == 10x_1 + 3{x_1}^2 + 15(12 - x_1) + 2(12 - x_1)^2 \tag{10.12}$$

$$\tag{10.13}$$

Differentiate the above equation to zero, and solving for x_1 and then x_2 leads to the solution x_1 = 5.3,x_2= 6.7 which gives Z(x) = 327.55

10.6 Other assignment methods

Let us discuss briefly some other assignments like incremental assignment, capacity restraint assignment, stochastic user equilibrium assignment and dynamic assignment.

10.6.1 Incremental assignment

Incremental assignment is a process in which fractions of traffic volumes are assigned in steps.In each step, a fixed proportion of total demand is assigned, based on all-or-nothing assignment. After each step, link travel times are recalculated based on link volumes. When there are many increments used, the flows may resemble an equilibrium assignment ; however, this method does not yield an equilibrium solution. Consequently, there will be inconsistencies between link volumes and travel times that can lead to errors in evaluation measures. Also, incremental assignment is influenced by the order in which volumes for O-D pairs are assigned, raising the possibility of additional bias in results.

10.6.2 Capacity restraint assignment

Capacity restraint assignment attempts to approximate an equilibrium solution by iterating between all-or-nothing traffic loadings and recalculating link travel times based on a congestion function that reflects link capacity. Unfortunately, this method does not converge and can flip-flop back and forth in loadings on some links.

10.6.3 Stochastic user equilibrium assignment

User equilibrium assignment procedures based on Wardrop's principle assume that all drivers perceive costs in an identical manner. A solution to assignment problem on this basis is an assignment such that no driver can reduce his journey cost by unilaterally changing route. Van Vilet considered as stochastic assignment models, all those models which explicitly allows non minimum cost routes to be selected. Virtually all such models assume that drivers perception of costs on any given route are not identical and that the trips between each O-D pair are divided among the routes with the most cheapest route attracting most trips. They have important advantage over other models because they load many routes between individual pairs of network nodes in a single pass through the tree building process,the assignments are more stable and less sensitive to slight variations in network definitions or link costs to be independent of flows and are thus most appropriate for use in uncongested traffic conditions such as in off peak periods or lightly trafficked rural areas.

10.6.4 Dynamic Assignment

Dynamic user equilibrium,expressed as an extension of Wardrop's user equilibrium principle, may be defined as the state of equilibrium which arises when no driver can reduce his disutility of travel by choosing a new route or departure time,where disutility includes, schedule delay in addition in to costs generally considered. Dynamic stochastic equilibrium may be similarly defined in terms of perceived utility of travel. The existence of such equilibrium in complex networks has not been proven theoretical and even if they exist the question of uniqueness remains open.

10.7 Summary

Traffic assignment is the last stage of traffic demand modeling. There are different types of traffic assignment models. All-or-nothing, User-equilibrium, and System-optimum assignment models are the commonly used models. All-or-nothing model is an unrealistic model since only one path between every O-D pair is utilised and they can give satisfactory results only when the network is least congested. User-equilibrium assignment is based on Wardrop's first principle and it's conditions are based on certain assumptions. Wardrop's second principle is utilized by System-optimum method and it tries to minimise the congestion by giving prior information to drivers regarding the respective routes to be chosen. Other assignment models are also briefly explained.

10.8 Problems

Calculate the system travel time and link flows by doing user equilibrium assignment for the network in the given figure 10:5. Verify that the flows are at user equilibrium.

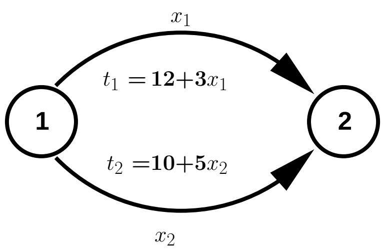


Figure 10:5: Two link problem with constant travel time

Solution Substituting the travel time in the respective equation yield to

$$\begin{aligned}\min Z(x) &= \int_0^{x_1} 12 + 3x_1 dx_1 + \int_0^{x_2} 10 + 5x_2 dx_2 \\ &= 12x_1 + \frac{3x_1^2}{2} + 10x_2 + \frac{5x_2^2}{2}\end{aligned}$$

subject to

$$x_1 + x_2 = 12$$

Substituting $x_2 = 12 - x_1$, in the above formulation will yield the unconstrained formulation as below :

$$\min\ Z(x) = 12x_1 + \frac{3x_1^2}{2} + 10(12 - x_1) + \frac{5(12 - x_1)^2}{2} \min\ Z(x) = 4x_1^2 - 58x_1 + 480$$

Differentiate the above equation x_1 and equating to zero,

$$\frac{dz(x)}{dx} = 0 - 58 + 8x_1 = 0 or x_1 = 7.25 Hence x_2 = 12 - 7,25 = 4.75$$

Travel times are

$$t_1 = 12 + 3 \times 7.25 = 33.75 t_2 = 10 + 5 \times 4.75 = 33.75 i.e. t_1 = t_2$$

It follows that the travel times are at user equilibrium.

Problems

1. Approximate length of National highway in India is:
   (a) 1000 km
   (b) 5000 km
   (c) 10000 km
   (d) 50000 km
   (e) 100000 km
2. The most accessible road is
   (a) National highway
   (b) State highway
   (c) Major District road
   (d) Other District road
   (e) Village road

Solutions

1. Approximate length of National highway in India is:
   (a) 1000 km
   (b) 5000 km
   (c) 10000 km
   (d) 50000 km
   (e) 100000 km
2. The most accessible road is
   (a) National highway
   (b) State highway
   (c) Major District road
   (d) Other District road
   (e) Village road$\surd$

Chapter 12

Cross sectional elements

12.1 Overview

The features of the cross-section of the pavement influences the life of the pavement as well as the riding comfort and safety. Of these, pavement surface characteristics affect both of these. Camber,kerbs, and geometry of various cross-sectional elements are important aspects to be considered in this regard. They are explained briefly in this chapter.

12.2 Pavement surface characteristics

For safe and comfortable driving four aspects of the pavement surface are important; the friction between the wheels and the pavement surface, smoothness of the road surface, the light reflection characteristics of the top of pavement surface, and drainage to water.

12.2.1 Friction

Friction between the wheel and the pavement surface is a crucial factor in the design of horizontal curves and thus the safe operating speed. Further, it also affect the acceleration and deceleration ability of vehicles. Lack of adequate friction can cause skidding or slipping of vehicles.

- Skidding happens when the path traveled along the road surface is more than the circumferential movement of the wheels due to friction
- Slip occurs when the wheel revolves more than the corresponding longitudinal movement along the road.

Various factors that affect friction are:

- Type of the pavement (like bituminous, concrete, or gravel),
- Condition of the pavement (dry or wet, hot or cold, etc),
- Condition of the tyre (new or old), and
- Speed and load of the vehicle.

The frictional force that develops between the wheel and the pavement is the load acting multiplied by a factor called the coefficient of friction and denoted as f. The choice of the value of f is a very complicated issue since it depends on many variables. IRC suggests the coefficient of longitudinal friction as **0.35-0.4** depending on the speed and coefficient of lateral friction as **0.15**. The former is useful in sight distance calculation and the latter in horizontal curve design.

12.2.2 Unevenness

It is always desirable to have an even surface, but it is seldom possible to have such a one. Even if a road is constructed with high quality pavers, it is possible to develop unevenness due to pavement failures. Unevenness affect the vehicle operating cost, speed, riding comfort, safety, fuel consumption and wear and tear of tyres.

Unevenness index is a measure of unevenness which is the cumulative measure of vertical undulations of the pavement surface recorded per unit horizontal length of the road. An unevenness index value less than 1500 mm/km is considered as good, a value less than 2500 mm.km is satisfactory up to speed of 100 kmph and values greater than 3200 mm/km is considered as uncomfortable even for 55 kmph.

12.2.3 Light reflection

- White roads have good visibility at night, but caused glare during day time.
- Black roads has no glare during day, but has poor visibility at night
- Concrete roads has better visibility and less glare

It is necessary that the road surface should be visible at night and reflection of light is the factor that answers it.

12.2.4 Drainage

The pavement surface should be absolutely impermeable to prevent seepage of water into the pavement layers. Further, both the geometry and texture of pavement surface should help in draining out the water from the surface in less time.

12.3 Camber

Camber or cant is the cross slope provided to raise middle of the road surface in the transverse direction to drain off rain water from road surface. The objectives of providing camber are:

- Surface protection especially for gravel and bituminous roads
- Sub-grade protection by proper drainage
- Quick drying of pavement which in turn increases safety

Too steep slope is undesirable for it will erode the surface. Camber is measured in *1 in n* or $n\%$ (Eg. 1 in 50 or 2%) and the value depends on the type of pavement surface. The values suggested by IRC for various categories of pavement is given in Table 12:1. The common types of camber are parabolic, straight, or combination of them (Figure 12:1)

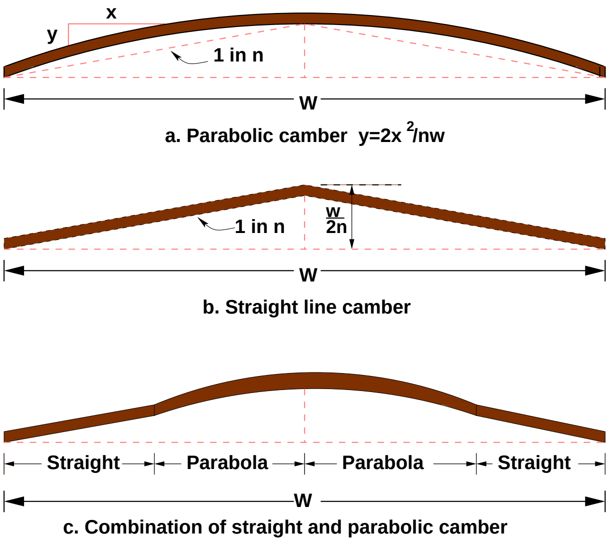


Figure 12:1: Different types of camber

Table 12:1: IRC Values for camber

Surface type	Heavy rain	Light rain
Concrete/Bituminous	2 %	1.7 %
Gravel/WBM	3 %	2.5 %
Earthen	4 %	3.0 %

Table 12:2: IRC Specification for carriage way width

Single lane	3.75
Two lane, no kerbs	7.0
Two lane, raised kerbs	7.5
Intermediate carriage	5.5
Multi-lane	3.5

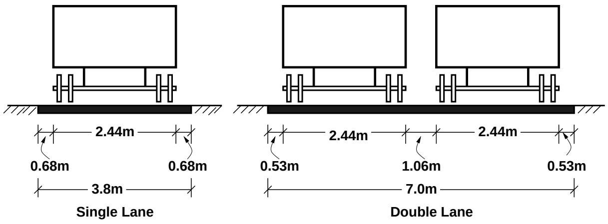


Figure 12:2: Lane width for single and two lane roads

12.4 Width of carriage way

Width of the carriage way or the width of the pavement depends on the width of the traffic lane and number of lanes. Width of a traffic lane depends on the width of the vehicle and the clearance. Side clearance improves operating speed and safety. The maximum permissible width of a vehicle is **2.44** and the desirable side clearance for single lane traffic is 0.68 m. This require minimum of lane width of 3.75 m for a single lane road (Figure 12:2a). However, the side clearance required is about 0.53 m, on either side and 1.06 m in the center. Therefore, a two lane road require minimum of 3.5 meter for each lane (Figure 12:2b). The desirable carriage way width recommended by IRC is given in Table 12:2

12.5 Kerbs

Kerbs indicate the boundary between the carriage way and the shoulder or islands or footpaths. Different types of kerbs are (Figure 12:3):

- Low or mountable kerbs : This type of kerbs are provided such that they encourage the traffic to remain in the through traffic lanes and also allow the driver to enter the shoulder area with little difficulty. The height of this kerb is about 10 cm above the pavement edge with a slope which allows the vehicle to climb easily. This is usually provided at medians and channelization schemes and also helps in longitudinal drainage.
- Semi-barrier type kerbs : When the pedestrian traffic is high, these kerbs are provided. Their height is 15 cm above the pavement edge. This type of kerb prevents encroachment of parking vehicles, but at acute emergency it is possible to drive over this kerb with some difficulty.
- Barrier type kerbs : They are designed to discourage vehicles from leaving the pavement. They are provided when there is considerable amount of pedestrian traffic. They are placed at a height of 20 cm above the pavement edge with a steep batter.

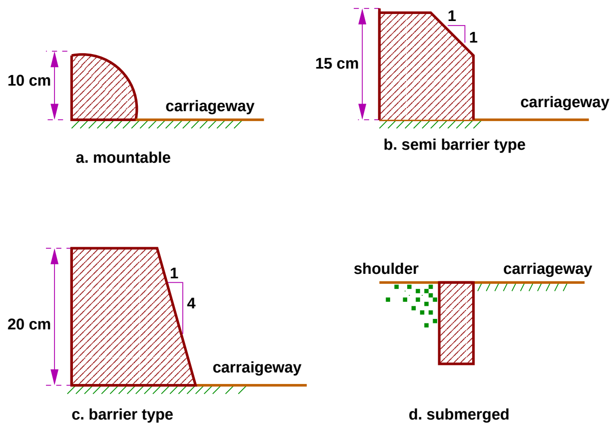


Figure 12:3: Different types of kerbs

- Submerged kerbs : They are used in rural roads. The kerbs are provided at pavement edges between the pavement edge and shoulders. They provide lateral confinement and stability to the pavement.

12.6 Road margins

The portion of the road beyond the carriageway and on the roadway can be generally called road margin. Various elements that form the road margins are given below.

12.6.1 Shoulders

Shoulders are provided along the road edge and is intended for accommodation of stopped vehicles, serve as an emergency lane for vehicles and provide lateral support for base and surface courses. The shoulder should be strong enough to bear the weight of a fully loaded truck even in wet conditions. The shoulder width should be adequate for giving working space around a stopped vehicle. It is desirable to have a width of 4.6 m for the shoulders. A minimum width of 2.5 m is recommended for 2-lane rural highways in India.

12.6.2 Parking lanes

Parking lanes are provided in urban lanes for side parking. Parallel parking is preferred because it is safe for the vehicles moving on the road. The parking lane should have a minimum of 3.0 m width in the case of parallel parking.

12.6.3 Bus-bays

Bus bays are provided by recessing the kerbs for bus stops. They are provided so that they do not obstruct the movement of vehicles in the carriage way. They should be at least 75 meters away from the intersection so that the traffic near the intersections is not affected by the bus-bay.

12.6.4 Service roads

Service roads or frontage roads give access to access controlled highways like freeways and expressways. They run parallel to the highway and will be usually isolated by a separator and access to the highway will be provided only at selected points. These roads are provided to avoid congestion in the expressways and also the speed of the traffic in those lanes is not reduced.

12.6.5 Cycle track

Cycle tracks are provided in urban areas when the volume of cycle traffic is high Minimum width of 2 meter is required, which may be increased by 1 meter for every additional track.

12.6.6 Footpath

Footpaths are exclusive right of way to pedestrians, especially in urban areas. They are provided for the safety of the pedestrians when both the pedestrian traffic and vehicular traffic is high. Minimum width is 1.5 meter and may be increased based on the traffic. The footpath should be either as smooth as the pavement or more smoother than that to induce the pedestrian to use the footpath.

12.6.7 Guard rails

They are provided at the edge of the shoulder usually when the road is on an embankment. They serve to prevent the vehicles from running off the embankment, especially when the height of the fill exceeds 3 m. Various designs of guard rails are there. Guard stones painted in alternate black and white are usually used. They also give better visibility of curves at night under headlights of vehicles.

12.7 Width of formation

Width of formation or roadway width is the sum of the widths of pavements or carriage way including separators and shoulders. This does not include the extra land in formation/cutting. The values suggested by IRC are given in Table 12:3.

Table 12:3: Width of formation for various classed of roads

Road classification	Roadway width in m	
	Plain and rolling terrain	Mountainous and steep terrain
NH/SH	12	6.25-8.8
MDR	9	4.75
ODR	7.5-9.0	4.75
VR	7.5	4.0

12.8 Right of way

Right of way (ROW) or land width is the width of land acquired for the road, along its alignment. It should be adequate to accommodate all the cross-sectional elements of the highway and may reasonably provide for future development. To prevent ribbon development along highways, control lines and building lines may be provided. Control line is a line which represents the nearest limits of future uncontrolled building activity in relation to a road. Building line represents a line on either side of the road, between which and the road no building activity is permitted at all. The right of way width is governed by:

- Width of formation: It depends on the category of the highway and width of roadway and road margins.
- Height of embankment or depth of cutting: It is governed by the topography and the vertical alignment.
- Side slopes of embankment or cutting: It depends on the height of the slope, soil type etc.
- Drainage system and their size which depends on rainfall, topography etc.
- Sight distance considerations : On curves etc. there is restriction to the visibility on the inner side of the curve due to the presence of some obstructions like building structures etc.
- Reserve land for future widening: Some land has to be acquired in advance anticipating future developments like widening of the road.

Table 12:4: Normal right of way for open areas

Road classification	Roadway width in m	
	Plain and rolling terrain	Mountainous and steep terrain
Open areas		
NH/SH	45	24
MDR	25	18
ODR	15	15
VR	12	9
Built-up areas		
NH/SH	30	20
MDR	20	15
ODR	15	12
VR	10	9

The importance of reserved land is emphasized by the following. Extra width of land is available for the construction of roadside facilities. Land acquisition is not possible later, because the land may be occupied for various other purposes (buildings, business etc.) The normal ROW requirements for built up and open areas as specified by IRC is given in Table 12:4 A typical cross section of a ROW is given in Figure 12:4.

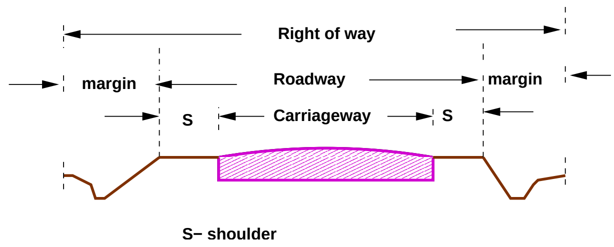


Figure 12:4: A typical Right of way (ROW)

12.9 Summary

The characteristics of cross-sectional elements are important in highway geometric design because they influence the safety and comfort. Camber provides for drainage, frictional resistance and reflectivity for safety etc. The road elements such as kerb, shoulders, carriageway width etc. should be adequate enough for smooth, safe and efficient movement of traffic. IRC has recommended the minimum values for all these cross-sectional elements.

12.10 Problems

1. IRC recommends the value for coefficient of lateral friction as
   (a) 0.05
   (b) 0.5
   (c) 0.15
   (d) 0.005
2. The height of semi-barrier type kerbs above the pavement edge is
   (a) 10cm
   (b) 15cm
   (c) 20cm
   (d) 25cm

12.11 Solutions

1. IRC recommends the value for coefficient of lateral friction as
   (a) 0.05
   (b) 0.5
   (c) 0.15$\surd$
   (d) 0.005
2. The height of semi-barrier type kerbs above the pavement edge is
   (a) 10cm

(b) 15cm$\surd$

(c) 20cm

(d) 25cm

Chapter 13

Sight distance

13.1 Overview

The safe and efficient operation of vehicles on the road depends very much on the visibility of the road ahead of the driver. Thus the geometric design of the road should be done such that any obstruction on the road length could be visible to the driver from some distance ahead . This distance is said to be the sight distance.

13.2 Types of sight distance

Sight distance available from a point is the actual distance along the road surface, over which a driver from a specified height above the carriage way has visibility of stationary or moving objects. Three sight distance situations are considered for design:

- Stopping sight distance (SSD) or the absolute minimum sight distance
- Intermediate sight distance (ISD) is defined as twice SSD
- Overtaking sight distance (OSD) for safe overtaking operation
- Head light sight distance is the distance visible to a driver during night driving under the illumination of head lights
- Safe sight distance to enter into an intersiection.

The most important consideration in all these is that at all times the driver traveling at the design speed of the highway must have sufficient carriageway distance within his line of vision to allow him to stop his vehicle before colliding with a slowly moving or stationary object appearing suddenly in his own traffic lane.

The computation of sight distance depends on:

- Reaction time of the driver

  Reaction time of a driver is the time taken from the instant the object is visible to the driver to the instant when the brakes are applied. The total reaction time may be split up into four components based on PIEV theory. In practice, all these times are usually combined into a total perception-reaction time suitable for design purposes as well as for easy measurement. Many of the studies shows that drivers require about 1.5 to 2 secs under normal conditions. However, taking into consideration the variability of driver characteristics, a higher value is normally used in design. For example, IRC suggests a reaction time of 2.5 secs.

- Speed of the vehicle

  The speed of the vehicle very much affects the sight distance. Higher the speed, more time will be required to stop the vehicle. Hence it is evident that, as the speed increases, sight distance also increases.

- Efficiency of brakes

  The efficiency of the brakes depends upon the age of the vehicle, vehicle characteristics etc. If the brake efficiency is 100%, the vehicle will stop the moment the brakes are applied. But practically, it is not possible to achieve 100% brake efficiency. Therefore the sight distance required will be more when the efficiency of brakes are less. Also for safe geometric design, we assume that the vehicles have only 50% brake efficiency.

- Frictional resistance between the tyre and the road

  The frictional resistance between the tyre and road plays an important role to bring the vehicle to stop. When the frictional resistance is more, the vehicles stop immediately. Thus sight required will be less. No separate provision for brake efficiency is provided while computing the sight distance. This is taken into account along with the factor of longitudinal friction. IRC has specified the value of longitudinal friction in between 0.35 to 0.4.

- Gradient of the road.

  Gradient of the road also affects the sight distance. While climbing up a gradient, the vehicle can stop immediately. Therefore sight distance required is less. While descending a gradient, gravity also comes into action and more time will be required to stop the vehicle. Sight distance required will be more in this case.

13.3 Stopping sight distance

Stopping sight distance (SSD) is the minimum sight distance available on a highway at any spot having sufficient length to enable the driver to stop a vehicle traveling at design speed, safely without collision with any other obstruction.

There is a term called *safe stopping distance* and is one of the important measures in traffic engineering. It is the distance a vehicle travels from the point at which a situation is first perceived to the time the deceleration is complete. Drivers must have adequate time if they are to suddenly respond to a situation. Thus in highway design, sight distance atleast equal to the safe stopping distance should be provided. The stopping sight distance is the sum of lag distance and the braking distance. Lag distance is the distance the vehicle traveled during the reaction time t and is given by vt, where v is the velocity in m/sec^2. Braking distance is the distance traveled by the vehicle during braking operation. For a level road this is obtained by equating the work done in stopping the vehicle and the kinetic energy of the vehicle. If F is the maximum frictional force developed and the braking distance is l, then work done against friction in stopping the vehicle is $Fl = fWl$ where W is the total weight of the vehicle. The kinetic energy at the design speed is

$$\begin{aligned} \frac{1}{2}mv^2 &= \frac{1}{2}\frac{Wv^2}{g} \\ fWl &= \frac{Wv^2}{2g} \end{aligned}$$

$$l = \frac{v^2}{2gf}$$

Therefore, the SSD = lag distance + braking distance and given by:

$$SSD = vt + \frac{v^2}{2gf} \tag{13.1}$$

where v is the design speed in m/sec^2, t is the reaction time in sec, g is the acceleration due to gravity and f is the coefficient of friction. The coefficient of friction f is given below for various design speed. When there is an

Table 13:1: Coefficient of longitudinal friction

Speed, kmph	<30	40	50	60	>80
f	0.40	0.38	0.37	0.36	0.35

ascending gradient of say $+n\%$, the component of gravity adds to braking action and hence braking distance is decreased. The component of gravity acting parallel to the surface which adds to the the braking force is equal to $W \sin\alpha \approx W \tan\alpha = Wn/100$. Equating kinetic energy and work done:

$$\left(fW + \frac{Wn}{100}\right) l = \frac{Wv^2}{2g}$$

$$l = \frac{v^2}{2g\left(f + \frac{n}{100}\right)}$$

Similarly the braking distance can be derived for a descending gradient. Therefore the general equation is given by Equation 13.2.

$$SSD = vt + \frac{v^2}{2g(f \pm 0.01n)} \tag{13.2}$$

13.4 Overtaking sight distance

The overtaking sight distance is the minimum distance open to the vision of the driver of a vehicle intending to overtake the slow vehicle ahead safely against the traffic in the opposite direction. The overtaking sight distance or passing sight distance is measured along the center line of the road over which a driver with his eye level 1.2 m above the road surface can see the top of an object 1.2 m above the road surface.

The factors that affect the OSD are:

- Velocities of the overtaking vehicle, overtaken vehicle and of the vehicle coming in the opposite direction.
- Spacing between vehicles, which in-turn depends on the speed
- Skill and reaction time of the driver
- Rate of acceleration of overtaking vehicle
- Gradient of the road

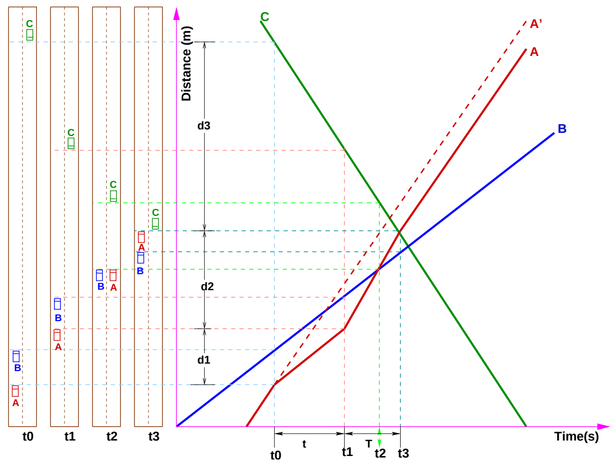


Figure 13:1: Time-space diagram: Illustration of overtaking sight distance

The dynamics of the overtaking operation is given in the figure which is a time-space diagram. The x-axis denotes the time and y-axis shows the distance traveled by the vehicles. The trajectory of the slow moving vehicle (B) is shown as a straight line which indicates that it is traveling at a constant speed. A fast moving vehicle (A) is traveling behind the vehicle B. The trajectory of the vehicle is shown initially with a steeper slope. The dotted line indicates the path of the vehicle A if B was absent. The vehicle A slows down to follow the vehicle B as shown in the figure with same slope from t_0 to t_1. Then it overtakes the vehicle B and occupies the left lane at time t_3. The time duration $T = t_3 - t_1$ is the actual duration of the overtaking operation. The snapshots of the road at time t_0, t_1, and t_3 are shown on the left side of the figure. From the Figure 13:1, the overtaking sight distance consists of three parts.

- d_1 the distance traveled by overtaking vehicle A during the reaction time $t = t_1 - t_0$
- d_2 the distance traveled by the vehicle during the actual overtaking operation $T = t3 - t1$
- d_3 is the distance traveled by on-coming vehicle C during the overtaking operation (T).

Therefore:

$$OSD = d_1 + d_2 + d_3 \tag{13.3}$$

It is assumed that the vehicle A is forced to reduce its speed to v_b, the speed of the slow moving vehicle B and travels behind it during the reaction time t of the driver. So d_1 is given by:

$$d_1 = v_b t \tag{13.4}$$

Then the vehicle A starts to accelerate, shifts the lane, overtake and shift back to the original lane. The vehicle A maintains the spacing s before and after overtaking. The spacing s in m is given by:

$$s = 0.7v_b + 6 \tag{13.5}$$

Let T be the duration of actual overtaking. The distance traveled by B during the overtaking operation is $2s + v_b T$. Also, during this time, vehicle A accelerated from initial velocity v_b and overtaking is completed while reaching final velocity v. Hence the distance traveled is given by:

$$\begin{aligned} d_2 &= v_b T + \frac{1}{2} a T^2 \\ 2s + v_b T &= v_b T + \frac{1}{2} a T^2 \\ 2s &= \frac{1}{2} a T^2 \\ T &= \sqrt{\frac{4s}{a}} \\ d_2 &= 2s + v_b \sqrt{\frac{4s}{a}} \end{aligned} \tag{13.6}$$

The distance traveled by the vehicle C moving at design speed v m/sec during overtaking operation is given by:

$$d_3 = vT \tag{13.7}$$

The the overtaking sight distance is (Figure 13:1)

$$OSD = v_b t + 2s + v_b \sqrt{\frac{4s}{a}} + vT \tag{13.8}$$

where v_b is the velocity of the slow moving vehicle in m/sec^2, t the reaction time of the driver in sec, s is the spacing between the two vehicle in m given by equation 13.5 and a is the overtaking vehicles acceleration in m/sec^2. In case the speed of the overtaken vehicle is not given, it can be assumed that it moves 16 kmph slower the the design speed.

The acceleration values of the fast vehicle depends on its speed and given in Table 13:2. Note that:

Table 13:2: Maximum overtaking acceleration at different speeds

Speed (kmph)	Maximum overtaking acceleration (m/sec^2)
25	1.41
30	1.30
40	1.24
50	1.11
65	0.92
80	0.72
100	0.53

- On divided highways, d_3 need not be considered
- On divided highways with four or more lanes, IRC suggests that it is not necessary to provide the OSD, but only SSD is sufficient.

Overtaking zones

Overtaking zones are provided when OSD cannot be provided throughout the length of the highway. These are zones dedicated for overtaking operation, marked with wide roads. The desirable length of overtaking zones is 5 time OSD and the minimum is three times OSD (Figure 13:2).

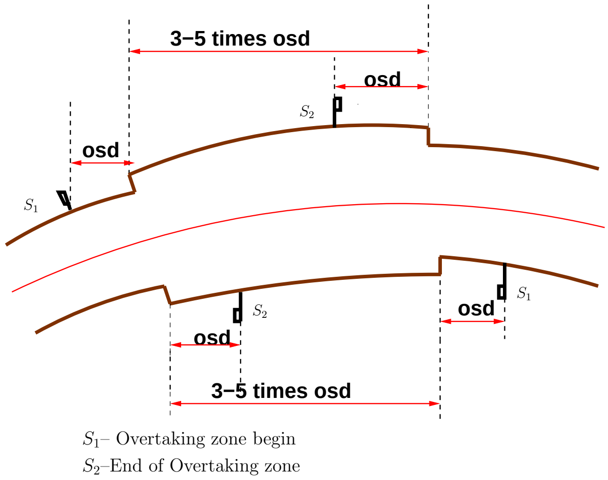


Figure 13:2: Overtaking zones

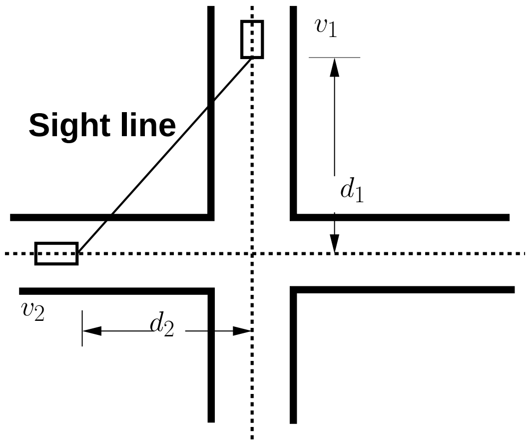


Figure 13:3: Sight distance at intersections

13.5 Sight distance at intersections

At intersections where two or more roads meet, visibility should be provided for the drivers approaching the intersection from either sides. They should be able to perceive a hazard and stop the vehicle if required. Stopping sight distance for each road can be computed from the design speed. The sight distance should be provided such that the drivers on either side should be able to see each other. This is illustrated in the figure 13:3.

Design of sight distance at intersections may be used on three possible conditions:

- Enabling approaching vehicle to change the speed
- Enabling approaching vehicle to stop
- Enabling stopped vehicle to cross a main road

13.6 Summary

One of the key factors for the safe and efficient operation of vehicles on the road is sight distance. Sight distances ensure overtaking and stopping operations at the right time. Different types of sight distances and the equations to find each of these had been discussed here.

13.7 Problems

1. Calculate SSD for V=50kmph for (a) two-way traffic in a two lane road (b) two-way traffic in single lane road. (Hint: f=0.37, t=2.5) [Ans: (a)61.4 m (b) 122.8 m.

   Given: V=50km/hr = 13.9m/s f=0.37 t= 2.5 sec stopping distance=lag distance + braking distance

   $$SD = vt + v^2/2gf$$

   Stopping Distance = 61.4 m
   Stopping sight distance when there are two lanes = stopping distance= 61.4m.
   Stopping sight distance for a two way traffic for a single lane = 2[stopping distance]=122.8m

2. Find minimum sight distance to avoid head-on collision of two cars approaching at 90 kmph and 60 kmph. Given t=2.5sec, f=0.7 and brake efficiency of 50 percent in either case. (Hint: brake efficiency reduces the coefficient of friction by 50 percent). [Ans: SD=153.6+82.2=235.8m]

   Given: V_1 =90 Km/hr. V_2 = 60 Km/hr. t = 2.5sec. Braking efficiency=50%. f=.7.
   Stopping distance for one of the cars

   $$SD = vt + v^2/2gf$$

   Coefficient of friction due to braking efficiency of 50% = 0.5*0.7=0.35. Stopping sight distance of first car= SD_1= 153.6m
   Stopping sight distance of second car= SD_2= 82.2m
   Stopping sight distance to avoid head on collision of the two approaching cars SD_1+ SD_2=235.8m.

3. Find SSD for a descending gradient of 2% for V=80kmph. [Ans: 132m].

   Given: Gradient(n) = -2V = 80 Km/hr.

   $$SD = vt + v^2/2g(f - n\%)$$

   SSD on road with gradient = 132m.

4. Find head light sight distance and intermediate sight distance for V=65 kmph. (Hint: f=0.36, t=2.5 s, HSD=SSD, ISD=2*SSD) [Ans: 91.4 and 182.8 m]

   Given: V=65km/hr f=0.36 t= 2.5 sec

$$SD = vt + v^2/2gf$$

   Headlight Sight distance = 91.4m.
   Intermediate Sight distance= 2[SSD]= 182.8m.

5. Overtaking and overtaken vehicles are at 70 and 40 kmph respectively. find (i) OSD (ii) min. and desirable length of overtaking zone (iii) show the sketch of overtaking zone with location of sign post (hint: a=0.99 m/sec2) [Ans: (i) 278 m (ii) 834 m/1390]

6. Calculate OSD for V=96 kmph. Assume all other data. (Hint: Vb=96-16kmph. a=0.72, t=2.5s) [Ans: OSD one way 342m, OSD two way 646m]

Chapter 14

Horizontal alignment I

14.1 Overview

Horizontal alignment is one of the most important features influencing the efficiency and safety of a highway. A poor design will result in lower speeds and resultant reduction in highway performance in terms of safety and comfort. In addition, it may increase the cost of vehicle operations and lower the highway capacity. Horizontal alignment design involves the understanding on the design aspects such as design speed and the effect of horizontal curve on the vehicles. The horizontal curve design elements include design of super elevation, extra widening at horizontal curves, design of transition curve, and set back distance. These will be discussed in this chapter and the following two chapters.

14.2 Design Speed

The design speed, as noted earlier, is the single most important factor in the design of horizontal alignment. The design speed also depends on the type of the road. For e.g, the design speed expected from a National highway will be much higher than a village road, and hence the curve geometry will vary significantly.

The design speed also depends on the type of terrain. A plain terrain can afford to have any geometry, but for the same standard in a hilly terrain requires substantial cutting and filling implying exorbitant costs as well as safety concern due to unstable slopes. Therefore, the design speed is normally reduced for terrains with steep slopes.

For instance, Indian Road Congress (IRC) has classified the terrains into four categories, namely plain, rolling, mountainous, and steep based on the cross slope as given in table 14:1. Based on the type of road and type of terrain the design speed varies. The IRC has suggested desirable or ruling speed as well as minimum suggested design speed and is tabulated in table 14:2. The recommended design speed is given in Table 14:2.

Table 14:1: Terrain classification

Terrain classification	Cross slope (%)
Plain	0-10
Rolling	10-25
Mountainous	25-60
Steep	> 60

Table 14:2: Design speed in km/hr as per IRC (ruling and minimum)

Type	Plain	Rolling	Hilly	Steep
NS&SH	100-80	80-65	50-40	40-30
MDR	80-65	65-50	40-30	30-20
ODR	65-50	50-40	30-25	25-20
VR	50-40	40-35	25-20	25-20

14.3 Horizontal curve

The presence of horizontal curve imparts centrifugal force which is a reactive force acting outward on a vehicle negotiating it. Centrifugal force depends on speed and radius of the horizontal curve and is counteracted to a certain extent by transverse friction between the tyre and pavement surface. On a curved road, this force tends to cause the vehicle to overrun or to slide outward from the centre of road curvature. For proper design of the curve, an understanding of the forces acting on a vehicle taking a horizontal curve is necessary. Various forces acting on the vehicle are illustrated in the figure 14:1.

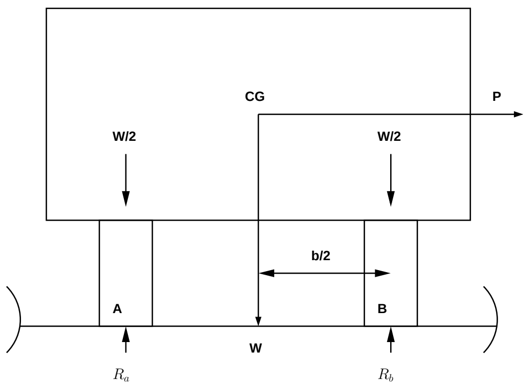


Figure 14:1: Effect of horizontal curve

They are the centrifugal force (P) acting outward, weight of the vehicle (W) acting downward, and the reaction of the ground on the wheels (R_A and R_B). The centrifugal force and the weight is assumed to be from the centre of gravity which is at h units above the ground. Let the wheel base be assumed as b units. The centrifugal force P in kg/m^2 is given by

$$P = \frac{Wv^2}{gR} \tag{14.1}$$

where W is the weight of the vehicle in kg, v is the speed of the vehicle in m/sec, g is the acceleration due to gravity in m/sec^2 and R is the radius of the curve in m.

The centrifugal ratio or the impact factor $\frac{P}{W}$ is given by:

$$\frac{P}{W} = \frac{v^2}{gR} \tag{14.2}$$

The centrifugal force has two effects: A tendency to overturn the vehicle about the outer wheels and a tendency for transverse skidding. Taking moments of the forces with respect to the outer wheel when the vehicle is just about to override,

$$Ph = W\frac{b}{2} \text{ or } \frac{\text{P}}{\text{W}} = \frac{\text{b}}{\text{2h}}$$

At the equilibrium over turning is possible when

$$\frac{v^2}{gR} = \frac{b}{2h}$$

and for safety the following condition must satisfy:

$$\frac{b}{2h} > \frac{v^2}{gR} \tag{14.3}$$

The second tendency of the vehicle is for transverse skidding. i.e. When the the centrifugal force P is greater than the maximum possible transverse skid resistance due to friction between the pavement surface and tyre. The transverse skid resistance (F) is given by:

$$\begin{aligned} F = & \quad F_A + F_B \\ = & \quad f(R_A + R_B) \\ = & \quad fW \end{aligned}$$

where F_A and F_B is the fractional force at tyre A and B, R_A and R_B is the reaction at tyre A and B, f is the lateral coefficient of friction and W is the weight of the vehicle. This is counteracted by the centrifugal force (P), and equating:

$$P = fW \text{ or } \frac{\text{P}}{\text{W}} = \text{f}$$

At equilibrium, when skidding takes place (from equation14.2)

$$\frac{P}{W} = f = \frac{v^2}{gR}$$

and for safety the following condition must satisfy:

$$f > \frac{v^2}{gR} \tag{14.4}$$

Equation 14.3 and 14.4 give the stable condition for design. If equation 14.3 is violated, the vehicle will overturn at the horizontal curve and if equation 14.4 is violated, the vehicle will skid at the horizontal curve

14.4 Analysis of super-elevation

Super-elevation or cant or banking is the transverse slope provided at horizontal curve to counteract the centrifugal force, by raising the outer edge of the pavement with respect to the inner edge, throughout the length of the horizontal curve. When the outer edge is raised, a component of the curve weight will be complimented in counteracting the effect of centrifugal force. In order to find out how much this raising should be, the following analysis may be done. The forces acting on a vehicle while taking a horizontal curve with superelevation is shown in figure 14:2.

Forces acting on a vehicle on horizontal curve of radius R m at a speed of v m/sec^2 are:

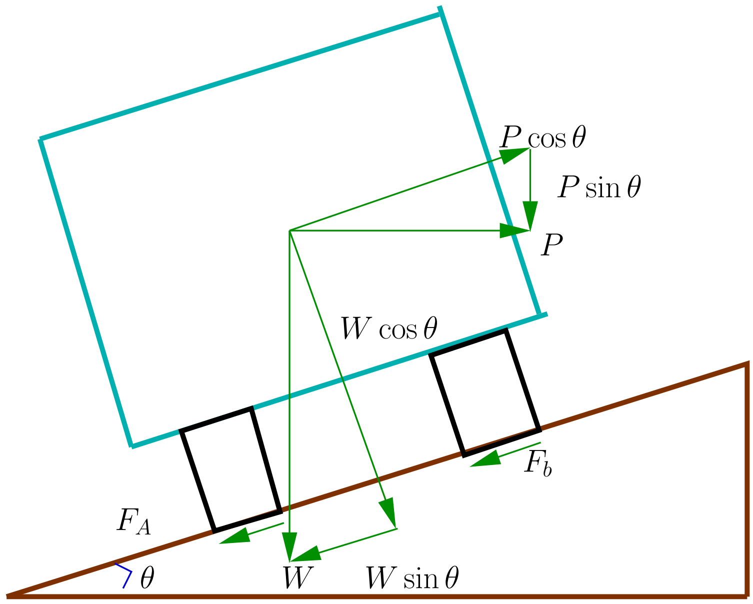


Figure 14:2: Analysis of super-elevation

- P the centrifugal force acting horizontally out-wards through the center of gravity,
- W the weight of the vehicle acting down-wards through the center of gravity, and
- F the friction force between the wheels and the pavement, along the surface inward.

At equilibrium, by resolving the forces parallel to the surface of the pavement we get,

$$\begin{aligned} P\cos\theta &= W\sin\theta + F_A + F_B \\ &= W\sin\theta + f(R_A + R_B) \\ &= W\sin\theta + f(W\cos\theta + P\sin\theta) \end{aligned}$$

where W is the weight of the vehicle, P is the centrifugal force, f is the coefficient of friction, θ is the transverse slope due to superelevation. Dividing by $W\cos\theta$, we get:

$$\begin{aligned} \frac{P\cos\theta}{W\cos\theta} &= \frac{W\sin\theta}{W\cos\theta} + \frac{fW\cos\theta}{W\cos\theta} + \frac{fPsin\theta}{W\cos\theta} \\ \frac{P}{W} &= \tan\theta + f + f\frac{P}{W}\tan\theta \\ \frac{P}{W}(1 - f\tan\theta) &= \tan\theta + f \\ \frac{P}{W} &= \frac{\tan\theta + f}{1 - f\tan\theta} \end{aligned} \tag{14.5}$$

We have already derived an expression for P/W.By substituting this in equation 14.5, we get:

$$\frac{v^2}{gR} = \frac{\tan\theta + f}{1 - f\tan\theta} \tag{14.6}$$

This is an exact expression for superelevation. But normally, $f = 0.15$ and $\theta < 4^o$, $1 - f\tan\theta \approx 1$ and for small θ, $\tan\theta \approx \sin\theta = E/B = e$, then equation 14.6 becomes:

$$e + f = \frac{v^2}{gR} \tag{14.7}$$

where, e is the rate of super elevation, f the coefficient of lateral friction 0.15, v the speed of the vehicle in m/sec^2, R the radius of the curve in m and $g = 9.8m/sec^2$.

Three specific cases that can arise from equation 14.7 are as follows:

1 If there is no friction due to some practical reasons, then $f = 0$ and equation 14.7 becomes $e = \frac{v^2}{gR}$. This results in the situation where the pressure on the outer and inner wheels are same; requiring very high super-elevation e.

2 If there is no super-elevation provided due to some practical reasons, then $e = 0$ and equation 14.7 becomes $f = \frac{v^2}{gR}$. This results in a very high coefficient of friction.

3 If $e = 0$ and $f = 0.15$ then for safe traveling speed from equation 14.7 is given by $v_b = \sqrt{fgR}$ where v_b is the restricted speed.

14.5 Summary

Design speed plays a major role in designing the elements of horizontal alignment. The most important element is superelevation which is influenced by speed, radius of curve and frictional resistance of pavement. Superelevation is necessary to balance centrifugal force. The design part is dealt in the next chapter.

14.6 Problems

1. The design speed recommended by IRC for National highways passign through rolling terrain is in the range of

   (a) 100-80

   (b) 80-65

   (c) 120-100

   (d) 50-40

2. For safety against skidding, the condition to be satisfied is

   (a) f¿$\frac{v^2}{gR}$

   (b) f¡$\frac{v^2}{gR}$

   (c) f¿$\frac{v}{gR}$

   (d) f=$\frac{v^2}{gR}$

14.7 Solutions

1. The design speed recommended by IRC for National highways passign through rolling terrain is in the range of
   (a) 100-80
   (b) 80-65$\surd$
   (c) 120-100
   (d) 50-40
2. For safety against skidding, the condition to be satisfied is
   (a) f¿$\frac{v^2}{gR}$$\surd$
   (b) f¡$\frac{v^2}{gR}$
   (c) f¿$\frac{v}{gR}$
   (d) f=$\frac{v^2}{gR}$

Chapter 15

Horizontal alignment II

15.1 Overview

This section discusses the design of superelevation and how it is attained. A brief discussion on pavement widening at curves is also given.

15.2 Guidelines on superelevation

While designing the various elements of the road like superelevation, we design it for a particular vehicle called design vehicle which has some standard weight and dimensions. But in the actual case, the road has to cater for mixed traffic. Different vehicles with different dimensions and varying speeds ply on the road. For example, in the case of a heavily loaded truck with high centre of gravity and low speed, superelevation should be less, otherwise chances of toppling are more. Taking into practical considerations of all such situations, IRC has given some guidelines about the maximum and minimum superelevation etc. These are all discussed in detail in the following sections.

15.2.1 Design of super-elevation

For fast moving vehicles, providing higher superelevation without considering coefficient of friction is safe, i.e. centrifugal force is fully counteracted by the weight of the vehicle or superelevation. For slow moving vehicles, providing lower superelevation considering coefficient of friction is safe, i.e.centrifugal force is counteracted by superelevation and coefficient of friction . IRC suggests following design procedure:

Step 1 Find e for 75 percent of design speed, neglecting f, i.e $e_1 = \frac{(0.75v)^2}{gR}$.

Step 2 If $e_1 \leq 0.07$, then $e = e_1 = \frac{(0.75v)^2}{gR}$, else if $e_1 > 0.07$ go to step 3.

Step 3 Find f_1 for the design speed and max e, i.e $f_1 = \frac{v^2}{gR} - e = \frac{v^2}{gR} - 0.07$. If $f_1 < 0.15$, then the maximum $e = 0.07$ is safe for the design speed, else go to step 4.

Step 4 Find the *allowable speed* v_a for the maximum $e = 0.07$ and $f = 0.15$, $v_a = \sqrt{0.22gR}$ If $v_a \geq v$ then the design is adequate, otherwise use speed adopt control measures or look for speed control measures.

15.2.2 Maximum and minimum super-elevation

Depends on (a) slow moving vehicle and (b) heavy loaded trucks with high CG. IRC specifies a maximum super-elevation of 7 percent for plain and rolling terrain, while that of hilly terrain is 10 percent and urban road is 4 percent. The minimum super elevation is 2-4 percent for drainage purpose, especially for large radius of the horizontal curve.

15.2.3 Attainment of super-elevation

1. Elimination of the crown of the cambered section by:
   (a) rotating the outer edge about the crown : The outer half of the cross slope is rotated about the crown at a desired rate such that this surface falls on the same plane as the inner half.
   (b) shifting the position of the crown: This method is also known as diagonal crown method. Here the position of the crown is progressively shifted outwards, thus increasing the width of the inner half of cross section progressively.
2. Rotation of the pavement cross section to attain full super elevation by:There are two methods of attaining superelevation by rotating the pavement
   (a) rotation about the center line : The pavement is rotated such that the inner edge is depressed and the outer edge is raised both by half the total amount of superelevation, i.e., by $E/2$ with respect to the centre.
   (b) rotation about the inner edge: Here the pavement is rotated raising the outer edge as well as the centre such that the outer edge is raised by the full amount of superelevation with respect to the inner edge.

15.3 Radius of Horizontal Curve

The radius of the horizontal curve is an important design aspect of the geometric design. The maximum comfortable speed on a horizontal curve depends on the radius of the curve. Although it is possible to design the curve with maximum superelevation and coefficient of friction, it is not desirable because re-alignment would be required if the design speed is increased in future. Therefore, a ruling minimum radius R_{ruling} can be derived by assuming maximum superelevation and coefficient of friction.

$$R_{\text{ruling}} = \frac{v^2}{g(e+f)} \tag{15.1}$$

Ideally, the radius of the curve should be higher than R_{ruling}. However, very large curves are also not desirable. Setting out large curves in the field becomes difficult. In addition, it also enhances driving strain.

15.4 Extra widening

Extra widening refers to the additional width of carriageway that is required on a curved section of a road over and above that required on a straight alignment. This widening is done due to two reasons: the first and most important is the additional width required for a vehicle taking a horizontal curve and the second is due to the tendency of the drivers to ply away from the edge of the carriageway as they drive on a curve. The first is referred as the mechanical widening and the second is called the psychological widening. These are discussed in detail below.

15.4.1 Mechanical widening

The reasons for the mechanical widening are: When a vehicle negotiates a horizontal curve, the rear wheels follow a path of shorter radius than the front wheels as shown in figure 15.5. This phenomenon is called off-tracking, and has the effect of increasing the effective width of a road space required by the vehicle. Therefore, to provide the same clearance between vehicles traveling in opposite direction on curved roads as is provided on straight sections, there must be extra width of carriageway available. This is an important factor when high proportion of vehicles are using the road. Trailor trucks also need extra carriageway, depending on the type of joint. In addition speeds higher than the design speed causes transverse skidding which requires additional width for safety purpose. The expression for extra width can be derived from the simple geometry of a vehicle at a horizontal curve as shown in figure 15.5. Let R_1 is the radius of the outer track line of the rear wheel, R_2 is the radius of the outer track line of the front wheel l is the distance between the front and rear wheel, n is the number of lanes, then the mechanical widening W_m (refer figure 15:1) is derived below:

$$\begin{aligned} R_2^2 &= R_1^2 + l^2 \\ &= (R_2 - W_m)^2 + l^2 \\ &= R_2^2 - 2R_2W_m + W_m^2 + l^2 \\ 2R_2W_m - W_m^2 &= l^2 \end{aligned}$$

Therefore the widening needed for a single lane road is:

$$W_m = \frac{l^2}{2R_2 - W_m} \tag{15.2}$$

If the road has n lanes, the extra widening should be provided on each lane. Therefore, the extra widening of a road with n lanes is given by,

$$W_m = \frac{nl^2}{2R_2 - W_m} \tag{15.3}$$

Please note that for large radius, $R_2 \approx R$, which is the mean radius of the curve,then W_m is given by:

$$W_m = \frac{nl^2}{2R} \tag{15.4}$$

15.4.2 Psychological widening

Widening of pavements has to be done for some psychological reasons also. There is a tendency for the drivers to drive close to the edges of the pavement on curves. Some extra space is to be provided for more clearance for the crossing and overtaking operations on curves. IRC proposed an empirical relation for the psychological widening at horizontal curves W_{ps}:

$$W_{ps} = \frac{v}{2.64\sqrt{R}} \tag{15.5}$$

Therefore, the total widening needed at a horizontal curve W_e is:

$$W_e = W_m + W_{ps} = \frac{nl^2}{2R} + \frac{v}{2.64\sqrt{R}} \tag{15.6}$$

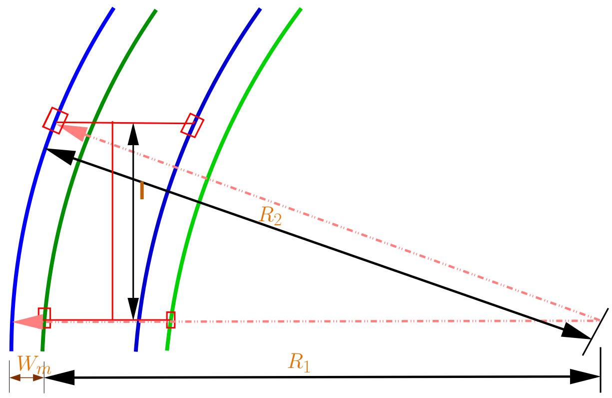


Figure 15:1: Extra-widening at a horizontal curve

15.5 Summary

In our country, the design of super-elevation follows IRC guidelines wherein the initial design is done by considering 75% of design speed and the safety of design is assessed. Pavement is to be given extra width at curves to account for mechanical and psychological aspects.

15.6 Problems

1. A national highway passing through a rolling terrain has two horizontal curves of radius 450 m and 150 m. Design the required super-elevation for the curves as per IRC guidelines.

   Solution

   Assumptions The ruling design speed for NH passing through a rolling terrain is 80 kmph. The coefficient of lateral friction f=0.15. The maximum permissible super elevation e=0.07.

   Case: Radius = 450m

   Step 1 Find e for 75 percent of design speed, neglecting f, i.e $e_1 = \frac{(0.75v)^2}{gR}$. $v = \frac{V}{3.6} = \frac{80}{3.6} = 22.22 m/sec$
   $e_1 = \frac{(0.75 \times 22.22)^2}{9.81 \times 450} = 0.0629$

   Step 2 $e_1 \leq 0.07$. Hence the design is sufficient.

   Answer: Design superelevation: 0.06.

Case: Radius = 150m

Step 1 Find e for 75 percent of design speed, neglecting f, i.e $e_1 = \frac{(0.75v)^2}{gR}$. $v = \frac{V}{3.6} = \frac{80}{3.6} = 22.22m/sec$
$e_1 = \frac{(0.75 \times 22.22)^2}{9.81 \times 150} = 0.188$ Max.e to be provided $= 0.07$

Step 3 Find f_1 for the design speed and max e, i.e $f_1 = \frac{v^2}{gR} - e = \frac{22.22^2}{9.81 \times 150} - 0.07 = 0.265$.

Step 4 Find the *allowable speed* v_a for the maximum $e = 0.07$ and $f = 0.15$, $v_a = \sqrt{0.22gR} = \sqrt{0.22 \times 9.81 \times 150} = 17.99m/sec = 17.99 \times 3.6 = 64kmph$

2. 2.Given R=100m, V=50 kmph, f=0.15. Find:
   (a) e if full lateral friction is assumed to develop [Ans: 0.047]
   (b) find f needed if no super elevation is provide [Ans: 0.197]
   (c) Find equilibrium super-elevation if pressure on inner and outer outer wheel should be equal (Hint: f=0) [Ans: 0.197]
3. 3.Two lane road, V=80 kmph, R=480 m, Width of the pavement at the horizontal curve=7.5 m. (i) Design super elevation for mixed traffic. (ii) By how much the outer edge of the pavement is to be raised with respect to the centerline, if the pavement is rotated with respect to centerline. [Ans:(i) 0.059 (ii) 0.22m]
4. 4.Design rate of super elevation for a horizontal highway curve of radius 500 m and speed 100 kmph. [Ans: e=0.07, f=0.087 and with in limits]
5. Given V=80 kmph, R=200m Design for super elevation. (Hint: f=0.15) [Ans: Allowable speed is 74.75 kmph and e=0.07]
6. 5.Calculate the ruling minimum and absolute minimum radius of horizontal curve of a NH in plain terrain. (Hint: V_{ruling}=100kmph, V_{min}=80kmph., e=0.07, f=0.15) [Ans: 360 and 230 m]
7. 6.Find the extra widening for W=7m, R=250m, longest wheel base, l=7m, V=70kmph. (Hint: n=2) [Ans:0.662m]
8. 7.Find the width of a pavement on a horizontal curve for a new NH on rolling terrain. Assume all data. (Hint: V=80kmph for rolling terrain, normal W=7.0m, n=2, l=6.0m, e=0.07, f=0.15). [Ans: R_{ruling}=230m, We=0.71, W at HC=7.71m]

Chapter 16

Horizontal alignment III

16.1 Overview

In this section we will deal with the design of transition curves and setback distances. Transition curve ensures a smooth change from straight road to circular curves. Setback distance looks in for safety at circular curves taking into consideration the sight distance aspects. A short note on curve resistance is also included.

16.2 Horizontal Transition Curves

Transition curve is provided to change the horizontal alignment from straight to circular curve gradually and has a radius which decreases from infinity at the straight end (*tangent point*) to the desired radius of the circular curve at the other end (*curve point*) There are five objectives for providing transition curve and are given below:

1. to introduce gradually the centrifugal force between the tangent point and the beginning of the circular curve, avoiding sudden jerk on the vehicle.This increases the comfort of passengers.
2. to enable the driver turn the steering gradually for his own comfort and security,
3. to provide gradual introduction of super elevation, and
4. to provide gradual introduction of extra widening.
5. to enhance the aesthetic appearance of the road.

16.2.1 Type of transition curve

Different types of transition curves are spiral or clothoid, cubic parabola, and Lemniscate. IRC recommends spiral as the transition curve because:

1. it fulfills the requirement of an ideal transition curve, that is;
   (a) rate of change or centrifugal acceleration is consistent (smooth) and
   (b) radius of the transition curve is ∞ at the straight edge and changes to R at the curve point ($L_s \propto \frac{1}{R}$) and calculation and field implementation is very easy.

16.2.2 Length of transition curve

The length of the transition curve should be determined as the maximum of the following three criteria: rate of change of centrifugal acceleration, rate of change of superelevation, and an empirical formula given by IRC.

1. Rate of change of centrifugal acceleration

At the tangent point, radius is infinity and hence centrifugal acceleration is zero. At the end of the transition, the radius R has minimum value R. The rate of change of centrifugal acceleration should be adopted such that the design should not cause discomfort to the drivers. If c is the rate of change of centrifugal acceleration, it can be written as:

$$\begin{aligned} c &= \frac{\frac{v^2}{R} - 0}{t}, \\ &= \frac{\frac{v^2}{R}}{\frac{L_s}{v}}, \\ &= \frac{v^3}{L_s R}. \end{aligned}$$

Therefore, the length of the transition curve L_{s_1} in m is

$$L_{s_1} = \frac{v^3}{cR}, \tag{16.1}$$

where c is the rate of change of centrifugal acceleration given by an empirical formula suggested by by IRC as below:

$$c = \frac{80}{75 + 3.6v}, \tag{16.2}$$

subject to :

$$\begin{aligned} c_{\min} &= 0.5, \\ c_{\max} &= 0.8. \end{aligned}$$

2. Rate of introduction of super-elevation

Raise (E) of the outer edge with respect to inner edge is given by $E = eB = e(W + W_e)$. The rate of change of this raise from 0 to E is achieved gradually with a gradient of 1 in N over the length of the transition curve (typical range of N is 60-150). Therefore, the length of the transition curve L_{s_2} is:

$$L_{s_2} = Ne(W + W_e) \tag{16.3}$$

3. By empirical formula

IRC suggest the length of the transition curve is minimum for a plain and rolling terrain:

$$L_{s_3} = \frac{35v^2}{R} \tag{16.4}$$

and for steep and hilly terrain is:

$$L_{s_3} = \frac{12.96v^2}{R} \tag{16.5}$$

and the shift s as:

$$s = \frac{L_s^2}{24R} \tag{16.6}$$

The length of the transition curve L_s is the maximum of equations 16.1, 16.3 and 16.4or16.5, i.e.

$$L_s = \text{Max}: \; (L_{s_1}, L_{s_2}, L_{s_3}) \tag{16.7}$$

16.3 Setback Distance

Setback distance m or the clearance distance is the distance required from the centerline of a horizontal curve to an obstruction on the inner side of the curve to provide adequate sight distance at a horizontal curve. The setback distance depends on:

1. sight distance (OSD, ISD and OSD),
2. radius of the curve, and
3. length of the curve.

Case (a) $L_s < L_c$

For single lane roads:

$$\begin{aligned}\alpha &= \frac{s}{R} \text{ radians} \\ &= \frac{180s}{\pi R} \text{ degrees} \\ \alpha/2 &= \frac{180s}{2\pi R} \text{ degrees}\end{aligned} \tag{16.8}$$

Therefore,

$$m = R - R\cos\left(\frac{\alpha}{2}\right) \tag{16.9}$$

For multi lane roads, if d is the distance between centerline of the road and the centerline of the inner lane, then

$$m = R - (R-d)\cos\left(\frac{180s}{2\pi(R-d)}\right) \tag{16.10}$$

$$m = R - R\cos\left(\frac{\alpha}{2}\right) \tag{16.11}$$

Case (b) $L_s > L_c$

For single lane:

$$\begin{aligned}m_1 &= R - R\cos(\alpha/2) \\ m_2 &= \frac{(S-L_c)}{2}\sin(\alpha/2)\end{aligned}$$

The set back is the sum of m_1 and m_2 given by:

$$m = R - R\cos(\alpha/2) + \frac{(S-L_c)}{2}\sin(\alpha/2) \tag{16.12}$$

where $\frac{\alpha}{2} = \frac{180L_c}{2\pi R}$. For multi-lane road $\frac{\alpha}{2} = \frac{180L_c}{2\pi(R-d)}$, and m is given by

$$m = R - (R-d)\cos(\alpha/2) + \frac{(S-L_c)}{2}\sin(\alpha/2) \tag{16.13}$$

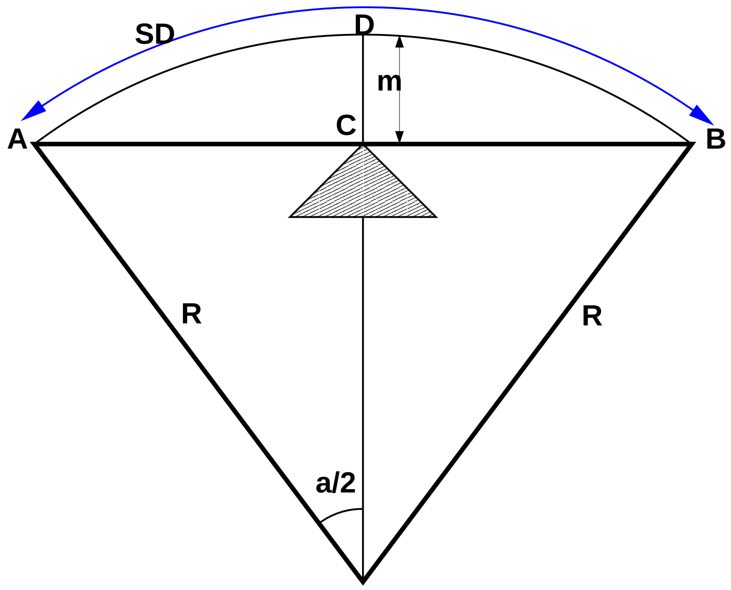


Figure 16:1: Set-back for single lane roads ($L_s < L_c$)

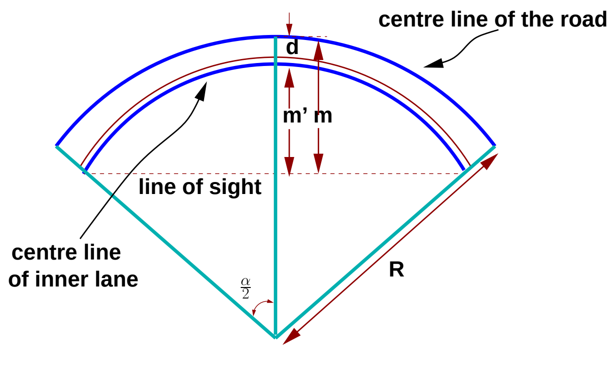


Figure 16:2: Set-back for multi-lane roads ($L_s < L_c$)

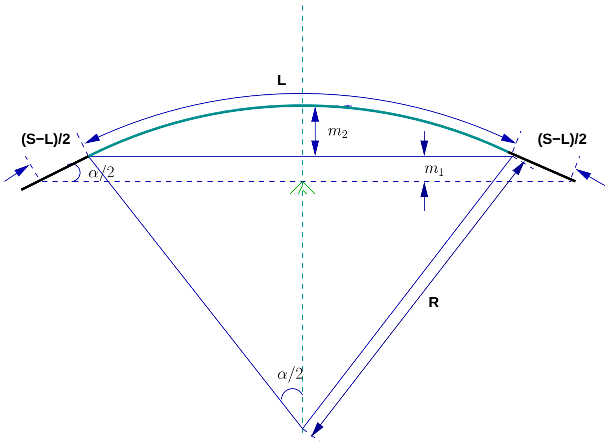


Figure 16:3: Set back for single lane roads ($L_s < L_c$)

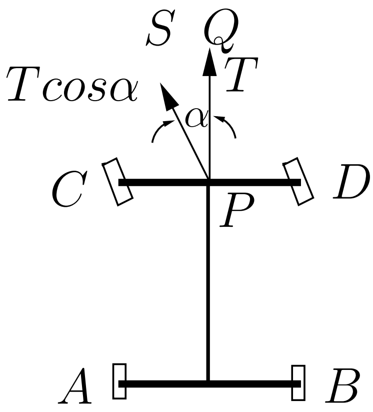


Figure 16:4: Curve resistance

16.4 Curve Resistance

When the vehicle negotiates a horizontal curve, the direction of rotation of the front and the r ear wheels are different. The front wheels are turned to move the vehicle along the curve, whereas the rear wheels seldom turn. This is illustrated in figure 16:4. The rear wheels exert a tractive force T in the PQ direction . The tractive force available on the front wheels is $T cos\alpha$ in the PS direction as shown in the figure 16:4. This is less than the actual tractive force, T applied. Hence, the loss of tractive force for a vehicle to negotiate a horizontal curve is:

$$CR = T - T\cos\alpha \tag{16.14}$$

16.5 Summary

Transition curves are introduced between straight road and circular curve. Setback distance controls alignment around obstacles at intersections and curves. Vehicles negotiating a curve are subjected to tractive resistances due to the curvature.

16.6 Problems

1. Calculate the length of transition curve and shift for V=65kmph, R=220m, rate of introduction of super elevation is 1 in 150, W+We=7.5 m. (Hint: c=0.57) [Ans: L_{s1}=47.1m, L_{s2}=39m (e=0.07, pav. rotated w.r.t centerline), L_{s3}=51.9m, s=0.51m, L_s=52m]

2. NH passing through rolling terrain of heavy rainfall area, R=500m. Design length of Transition curve. (Hint: Heavy rainfall. Pavement surface rotated w.r.t to inner edge. V=80kmph, W=7.0m, N=1 in 150) [Ans: c=0.52, L_{s1}=42.3, L_{s2}=63.7m (e=0.057, W+We=7.45), L_{s3}=34.6m, L_s=64m]

3. Horizontal curve of R=400m, L=200 m. Compute setback distance required to provide (a) SSD of 90m, (b) OSD of 300 m. Distance between center line of road and inner land (d) is 1.9m. [Ans: (a) $\alpha/2 \approx 6.5^o$, m=4.4 m (b)OSD>L, for multi lane, with d=1.9, m=26.8 m]

Chapter 17

Vertical alignment-I

17.1 Overview

The vertical alignment of a road consists of gradients(straight lines in a vertical plane) and vertical curves. The vertical alignment is usually drawn as a profile, which is a graph with elevation as vertical axis and the horizontal distance along the centre line of the road as the the horizontal axis. Just as a circular curve is used to connect horizontal straight stretches of road, vertical curves connect two gradients. When these two curves meet, they form either convex or concave. The former is called a summit curve, while the latter is called a valley curve. This section covers a discussion on gradient and summit curves.

17.2 Gradient

Gradient is the rate of rise or fall along the length of the road with respect to the horizontal. While aligning a highway, the gradient is decided for designing the vertical curve. Before finalizing the gradients, the construction cost, vehicular operation cost and the practical problems in the site also has to be considered. Usually steep gradients are avoided as far as possible because of the difficulty to climb and increase in the construction cost. More about gradients are discussed below.

17.2.1 Effect of gradient

The effect of long steep gradient on the vehicular speed is considerable. This is particularly important in roads where the proportion of heavy vehicles is significant. Due to restrictive sight distance at uphill gradients the speed of traffic is often controlled by these heavy vehicles. As a result, not only the operating costs of the vehicles are increased, but also capacity of the roads will have to be reduced. Further, due to high differential speed between heavy and light vehicles, and between uphill and downhill gradients, accidents abound in gradients.

17.2.2 Representation of gradient

The positive gradient or the ascending gradient is denoted as $+n$ and the negative gradient as $-n$. The deviation angle N is: when two grades meet, the angle which measures the change of direction and is given by the algebraic difference between the two grades $(n_1 - (-n_2)) \;=\; n_1 + n_2 \;=\; \alpha_1 + \alpha_2$. Example: 1 in 30 = 3.33% $\approx 2^o$ is a steep gradient, while 1 in 50 = 2% $\approx 1^o10^{'}$ is a flatter gradient. The gradient representation is illustrated in the figure 17:1.

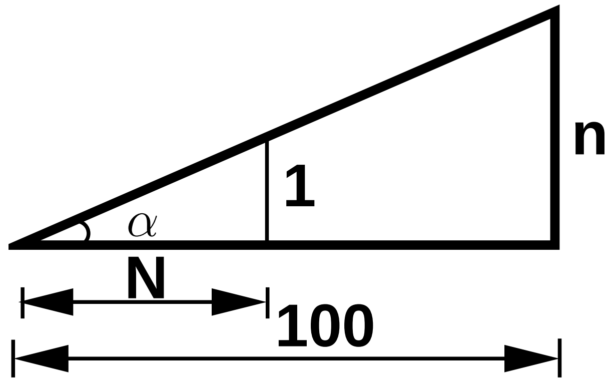


Figure 17:1: Representation of gradient

Table 17:1: IRC Specifications for gradients for different roads

Terrain	Ruling	Limitings	Exceptional
Plain/Rolling	3.3	5.0	6.7
Hilly	5.0	6.0	7.0
Steep	6.0	7.0	8.0

17.2.3 Types of gradient

Many studies have shown that gradient upto seven percent can have considerable effect on the speeds of the passenger cars. On the contrary, the speeds of the heavy vehicles are considerably reduced when long gradients as flat as two percent is adopted. Although, flatter gradients are desirable, it is evident that the cost of construction will also be very high. Therefore, IRC has specified the desirable gradients for each terrain. However, it may not be economically viable to adopt such gradients in certain locations, steeper gradients are permitted for short duration. Different types of grades are discussed below and the recommended type of gradients for each type of terrain and type of gradient is given in table 17:1.

Ruling gradient, limiting gradient, exceptional gradient and minimum gradient are some types of gradients which are discussed below.

Ruling gradient

The ruling gradient or the design gradient is the maximum gradient with which the designer attempts to design the vertical profile of the road. This depends on the terrain, length of the grade, speed, pulling power of the vehicle and the presence of the horizontal curve. In flatter terrain, it may be possible to provide flat gradients, but in hilly terrain it is not economical and sometimes not possible also. The ruling gradient is adopted by the designer by considering a particular speed as the design speed and for a design vehicle with standard dimensions. But our country has a heterogeneous traffic and hence it is not possible to lay down precise standards for the country as a whole. Hence IRC has recommended some values for ruling gradient for different types of terrain.

Limiting gradient

This gradient is adopted when the ruling gradient results in enormous increase in cost of construction. On rolling terrain and hilly terrain it may be frequently necessary to adopt limiting gradient. But the length of the limiting gradient stretches should be limited and must be sandwiched by either straight roads or easier grades.

Exceptional gradient

Exceptional gradient are very steeper gradients given at unavoidable situations. They should be limited for short stretches not exceeding about 100 metres at a stretch. In mountainous and steep terrain, successive exceptional gradients must be separated by a minimum 100 metre length gentler gradient. At hairpin bends, the gradient is restricted to 2.5%.

Critical length of the grade

The maximum length of the ascending gradient which a loaded truck can operate without undue reduction in speed is called critical length of the grade. A speed of 25 kmph is a reasonable value. This value depends on the size, power, load, grad-ability of the truck, initial speed, final desirable minimum speed etc.

Minimum gradient

This is important only at locations where surface drainage is important. Camber will take care of the lateral drainage. But the longitudinal drainage along the side drains require some slope for smooth flow of water. Therefore minimum gradient is provided for drainage purpose and it depends on the rain fall, type of soil and other site conditions. A minimum of 1 in 500 may be sufficient for concrete drain and 1 in 200 for open soil drains are found to give satisfactory performance..

17.2.4 Creeper lane

When the uphill climb is extremely long, it may be desirable to introduce an additional lane so as to allow slow ascending vehicles to be removed from the main stream so that the fast moving vehicles are not affected. Such a newly introduced lane is called *creeper lane*. There are no hard and fast rules as when to introduce a creeper lane. But generally, it can be said that it is desirable to provide a creeper lane when the speed of the vehicle gets reduced to half the design speed. When there is no restrictive sight distance to reduce the speed of the approaching vehicle, the additional lane may be initiated at some distance uphill from the beginning of the slope. But when the restrictions are responsible for the lowering of speeds, obviously the lane should be initiated at a point closer to the bottom of the hill. Also the creeper lane should end at a point well beyond the hill crest, so that the slow moving vehicles can return back to the normal lane without any danger. In addition, the creeper lane should not end suddenly, but only in a tapered manner for efficient as well as safer transition of vehicles to the normal lane.

17.2.5 Grade compensation

While a vehicle is negotiating a horizontal curve, if there is a gradient also, then there will be increased resistance to traction due to both curve and the gradient. In such cases, the total resistance should not exceed the resistance due to gradient specified. For the design, in some cases this maximum value is limited to the ruling gradient

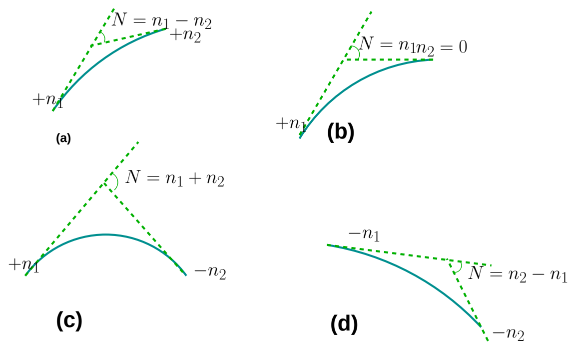


Figure 17:2: Types of summit curves

and in some cases as limiting gradient. So if a curve need to be introduced in a portion which has got the maximum permissible gradient, then some compensation should be provided so as to decrease the gradient for overcoming the tractive loss due to curve. Thus grade compensation can be defined as the reduction in gradient at the horizontal curve because of the additional tractive force required due to curve resistance $(T - Tcos\alpha)$, which is intended to offset the extra tractive force involved at the curve. IRC gave the following specification for the grade compensation.

1. Grade compensation is not required for grades flatter than 4% because the loss of tractive force is negligible.
2. Grade compensation is $\frac{30+R}{R}$%, where R is the radius of the horizontal curve in meters.
3. The maximum grade compensation is limited to $\frac{75}{R}$%.

17.3 Summit curve

Summit curves are vertical curves with gradient upwards. They are formed when two gradients meet as illustrated in figure 17:2 in any of the following four ways:

1. when a positive gradient meets another positive gradient [figure 17:2a].
2. when positive gradient meets a flat gradient [figure 17:2b]. .
3. when an ascending gradient meets a descending gradient [figure 17:2c]. .
4. when a descending gradient meets another descending gradient [figure 17:2d]. .

17.3.1 Type of Summit Curve

Many curve forms can be used with satisfactory results, the common practice has been to use parabolic curves in summit curves. This is primarily because of the ease with it can be laid out as well as allowing a comfortable transition from one gradient to another. Although a circular curve offers equal sight distance at every point on the curve, for very small deviation angles a circular curve and parabolic curves are almost congruent. Furthermore, the use of parabolic curves were found to give excellent riding comfort.

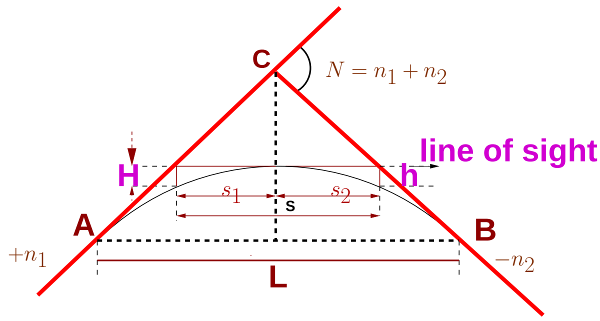


Figure 17:3: Length of summit curve ($L > S$)

17.3.2 Design Consideration

In determining the type and length of the vertical curve, the design considerations are comfort and security of the driver, and the appearance of the profile alignment. Among these, sight distance requirements for the safety is most important on summit curves. The stopping sight distance or absolute minimum sight distance should be provided on these curves and where overtaking is not prohibited, overtaking sight distance or intermediate sight distance should be provided as far as possible. When a fast moving vehicle travels along a summit curve, there is less discomfort to the passengers. This is because the centrifugal force will be acting upwards while the vehicle negotiates a summit curve which is against the gravity and hence a part of the tyre pressure is relieved. Also if the curve is provided with adequate sight distance, the length would be sufficient to ease the shock due to change in gradient. Circular summit curves are identical since the radius remains same throughout and hence the sight distance. From this point of view, transition curves are not desirable since it has varying radius and so the sight distance will also vary. The deviation angle provided on summit curves for highways are very large, and so the a simple parabola is almost congruent to a circular arc, between the same tangent points. Parabolic curves is easy for computation and also it had been found out that it provides good riding comfort to the drivers. It is also easy for field implementation. Due to all these reasons, a simple parabolic curve is preferred as summit curve.

17.3.3 Length of the summit curve

The important design aspect of the summit curve is the determination of the length of the curve which is parabolic. As noted earlier, the length of the curve is guided by the sight distance consideration. That is, a driver should be able to stop his vehicle safely if there is an obstruction on the other side of the road. Equation of the parabola is given by $y = ax^2$, where $a = \frac{N}{2L}$, where N is the deviation angle and L is the length of the In deriving the length of the curve, two situations can arise depending on the uphill and downhill gradients when the length of the curve is greater than the sight distance and the length of the curve is greater than the sight distance.

Let$L$ is the length of the summit curve, S is the SSD/ISD/OSD, N is the deviation angle, h_1 driver's eye height (1.2 m), and h_2 the height of the obstruction, then the length of the summit curve can be derived for the following two cases. The length of the summit curve can be derived from the simple geometry as shown below:

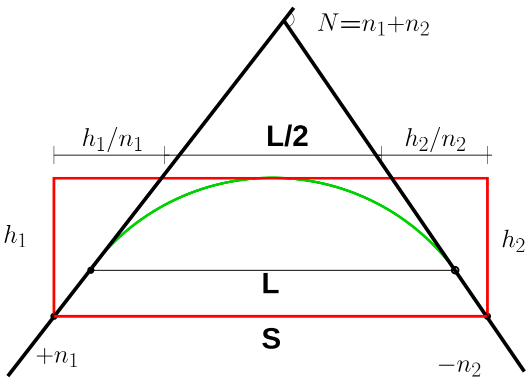


Figure 17:4: Length of summit curve ($L < S$)

Case a. Length of summit curve greater than sight distance($L > S$)

The situation when the sight distance is less than the length of the curve is shown in figure 17:3.

$$
\begin{aligned}
y &= ax^2 \\
a &= \frac{N}{2L} \\
h_1 &= aS_1^2 \\
h_2 &= aS_2^2 \\
S_1 &= \sqrt{\frac{h_1}{a}} \\
S_2 &= \sqrt{\frac{h_2}{a}} \\
S_1 + S_2 &= \sqrt{\frac{h_1}{a}} + \sqrt{\frac{h_2}{a}} \\
S^2 &= \left(\frac{1}{\sqrt{a}}\right)^2 \left(\sqrt{h_1} + \sqrt{h_2}\right)^2 \\
S^2 &= \frac{2L}{N} \left(\sqrt{h_1} + \sqrt{h_2}\right)^2 \\
L &= \frac{NS^2}{2\left(\sqrt{h_1} + \sqrt{h_2}\right)^2} \qquad (17.1)
\end{aligned}
$$

Case b. Length of summit curve less than sight distance

The second case is illustrated in figure 17:4

From the basic geometry, one can write

$$S = \frac{L}{2} + \frac{h_1}{n_1} + \frac{h_2}{n_2} = \frac{L}{2} + \frac{h_1}{n_1} + \frac{h_2}{N - n_2} \qquad (17.2)$$

Therefore for a given L, h_1 and h_2 to get minimum S, differentiate the above equation with respect to h_1 and equate it to zero. Therefore,

$$\frac{dS}{dh_1} = \frac{-h_1}{n_1^2} + \frac{h_2}{N - {n_1}^2} = 0 h_1 \left(N - n_1\right)^2 = \quad h_2 n_1^2$$

$$
\begin{aligned}
h_1\left(N^2+n_1^2-2Nn_1\right) &= h_2 n_1^2 \\
h_1N^2+h_1n_1^2-2Nn_1h_1 &= h_2n_1^2 \\
\left(h_2-h_1\right)n_1^2+2Nh_1n_1-h_1N^2 &= 0
\end{aligned} \tag{17.3}
$$

Solving the quadratic equation for n_1,

$$
\begin{aligned}
n_1 &= \frac{-2Nh_1 \pm \sqrt{\left(2Nh_1\right)^2-4\left(h_2-h_1\right)\left(-h_1N^2\right)}}{2\left(h_2-h_1\right)} \\
&= \frac{-2Nh_1+\sqrt{4N^2h_1^2+4h_1N^2h_2-4h_1^2N^2}}{2\left(h_2-h_1\right)} \\
&= \frac{-2Nh_1+2N\sqrt{h_1h_2}}{2\left(h_2-h_1\right)} \\
n_1 &= \frac{N\sqrt{h_1h_2}-h_1N}{h_2-h_1}
\end{aligned} \tag{17.4}
$$

Now we can substitute n back to get the value of minimum value of L for a given n_1, n_2, h_1 and h_2. Therefore,

$$
S = \frac{L}{2} + \frac{h_1}{\frac{N\sqrt{h_1h_2}-Nh_1}{h_2-h_1}} + \frac{h_2}{N-\frac{N\sqrt{h_1h_2}-Nh_1}{h_2-h_1}} \tag{17.5}
$$

Solving for L,

$$
\begin{aligned}
&= \frac{L}{2} + \frac{h_1\left(h_2-h_1\right)}{N\left(\sqrt{h_1h_2}-h_1\right)} + \frac{h_2\left(h_2-h_1\right)}{Nh_2-Nh_1-N\sqrt{h_1h_2}+Nh_1} \\
&= \frac{L}{2} + \frac{h_1\left(h_2-h_1\right)}{N\left(\sqrt{h_1h_2}-h_1\right)} + \frac{h_2\left(h_2-h_1\right)}{N\left(h_2-\sqrt{h_1h_2}\right)} \\
&= \frac{L}{2} + \frac{h_1\left(h_2-h_1\right)\left(h_2-\sqrt{h_1h_2}\right)+\left(h_2-h_1\right)h_2\left(\sqrt{h_1h_2}-h_1\right)}{N\left(\sqrt{h_1h_2}-h_1\right)\left(h_2-\sqrt{h_1h_2}\right)} \\
&= \frac{L}{2} + \frac{\left(h_2-h_1\right)\left(h_1h_2-h_1\sqrt{h_1h_2}+h_2\sqrt{h_1h_2}-h_1h_2\right)}{N\left(\sqrt{h_1h_2}-h_1\right)\left(h_2-\sqrt{h_1h_2}\right)} \\
&= \frac{L}{2} + \frac{\left(h_2-h_1\right)\left(\sqrt{h_1h_2}\left(h_2-h_1\right)\right)}{N\left(h_2\sqrt{h_1h_2}-h_1h_2+h_1\sqrt{h_1h_2}-h_1h_2\right)} \\
&= \frac{L}{2} + \frac{\left(h_2-h_1\right)\sqrt{h_1h_2}\left(\sqrt{h_2}+\sqrt{h_1}\right)\left(\sqrt{h_2}-\sqrt{h_1}\right)}{N\sqrt{h_1h_2}\left(h_2-2\sqrt{h_1h_2}+h_2\right)} \\
&= \frac{L}{2} + \frac{\left(h_2-h_1\right)\left(\sqrt{h_2}+\sqrt{h_1}\right)\left(\sqrt{h_2}-\sqrt{h_1}\right)}{N\left(\sqrt{h_2}-\sqrt{h_1}\right)^2} \\
&= \frac{L}{2} + \frac{\left(h_2-h_1\right)\left(\sqrt{h_2}+\sqrt{h_1}\right)}{N\left(\sqrt{h_2}-\sqrt{h_1}\right)} \\
&= \frac{L}{2} + \frac{\left(\sqrt{h_2}+\sqrt{h_1}\right)\left(\sqrt{h_2}-\sqrt{h_1}\right)\left(\sqrt{h_2}+\sqrt{h_1}\right)}{N\left(\sqrt{h_2}-\sqrt{h_1}\right)} \\
&= \frac{L}{2} + \frac{\left(\sqrt{h_2}+\sqrt{h_1}\right)^2}{N} \\
L &= 2S - \frac{2\left(\sqrt{h_2}+\sqrt{h_1}\right)^2}{N}
\end{aligned} \tag{17.6}
$$

$$L = 2S - \frac{\left(\sqrt{2h_1} + \sqrt{2h_2}\right)^2}{N} \tag{17.7}$$

When stopping sight distance is considered the height of driver's eye above the road surface (h_1) is taken as 1.2 metres, and height of object above the pavement surface (h_2) is taken as 0.15 metres. If overtaking sight distance is considered, then the value of driver's eye height (h_1) and the height of the obstruction (h_2) are taken equal as 1.2 metres.

17.4 Summary

Different types of gradients and IRC recommendations for their maximum and minimum limit were discussed. At points of combination of horizontal curve and gradient, grade compensation has to be provided. Due to changes in grade in the vertical alignment of the highway, vertical curves become essential. Summit curve, which is a type of vertical curve was discussed in detail in the chapter. One of the application of summit curves that can be seen usually in the urban areas are where fly-overs come.

17.5 Problems

1. A vertical summit curve is formed by $n_1 = +3.0\%$ and $n_2 = -5.0\%$. Design the length of the summit curve for V=80 kmph. (Hint: SSD=128m). [Ans: 298m]

2. n_1 ==1 in 1–, n_2 = 1 in 120. Design summit curve for V=80kmph, OSD=470m. [Ans: L=417m]

3. $n_1 = +1/50$ and $n_2 = -1/80$, SSD=180m, OSD=640m. Due to site constraints, L is limited to 500m. Calculate the length of summit curve to meet SSD, ISD and OSD. Discuss results. [Ans: L for SSD=240m, okay, L for OSD=1387m, > 500m not ok, L for ISD=439m ok.]

Chapter 18

Vertical alignment -II

18.1 Overview

As discussed earlier, changes in topography necessitate the introduction of vertical curves. The second curve of this type is the valley curve. This section deals with the types of valley curve and their geometrical design.

18.2 Valley curve

Valley curve or sag curves are vertical curves with convexity downwards. They are formed when two gradients meet as illustrated in figure 18:1 in any of the following four ways:

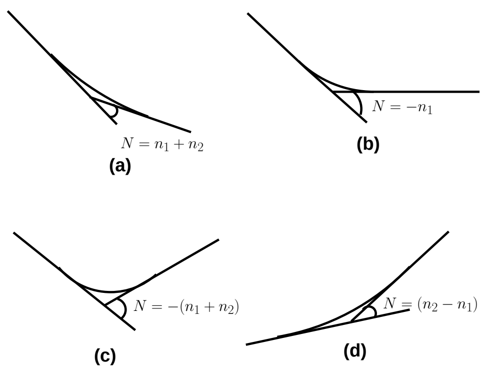


Figure 18:1: Types of valley curve

1. when a descending gradient meets another descending gradient [figure 18:1a].
2. when a descending gradient meets a flat gradient [figure 18:1b].
3. when a descending gradient meets an ascending gradient [figure 18:1c].
4. when an ascending gradient meets another ascending gradient [figure 18:1d].

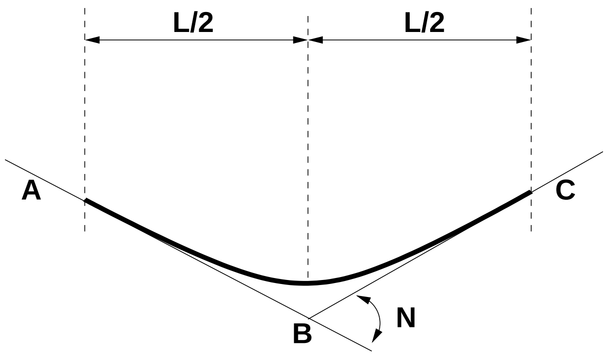


Figure 18:2: Valley curve details

18.2.1 Design considerations

There is no restriction to sight distance at valley curves during day time. But visibility is reduced during night. In the absence or inadequacy of street light, the only source for visibility is with the help of headlights. Hence valley curves are designed taking into account of headlight distance. In valley curves, the centrifugal force will be acting downwards along with the weight of the vehicle, and hence impact to the vehicle will be more. This will result in jerking of the vehicle and cause discomfort to the passengers. Thus the most important design factors considered in valley curves are: (1) impact-free movement of vehicles at design speed and (2) availability of stopping sight distance under headlight of vehicles for night driving.

For gradually introducing and increasing the centrifugal force acting downwards, the best shape that could be given for a valley curve is a transition curve. Cubic parabola is generally preferred in vertical valley curves. See figure 18:2.

During night, under headlight driving condition, sight distance reduces and availability of stopping sight distance under head light is very important. The head light sight distance should be at least equal to the stopping sight distance. There is no problem of overtaking sight distance at night since the other vehicles with headlights could be seen from a considerable distance.

18.2.2 Length of the valley curve

The valley curve is made fully transitional by providing two similar transition curves of equal length The transitional curve is set out by a cubic parabola $y = bx^3$ where $b = \frac{2N}{3L^2}$ The length of the valley transition curve is designed based on two criteria:

1. comfort criteria; that is allowable rate of change of centrifugal acceleration is limited to a comfortable level of about $0.6 m/sec^3$.

2. safety criteria; that is the driver should have adequate headlight sight distance at any part of the country.

Comfort criteria

The length of the valley curve based on the rate of change of centrifugal acceleration that will ensure comfort: Let c is the rate of change of acceleration, R the minimum radius of the curve, v is the design speed and t is the time, then c is given as:

$$\begin{aligned} c &= \frac{\frac{v^2}{R} - 0}{t} \\ &= \frac{\frac{v^2}{R} - 0}{\frac{L}{v}} \\ &= \frac{v^3}{LR} \\ L &= \frac{v^3}{cR} \end{aligned} \tag{18.1}$$

For a cubic parabola, the value of R for length L_s is given by:

$$R = \frac{L}{N} \tag{18.2}$$

Therefore,

$$\begin{aligned} L_s &= \frac{v^3}{\frac{cL_s}{N}} \\ L_s &= \sqrt[2]{\frac{Nv^3}{c}} \\ L &= 2\sqrt[2]{\frac{Nv^3}{c}} \end{aligned} \tag{18.3}$$

where L is the total length of valley curve, N is the deviation angle in radians or tangent of the deviation angle or the algebraic difference in grades, and c is the allowable rate of change of centrifugal acceleration which may be taken as $0.6 m/sec^3$.

Safety criteria

Length of the valley curve for headlight distance may be determined for two conditions: (1) length of the valley curve greater than stopping sight distance and (2) length of the valley curve less than the stopping sight distance.

Case 1 Length of valley curve greater than stopping sight distance ($L > S$)

The total length of valley curve L is greater than the stopping sight distance SSD. The sight distance available will be minimum when the vehicle is in the lowest point in the valley. This is because the beginning of the curve will have infinite radius and the bottom of the curve will have minimum radius which is a property of the transition curve. The case is shown in figure 18:3. From the geometry of the figure, we have:

$$\begin{aligned} h_1 + S\tan\alpha &= aS^2 \\ &= \frac{NS^2}{2L} \\ L &= \frac{NS^2}{2h_1 + 2S\tan\alpha} \end{aligned} \tag{18.4}$$

where N is the deviation angle in radians, h_1 is the height of headlight beam, α is the head beam inclination in degrees and S is the sight distance. The inclination α is ≈ 1 degree.

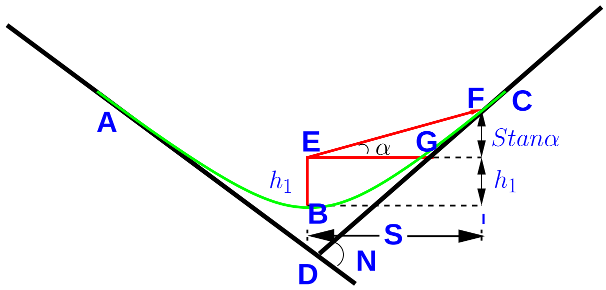


Figure 18:3: Valley curve, case 1, $L > S$

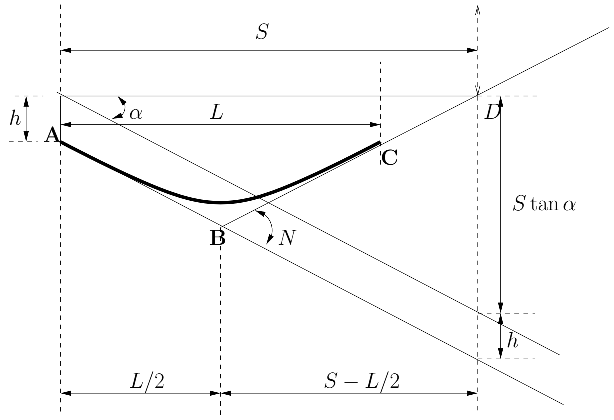


Figure 18:4: Valley curve, case 2, $S > L$

Case 2 Length of valley curve less than stopping sight distance ($L < S$)

The length of the curve L is less than SSD. In this case the minimum sight distance is from the beginning of the curve. The important points are the beginning of the curve and the bottom most part of the curve. If the vehicle is at the bottom of the curve, then its headlight beam will reach far beyond the endpoint of the curve whereas, if the vehicle is at the beginning of the curve, then the headlight beam will hit just outside the curve. Therefore, the length of the curve is derived by assuming the vehicle at the beginning of the curve . The case is shown in figure 18:4.

From the figure,

$$\begin{aligned} h_1 + s\tan\alpha &= \left(S - \frac{L}{2}\right) N \\ L &= 2S - \frac{2h_1 + 2S\tan\alpha}{N} \end{aligned} \tag{18.5}$$

Note that the above expression is approximate and is satisfactory because in practice, the gradients are very small and is acceptable for all practical purposes. We will not be able to know prior to which case to be adopted. Therefore both has to be calculated and the one which satisfies the condition is adopted.

18.3 Summary

The valley curve should be designed such that there is enough headlight sight distance. Improperly designed valley curves results in extreme riding discomfort as well as accident risks especially at nights. The length of valley curve for various cases were also explained in the section. The concept of valley curve is used in underpasses.

18.4 Problems

1. A valley curve is formed by descending gradient n_1= 1 in 25 and ascending gradient n_2= 1 in 30. Design the length of the valley curve for V=80kmph. (Hint: c=0.6 m/cm^3, SSD=127.3m) [Ans: L=max(73.1,199.5)]

Chapter 19

Introduction to pavement design

19.1 Overview

A highway pavement is a structure consisting of superimposed layers of processed materials above the natural soil sub-grade, whose primary function is to distribute the applied vehicle loads to the sub-grade. The pavement structure should be able to provide a surface of acceptable riding quality, adequate skid resistance, favorable light reflecting characteristics, and low noise pollution. The ultimate aim is to ensure that the transmitted stresses due to wheel load are sufficiently reduced, so that they will not exceed bearing capacity of the sub-grade. Two types of pavements are generally recognized as serving this purpose, namely flexible pavements and rigid pavements. This chapter gives an overview of pavement types, layers, and their functions, and pavement failures. Improper design of pavements leads to early failure of pavements affecting the riding quality.

19.2 Requirements of a pavement

An ideal pavement should meet the following requirements:

- Sufficient thickness to distribute the wheel load stresses to a safe value on the sub-grade soil,
- Structurally strong to withstand all types of stresses imposed upon it,
- Adequate coefficient of friction to prevent skidding of vehicles,
- Smooth surface to provide comfort to road users even at high speed,
- Produce least noise from moving vehicles,
- Dust proof surface so that traffic safety is not impaired by reducing visibility,
- Impervious surface, so that sub-grade soil is well protected, and
- Long design life with low maintenance cost.

19.3 Types of pavements

The pavements can be classified based on the structural performance into two, flexible pavements and rigid pavements. In flexible pavements, wheel loads are transferred by grain-to-grain contact of the aggregate through the granular structure. The flexible pavement, having less flexural strength, acts like a flexible sheet (e.g.

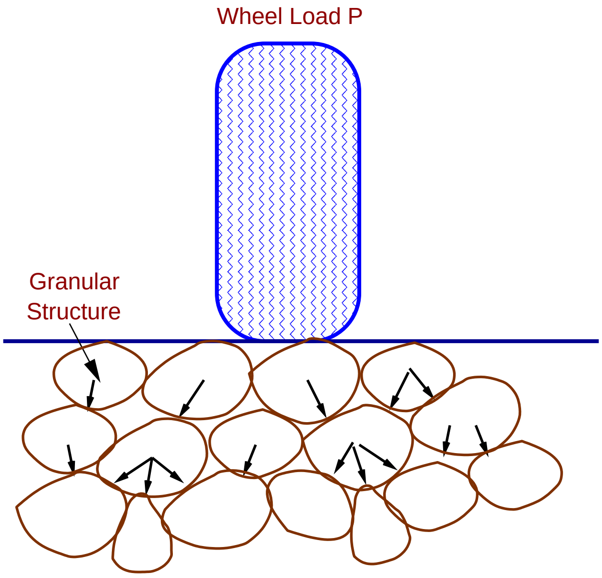


Figure 19:1: Load transfer in granular structure

bituminous road). On the contrary, in rigid pavements, wheel loads are transferred to sub-grade soil by flexural strength of the pavement and the pavement acts like a rigid plate (e.g. cement concrete roads). In addition to these, composite pavements are also available. A thin layer of flexible pavement over rigid pavement is an ideal pavement with most desirable characteristics. However, such pavements are rarely used in new construction because of high cost and complex analysis required.

19.4 Flexible pavements

Flexible pavements will transmit wheel load stresses to the lower layers by grain-to-grain transfer through the points of contact in the granular structure (see Figure 19:1). The wheel load acting on the pavement will be distributed to a wider area, and the stress decreases with the depth. Taking advantage of this stress distribution characteristic, flexible pavements normally has many layers. Hence, the design of flexible pavement uses the concept of layered system. Based on this, flexible pavement may be constructed in a number of layers and the top layer has to be of best quality to sustain maximum compressive stress, in addition to wear and tear. The lower layers will experience lesser magnitude of stress and low quality material can be used. Flexible pavements are constructed using bituminous materials. These can be either in the form of surface treatments (such as bituminous surface treatments generally found on low volume roads) or, asphalt concrete surface courses (generally used on high volume roads such as national highways). Flexible pavement layers reflect the deformation of the lower layers on to the surface layer (e.g., if there is any undulation in sub-grade then it will be transferred to the surface layer). In the case of flexible pavement, the design is based on overall performance of flexible pavement, and the stresses produced should be kept well below the allowable stresses of each pavement layer.

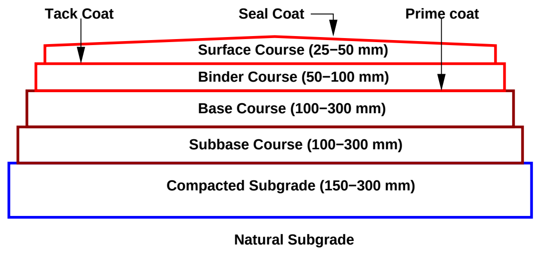


Figure 19:2: Typical cross section of a flexible pavement

19.4.1 Types of Flexible Pavements

The following types of construction have been used in flexible pavement:

- Conventional layered flexible pavement,
- Full - depth asphalt pavement, and
- Contained rock asphalt mat (CRAM).

Conventional flexible pavements are layered systems with high quality expensive materials are placed in the top where stresses are high, and low quality cheap materials are placed in lower layers.

Full - depth asphalt pavements are constructed by placing bituminous layers directly on the soil sub-grade. This is more suitable when there is high traffic and local materials are not available.

Contained rock asphalt mats are constructed by placing dense/open graded aggregate layers in between two asphalt layers. Modified dense graded asphalt concrete is placed above the sub-grade will significantly reduce the vertical compressive strain on soil sub-grade and protect from surface water.

19.4.2 Typical layers of a flexible pavement

Typical layers of a conventional flexible pavement includes seal coat, surface course, tack coat, binder course, prime coat, base course, sub-base course, compacted sub-grade, and natural sub-grade (Figure 19:2).

Seal Coat: Seal coat is a thin surface treatment used to water-proof the surface and to provide skid resistance.

Tack Coat: Tack coat is a very light application of asphalt, usually asphalt emulsion diluted with water. It provides proper bonding between two layer of binder course and must be thin, uniformly cover the entire surface, and set very fast.

Prime Coat: Prime coat is an application of low viscous cutback bitumen to an absorbent surface like granular bases on which binder layer is placed. It provides bonding between two layers. Unlike tack coat, prime coat penetrates into the layer below, plugs the voids, and forms a water tight surface.

Surface course

Surface course is the layer directly in contact with traffic loads and generally contains superior quality materials. They are usually constructed with dense graded asphalt concrete(AC). The functions and requirements of this layer are:

- It provides characteristics such as friction, smoothness, drainage, etc. Also it will prevent the entrance of excessive quantities of surface water into the underlying base, sub-base and sub-grade,
- It must be tough to resist the distortion under traffic and provide a smooth and skid- resistant riding surface,
- It must be water proof to protect the entire base and sub-grade from the weakening effect of water.

Binder course

This layer provides the bulk of the asphalt concrete structure. It's chief purpose is to distribute load to the base course The binder course generally consists of aggregates having less asphalt and doesn't require quality as high as the surface course, so replacing a part of the surface course by the binder course results in more economical design.

Base course

The base course is the layer of material immediately beneath the surface of binder course and it provides additional load distribution and contributes to the sub-surface drainage It may be composed of crushed stone, crushed slag, and other untreated or stabilized materials.

Sub-Base course

The sub-base course is the layer of material beneath the base course and the primary functions are to provide structural support, improve drainage, and reduce the intrusion of fines from the sub-grade in the pavement structure If the base course is open graded, then the sub-base course with more fines can serve as a filler between sub-grade and the base course A sub-base course is not always needed or used. For example, a pavement constructed over a high quality, stiff sub-grade may not need the additional features offered by a sub-base course. In such situations, sub-base course may not be provided.

Sub-grade

The top soil or sub-grade is a layer of natural soil prepared to receive the stresses from the layers above. It is essential that at no time soil sub-grade is overstressed. It should be compacted to the desirable density, near the optimum moisture content.

19.4.3 Failure of flexible pavements

The major flexible pavement failures are fatigue cracking, rutting, and thermal cracking. The fatigue cracking of flexible pavement is due to horizontal tensile strain at the bottom of the asphaltic concrete. The failure criterion relates allowable number of load repetitions to tensile strain and this relation can be determined in the laboratory *fatigue test* on asphaltic concrete specimens. Rutting occurs only on flexible pavements as indicated by permanent deformation or rut depth along wheel load path. Two design methods have been used to control rutting: one to limit the vertical compressive strain on the top of subgrade and other to limit rutting to a tolerable amount (12 mm normally). Thermal cracking includes both low-temperature cracking and thermal fatigue cracking.

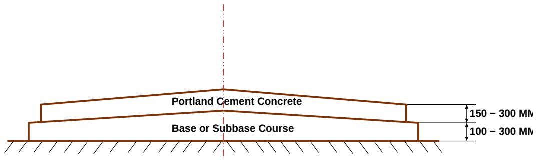


Figure 19:3: Typical Cross section of Rigid pavement

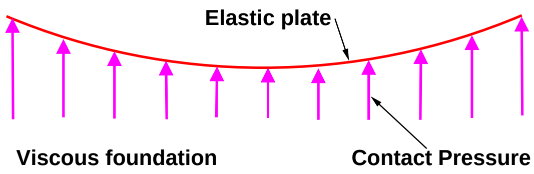


Figure 19:4: Elastic plate resting on Viscous foundation

19.5 Rigid pavements

Rigid pavements have sufficient flexural strength to transmit the wheel load stresses to a wider area below. A typical cross section of the rigid pavement is shown in Figure 19:3. Compared to flexible pavement, rigid pavements are placed either directly on the prepared sub-grade or on a single layer of granular or stabilized material. Since there is only one layer of material between the concrete and the sub-grade, this layer can be called as base or sub-base course.

In rigid pavement, load is distributed by the slab action, and the pavement behaves like an elastic plate resting on a viscous medium (Figure 19:4). Rigid pavements are constructed by Portland cement concrete (PCC) and should be analyzed by plate theory instead of layer theory, assuming an elastic plate resting on viscous foundation. Plate theory is a simplified version of layer theory that assumes the concrete slab as a medium thick plate which is plane before loading and to remain plane after loading. Bending of the slab due to wheel load and temperature variation and the resulting tensile and flexural stress.

19.5.1 Types of Rigid Pavements

Rigid pavements can be classified into four types:

- Jointed plain concrete pavement (JPCP),
- Jointed reinforced concrete pavement (JRCP),
- Continuous reinforced concrete pavement (CRCP), and
- Pre-stressed concrete pavement (PCP).

Jointed Plain Concrete Pavement: are plain cement concrete pavements constructed with closely spaced contraction joints. Dowel bars or aggregate interlocks are normally used for load transfer across joints. They normally has a joint spacing of 5 to 10m.

Jointed Reinforced Concrete Pavement: Although reinforcements do not improve the structural capacity significantly, they can drastically increase the joint spacing to 10 to 30m. Dowel bars are required for load transfer. Reinforcements help to keep the slab together even after cracks.

Continuous Reinforced Concrete Pavement: Complete elimination of joints are achieved by reinforcement.

19.5.2 Failure criteria of rigid pavements

Traditionally fatigue cracking has been considered as the major, or only criterion for rigid pavement design. The allowable number of load repetitions to cause fatigue cracking depends on the stress ratio between flexural tensile stress and concrete modulus of rupture. Of late, pumping is identified as an important failure criterion. Pumping is the ejection of soil slurry through the joints and cracks of cement concrete pavement, caused during the downward movement of slab under the heavy wheel loads. Other major types of distress in rigid pavements include faulting, spalling, and deterioration.

19.6 Summary

Pavements form the basic supporting structure in highway transportation. Each layer of pavement has a multitude of functions to perform which has to be duly considered during the design process. Different types of pavements can be adopted depending upon the traffic requirements. Improper design of pavements leads to early failure of pavements affecting the riding quality also.

19.7 Problems

1. The thin layer of bitumen coating between an existing bituminous layer and a new bituminous layer is:
   (a) Seal coat
   (b) Intermediate coat
   (c) Tack coat
   (d) Prime coat
2. Rigid pavements are designed by
   (a) Rigid plate theory
   (b) Elastic plate theory
   (c) Infinite layer theory
   (d) Interlocking of aggregates

19.8 Solutions

1. The thin layer of bitumen coating between an existing bituminous layer and a new bituminous layer is:
   (a) Seal coat
   (b) Intermediate coat
   (c) Tack coat$\surd$
   (d) Prime coat

2. Rigid pavements are designed by

   (a) Rigid plate theory

   (b) Elastic plate theory$\surd$

   (c) Infinite layer theory

   (d) Interlocking of aggregates

Chapter 20

Factors affecting pavement design

20.1 Overview

In the previous chapter we had discussed about the types of pavements and their failure criteria. There are many factors that affect pavement design which can be classified into four categories as traffic and loading, structural models, material characterization, environment. They will be discussed in detail in this chapter.

20.2 Traffic and loading

Traffic is the most important factor in the pavement design. The key factors include contact pressure, wheel load, axle configuration, moving loads, load, and load repetitions.

Contact pressure: The tyre pressure is an important factor, as it determine the contact area and the contact pressure between the wheel and the pavement surface. Even though the shape of the contact area is elliptical, for sake of simplicity in analysis, a circular area is often considered.

Wheel load: The next important factor is the wheel load which determines the depth of the pavement required to ensure that the subgrade soil is not failed. Wheel configuration affect the stress distribution and deflection within a pavemnet. Many commercial vehicles have dual rear wheels which ensure that the contact pressure is within the limits. The normal practice is to convert dual wheel into an equivalent single wheel load so that the analysis is made simpler.

Axle configuration: The load carrying capacity of the commercial vehicle is further enhanced by the introduction of multiple axles.

Moving loads: The damage to the pavement is much higher if the vehicle is moving at creep speed. Many studies show that when the speed is increased from 2 km/hr to 24 km/hr, the stresses and deflection reduced by 40 per cent.

Repetition of Loads: The influence of traffic on pavement not only depend on the magnitude of the wheel load, but also on the frequency of the load applications. Each load application causes some deformation and the total deformation is the summation of all these. Although the pavement deformation due to single axle load is very small, the cumulative effect of number of load repetition is significant. Therefore, modern design is based on total number of standard axle load (usually 80 kN single axle).

20.3 Structural models

The structural models are various analysis approaches to determine the pavement responses (stresses, strains, and deflections) at various locations in a pavement due to the application of wheel load. The most common structural models are layered elastic model and visco-elastic models.

Layered elastic model: A layered elastic model can compute stresses, strains, and deflections at any point in a pavement structure resulting from the application of a surface load. Layered elastic models assume that each pavement structural layer is homogeneous, isotropic, and linearly elastic. In other words, the material properties are same at every point in a given layer and the layer will rebound to its original form once the load is removed. The layered elastic approach works with relatively simple mathematical models that relates stress, strain, and deformation with wheel loading and material properties like modulus of elasticity and poissons ratio.

20.4 Material characterization

The following material properties are important for both flexible and rigid pavements.

- When pavements are considered as linear elastic, the elastic moduli and poisson ratio of subgrade and each component layer must be specified.
- If the elastic modulus of a material varies with the time of loading, then the resilient modulus, which is elastic modulus under repeated loads, must be selected in accordance with a load duration corresponding to the vehicle speed.
- When a material is considered non-linear elastic, the constitutive equation relating the resilient modulus to the state of the stress must be provided.

However, many of these material properties are used in visco-elastic models which are very complex and in the development stage. This book covers the layered elastic model which require the modulus of elasticity and poisson ratio only.

20.5 Environmental factors

Environmental factors affect the performance of the pavement materials and cause various damages. Environmental factors that affect pavement are of two types, temperature and precipitation and they are discussed below:

20.5.1 Temperature

The effect of temperature on asphalt pavements is different from that of concrete pavements. Temperature affects the resilient modulus of asphalt layers, while it induces curling of concrete slab. In rigid pavements, due to difference in temperatures of top and bottom of slab, temperature stresses or frictional stresses are developed. While in flexible pavement, dynamic modulus of asphaltic concrete varies with temperature. Frost heave causes differential settlements and pavement roughness. Most detrimental effect of frost penetration occurs during the spring break up period when the ice melts and subgrade is a saturated condition.

20.5.2 Precipitation

The precipitation from rain and snow affects the quantity of surface water infiltrating into the subgrade and the depth of ground water table. Poor drainage may bring lack of shear strength, pumping, loss of support, etc.

20.6 Summary

Several factors affecting pavement design were discussed, the most important being wheel load. Since pavements are designed to take moving loads, slow moving loads and static loads can be detrimental to the pavement. Temperature also influences pavement design especially the frost action which is very important in cold countries.

20.7 Problems

1. Factor that least affect the pavement is
   (a) Speed of vehicles
   (b) Wheel load
   (c) Axle configuration
   (d) Load repetition

2. Standard axle load is
   (a) 40kN
   (b) 60kN
   (c) 80kN
   (d) 10kN

20.8 Solutions

1. Factor that least affect the pavement is
   (a) Speed of vehicles$\surd$
   (b) Wheel load
   (c) Axle configuration
   (d) Load repetition

2. Standard axle load is
   (a) 40kN
   (b) 60kN
   (c) 80kN$\surd$
   (d) 10kN

Chapter 21

Pavement materials: Soil

21.1 Overview

Pavements are a conglomeration of materials. These materials, their associated properties, and their interactions determine the properties of the resultant pavement. Thus, a good understanding of these materials, how they are characterized, and how they perform is fundamental to understanding pavement. The materials which are used in the construction of highway are of intense interest to the highway engineer. This requires not only a thorough understanding of the soil and aggregate properties which affect pavement stability and durability, but also the binding materials which may be added to improve these pavement features.

21.2 Sub grade soil

Soil is an accumulation or deposit of earth material, derived naturally from the disintegration of rocks or decay of vegetation, that can be excavated readily with power equipment in the field or disintegrated by gentle mechanical means in the laboratory. The supporting soil beneath pavement and its special under courses is called sub grade. Undisturbed soil beneath the pavement is called natural sub grade. Compacted sub grade is the soil compacted by controlled movement of heavy compactors.

21.2.1 Desirable properties

The desirable properties of sub grade soil as a highway material are

- Stability
- Incompressibility
- Permanency of strength
- Minimum changes in volume and stability under adverse conditions of weather and ground water
- Good drainage, and
- Ease of compaction

21.2.2 Soil Types

The wide range of soil types available as highway construction materials have made it obligatory on the part of the highway engineer to identify and classify different soils. A survey of locally available materials and soil types conducted in India revealed wide variety of soil types, gravel, moorum and naturally occurring soft aggregates, which can be used in road construction. Broadly, the soil types can be categorized as Laterite soil, Moorum / red soil, Desert sands, Alluvial soil, Clay including Black cotton soil.

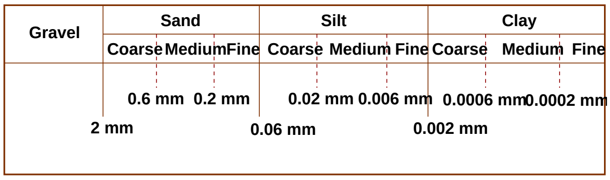


Figure 21:1: Indian standard grain size soil classification system

- **Gravel:** These are coarse materials with particle size under 2.36 mm with little or no fines contributing to cohesion of materials.
- **Moorum:** These are products of decomposition and weathering of the pavement rock. Visually these are similar to gravel except presence of higher content of fines.
- **Silts:** These are finer than sand, brighter in color as compared to clay, and exhibit little cohesion. When a lump of silty soil mixed with water, alternately squeezed and tapped a shiny surface makes its appearance, thus dilatancy is a specific property of such soil.
- **Clays:** These are finer than silts. Clayey soils exhibit stickiness, high strength when dry, and show no dilatancy. Black cotton soil and other expansive clays exhibit swelling and shrinkage properties. Paste of clay with water when rubbed in between fingers leaves stain, which is not observed for silts.

21.2.3 Tests on soil

Sub grade soil is an integral part of the road pavement structure as it provides the support to the pavement from beneath. The sub grade soil and its properties are important in the design of pavement structure. The main function of the sub grade is to give adequate support to the pavement and for this the sub grade should possess sufficient stability under adverse climatic and loading conditions. Therefore, it is very essential to evaluate the sub grade by conducting tests.

The tests used to evaluate the strength properties of soils may be broadly divided into three groups:

- Shear tests
- Bearing tests
- Penetration tests

Shear tests are usually carried out on relatively small soil samples in the laboratory. In order to find out the strength properties of soil, a number of representative samples from different locations are tested. Some of the commonly known shear tests are direct shear test, triaxial compression test, and unconfined compression test.

Bearing tests are loading tests carried out on sub grade soils in-situ with a load bearing area. The results of the bearing tests are influenced by variations in the soil properties within the stressed soil mass underneath and hence the overall stability of the part of the soil mass stressed could be studied.

Penetration tests may be considered as small scale bearing tests in which the size of the loaded area is relatively much smaller and ratio of the penetration to the size of the loaded area is much greater than the ratios in bearing tests. The penetration tests are carried out in the field or in the laboratory.

21.2.4 California Bearing Ratio Test

California Bearing Ratio (CBR) test was developed by the California Division of Highway as a method of classifying and evaluating soil-sub grade and base course materials for flexible pavements. CBR test, an empirical test, has been used to determine the material properties for pavement design. Empirical tests measure the strength of the material and are not a true representation of the resilient modulus. It is a penetration test wherein a standard piston, having an area of 3 in^2 (or 50 mm diameter), is used to penetrate the soil at a standard rate of 1.25 mm/minute. The pressure up to a penetration of 12.5 mm and it's ratio to the bearing value of a standard crushed rock is termed as the CBR.

In most cases, CBR decreases as the penetration increases. The ratio at 2.5 mm penetration is used as the CBR. In some case, the ratio at 5 mm may be greater than that at 2.5 mm. If this occurs, the ratio at 5 mm should be used. The CBR is a measure of resistance of a material to penetration of standard plunger under controlled density and moisture conditions. The test procedure should be strictly adhered if high degree of reproducibility is desired. The CBR test may be conducted in re-moulded or undisturbed specimen in the laboratory. The test is simple and has been extensively investigated for field correlations of flexible pavement thickness requirement.

Test Procedure

- The laboratory CBR apparatus consists of a mould 150 mm diameter with a base plate and a collar, a loading frame and dial gauges for measuring the penetration values and the expansion on soaking.
- The specimen in the mould is soaked in water for four days and the swelling and water absorption values are noted. The surcharge weight is placed on the top of the specimen in the mould and the assembly is placed under the plunger of the loading frame.
- Load is applied on the sample by a standard plunger with dia of 50 mm at the rate of 1.25 mm/min. A load penetration curve is drawn. The load values on standard crushed stones are 1370 kg and 2055 kg at 2.5 mm and 5.0 mm penetrations respectively.
- CBR value is expressed as a percentage of the actual load causing the penetrations of 2.5 mm or 5.0 mm to the standard loads mentioned above. Therefore,

$$CBR = \frac{\text{load carries by specimen}}{\text{load carries by standard specimen}} \times 100$$

- Two values of CBR will be obtained. If the value of 2.5 mm is greater than that of 5.0 mm penetration, the former is adopted. If the CBR value obtained from test at 5.0 mm penetration is higher than that at 2.5 mm, then the test is to be repeated for checking. If the check test again gives similar results, then higher value obtained at 5.0 mm penetration is reported as the CBR value. The average CBR value of three test specimens is reported as the CBR value of the sample.

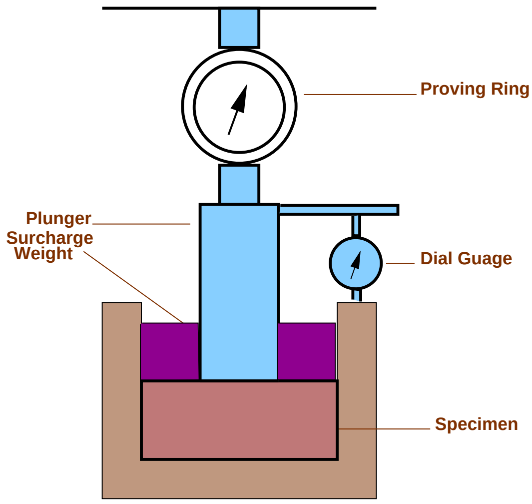


Figure 21:2: CBR Test

21.2.5 Plate Bearing Test

Plate bearing test is used to evaluate the support capability of sub-grades, bases and in some cases, complete pavement. Data from the tests are applicable for the design of both flexible and rigid pavements. In plate bearing test, a compressive stress is applied to the soil or pavement layer through rigid plates relatively large size and the deflections are measured for various stress values. The deflection level is generally limited to a low value, in the order of 1.25 to 5 mm and so the deformation caused may be partly elastic and partly plastic due to compaction of the stressed mass with negligible plastic deformation. The plate-bearing test has been devised to evaluate the supporting power of sub grades or any other pavement layer by using plates of larger diameter. The plate-bearing test was originally meant to find the modulus of sub grade reaction in the Westergaard's analysis for wheel load stresses in cement concrete pavements.

Test Procedure

- The test site is prepared and loose material is removed so that the 75 cm diameter plate rests horizontally in full contact with the soil sub-grade. The plate is seated accurately and then a seating load equivalent to a pressure of 0.07 kg/cm2 (320 kg for 75 cm diameter plate) is applied and released after a few seconds. The settlement dial gauge is now set corresponding to zero load.
- A load is applied by means of jack, sufficient to cause an average settlement of about 0.25 cm. When there is no perceptible increase in settlement or when the rate of settlement is less than 0.025 mm per minute (in the case of soils with high moisture content or in clayey soils) the load dial reading and the settlement dial readings are noted.
- Deflection of the plate is measured by means of deflection dials; placed usually at one-third points of the plate near it's outer edge.
- To minimize bending, a series of stacked plates should be used.
- Average of three or four settlement dial readings is taken as the settlement of the plate corresponding to the applied load. Load is then increased till the average settlement increase to a further amount of about

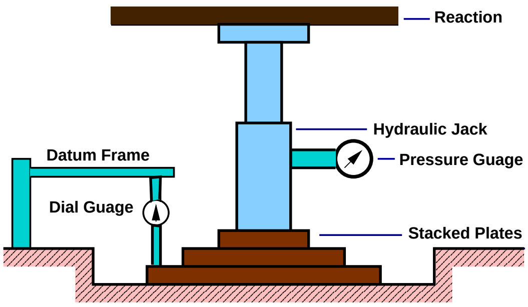


Figure 21:3: Plate load test

0.25 mm, and the load and average settlement readings are noted as before. The procedure is repeated till the settlement is about 1.75 mm or more.

- Allowance for worst subgrade moisture and correction for small plate size should be dealt properly.
- **Calculation** A graph is plotted with the mean settlement versus bearing pressure (load per unit area) as shown in Figure 21:3. The pressure corresponding to a settlement is obtained from this graph. The modulus of subgrade reaction is calculated from the relation.

$$K = \frac{P}{0.125} kg/cm^2/cm. \tag{21.1}$$

21.3 Summary

The quality of any pavement is affected by the materials used for construction. Coming to the subgrade, soil is the most important material. Here we have seen various tests used for finding the strength of soil, the prominent ones being CBR and plate load test. CBR test assesses the strength of soil, whereas plate load test is used to evaluate its support capability.

21.4 Problems

1. The load value on standard crushed stone for 2.5mm penetration in CBR test is
   (a) 1370kg
   (b) 1730 kg
   (c) 2055kg
   (d) 1500kg
2. Modulus of subgrade reaction is
   (a) $K = \frac{P}{2.15} kg/cm^2/cm$
   (b) $K = \frac{P}{0.125} kg/cm^2/cm$
   (c) $K = \frac{P^2}{2.15} kg/cm^2/cm$
   (d) $K = \frac{P}{2.5} kg/cm^2/cm$

21.5 Solutions

1. The load value on standard crushed stone for 2.5mm penetration in CBR test is
   (a) 1370kg$\surd$
   (b) 1730 kg
   (c) 2055kg
   (d) 1500kg
2. Modulus of subgrade reaction is
   (a) $K = \frac{P}{2.15} kg/cm^2/cm$
   (b) $K = \frac{P}{0.125} kg/cm^2/cm\surd$
   (c) $K = \frac{P^2}{2.15} kg/cm^2/cm$
   (d) $K = \frac{P}{2.5} kg/cm^2/cm$

Chapter 22

Pavement materials: Aggregates

22.1 Overview

Aggregate is a collective term for the mineral materials such as sand, gravel, and crushed stone that are used with a binding medium (such as water, bitumen, Portland cement, lime, etc.) to form compound materials (such as bituminous concrete and Portland cement concrete). By volume, aggregate generally accounts for 92 to 96 percent of Bituminous concrete and about 70 to 80 percent of Portland cement concrete. Aggregate is also used for base and sub-base courses for both flexible and rigid pavements. Aggregates can either be natural or manufactured. Natural aggregates are generally extracted from larger rock formations through an open excavation (quarry). Extracted rock is typically reduced to usable sizes by mechanical crushing. Manufactured aggregate is often a bye product of other manufacturing industries. The requirements of the aggregates in pavement are also discussed in this chapter.

22.2 Desirable properties

22.2.1 Strength

The aggregates used in top layers are subjected to (i) Stress action due to traffic wheel load, (ii) Wear and tear, (iii) crushing. For a high quality pavement, the aggregates should posses high resistance to crushing, and to withstand the stresses due to traffic wheel load.

22.2.2 Hardness

The aggregates used in the surface course are subjected to constant rubbing or abrasion due to moving traffic. The aggregates should be hard enough to resist the abrasive action caused by the movements of traffic. The abrasive action is severe when steel tyred vehicles moves over the aggregates exposed at the top surface.

22.2.3 Toughness

Resistance of the aggregates to impact is termed as toughness. Aggregates used in the pavement should be able to resist the effect caused by the jumping of the steel tyred wheels from one particle to another at different levels causes severe impact on the aggregates.

22.2.4 Shape of aggregates

Aggregates which happen to fall in a particular size range may have rounded, cubical, angular, flaky or elongated particles. It is evident that the flaky and elongated particles will have less strength and durability when compared with cubical, angular or rounded particles of the same aggregate. Hence too flaky and too much elongated aggregates should be avoided as far as possible.

22.2.5 Adhesion with bitumen

The aggregates used in bituminous pavements should have less affinity with water when compared with bituminous materials, otherwise the bituminous coating on the aggregate will be stripped off in presence of water.

22.2.6 Durability

The property of aggregates to withstand adverse action of weather is called soundness. The aggregates are subjected to the physical and chemical action of rain and bottom water, impurities there-in and that of atmosphere, hence it is desirable that the road aggregates used in the construction should be sound enough to withstand the weathering action

22.2.7 Freedom from deleterious particles

Specifications for aggregates used in bituminous mixes usually require the aggregates to be clean, tough and durable in nature and free from excess amount of flat or elongated pieces, dust, clay balls and other objectionable material. Similarly aggregates used in Portland cement concrete mixes must be clean and free from deleterious substances such as clay lumps, chert, silt and other organic impurities.

22.3 Aggregate tests

In order to decide the suitability of the aggregate for use in pavement construction, following tests are carried out:

- Crushing test
- Abrasion test
- Impact test
- Soundness test
- Shape test
- Specific gravity and water absorption test
- Bitumen adhesion test

22.3.1 Crushing test

One of the model in which pavement material can fail is by crushing under compressive stress. A test is standardized by IS:2386 part-IV and used to determine the crushing strength of aggregates. The aggregate crushing value provides a relative measure of resistance to crushing under gradually applied crushing load. The test consists of subjecting the specimen of aggregate in standard mould to a compression test under standard load conditions (Figure 22:1). Dry aggregates passing through 12.5 mm sieves and retained 10 mm sieves are filled in a cylindrical measure of 11.5 mm diameter and 18 cm height in three layers. Each layer is tampered 25 times with at standard tamping rod. The test sample is weighed and placed in the test cylinder in three layers each layer being tampered again. The specimen is subjected to a compressive load of 40 tonnes gradually applied at the rate of 4 tonnes per minute. Then crushed aggregates are then sieved through 2.36 mm sieve and weight of passing material (W_2) is expressed as percentage of the weight of the total sample (W_1) which is the aggregate crushing value.

$$\text{Aggregate crushing value} = \frac{W_1}{W_2} \times 100$$

A value less than 10 signifies an exceptionally strong aggregate while above 35 would normally be regarded as weak aggregates.

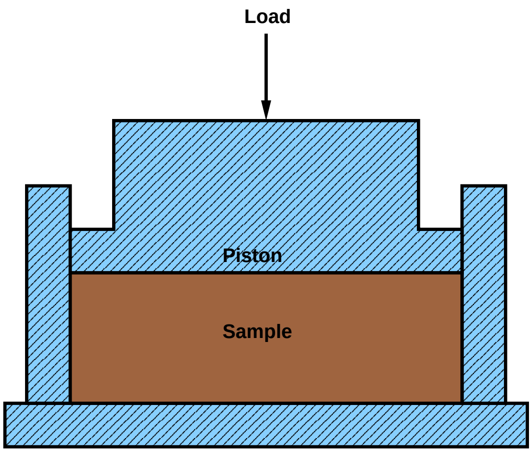


Figure 22:1: Crushing test setup

22.3.2 Abrasion test

Abrasion test is carried out to test the hardness property of aggregates and to decide whether they are suitable for different pavement construction works. Los Angeles abrasion test is a preferred one for carrying out the hardness property and has been standardized in India (IS:2386 part-IV). The principle of Los Angeles abrasion test is to find the percentage wear due to relative rubbing action between the aggregate and steel balls used as abrasive charge.

Los Angeles machine consists of circular drum of internal diameter 700 mm and length 520 mm mounted on horizontal axis enabling it to be rotated (see Figure 22:2). An abrasive charge consisting of cast iron spherical balls of 48 mm diameters and weight 340-445 g is placed in the cylinder along with the aggregates. The number of the abrasive spheres varies according to the grading of the sample. The quantity of aggregates to be used depends upon the gradation and usually ranges from 5-10 kg. The cylinder is then locked and rotated at the speed of 30-33 rpm for a total of 500 -1000 revolutions depending upon the gradation of aggregates.

After specified revolutions, the material is sieved through 1.7 mm sieve and passed fraction is expressed as percentage total weight of the sample. This value is called Los Angeles abrasion value.

A maximum value of 40 percent is allowed for WBM base course in Indian conditions. For bituminous concrete, a maximum value of 35 is specified.

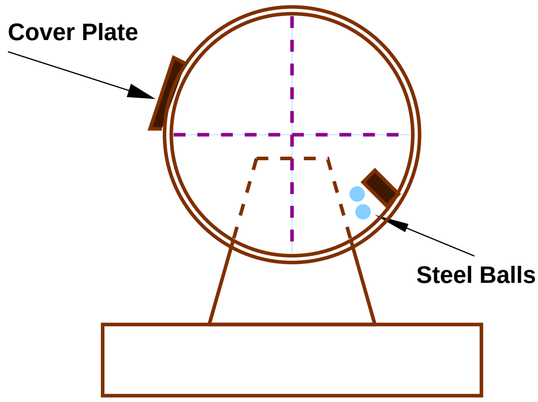


Figure 22:2: Los Angeles abrasion test setup

22.3.3 Impact test

The aggregate impact test is carried out to evaluate the resistance to impact of aggregates. Aggregates passing 12.5 mm sieve and retained on 10 mm sieve is filled in a cylindrical steel cup of internal dia 10.2 mm and depth 5 cm which is attached to a metal base of impact testing machine. The material is filled in 3 layers where each layer is tamped for 25 number of blows. Metal hammer of weight 13.5 to 14 Kg is arranged to drop with a free fall of 38.0 cm by vertical guides and the test specimen is subjected to 15 number of blows. The crushed aggregate is allowed to pass through 2.36 mm IS sieve. And the impact value is measured as percentage of aggregates passing sieve (W_2) to the total weight of the sample (W_1).

$$\text{Aggregate impact value} = \frac{W_1}{W_2} \times 100$$

Aggregates to be used for wearing course, the impact value shouldn't exceed 30 percent. For bituminous macadam the maximum permissible value is 35 percent. For Water bound macadam base courses the maximum permissible value defined by IRC is 40 percent

22.3.4 Soundness test

Soundness test is intended to study the resistance of aggregates to weathering action, by conducting accelerated weathering test cycles. The Porous aggregates subjected to freezing and thawing are likely to disintegrate prematurely. To ascertain the durability of such aggregates, they are subjected to an accelerated soundness test as specified in IS:2386 part-V. Aggregates of specified size are subjected to cycles of alternate wetting in a saturated solution of either sodium sulphate or magnesium sulphate for 16 - 18 hours and then dried in oven at $105 - 110^oC$ to a constant weight. After five cycles, the loss in weight of aggregates is determined by sieving

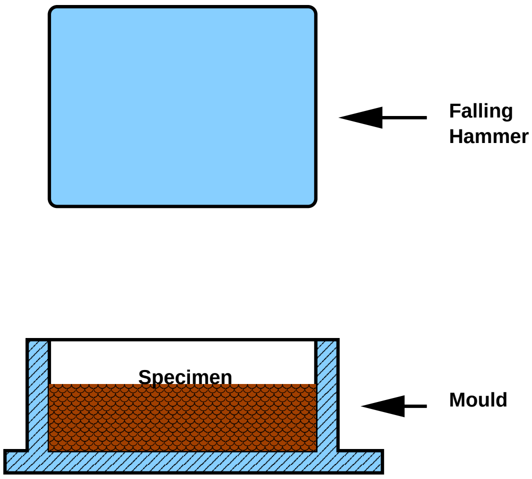


Figure 22:3: Impact test setup

out all undersized particles and weighing. And the loss in weight should not exceed 12 percent when tested with sodium sulphate and 18 percent with magnesium sulphate solution.

22.3.5 Shape tests

The particle shape of the aggregate mass is determined by the percentage of flaky and elongated particles in it. Aggregates which are flaky or elongated are detrimental to higher workability and stability of mixes.

The flakiness index is defined as the percentage by weight of aggregate particles whose least dimension is less than 0.6 times their mean size. Test procedure had been standardized in India (IS:2386 part-I)

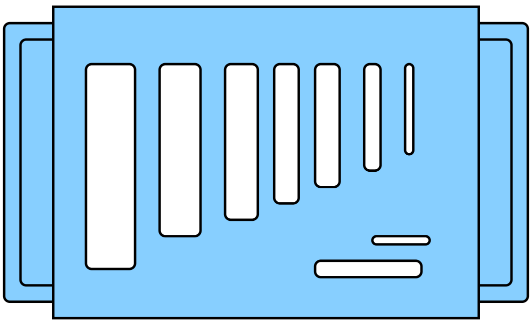


Figure 22:4: Flakiness gauge

The elongation index of an aggregate is defined as the percentage by weight of particles whose greatest dimension (length) is 1.8 times their mean dimension. This test is applicable to aggregates larger than 6.3 mm. This test is also specified in (IS:2386 Part-I). However there are no recognized limits for the elongation index.

22.3.6 Specific Gravity and water absorption

The specific gravity and water absorption of aggregates are important properties that are required for the design of concrete and bituminous mixes. The specific gravity of a solid is the ratio of its mass to that of an equal

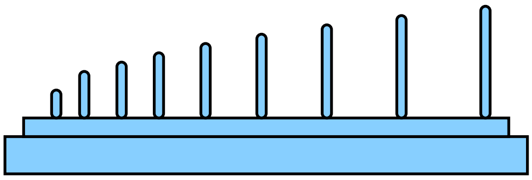


Figure 22:5: Elongation gauge

volume of distilled water at a specified temperature. Because the aggregates may contain water-permeable voids, so two measures of specific gravity of aggregates are used: *apparent* specific gravity and *bulk* specific gravity.

- Apparent Specific Gravity, G_{app}, is computed on the basis of the net volume of aggregates i.e the volume excluding water-permeable voids. Thus

$$G_{app} = \frac{M_D/V_N}{W} \tag{22.1}$$

  where, M_D is the dry mass of the aggregate, V_N is the net volume of the aggregates excluding the volume of the absorbed matter, W is the density of water.

- Bulk Specific Gravity, G_{bulk}, is computed on the basis of the total volume of aggregates including water permeable voids.Thus

$$G_{bulk} = \frac{M_D/V_B}{W} \tag{22.2}$$

  where, V_B is the total volume of the aggregates including the volume of absorbed water.

- Water absorption, The difference between the apparent and bulk specific gravities is nothing but the water-permeable voids of the aggregates. We can measure the volume of such voids by weighing the aggregates dry and in a saturated, surface dry condition, with all permeable voids filled with water. The difference of the above two is M_W. M_W is the weight of dry aggregates minus weight of aggregates saturated surface dry condition. Thus

$$\text{water absorption} = \frac{\mathrm{M_W}}{\mathrm{M_D}} \times 100 \tag{22.3}$$

  The specific gravity of aggregates normally used in road construction ranges from about 2.5 to 2.9. Water absorption values ranges from 0.1 to about 2.0 percent for aggregates normally used in road surfacing.

22.3.7 Bitumen adhesion test

Bitumen adheres well to all normal types of road aggregates provided they are dry and free from dust. In the absence of water there is practically no adhesion problem of bituminous construction. Adhesion problem occurs when the aggregate is wet and cold. This problem can be dealt with by removing moisture from the aggregate by drying and increasing the mixing temperature. Further, the presence of water causes stripping of binder from the coated aggregates. This problems occur when bitumen mixture is permeable to water. Several laboratory tests are conducted to arbitrarily determine the adhesion of bitumen binder to an aggregate in the presence of water. Static immersion test is one specified by IRC and is quite simple. The principle of the test is by immersing aggregate fully coated with binder in water maintained at 40^0C temperature for 24 hours. IRC has specified maximum stripping value of aggregates should not exceed 5%.

Property of aggregate	**Type of Test**	**Test Method**
Crushing strength	Crushing test	IS : 2386 (part 4) -1963
Hardness	Los Angeles abrasion test	IS : 2386 (Part 5)-1963
Toughness	Aggregate impact test	IS : 2386 (Part 4)-1963
Durability	Soundness test- accelerated durability test	IS : 2386 (Part 5)-1963
Shape factors	Shape test	IS : 2386 (Part 1)-1963
Specific gravity and porosity	Specific gravity test and water absorption test	IS : 2386 (Part 3)-1963
Adhesion to bitumen	Stripping value of aggregate	IS : 6241-1971

Table 22:1: Tests for Aggregates with IS codes

22.4 Summary

Aggregates influence, to a great extent, the load transfer capability of pavements. Hence it is essential that they should be thoroughly tested before using for construction. Not only that aggregates should be strong and durable, they should also possess proper shape and size to make the pavement act monolithically. Aggregates are tested for strength, toughness, hardness, shape, and water absorption.

22.5 Problems

1. IRC has specified the maximum value of stripping value of bitumen not to exceed
   (a) 2%
   (b) 3%
   (c) 4%
   (d) 5%
2. Which property of aggregate is tested by conducting aggregate impact test?
   (a) Durability
   (b) Hardness
   (c) Toughness
   (d) Porosity

22.6 Solutions

1. IRC has specified the maximum value of stripping value of bitumen not to exceed
   (a) 2%$\surd$
   (b) 3%
   (c) 4%
   (d) 5%

2. Which property of aggregate is tested by conducting aggregate impact test?
   (a) Durability
   (b) Hardness
   (c) Toughness$\surd$
   (d) Porosity

Chapter 23

Pavement materials: Bitumen

23.1 Overview

Bituminous materials or asphalts are extensively used for roadway construction, primarily because of their excellent binding characteristics and water proofing properties and relatively low cost. Bituminous materials consists of bitumen which is a black or dark coloured solid or viscous cementitious substances consists chiefly high molecular weight hydrocarbons derived from distillation of petroleum or natural asphalt, has adhesive properties, and is soluble in carbon disulphide. Tars are residues from the destructive distillation of organic substances such as coal, wood, or petroleum and are temperature sensitive than bitumen. Bitumen will be dissolved in petroleum oils where unlike tar.

23.1.1 Production of Bitumen

bitumen is the residue or by-product when the crude petrolium is refined. A wide variety of refinery processes, such as the straight distillation process, solvent extraction process etc. may be used to produce bitumen of different consistency and other desirable properties. Depending on the sources and characteristics of the crude oils and on the properties of bitumen required, more than one processing method may be employed.

23.1.2 Vacuum steam distillation of petroleum oils

In the vacuum-steam distillation process the crude oil is heated and is introduced into a large cylindrical still. Steam is introduced into the still to aid in the vapourisation of the more volatile constituents of the petroleum and to minimise decomposition of the distillates and residues. The volatile constituents are collected, condensed, and the various fractions stored for further refining, if needed. The residues from this distillation are then fed into a vacuum distillation unit, where residue pressure and steam will further separate out heavier gas oils. The bottom fraction from this unit is the vacuum-steam-refined asphalt cement. The consistency of asphalt cement from this process can be controlled by the amount of heavy gas oil removed. Normally, asphalt produced by this process is softer. As the asphalt cools down to room temperature, it becomes a semi solid viscous material.

23.2 Different forms of bitumen

23.2.1 Cutback bitumen

Normal practice is to heat bitumen to reduce its viscosity. In some situations preference is given to use liquid binders such as cutback bitumen. In cutback bitumen suitable solvent is used to lower the viscosity of the bitumen. From the environmental point of view also cutback bitumen is preferred. The solvent from the bituminous material will evaporate and the bitumen will bind the aggregate. Cutback bitumen is used for cold weather bituminous road construction and maintenance. The distillates used for preparation of cutback bitumen are naphtha, kerosene, diesel oil, and furnace oil. There are different types of cutback bitumen like rapid curing (RC), medium curing (MC), and slow curing (SC). RC is recommended for surface dressing and patchwork. MC is recommended for premix with less quantity of fine aggregates. SC is used for premix with appreciable quantity of fine aggregates.

23.2.2 Bitumen Emulsion

Bitumen emulsion is a liquid product in which bitumen is suspended in a finely divided condition in an aqueous medium and stabilised by suitable material. Normally cationic type emulsions are used in India. The bitumen content in the emulsion is around 60% and the remaining is water. When the emulsion is applied on the road it breaks down resulting in release of water and the mix starts to set. The time of setting depends upon the grade of bitumen. The viscosity of bituminous emulsions can be measured as per IS: 8887-1995. Three types of bituminous emulsions are available, which are Rapid setting (RS), Medium setting (MS), and Slow setting (SC). Bitumen emulsions are ideal binders for hill road construction. Where heating of bitumen or aggregates are difficult. Rapid setting emulsions are used for surface dressing work. Medium setting emulsions are preferred for premix jobs and patch repairs work. Slow setting emulsions are preferred in rainy season.

23.2.3 Bituminous primers

In bituminous primer the distillate is absorbed by the road surface on which it is spread. The absorption therefore depends on the porosity of the surface. Bitumen primers are useful on the stabilised surfaces and water bound macadam base courses. Bituminous primers are generally prepared on road sites by mixing penetration bitumen with petroleum distillate.

23.2.4 Modified Bitumen

Certain additives or blend of additives called as bitumen modifiers can improve properties of Bitumen and bituminous mixes. Bitumen treated with these modifiers is known as modified bitumen. Polymer modified bitumen (PMB)/ crumb rubber modified bitumen (CRMB) should be used only in wearing course depending upon the requirements of extreme climatic variations. The detailed specifications for modified bitumen have been issued by IRC: SP: 53-1999. It must be noted that the performance of PMB and CRMB is dependent on strict control on temperature during construction. The advantages of using modified bitumen are as follows

- Lower susceptibility to daily and seasonal temperature variations
- Higher resistance to deformation at high pavement temperature
- Better age resistance properties
- Higher fatigue life for mixes
- Better adhesion between aggregates and binder
- Prevention of cracking and reflective cracking

23.3 Requirements of Bitumen

The desirable properties of bitumen depend on the mix type and construction. In general, Bitumen should posses following desirable properties.

- The bitumen should not be highly temperature susceptible: during the hottest weather the mix should not become too soft or unstable, and during cold weather the mix should not become too brittle causing cracks.
- The viscosity of the bitumen at the time of mixing and compaction should be adequate. This can be achieved by use of cutbacks or emulsions of suitable grades or by heating the bitumen and aggregates prior to mixing.
- There should be adequate affinity and adhesion between the bitumen and aggregates used in the mix.

23.4 Tests on bitumen

There are a number of tests to assess the properties of bituminous materials. The following tests are usually conducted to evaluate different properties of bituminous materials.

1. Penetration test
2. Ductility test
3. Softening point test
4. Specific gravity test
5. Viscosity test
6. Flash and Fire point test
7. Float test
8. Water content test
9. Loss on heating test

23.4.1 Penetration test

It measures the hardness or softness of bitumen by measuring the depth in tenths of a millimeter to which a standard loaded needle will penetrate vertically in 5 seconds. BIS had standardised the equipment and test procedure. The penetrometer consists of a needle assembly with a total weight of 100g and a device for releasing and locking in any position. The bitumen is softened to a pouring consistency, stirred thoroughly and poured into containers at a depth at least 15 mm in excess of the expected penetration. The test should be conducted at a specified temperature of 25^o C. It may be noted that penetration value is largely influenced by any inaccuracy with regards to pouring temperature, size of the needle, weight placed on the needle and the test temperature. A grade of 40/50 bitumen means the penetration value is in the range 40 to 50 at standard test conditions. In hot climates, a lower penetration grade is preferred. The Figure 23.4.1 shows a schematic Penetration Test setup.

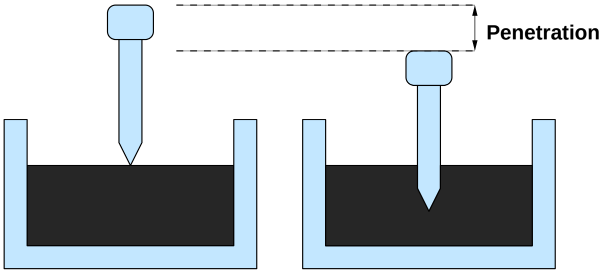


Figure 23:1: Penetration Test Setup

23.4.2 Ductility test

Ductility is the property of bitumen that permits it to undergo great deformation or elongation. Ductility is defined as the distance in cm, to which a standard sample or briquette of the material will be elongated without breaking. Dimension of the briquette thus formed is exactly 1 cm square. The bitumen sample is heated and poured in the mould assembly placed on a plate. These samples with moulds are cooled in the air and then in water bath at 27^o C temperature. The excess bitumen is cut and the surface is leveled using a hot knife. Then the mould with assembly containing sample is kept in water bath of the ductility machine for about 90 minutes. The sides of the moulds are removed, the clips are hooked on the machine and the machine is operated. The distance up to the point of breaking of thread is the ductility value which is reported in cm. The ductility value gets affected by factors such as pouring temperature, test temperature, rate of pulling etc. A minimum ductility value of 75 cm has been specified by the BIS. Figure 23.4.2 shows ductility moulds to be filled with bitumen.

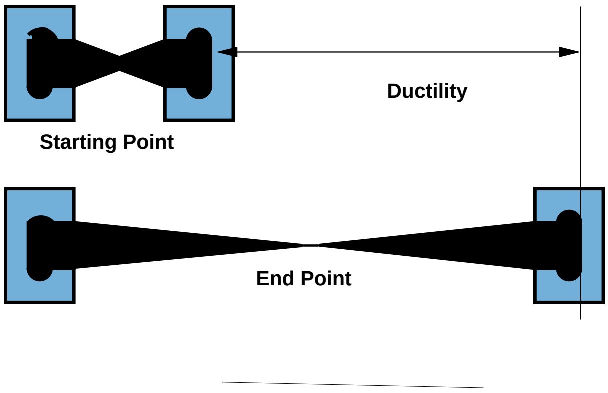


Figure 23:2: Ductility Test

23.4.3 Softening point test

Softening point denotes the temperature at which the bitumen attains a particular degree of softening under the specifications of test. The test is conducted by using Ring and Ball apparatus. A brass ring containing test sample of bitumen is suspended in liquid like water or glycerin at a given temperature. A steel ball is placed upon the bitumen sample and the liquid medium is heated at a rate of 5^o C per minute. Temperature is noted when the softened bitumen touches the metal plate which is at a specified distance below. Generally, higher softening point indicates lower temperature susceptibility and is preferred in hot climates. Figure 23.4.3 shows Softening Point test setup.

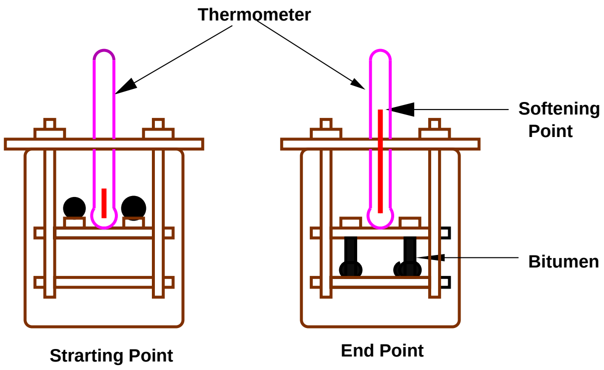


Figure 23:3: Softening Point Test Setup

23.4.4 Specific gravity test

In paving jobs, to classify a binder, density property is of great use. In most cases bitumen is weighed, but when used with aggregates, the bitumen is converted to volume using density values. The density of bitumen is greatly influenced by its chemical composition. Increase in aromatic type mineral impurities cause an increase in specific gravity.

The specific gravity of bitumen is defined as the ratio of mass of given volume of bitumen of known content to the mass of equal volume of water at 27^o C. The specific gravity can be measured using either pycnometer or preparing a cube specimen of bitumen in semi solid or solid state. The specific gravity of bitumen varies from 0.97 to 1.02.

23.4.5 Viscosity test

Viscosity denotes the fluid property of bituminous material and it is a measure of resistance to flow. At the application temperature, this characteristic greatly influences the strength of resulting paving mixes. Low or high viscosity during compaction or mixing has been observed to result in lower stability values. At high viscosity, it resist the compactive effort and thereby resulting mix is heterogeneous, hence low stability values. And at low viscosity instead of providing a uniform film over aggregates, it will lubricate the aggregate particles. Orifice type viscometers are used to indirectly find the viscosity of liquid binders like cutbacks and emulsions. The viscosity expressed in seconds is the time taken by the 50 ml bitumen material to pass through the orifice of a cup, under standard test conditions and specified temperature. Viscosity of a cutback can be measured with either 4.0 mm orifice at 25^o C or 10 mm orifice at 25 or 40^o C.

23.4.6 Flash and fire point test

At high temperatures depending upon the grades of bitumen materials leave out volatiles. And these volatiles catches fire which is very hazardous and therefore it is essential to qualify this temperature for each bitumen grade. BIS defined the flash point as the temperature at which the vapour of bitumen momentarily catches fire

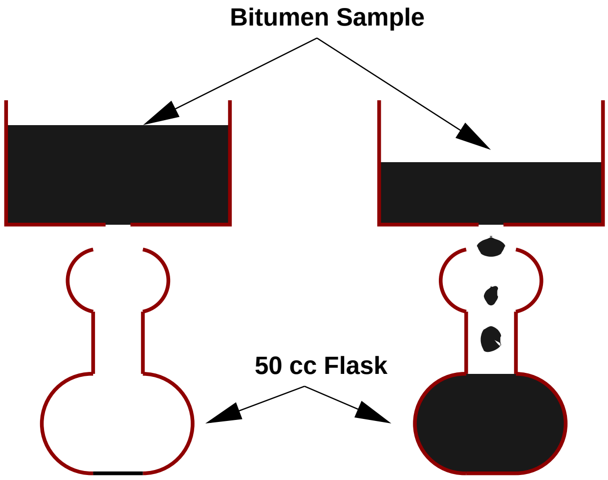


Figure 23:4: Viscosity Test

in the form of flash under specified test conditions. The fire point is defined as the lowest temperature under specified test conditions at which the bituminous material gets ignited and burns.

23.4.7 Float test

Normally the consistency of bituminous material can be measured either by penetration test or viscosity test. But for certain range of consistencies, these tests are not applicable and Float test is used. The apparatus consists of an aluminum float and a brass collar filled with bitumen to be tested. The specimen in the mould is cooled to a temperature of 5^{o}C and screwed in to float. The total test assembly is floated in the water bath at 50^{o}C and the time required for water to pass its way through the specimen plug is noted in seconds and is expressed as the float value.

23.4.8 Water content test

It is desirable that the bitumen contains minimum water content to prevent foaming of the bitumen when it is heated above the boiling point of water. The water in a bitumen is determined by mixing known weight of specimen in a pure petroleum distillate free from water, heating and distilling of the water. The weight of the water condensed and collected is expressed as percentage by weight of the original sample. The allowable maximum water content should not be more than 0.2% by weight.

23.4.9 Loss on heating test

When the bitumen is heated it loses the volatility and gets hardened. About 50gm of the sample is weighed and heated to a temperature of 163^{0}C for 5hours in a specified oven designed for this test. The sample specimen is weighed again after the heating period and loss in weight is expressed as percentage by weight of the original sample. Bitumen used in pavement mixes should not indicate more than 1% loss in weight, but for bitumen having penetration values 150-200 up to 2% loss in weight is allowed.

Table 23:1: Tests for Bitumen with IS codes

Type of test	**Test Method**
Penetration Test	IS: 1203-1978
Ductility test	IS: 1208-1978
Softening Point test	IS: 1205-1978
Specific gravity test	IS: 1202-1978
Viscosity test	IS: 1206-1978
Flash and Fire Point test	IS: 1209-1978
Float Test	IS: 1210-1978
Determination of water content	IS: 1211-1978
Determination of Loss on heating	IS:1212-1978

23.5 Summary

Requirements of bitumen as a binding material and its differrent forms were discussed. Various tests are conducted on bitumen to assess its consistency, gradation, viscosity, temperature susceptibility, and safety. Standard test procedures on bitumen were also covered in this chapter.

23.6 Problems

The minimum ductility value specified by BIS for bitumen is

1. 50 cm
2. 25 cm
3. 75 cm
4. 100 cm

The allowable maximum water content in bitumen should not be more than

1. 2%by weight
2. 0.2% by weight
3. 2.5% by weight
4. 5% by weight

23.7 Solutions

The minimum ductility value specified by BIS for bitumen is

1. 50 cm
2. 25 cm
3. 75 cm $\surd$

4. 100 cm

The allowable maximum water content in bitumen should not be more than

1. 2% by weight
2. 0.2% by weight $\surd$
3. 2.5% by weight
4. 5% by weight

Chapter 24

Bituminous mix design

24.1 Overview

The bituminous mix design aims to determine the proportion of bitumen, filler, fine aggregates, and coarse aggregates to produce a mix which is workable, strong, durable and economical. The requirements of the mix design and the two major stages of the mix design, i.e dry mix design and wet mix design will be discussed.

24.2 Evolution of road surface

- Unsurfaced earthen roads, or cart-track
- Unsurfaced earthen roads upgrades with natural soil from borrow pits and attention to drainage, and compaction is by traffic
- Dry aggregate and sand-clays mix, in which the the former act as wear resistant and the latter as natural binder
- Water-bound macadam, the above constituents, mixed together (pre-mix or in-situ) with water and compacted to improve the strength
- Oiled roads, introduced to reduce dust by bitumen stabilized soils
- Seal coat: the base course is protected from traffic and moisture by sealing the surface with a thin film of bitumen aggregate mix, which is structurally strong surface for pneumatic-tyred traffic. This is provided on firm and smooth base course after a tack coat using cutback bitumen or bitumen emulsions with a penetration of 5 mm.
- Asphaltic concrete: Traffic and the axle configuration are increasing very much which raises demand for the new type of pavement which can meet the above requirements. The asphaltic concrete is one which is the high dense graded premix and it is termed as the highest quality pavement surface course.
- Bitumen mix or asphalt mix overlays of minimum 20 - 40 mm to as high as 300 - 500 mm or even more.

24.2.1 Objectives of mix design

The objective of the mix design is to produce a bituminous mix by proportionating various components so as to have:

1. sufficient bitumen to ensure a durable pavement,
2. sufficient strength to resist shear deformation under traffic at higher temperature,
3. sufficient air voids in the compacted bitumen to allow for additional compaction by traffic,
4. sufficient workability to permit easy placement without segregation,
5. sufficient flexibility to avoid premature cracking due to repeated bending by traffic,and
6. sufficient flexibility at low temperature to prevent shrinkage cracks.

24.2.2 Constituents of a mix

- **Coarse aggregates:** Offer compressive and shear strength and shows good interlocking properties. E.g. *Granite*
- **Fine aggregates:** Fills the voids in the coarse aggregate and stiffens the binder. E.g. *Sand, Rock dust*
- **Filler:** Fills the voids, stiffens the binder and offers permeability. E.g. *Rock dust, cement, lime*
- **Binder:** Fills the voids, cause particle adhesion and gluing and offers impermeability. E.g. *Bitumen, Asphalt, Tar*

24.2.3 Types of mix

- **Well-graded mix:-** *Dense mix, bituminous concrete* has good proportion of all constituents and are called dense bituminous macadam, offers good compressive strength and some tensile strength
- **Gap-graded mix:-** Some large coarse aggregates are missing and has good fatigue and tensile strength.
- **Open-graded mix:-** Fine aggregate and filler are missing, it is porous and offers good friction, low strength and for high speed.
- **Unbounded:-** Binder is absent and behaves under loads as if its components were not linked together, though good interlocking exists. Very low tensile strength and needs kerb protection.

24.2.4 Different layers in a pavement

- **Bituminous base course** Consist of mineral aggregate such as stone, gravel, or sand bonded together by a bituminous material and used as a foundation upon which to place a binder or surface course.
- **Bituminous binder course** A bituminous-aggregate mixture used as an intermediate coarse between the base and surface courses or as the first bituminous layer in a two-layer bituminous resurfacing. It is sometimes called a leveling course.
- **Asphaltic/Bituminous concrete** Bituminous concrete consists of a mixture of aggregates continuously graded from maximum size , typically less than 25 mm, through fine filler that is smaller than 0.075 mm. Sufficient bitumen is added to the mix so that the compacted mix is *effectively impervious* and will have acceptable dissipative and elastic properties.

Table 24:1: Aggregate gradation for BC

Sieve size		Passing (%)
26.5	mm	- 100
19.	mm	90 - 100
9.5	mm	56 - 80
4.75	mm	35 - 65
2.36	mm	23 - 49
300	micron	5 - 19
75	micron	2 - 8

24.3 Requirements of Bituminous mixes

24.3.1 Stability

Stability is defined as the resistance of the paving mix to deformation under traffic load. Two examples of failure are (i) *shoving* - a transverse rigid deformation which occurs at areas subject to severe acceleration and (ii) *grooving* - longitudinal ridging due to channelization of traffic. Stability depend on the inter-particle friction, primarily of the aggregates and the cohesion offered by the bitumen. Sufficient binder must be available to coat all the particles at the same time should offer enough liquid friction. However, the stability decreases when the binder content is high and when the particles are kept apart.

24.3.2 Durability

Durability is defined as the resistance of the mix against weathering and abrasive actions. Weathering causes hardening due to loss of volatiles in the bitumen. Abrasion is due to wheel loads which causes tensile strains. Typical examples of failure are (i) *pot-holes,* - deterioration of pavements locally and (ii) *stripping,* lost of binder from the aggregates and aggregates are exposed. Disintegration is minimized by high binder content since they cause the mix to be air and waterproof and the bitumen film is more resistant to hardening.

24.3.3 Flexibility

Flexibility is a measure of the level of bending strength needed to counteract traffic load and prevent cracking of surface. Fracture is the cracks formed on the surface (hairline-cracks, alligator cracks), main reasons are shrinkage and brittleness of the binder. Shrinkage cracks are due to volume change in the binder due to aging. Brittleness is due to repeated bending of the surface due to traffic loads. Higher bitumen content will give better flexibility and less fracture.

24.3.4 Skid resistance

It is the resistance of the finished pavement against skidding which depends on the surface texture and bitumen content. It is an important factor in high speed traffic. Normally, an open graded coarse surface texture is desirable.

24.3.5 Workability

Workability is the ease with which the mix can be laid and compacted, and formed to the required condition and shape. This depends on the gradation of aggregates, their shape and texture, bitumen content and its type. Angular, flaky, and elongated aggregates workability. On the other hand, rounded aggregates improve workability.

24.3.6 Desirable properties

From the above discussion, the desirable properties of a bituminous mix can be summarized as follows:

- Stability to meet traffic demand
- Bitumen content to ensure proper binding and water proofing
- Voids to accommodate compaction due to traffic
- Flexibility to meet traffic loads, esp. in cold season
- Sufficient workability for construction
- Economical mix

24.4 Summary

Bituminous mixes should be stable, durable, flexible, workable and should offer sufficient skid resistance. The mix consists of coarse and fine aggregates, filler and binder. It may be well graded, open graded, gap graded or unbounded as per the requirements. As far as possible, it should be economical also.

24.5 Problems

1. Granite is an example for
   (a) Coarse aggregate
   (b) Fine aggregate
   (c) Filler
   (d) none of these
2. Grooving is
   (a) deterioration of pavements locally
   (b) exposure of aggregate due to losing of bitumen
   (c) longitudinal ridging due to channelization of traffic
   (d) none of these

24.6 Solutions

1. Granite is an example for
   (a) Coarse aggregate$\surd$
   (b) Fine aggregate
   (c) Filler
   (d) none of these
2. Grooving is
   (a) deterioration of pavements locally
   (b) exposure of aggregate due to losing of bitumen
   (c) longitudinal ridging due to channelization of traffic$\surd$
   (d) none of these

Chapter 25

Dry Mix Design

25.1 Overview

The objective of dry mix design is to determine the amount of various sizes of mineral aggregates to use to get a mix of maximum density. The dry mix design involves three important steps, viz. selection of aggregates, aggregates gradation, and proportion of aggregates, which are discussed below.

25.2 Selection of aggregates

The desirable qualities of a bituminous paving mixture are dependent to a considerable degree on the nature of the aggregates used. Aggregates are classified as coarse, fine, and filler. The function of the coarse aggregates in contributing to the stability of a bituminous paving mixture is largely due to interlocking and frictional resistance of adjacent particles. Similarly, fines or sand contributes to stability failure function in filling the voids between coarse aggregates. Mineral filler is largely visualized as a void filling agent. Crushed aggregates and sharp sands produce higher stability of the mix when compared with gravel and rounded sands.

25.3 Aggregate gradation

The properties of the bituminous mix including the density and stability are very much dependent on the aggregates and their grain size distribution. Gradation has a profound effect on mix performance. It might be reasonable to believe that the best gradation is one that produces maximum density. This would involve a particle arrangement where smaller particles are packed between larger particles, thus reducing the void space between particles. This create more particle-to-particle contact, which in bituminous pavements would increase stability and reduce water infiltration. However, some minimum amount of void space is necessary to:

- provide adequate volume for the binder to occupy,
- promote rapid drainage, and
- provide resistance to frost action for base and sub base courses.

A dense mixture may be obtained when this particle size distribution follows Fuller law which is expressed as:

$$p = 100\left(\frac{d}{D}\right)^n \tag{25.1}$$

where, p is the percent by weight of the total mixture passing any given sieve sized, D is the size of the largest particle in that mixture, and n is the parameter depending on the shape of the aggregate (0.5 for perfectly rounded particles). Based on this law Fuller-Thompson gradation charts were developed by adjusting the parameter n for fineness or coarseness of aggregates. Practical considerations like construction, layer thickness, workability, etc, are also considered. For example Table 25:1 provides a typical gradation for bituminous concrete for a thickness of 40 mm.

Table 25:1: Specified gradation of aggregates for BC surface course of 40 mm

Sieve size (mm)	Wt passing (%) Grade 1	Wt passing (%) Grade 2
20	-	100
12.5	100	80-100
10.0	80 - 100	70 - 90
4.75	55 - 75	50 - 70
2.36	35 - 50	35 - 50
0.60	18 - 29	18 - 29
0.30	13 - 23	13 - 23
0.15	8 - 16	8 - 16
0.075	4 - 10	4 - 10
Binder*	5 - 7.5	5 - 7.5

Bitumen content in percent by weight of the mix

25.4 Proportioning of aggregates

After selecting the aggregates and their gradation, proportioning of aggregates has to be done and following are the common methods of proportioning of aggregates:

- **Trial and error procedure**: Vary the proportion of materials until the required aggregate gradation is achieved.
- **Graphical Methods**: Two graphical methods in common use for proportioning of aggregates are, Triangular chart method and Roch's method. The former is used when only three materials are to be mixed.
- **Analytical Method**: In this method a system of equations are developed based on the gradation of each aggregates, required gradation, and solved by numerical methods. With the advent of computer, this method is becoming popular and is discussed below. The resulting solution gives the proportion of each type of material required for the given aggregate gradation.

25.5 Example 1

The gradation required for a typical mix is given in Table 25:2 in column 1 and 2. The gradation of available for three types of aggregate A, B, and C are given in column 3, 4, and 5. Determine the proportions of A,B and C if mixed will get the required gradation in column 2.

Table 25:2: Gradation

Sieve size (mm) (1)	Required Gradation Range (2)	Filler (A) (3)	Fine Aggr. (B) (4)	Coarse Aggr. (C) (5)
25.4	100.0	100.0	100.0	100.0
12.7	90-100	100.0	100.0	94.0
4.76	60-75	100.0	100.0	54.0
1.18	40-55	100.0	66.4	31.3
0.3	20-35	100.0	26.0	22.8
0.15	12-22	73.6	17.6	9.0
0.075	5-10	40.1	5.0	3.1

Solution The solution is obtained by constructing a set of equations considering the lower and upper limits of the required gradation as well as the percentage passing of each type of aggregate. The decision need to take is the proportion of aggregate A, B, C need to be blended to get the gradation of column 2. Let x_1, x_2, x_3 represent the proportion of A, B, and C respectively. Equation of the form $ax_1 + bx_2 + cx_3 \leq p_l$ or $\geq p_v$ can be written for each sieve size, where a, b, c is the proportion of aggregates A, B, and C passing for that sieve size and p_l and p_v are the required gradation for that sieve size. This will lead to the following system of equations:

$$
\begin{array}{rrrcl}
x_1+ & x_2+ & x_3 & = & 1 \\
x_1+ & x_2+ & 0.94x_3 & \geq & 0.90 \\
x_1+ & x_2+ & 0.94x_3 & \leq & 1.0 \\
x_1+ & x_2+ & 0.54x_3 & \geq & 0.6 \\
x_1+ & x_2+ & 0.54x_3 & \leq & 0.75 \\
x_1+ & 0.664x_2+ & 0.313x_3 & \geq & 0.4 \\
x_1+ & 0.664x_2+ & 0.313x_3 & \leq & 0.55 \\
x_1+ & 0.260x_2+ & 0.228x_3 & \geq & 0.2 \\
x_1+ & 0.260x_2+ & 0.228x_3 & \leq & 0.35 \\
0.736x_1+ & 0.176x_2+ & 0.09x_3 & \geq & 0.12 \\
0.736x_1+ & 0.176x_2+ & 0.09x_3 & \leq & 0.22 \\
0.401x_1+ & 0.050x_2+ & 0.031x_3 & \geq & 0.05 \\
0.401x_1+ & 0.050x_2+ & 0.031x_3 & \leq & 0.10
\end{array}
\tag{25.2}
$$

Solving the above system of equations manually is extremely difficult. Good computer programs are required to solve this. Software like solver in Excel and Matlab can be used. Soving this set of equations is outside the scope of this book. Suppose the solution to this problem is $x_1 = 0.05$, $x_2 = 0.3$, $x_3 = 0.65$. Then Table 25:3 shows how when these proportions of aggregates A, B, and C are combined, produces the required gradation.

Table 25:3: Result of mix design

Sieve size (mm) (1)	Filler (A) (2)	Fine Aggr. (B) (3)	Coarse Aggr. (C) (4)	Combined Gradation Obtained (5)	Required Gradation Range (6)
25.4	100x0.05=5.0	100x0.3=30.0	100x.65=65	100	100
12.7	100x0.05=5.0	100x0.3=30.0	94x0.65=61	96	90-100
4.76	100x0.05=5.0	100x0.3=30.0	54x0.65=35.1	70.1	60-75
1.18	100x0.05=5.0	66.4x0.3=19.8	31.3x0.65=20.4	45.2	40-55
0.3	100x0.05=5.0	26.3x0.3=07.8	22.8x.65=14.8	27.6	20-35
0.15	73.6x0.05=3.7	17.6x0.3=05.3	9x0.65=5.9	14.9	12-22
0.75	40.1x0.05=2.0	5x0.3=01.5	3.1x0.65=2.0	5.5	5-10

25.6 Summary

Various steps involved in the dry mix design were discussed. Gradation aims at reducing the void space, thus improving the performance of the mix. Proportioning is done by trial and error and graphical methods.

25.7 Problems

1. Fullers law is expressed as
   (a) p=100$\times \left[\frac{d}{D}\right]^n$
   (b) p=100$\times \left[\frac{D}{d}\right]^n$
   (c) p=100$\times \left[\frac{d^2}{D}\right]^n$
   (d) p=100$\times \left[\frac{d}{D^2}\right]^n$

Solutions

1. Fullers law is expressed as
   (a) p=100$\times \left[\frac{d}{D}\right]^n \surd$
   (b) p=100$\times \left[\frac{D}{d}\right]^n$
   (c) p=100$\times \left[\frac{d^2}{D}\right]^n$
   (d) p=100$\times \left[\frac{d}{D^2}\right]^n$

-

Chapter 26

Marshall Mix Design

26.1 Overview

The mix design (wetmix) determines the optimum bitumen content. This is preceded by the dry mix design discussed in the previous chapter. There are many methods available for mix design which vary in the size of the test specimen, compaction, and other test specifications. Marshall method of mix design is the most popular one and is discussed below.

26.2 Marshall mix design

The Marshall stability and flow test provides the performance prediction measure for the Marshall mix design method. The stability portion of the test measures the maximum load supported by the test specimen at a loading rate of 50.8 mm/minute. Load is applied to the specimen till failure, and the maximum load is designated as stability. During the loading, an attached dial gauge measures the specimen's plastic flow (deformation) due to the loading. The flow value is recorded in 0.25 mm (0.01 inch) increments at the same time when the maximum load is recorded. The important steps involved in marshal mix design are summarized next.

26.3 Specimen preparation

Approximately 1200gm of aggregates and filler is heated to a temperature of $175 - 190^oC$. Bitumen is heated to a temperature of $121 - 125^oC$ with the first trial percentage of bitumen (say 3.5 or 4% by weight of the mineral aggregates). The heated aggregates and bitumen are thoroughly mixed at a temperature of $154 - 160^oC$. The mix is placed in a preheated mould and compacted by a rammer with 50 blows on either side at temperature of 138^oC to 149^oC. The weight of mixed aggregates taken for the preparation of the specimen may be suitably altered to obtain a compacted thickness of 63.5+/-3 mm. Vary the bitumen content in the next trial by +0.5% and repeat the above procedure. Number of trials are predetermined. The prepared mould is loaded in the Marshall test setup as shown in the figure 26:1.

26.4 Properties of the mix

The properties that are of interest include the theoretical specific gravity G_t, the bulk specific gravity of the mix G_m, percent air voids V_v, percent volume of bitumen V_b, percent void in mixed aggregate VMA and percent

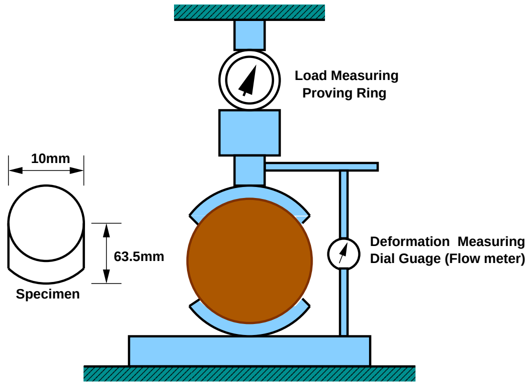


Figure 26:1: Marshall test setup

voids filled with bitumen VFB. These calculations are discussed next. To understand these calculation a phase diagram is given in Figure 26:2.

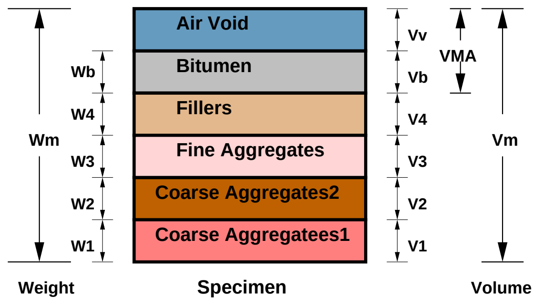


Figure 26:2: Phase diagram of a bituminous mix

26.4.1 Theoretical specific gravity of the mix G_t

Theoretical specific gravity G_t is the specific gravity without considering air voids, and is given by:

$$G_t = \frac{W_1 + W_2 + W_3 + W_b}{\frac{W_1}{G_1} + \frac{W_2}{G_2} + \frac{W_3}{G_3} + \frac{W_b}{G_b}} \tag{26.1}$$

where, W_1 is the weight of coarse aggregate in the total mix, W_2 is the weight of fine aggregate in the total mix, W_3 is the weight of filler in the total mix, W_b is the weight of bitumen in the total mix, G_1 is the apparent specific gravity of coarse aggregate, G_2 is the apparent specific gravity of fine aggregate, G_3 is the apparent specific gravity of filler and G_b is the apparent specific gravity of bitumen,

26.4.2 Bulk specific gravity of mix G_m

The bulk specific gravity or the actual specific gravity of the mix G_m is the specific gravity considering air voids and is found out by:

$$G_m = \frac{W_m}{W_m - W_w} \tag{26.2}$$

where, W_m is the weight of mix in air, W_w is the weight of mix in water, Note that $W_m - W_w$ gives the volume of the mix. Sometimes to get accurate bulk specific gravity, the specimen is coated with thin film of paraffin wax, when weight is taken in the water. This, however requires to consider the weight and volume of wax in the calculations.

26.4.3 Air voids percent V_v

Air voids V_v is the percent of air voids by volume in the specimen and is given by:

$$V_v = \frac{(G_t - G_m)100}{G_t} \tag{26.3}$$

where G_t is the theoretical specific gravity of the mix, given by equation 26.1. and G_m is the bulk or actual specific gravity of the mix given by equation 26.2.

26.4.4 Percent volume of bitumen V_b

The volume of bitumen V_b is the percent of volume of bitumen to the total volume and given by:

$$V_b = \frac{\frac{W_b}{G_b}}{\frac{W_1+W_2+W_3+W_b}{G_m}} \tag{26.4}$$

where, W_1 is the weight of coarse aggregate in the total mix, W_2 is the weight of fine aggregate in the total mix, W_3 is the weight of filler in the total mix, W_b is the weight of bitumen in the total mix, G_b is the apparent specific gravity of bitumen, and G_m is the bulk specific gravity of mix given by equation 26.2.

26.4.5 Voids in mineral aggregateVMA

Voids in mineral aggregate VMA is the volume of voids in the aggregates, and is the sum of air voids and volume of bitumen, and is calculated from

$$VMA = V_v + V_b \tag{26.5}$$

where, V_v is the percent air voids in the mix, given by equation 26.3. and V_b is percent bitumen content in the mix, given by equation 26.4. (26.4).

26.4.6 Voids filled with bitumen VFB

Voids filled with bitumen VFB is the voids in the mineral aggregate frame work filled with the bitumen, and is calculated as:

$$VFB = \frac{V_b \times 100}{VMA} \tag{26.6}$$

where, V_b is percent bitumen content in the mix, given by equation 26.4. and VMA is the percent voids in the mineral aggregate, given by equation 26.5.

26.5 Determine Marshall stability and flow

Marshall stability of a test specimen is the maximum load required to produce failure when the specimen is preheated to a prescribed temperature placed in a special test head and the load is applied at a constant strain

(5 cm per minute). While the stability test is in progress dial gauge is used to measure the vertical deformation of the specimen. The deformation at the failure point expressed in units of 0.25 mm is called the Marshall flow value of the specimen.

26.6 Apply stability correction

It is possible while making the specimen the thickness slightly vary from the standard specification of 63.5 mm. Therefore, measured stability values need to be corrected to those which would have been obtained if the specimens had been exactly 63.5 mm. This is done by multiplying each measured stability value by an appropriated correlation factors as given in Table below.

Table 26:1: Correction factors for Marshall stability values

Volume of specimen (cm^3)	Thickness of specimen (mm)	Correction Factor
457 - 470	57.1	1.19
471 - 482	68.7	1.14
483 - 495	60.3	1.09
496 - 508	61.9	1.04
509 - 522	63.5	1.00
523 - 535	65.1	0.96
536 - 546	66.7	0.93
547 - 559	68.3	0.89
560 - 573	69.9	0.86

26.7 Prepare graphical plots

The average value of the above properties are determined for each mix with different bitumen content and the following graphical plots are prepared:

1. Binder content versus corrected Marshall stability
2. Binder content versus Marshall flow
3. Binder content versus percentage of void (V_v) in the total mix
4. Binder content versus voids filled with bitumen (VFB)
5. Binder content versus unit weight or bulk specific gravity (G_m)

26.8 Determine optimum bitumen content

Determine the optimum binder content for the mix design by taking average value of the following three bitumen contents found form the graphs obtained in the previous step.

1. Binder content corresponding to maximum stability
2. Binder content corresponding to maximum bulk specific gravity (G_m)
3. Binder content corresponding to the median of designed limits of percent air voids (V_v) in the total mix (i.e. 4%)

The stability value, flow value, and VFB are checked with Marshall mix design specification chart given in Table below. Mixes with very high stability value and low flow value are not desirable as the pavements constructed with such mixes are likely to develop cracks due to heavy moving loads.

Table 26:2: Marshall mix design specification

Test Property	Specified Value
Marshall stability, kg	340 (minimum)
Flow value, 0.25 mm units	8 - 17
Percent air voids in the mix V_v%	3 - 5
Voids filled with bitumen VFB%	75 - 85

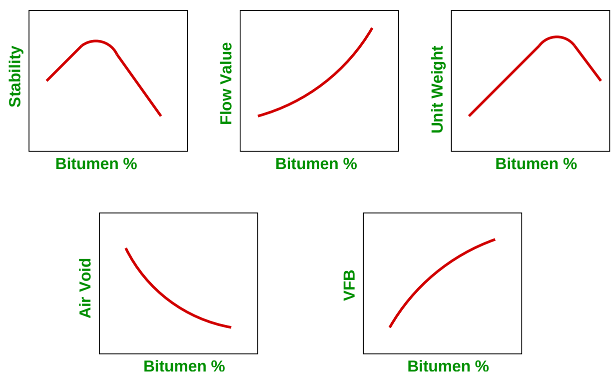


Figure 26:3: Marshal graphical plots

26.9 Numerical example - 1

The specific gravities and weight proportions for aggregate and bitumen are as under for the preparation of Marshall mix design. The volume and weight of one Marshall specimen was found to be 475 cc and 1100 gm. Assuming absorption of bitumen in aggregate is zero, find V_v, V_b, VMA and VFB;

Item	A_1	A_2	A_3	A_4	B
Wt (gm)	825	1200	325	150	100
Sp. Gr	2.63	2.51	2.46	2.43	1.05

Solution

$$
\begin{aligned}
G_t &= \frac{825+1200+325+150+100}{\frac{825}{2.63}+\frac{1200}{2.51}+\frac{325}{2.46}+\frac{150}{2.43}+\frac{100}{1.05}} \\
&= \frac{2600}{1080.86} \\
&= 2.406 \\
G_m &= \frac{1100}{475} \\
&= 2.316 \\
V_v &= \frac{2.406-2.316}{2.406}\times 100 \\
&= 3.741\ \% \\
V_b &= \frac{100}{1.05}\times\frac{2.316}{1100} \\
&= 20.052\ \% \\
VMA &= (3.741+20.05) \\
&= 23.793\ \% \\
VFB &= \frac{20.052}{23.793}\times 100 \\
&= 84.277\ \%
\end{aligned}
$$

26.10 Numerical example - 2

The results of Marshall test for five specimen is given below. Find the optimum bitumen content of the mix.

Bitumen content	Stability (kg)	Flow (units)	V_v (%)	VFB (%)	G_m
3	499.4	9.0	12.5	34	2.17
4	717.3	9.6	7.2	65	2.21
5	812.7	12.0	3.9	84	2.26
6	767.3	14.8	2.4	91	2.23
7	662.8	19.5	1.9	93	2.18

Solution Plot the graphs and find bitumen content corresponding to

1. Max stability = 5 percent bitumen content.

2. Max G_m = 5 percent bitumen content.
3. 4% percent air void = 3 percent bitumen content.

The optimum bitumen extent is the average of above = 4.33 percent.

26.11 Summary

Marshal stability test is the performance prediction measure conducted on the bituminous nix. The procedure consists of determination of properties of mix, Marshal stability and flow analysis and finally determination of optimum bitumen content. The concept of phase diagram is used for the calculations.

26.12 Problems

1. In Marshall stability test, the sample is compacted using a rammer giving
   (a) 50 blows
   (b) 20 blows
   (c) 25 blows
   (d) 75 blows
2. The Marshall flow value is expressed in units of
   (a) 25 mm
   (b) 2.5mm
   (c) 5mm
   (d) 3mm

26.13 Solutions

1. In Marshall stability test, the sample is compacted using a rammer giving
   (a) 50 blows$\surd$
   (b) 20 blows
   (c) 25 blows
   (d) 75 blows
2. The Marshall flow value is expressed in units of
   (a) 25 mm
   (b) 2.5mm$\surd$
   (c) 5mm
   (d) 3mm

Chapter 27

Flexible pavement design

27.1 Overview

Flexible pavements are so named because the total pavement structure deflects, or flexes, under loading. A flexible pavement structure is typically composed of several layers of materials. Each layer receives loads from the above layer, spreads them out, and passes on these loads to the next layer below. Thus the stresses will be reduced, which are maximum at the top layer and minimum on the top of subgrade. In order to take maximum advantage of this property, layers are usually arranged in the order of descending load bearing capacity with the highest load bearing capacity material (and most expensive) on the top and the lowest load bearing capacity material (and least expensive) on the bottom.

27.2 Design procedures

For flexible pavements, structural design is mainly concerned with determining appropriate layer thickness and composition. The main design factors are stresses due to traffic load and temperature variations. Two methods of flexible pavement structural design are common today: Empirical design and mechanistic empirical design.

27.2.1 Empirical design

An empirical approach is one which is based on the results of experimentation or experience. Some of them are either based on physical properties or strength parameters of soil subgrade. An empirical approach is one which is based on the results of experimentation or experience. An empirical analysis of flexible pavement design can be done with or with out a soil strength test. An example of design without soil strength test is by using HRB soil classification system, in which soils are grouped from A-1 to A-7 and a group index is added to differentiate soils within each group. Example with soil strength test uses McLeod, Stabilometer, California Bearing Ratio (CBR) test. CBR test is widely known and will be discussed.

27.2.2 Mechanistic-Empirical Design

Empirical-Mechanistic method of design is based on the mechanics of materials that relates input, such as wheel load, to an output or pavement response. In pavement design, the responses are the stresses, strains, and deflections within a pavement structure and the physical causes are the loads and material properties of the pavement structure. The relationship between these phenomena and their physical causes are typically described using some mathematical models. Along with this mechanistic approach, empirical elements are used when defining what value of the calculated stresses, strains, and deflections result in pavement failure. The relationship between physical phenomena and pavement failure is described by empirically derived equations that compute the number of loading cycles to failure.

27.3 Traffic and Loading

There are three different approaches for considering vehicular and traffic characteristics, which affects pavement design.

Fixed traffic: Thickness of pavement is governed by single load and number of load repetitions is not considered. The heaviest wheel load anticipated is used for design purpose. This is an old method and is rarely used today for pavement design.

Fixed vehicle: In the fixed vehicle procedure, the thickness is governed by the number of repetitions of a standard axle load. If the axle load is not a standard one, then it must be converted to an equivalent axle load by number of repetitions of given axle load and its equivalent axle load factor.

Variable traffic and vehicle: In this approach, both traffic and vehicle are considered individually, so there is no need to assign an equivalent factor for each axle load. The loads can be divided into a number of groups and the stresses, strains, and deflections under each load group can be determined separately; and used for design purposes. The traffic and loading factors to be considered include axle loads, load repetitions, and tyre contact area.

27.3.1 Equivalent single wheel load

To carry maximum load with in the specified limit and to carry greater load, dual wheel, or dual tandem assembly is often used. Equivalent single wheel load (ESWL) is the single wheel load having the same contact pressure, which produces same value of maximum stress, deflection, tensile stress or contact pressure at the desired depth. The procedure of finding the ESWL for equal stress criteria is provided below. This is a semi-rational method, known as Boyd and Foster method, based on the following assumptions:

- equalancy concept is based on equal stress;
- contact area is circular;
- influence angle is 45^o; and
- soil medium is elastic, homogeneous, and isotropic half space.

The ESWL is given by:

$$\log_{10} ESWL = \log_{10} P + \frac{0.301 \log_{10}\left(\frac{z}{d/2}\right)}{\log_{10}\left(\frac{2S}{d/2}\right)} \tag{27.1}$$

where P is the wheel load, S is the center to center distance between the two wheels, d is the clear distance between two wheels, and z is the desired depth.

Example 1

Find ESWL at depths of 5cm, 20cm and 40cm for a dual wheel carrying 2044 kg each. The center to center tyre spacing is 20cm and distance between the walls of the two tyres is 10cm.

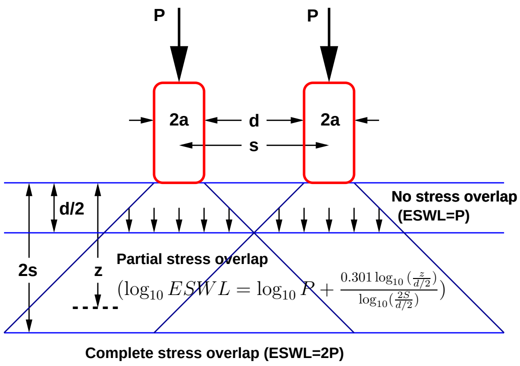


Figure 27:1: ESWL-Equal stress concept

Solution For desired depth z=40cm, which is twice the tyre spacing, ESWL = 2P=2×2044 = 4088 kN. For z=5cm, which is half the distance between the walls of the tyre, ESWL = P = 2044kN. For z=20cm, $\log_{10} ESWL = \log_{10} P + \frac{0.301 \log_{10}(\frac{z}{d/2})}{\log_{10}(\frac{2S}{d/2})}$ =$\log_{10} ESWL = \log_{10} 2044 + \frac{0.301 \log_{10}(\frac{20}{10/2})}{\log_{10}(\frac{2\times 20}{10/2})}$ =3.511. Therefore, ESWL = antilog(3.511)= 3244.49 kN

27.3.2 Equivalent single axle load

Vehicles can have many axles which will distribute the load into different axles, and in turn to the pavement through the wheels. A standard truck has two axles, front axle with two wheels and rear axle with four wheels. But to carry large loads multiple axles are provided. Since the design of flexible pavements is by layered theory, only the wheels on one side needed to be considered. On the other hand, the design of rigid pavement is by plate theory and hence the wheel load on both sides of axle need to be considered. *Legal axle load:* The maximum allowed axle load on the roads is called legal axle load. For highways the maximum legal axle load in India, specified by IRC, is 10 tonnes. *Standard axle load:* It is a single axle load with dual wheel carrying 80 KN load and the design of pavement is based on the standard axle load.

Repetition of axle loads: The deformation of pavement due to a single application of axle load may be small but due to repeated application of load there would be accumulation of unrecovered or permanent deformation which results in failure of pavement. If the pavement structure fails with N_1 number of repetition of load W_1 and for the same failure criteria if it requires N_2 number of repetition of load W_2, then W_1N_1 and W_2N_2 are considered equivalent. Note that, W_1N_1 and W_2N_2 equivalency depends on the failure criterion employed.

Equivalent axle load factor: An equivalent axle load factor (EALF) defines the damage per pass to a pavement by the i^{th} type of axle relative to the damage per pass of a standard axle load. While finding the EALF, the failure criterion is important. Two types of failure criterias are commonly adopted: fatigue cracking and ruttings. The fatigue cracking model has the following form:

$$N_f = f_1 \left(\epsilon_t\right)^{-f_2} \times \left(E\right)^{-f_3} or N_f \propto {\epsilon_t}^{-f_2} \tag{27.2}$$

where, N_f is the number of load repetition for a certain percentage of cracking, ϵ_t is the tensile strain at the bottom of the binder course, E is the modulus of elasticity, and f_1, f_2, f_3 are constants. If we consider fatigue cracking as failure criteria, and a typical value of 4 for f_2, then:

$$EALF = \left(\frac{\epsilon_i}{\epsilon_{std}}\right)^4 \tag{27.3}$$

where, i indicate i^{th} vehicle, and std indicate the standard axle. Now if we assume that the strain is proportional to the wheel load,

$$EALF = \left(\frac{W_i}{W_{std}}\right)^4 \tag{27.4}$$

Similar results can be obtained if rutting model is used, which is:

$$N_d = f_4\left(\epsilon_c\right)^{-f_5} \tag{27.5}$$

where N_d is the permissible design rut depth (say 20mm), ϵ_c is the compressive strain at the top of the subgrade, and f_4, f_5 are constants. Once we have the EALF, then we can get the ESAL as given below.

$$\text{Equivalent single axle load, ESAL} = \sum_{i=1}^{m} F_i n_i \tag{27.6}$$

where,m is the number of axle load groups, F_i is the $EALF$ for i^{th} axle load group, and n_i is the number of passes of i^{th} axle load group during the design period.

Example 1

Let number of load repetition expected by 80 KN standard axle is 1000, 160 KN is 100 and 40 KN is 10000. Find the equivalent axle load.

Solution: Refer the Table 27:1. The ESAL is given as $\sum F_i n_i = 3225\ kN$

Table 27:1: Example 1 Solution

i	Axle Load (KN)	No.of Load Repetition (n_i)	EALF (F_i)	$F_i n_i$
1	40	10000	$(40/80)^4 = 0.0625$	625
2	80	1000	$(80/80)^4 = 1$	1000
3	160	100	$(160/80)^4 = 16$	1600

Example 2

Let the number of load repetition expected by 120 kN axle is 1000, 160 kN is 100, and 40 kN is 10,000. Find the equivalent standard axle load if the equivalence criteria is rutting. Assume 80 kN as standard axle load and the rutting model is $N_r = f_4 \epsilon_c^{-f_5}$ where f_4 =4.2 and f_5 =4.5.

Solution Refer the Table 27:2. The ESAL is given as $\sum F_i n_i = 8904.94\ kN$

Table 27:2: Example 2 Solution

i	Axle Load (KN)	No.of Load Repetition (n_i)	EALF (F_i)	$F_i n_i$
1	120	1000	$(120/80)^{4.5} = 6.200$	6200
2	160	100	$(160/80)^{4.5} = 22.63$	2263
3	40	10000	$(40/80)^{4.5} = 0.04419$	441.9

Example 3

Let number of load repetition expected by 60kN standard axle is 1000, 120kN is 200 and 40 kN is 10000. Find the equivalent axle load using fatigue cracking as failure criteria according to IRC. Hint:$N_f = 2.21 \times 10^{-4}(\epsilon_t)^{-3.89}(E)^0.854$

Solution Refer the Table 27:3. The ESAL is given as $\sum F_i n_i = 6030.81\ kN$

Table 27:3: Example 3 Solution

i	Axle Load (KN)	No.of Load Repetition (n_i)	EALF (F_i)	$F_i n_i$
1	40	10000	$(40/60)^{3.89} = 0.2065$	2065
2	60	1000	$(60/60)^{3.89} = 1$	1000
3	120	200	$(120/60)^{3.89} = 14.825$	2965.081

27.4 Material characterization

It is well known that the pavement materials are not perfectly elastic but experiences some permanent deformation after each load repetitions. It is well known that most paving materials are not elastic but experience some permanent deformation after each load application. However, if the load is small compared to the strength of the material and the deformation under each load repetition is almost completely recoverable then the material can be considered as elastic.

The Figure 27:2 shows straining of a specimen under a repeated load test. At the initial stage of load applications, there is considerable permanent deformation as indicated by the plastic strain in the Figure 27:2. As the number of repetition increases, the plastic strain due to each load repetition decreases. After 100 to 200 repetitions, the strain is practically all-recoverable, as indicated by ϵ_r in the figure.

27.4.1 Resilient modulus of soil

The elastic modulus based on the recoverable strain under repeated loads is called the resilient modulus M_R, defined as $M_R = \frac{\sigma_d}{\epsilon_r}$. In which σ_d is the deviator stress, which is the axial stress in an unconfined compression

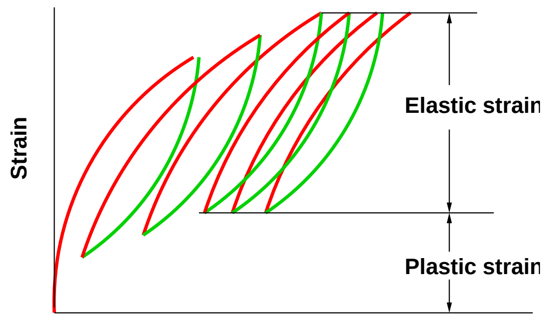


Figure 27:2: Recoverable strain under repeated loads

test or the axial stress in excess of the confining pressure in a triaxial compression test.

In pavements the load applied are mostly transient and the type and duration of loading used in the repeated load test should simulate that actually occurring in the field. When a load is at a considerable distance from a given point, the stress at that point is maximum. It is therefore reasonable to assume the stress pulse to be a haversine or triangular loading, and the duration of loading depends on the vehicle speed and the depth of the point below the pavement surface. Resilient modulus test can be conducted on all types of pavement materials ranging from cohesive to stabilized materials. The test is conducted in a triaxial device equipped for repetitive load conditions.

27.4.2 Dynamic complex modulus

When the loading wave form is sinusoidal and if there is no rest period, then, the modulus obtained is called dynamic complex modulus. This is one of the way of explaining the stress-strain relationship of visco-elastic materials. This modulus is a complex quantity and the absolute value of the complex modulus is called the dynamic modulus. This complex modulus test is usually conducted on cylindrical specimens subjected to a compressive haversine loading. The test setup is similar to resiliant modulus. The dynamic modulus varies with the loading frequency. Therefore, a frequency that most closely simulates the actual traffic load should be selected for the test.

27.4.3 Correlations with other tests

Determination of resiliant modulus is often cumbersome. Therefore, various empirical tests have been used to determine the material properties for pavement design. Most of these test measure the strength of the material and are not a true representation of the resiliant modulus. Accordingly, various studies has related empirical tests like CBR test, Tri-axial test etc are correlated to resiliant modulus.

27.5 Mechanistic-empirical analysis

Mechanics is the science of motion and action of forces on bodies. In pavement design these phenomena are stresses, strains, and deflections within a pavement structure and the physical causes are loads and material properties of the pavements structure. The relationship between these phenomena and their physical causes is described by a mathematical model. The most common of them is layered elastic model.

27.5.1 Advantages

The basic advantages of the Mechanistic-Empirical pavement design method over a purely empirical one are:

1. It can be used for both existing pavement rehabilitation and new pavement construction
2. It can accommodate changing load types
3. It can better characterize materials allowing for
   - better utilization of available materials
   - accommodation of new materials
   - improved definition of existing layer proportion
4. It uses material proportion that relates better with actual pavement performance
5. It provides more reliable performance predictions
6. It defines role of construction in a better way
7. It accommodates environment and aging effect of materials in the pavement

27.5.2 Mechanistic model

Mechanistic models are used to mathematically model pavement physics. There are a number of different types of models available today (e.g., layered elastic, dynamic, viscoelastic) but this section will present the layered elastic model.

Layered elastic model

A layered elastic model can compute stresses, strains and deflections at any point in a pavement structure resulting from the application of a surface load. Layered elastic models assume that each pavement structural layer is homogeneous, isotropic, and linearly elastic. In other words, it is the same everywhere and will rebound to its original form once the load is removed. This section covers the basic assumptions, inputs and outputs from a typical layered elastic model.

Assumptions in layered elastic model

The layered elastic approach works with relatively simple mathematical models and thus requires following assumptions

- Pavement layer extends infinitely in the horizontal direction
- The bottom layer (usually the subgrade) extends infinitely downwards
- Materials are not stressed beyond their elastic ranges

Inputs

A layered elastic model requires a minimum number of inputs to adequately characterize a pavement structure and its response to loading. These inputs are:

- Material properties of each layer, like modulus of elasticity (E), Poisson's ratio (ν),
- Pavement layer thicknesses, and
- Loading conditions which include the total wheel load (P) and load repetitions.

Output

The outputs of the layered elastic model are the stresses, strains and deflections in the pavements.

- **Stress**. The intensity of internally distributed forces experienced within the pavement structure at various points. Stress has units of force per unit area(pa)
- **Strain**. The unit displacement due to stress , usually expressed as a ratio of change in dimension to the original dimension (mm/mm)
- **Deflection**. The linear change in dimension. Deflection is expressed in units of length (mm)

Failure criteria

The main empirical portions of the mechanistic-empirical design process are the equations used to compute the number of loading cycles to failure. These equations are derived by observing the performance of pavements and relating the type and extent of observed failure to an initial strain under various loads. Currently, two types of failure criteria are widely recognized, one relating to fatigue cracking and the other to rutting initiating in the subgrade.

27.6 Summary

Basic concepts of flexible pavement design were discussed. There are two main design procedures- empirical and mechanistic empirical design. For design purposes, equivalent single wheel load and equivalent single axle load concepts are used.

27.7 Problems

1. A set of dual tyres has a total load of 4090 kg, a contact radius a of 11.4 cm and a center to center tyre spacing of 34.3 cm. Find the ESWL by Boyd & Foster method for a depth of 34.3 cm. [Ans: 3369.3 kg]
2. Calculate ESWL by equal stress criteria for a dual wheel assembly carrying 2044 kg each for a pavement thickness of 5, 15, 20, 25 and 30 cms. The distance between walls of the tyre is 11 cm. Use either graphical or functional methods. (Hint: P=2044kg, d=11cm, s=27cm). [Ans: 2044, 2760, 3000, 3230 and 4088]
3. Let number of load repetition expected by 60kN standard axle is 1000, 120kN is 200 and 40 kN is 10000. Find the equivalent axle load using fatigue cracking as failure criteria according to IRC. Hint:$N_f = 2.21 \times 10^{-4}(\epsilon_t)^{-3.89}(E)^{0.854}$

Chapter 28

IRC method of design of flexible pavements

28.1 Overview

Indian roads congress has specified the design procedures for flexible pavements based on CBR values. The Pavement designs given in the previous edition IRC:37-1984 were applicable to design traffic upto only 30 million standard axles (msa). The earlier code is empirical in nature which has limitations regarding applicability and extrapolation. This guidelines follows analytical designs and developed new set of designs up to 150 msa in IRC:37-2001.

28.2 Scope

These guidelines will apply to design of flexible pavements for Expressway, National Highways, State Highways, Major District Roads, and other categories of roads. Flexible pavements are considered to include the pavements which have bituminous surfacing and granular base and sub-base courses conforming to IRC/ MOST standards. These guidelines apply to new pavements.

28.3 Design criteria

The flexible pavements has been modeled as a three layer structure and stresses and strains at critical locations have been computed using the linear elastic model. To give proper consideration to the aspects of performance, the following three types of pavement distress resulting from repeated (cyclic) application of traffic loads are considered:

1. vertical compressive strain at the top of the sub-grade which can cause sub-grade deformation resulting in permanent deformation at the pavement surface.
2. horizontal tensile strain or stress at the bottom of the bituminous layer which can cause fracture of the bituminous layer.
3. pavement deformation within the bituminous layer.

While the permanent deformation within the bituminous layer can be controlled by meeting the mix design requirements, thickness of granular and bituminous layers are selected using the analytical design approach so

!h

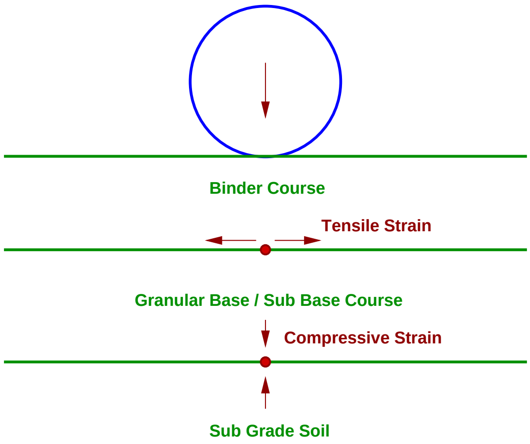


Figure 28:1: Critical Locations in Pavement

that strains at the critical points are within the allowable limits. For calculating tensile strains at the bottom of the bituminous layer, the stiffness of dense bituminous macadam (DBM) layer with 60/70 bitumen has been used in the analysis.

28.4 Failure Criteria

A and B are the critical locations for tensile strains (ϵ_t). Maximum value of the strain is adopted for design. C is the critical location for the vertical subgrade strain (ϵ_z) since the maximum value of the (ϵ_z) occurs mostly at C.

Fatigue Criteria:

Bituminous surfacings of pavements display flexural fatigue cracking if the tensile strain at the bottom of the bituminous layer is beyond certain limit. The relation between the fatigue life of the pavement and the tensile strain in the bottom of the bituminous layer was obtained as

$$N_f = 2.21 \times 10^{-4} \times \left(\frac{1}{\epsilon_t}\right)^{3.89} \times \left(\frac{1}{E}\right)^{0.854} \tag{28.1}$$

in which, N_f is the allowable number of load repetitions to control fatigue cracking and E is the Elastic modulus of bituminous layer. The use of equation 28.1 would result in fatigue cracking of 20% of the total area.

Rutting Criteria

The allowable number of load repetitions to control permanent deformation can be expressed as

$$N_r = 4.1656 \times 10^{-8} \times \left(\frac{1}{\epsilon_z}\right)^{4.5337} \tag{28.2}$$

N_r is the number of cumulative standard axles to produce rutting of 20 mm.

28.5 Design procedure

Based on the performance of existing designs and using analytical approach, simple design charts and a catalogue of pavement designs are added in the code. The pavement designs are given for subgrade CBR values ranging from 2% to 10% and design traffic ranging from 1 msa to 150 msa for an average annual pavement temperature of 35 C. The later thicknesses obtained from the analysis have been slightly modified to adapt the designs to stage construction. Using the following simple input parameters, appropriate designs could be chosen for the given traffic and soil strength:

- Design traffic in terms of cumulative number of standard axles; and
- CBR value of subgrade.

28.6 Design traffic

The method considers traffic in terms of the cumulative number of standard axles (8160 kg) to be carried by the pavement during the design life. This requires the following information:

1. Initial traffic in terms of CVPD
2. Traffic growth rate during the design life
3. Design life in number of years
4. Vehicle damage factor (VDF)
5. Distribution of commercial traffic over the carriage way.

Initial traffic
Initial traffic is determined in terms of commercial vehicles per day (CVPD). For the structural design of the pavement only commercial vehicles are considered assuming laden weight of three tonnes or more and their axle loading will be considered. Estimate of the initial daily average traffic flow for any road should normally be based on 7-day 24-hour classified traffic counts (ADT). In case of new roads, traffic estimates can be made on the basis of potential land use and traffic on existing routes in the area.

Traffic growth rate
Traffic growth rates can be estimated (i) by studying the past trends of traffic growth, and (ii) by establishing econometric models. If adequate data is not available, it is recommended that an average annual growth rate of 7.5 percent may be adopted.

Design life
For the purpose of the pavement design, the design life is defined in terms of the cumulative number of standard axles that can be carried before strengthening of the pavement is necessary. It is recommended that pavements for arterial roads like NH, SH should be designed for a life of 15 years, EH and urban roads for 20 years and other categories of roads for 10 to 15 years.

Vehicle Damage Factor

The vehicle damage factor (VDF) is a multiplier for converting the number of commercial vehicles of different axle loads and axle configurations to the number of standard axle-load repetitions. It is defined as equivalent number of standard axles per commercial vehicle. The VDF varies with the axle configuration, axle loading, terrain, type of road, and from region to region. The axle load equivalency factors are used to convert different axle load repetitions into equivalent standard axle load repetitions. For these equivalency factors refer IRC:37 2001. The exact VDF values are arrived after extensive field surveys.

Vehicle distribution
A realistic assessment of distribution of commercial traffic by direction and by lane is necessary as it directly affects the total equivalent standard axle load application used in the design. Until reliable data is available, the following distribution may be assumed.

- **Single lane roads:** Traffic tends to be more channelized on single roads than two lane roads and to allow for this concentration of wheel load repetitions, the design should be based on total number of commercial vehicles in both directions.
- **Two-lane single carriageway roads:** The design should be based on 75 % of the commercial vehicles in both directions.
- **Four-lane single carriageway roads:** The design should be based on 40 % of the total number of commercial vehicles in both directions.
- **Dual carriageway roads:** For the design of dual two-lane carriageway roads should be based on 75 % of the number of commercial vehicles in each direction. For dual three-lane carriageway and dual four-lane carriageway the distribution factor will be 60 % and 45 % respectively.

28.7 Pavement thickness design charts

For the design of pavements to carry traffic in the range of 1 to 10 msa, use chart 1 and for traffic in the range 10 to 150 msa, use chart 2 of IRC:37 2001. The design curves relate pavement thickness to the cumulative number of standard axles to be carried over the design life for different sub-grade CBR values ranging from 2 % to 10 %. The design charts will give the total thickness of the pavement for the above inputs. The total thickness consists of granular sub-base, granular base and bituminous surfacing. The individual layers are designed based on the the recommendations given below and the subsequent tables.

28.8 Pavement composition

Sub-base
Sub-base materials comprise natural sand, gravel, laterite, brick metal, crushed stone or combinations thereof meeting the prescribed grading and physical requirements. The sub-base material should have a minimum CBR of 20 % and 30 % for traffic upto 2 msa and traffic exceeding 2 msa respectively. Sub-base usually consist of granular or WBM and the thickness should not be less than 150 mm for design traffic less than 10 msa and 200 mm for design traffic of 1:0 msa and above.
Base
The recommended designs are for unbounded granular bases which comprise conventional water bound macadam

(WBM) or wet mix macadam (WMM) or equivalent confirming to MOST specifications. The materials should be of good quality with minimum thickness of 225 mm for traffic up to 2 msa an 150 mm for traffic exceeding 2 msa.

Bituminous surfacing

The surfacing consists of a wearing course or a binder course plus wearing course. The most commonly used wearing courses are surface dressing, open graded premix carpet, mix seal surfacing, semi-dense bituminous concrete and bituminous concrete. For binder course, MOST specifies, it is desirable to use bituminous macadam (BM) for traffic upto o 5 msa and dense bituminous macadam (DBM) for traffic more than 5 msa.

28.9 Numerical example

Design the pavement for construction of a new bypass with the following data:

1. Two lane carriage way
2. Initial traffic in the year of completion of construction = 400 CVPD (sum of both directions)
3. Traffic growth rate = 7.5 %
4. Design life = 15 years
5. Vehicle damage factor based on axle load survey = 2.5 standard axle per commercial vehicle
6. Design CBR of subgrade soil = 4%.

Solution

1. Distribution factor = 0.75
2. $$\begin{aligned} N &= \frac{365 \times \left[(1+0.075)^{15}-1)\right]}{0.075} \times 400 \times 0.75 \times 2.5 \\ &= 7200000 \\ &= 7.2\ msa \end{aligned}$$
3. Total pavement thickness for CBR 4% and traffic 7.2 msa from IRC:37 2001 chart1 = 660 mm
4. Pavement composition can be obtained by interpolation from Pavement Design Catalogue (IRC:37 2001).
   (a) Bituminous surfacing = 25 mm SDBC + 70 mm DBM
   (b) Road-base = 250 mm WBM
   (c) sub-base = 315 mm granular material of CBR not less than 30 %

28.10 Summary

The design procedure given by IRC makes use of the CBR value, million standard axle concept, and vehicle damage factor. Traffic distribution along the lanes are taken into account. The design is meant for design traffic which is arrived at using a growth rate.

28.11 Problems

1. Design the pavement for construction of a new two lane carriageway for design life 15 years using IRC method. The initial traffic in the year of completion in each direction is 150 CVPD and growth rate is 5%. Vehicle damage factor based on axle load survey = 2.5 std axle per commercial vehicle. Design CBR of subgrade soil=4%.

28.12 Solutions

1. Distribution factor = 0.75

2.

$$\begin{aligned} N &= \frac{365 \times \left[(1+0.05)^{15}-1)\right]}{0.05} \times 300 \times 0.75 \times 2.5 \\ &= 4430348.837 \\ &= 4.4\ msa \end{aligned}$$

3. Total pavement thickness for CBR 4% and traffic 4.4 msa from IRC:37 2001 chart1 = 580 mm

4. Pavement composition can be obtained by interpolation from Pavement Design Catalogue (IRC:37 2001).

   (a) Bituminous surfacing = 20 mm PC + 50 mm BM

   (b) Road-base = 250 mm Granular base

   (c) sub-base = 280 mm granular material.

Chapter 29

Rigid pavement design

29.1 Overview

As the name implies, rigid pavements are rigid i.e, they do not flex much under loading like flexible pavements. They are constructed using cement concrete. In this case, the load carrying capacity is mainly due to the rigidity ad high modulus of elasticity of the slab (slab action). H. M. Westergaard is considered the pioneer in providing the rational treatment of the rigid pavement analysis.

29.1.1 Modulus of sub-grade reaction

Westergaard considered the rigid pavement slab as a thin elastic plate resting on soil sub-grade, which is assumed as a dense liquid. The upward reaction is assumed to be proportional to the deflection. Base on this assumption, Westergaard defined a *modulus of sub-grade reaction* K in kg/cm^3 given by $K = \frac{p}{\Delta}$ where Δ is the displacement level taken as 0.125 cm and p is the pressure sustained by the rigid plate of 75 cm diameter at a deflection of 0.125 cm.

29.1.2 Relative stiffness of slab to sub-grade

A certain degree of resistance to slab deflection is offered by the sub-grade. The sub-grade deformation is same as the slab deflection. Hence the slab deflection is direct measurement of the magnitude of the sub-grade pressure. This pressure deformation characteristics of rigid pavement lead Westergaard to the define the term *radius of relative stiffness* l in cm is given by the equation 29.1.

$$l = \sqrt[4]{\frac{Eh^3}{12K(1-\mu^2)}} \tag{29.1}$$

where E is the modulus of elasticity of cement concrete in kg/cm^2 (3.0×10^5), μ is the Poisson's ratio of concrete (0.15), h is the slab thickness in cm and K is the modulus of sub-grade reaction.

29.1.3 Critical load positions

Since the pavement slab has finite length and width, either the character or the intensity of maximum stress induced by the application of a given traffic load is dependent on the location of the load on the pavement surface. There are three typical locations namely the *interior, edge* and *corner*, where differing conditions of slab continuity exist. These locations are termed as critical load positions.

29.1.4 Equivalent radius of resisting section

When the interior point is loaded, only a small area of the pavement is resisting the bending moment of the plate. Westergaard's gives a relation for equivalent radius of the resisting section in cm in the equation 29.2.

$$b = \begin{cases} \sqrt{1.6a^2 + h^2} - 0.675\ h & \text{if a} < 1.724 \text{ h} \\ a & \text{otherwise} \end{cases} \tag{29.2}$$

where a is the radius of the wheel load distribution in cm and h is the slab thickness in cm.

29.2 Wheel load stresses - Westergaard's stress equation

The cement concrete slab is assumed to be homogeneous and to have uniform elastic properties with vertical sub-grade reaction being proportional to the deflection. Westergaard developed relationships for the stress at interior, edge and corner regions, denoted as σ_i, σ_e, σ_c in kg/cm^2 respectively and given by the equation 29.3-29.5.

$$\sigma_i = \frac{0.316\ P}{h^2}\left[4\ \log_{10}\left(\frac{l}{b}\right) + 1.069\right] \tag{29.3}$$

$$\sigma_e = \frac{0.572\ P}{h^2}\left[4\ \log_{10}\left(\frac{l}{b}\right) + 0.359\right] \tag{29.4}$$

$$\sigma_c = \frac{3\ P}{h^2}\left[1 - \left(\frac{a\sqrt{2}}{l}\right)^{0.6}\right] \tag{29.5}$$

where h is the slab thickness in cm, P is the wheel load in kg, a is the radius of the wheel load distribution in cm, l the radius of the relative stiffness in cm and b is the radius of the resisting section in cm

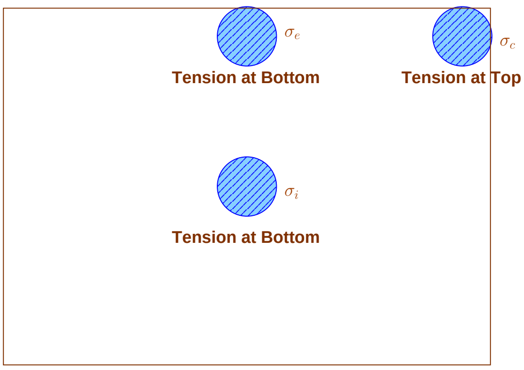


Figure 29:1: Critical stress locations

29.3 Temperature stresses

Temperature stresses are developed in cement concrete pavement due to variation in slab temperature. This is caused by (i) *daily variation* resulting in a temperature gradient across the thickness of the slab and (ii) *seasonal variation* resulting in overall change in the slab temperature. The former results in **warping** stresses and the later in **frictional** stresses.

29.3.1 Warping stress

The warping stress at the interior, edge and corner regions, denoted as σ_{t_i}, σ_{t_e}, σ_{t_c} in kg/cm^2 respectively and given by the equation 29.7-29.8.

$$\sigma_{t_i} = \frac{E\epsilon t}{2}\left(\frac{C_x + \mu C_y}{1-\mu^2}\right) \tag{29.6}$$

$$\sigma_{t_e} = \text{Max}\left(\frac{\mathrm{C_x E\epsilon t}}{2}, \frac{\mathrm{C_y E\epsilon t}}{2}\right) \tag{29.7}$$

$$\sigma_{t_c} = \frac{E\epsilon t}{3(1-\mu)}\sqrt{\frac{a}{l}} \tag{29.8}$$

where E is the modulus of elasticity of concrete in kg/cm^2 ($3{\times}10^5$), ϵ is the thermal coefficient of concrete per ^{o}C ($1{\times}10^{-7}$) t is the temperature difference between the top and bottom of the slab, C_x and C_y are the coefficient based on L_x/l in the desired direction and L_y/l right angle to the desired direction, μ is the Poisson's ration (0.15), a is the radius of the contact area and l is the radius of the relative stiffness.

29.3.2 Frictional stresses

The frictional stress σ_f in kg/cm^2 is given by the equation

$$\sigma_f = \frac{WLf}{2 \times 10^4} \tag{29.9}$$

where W is the unit weight of concrete in kg/cm^2 (2400), f is the coefficient of sub grade friction (1.5) and L is the length of the slab in meters.

29.4 Combination of stresses

The cumulative effect of the different stress give rise to the following thee critical cases

- *Summer, mid-day*: The critical stress is for edge region given by $\sigma_{critical} = \sigma_e + \sigma_{t_e} - \sigma_f$
- *Winter, mid-day*: The critical combination of stress is for the edge region given by $\sigma_{critical} = \sigma_e + \sigma_{t_e} + \sigma_f$
- *Mid-nights*: The critical combination of stress is for the corner region given by $\sigma_{critical} = \sigma_c + \sigma_{t_c}$

29.5 Design of joints

29.5.1 Expansion joints

The purpose of the expansion joint is to allow the expansion of the pavement due to rise in temperature with respect to construction temperature. The design consideration are:

- Provided along the longitudinal direction,
- design involves finding the joint spacing for a given expansion joint thickness (say 2.5 cm specified by IRC) subjected to some maximum spacing (say 140 as per IRC)

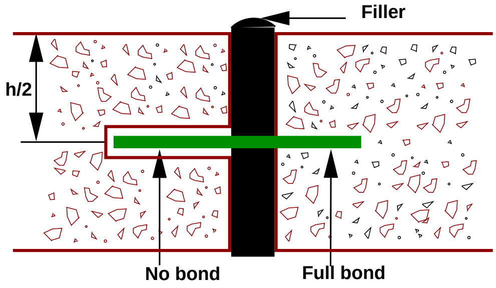


Figure 29:2: Expansion joint

29.5.2 Contraction joints

The purpose of the contraction joint is to allow the contraction of the slab due to fall in slab temperature below the construction temperature. The design considerations are:

- The movement is restricted by the sub-grade friction
- Design involves the length of the slab given by:

$$L_c = \frac{2 \times 10^4 S_c}{W.f} \tag{29.10}$$

  where, S_c is the allowable stress in tension in cement concrete and is taken as 0.8 kg/cm^2, W is the unit weight of the concrete which can be taken as 2400 kg/cm^3 and f is the coefficient of sub-grade friction which can be taken as 1.5.

- Steel reinforcements can be use, however with a maximum spacing of 4.5 m as per IRC.

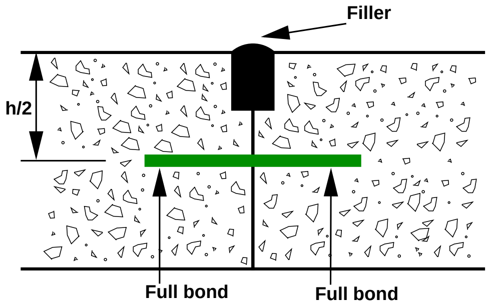


Figure 29:3: Contraction joint

29.5.3 Dowel bars

The purpose of the dowel bar is to effectively transfer the load between two concrete slabs and to keep the two slabs in same height. The dowel bars are provided in the direction of the traffic (longitudinal). The design considerations are:

- Mild steel rounded bars,
- bonded on one side and free on other side

Bradbury's analysis

Bradbury's analysis gives load transfer capacity of single dowel bar in shear, bending and bearing as follows:

$$P_s = 0.785\ d^2\ F_s \tag{29.11}$$

$$P_f = \frac{2\ d^3\ F_f}{L_d + 8.8\delta} \tag{29.12}$$

$$P_b = \frac{F_b\ L_d^2\ d}{12.5\ (L_d + 1.5\delta)} \tag{29.13}$$

where, P is the load transfer capacity of a single dowel bar in shear s, bending f and bearing b, d is the diameter of the bar in cm, L_d is the length of the embedment of dowel bar in cm, δ is the joint width in cm, $F_s,\ Ff,\ F_b$ are the permissible stress in shear, bending and bearing for the dowel bar in kg/cm^2.

Design procedure

Step 1 Find the length of the dowel bar embedded in slab L_d by equating Eq. 29.12=Eq. 29.13, i.e.

$$L_d = 5d\sqrt{\frac{F_f}{F_b}\frac{(L_d + 1.5\delta)}{(L_d + 8.8\delta)}} \tag{29.14}$$

Step 2 Find the load transfer capacities P_s, P_f, and P_b of single dowel bar with the L_d

Step 3 Assume load capacity of dowel bar is 40 percent wheel load, find the load capacity factor f as

$$\max\left\{\frac{0.4P}{P_s},\ \frac{0.4P}{P_f},\ \frac{0.4P}{P_b}\right\} \tag{29.15}$$

Step 4 Spacing of the dowel bars.

- Effective distance upto which effective load transfer take place is given by 1.8 l, where l is the radius of relative stiffness.
- Assume a linear variation of capacity factor of 1.0 under load to 0 at 1.8 l.
- Assume a dowel spacing and find the capacity factor of the above spacing.
- Actual capacity factor should be greater than the required capacity factor.
- If not, do one more iteration with new spacing.

Example

Design size and spacing of dowel bars at an expansion joint of concrete pavement of thickness 25 cm. Given the radius of relative stiffness of 80 cm. design wheel load 5000 kg. Load capacity of the dowel system is 40 percent of design wheel load. Joint width is 2.0 cm and the permissible stress in shear, bending and bearing stress in dowel bars are 1000,1400 and 100 kg/cm^2 respectively.

Solution: Given, $P = 5000\ kg$, $l = 80\ cm$, $h = 25\ cm$, $\delta = 2\ cm$, $F_s = 1000\ kg/cm^2$, $F_f = 1400\ kg/cm^2$ and $F_b = 100\ kg/cm^2$; and assume $d = 2.5\ cm$ diameter.
Step-1: length of the dowel bar L_d

$$\begin{aligned} L_d &= 5 \times 2.5 \sqrt{\frac{1400}{100}\frac{(L_d + 1.5 \times 2)}{(L_d + 8.8 \times 2)}} \\ &= 12.5 \times \sqrt{14\ \frac{(L_d + 3)}{(L_d + 17.6)}} \end{aligned}$$

Solve for L_d by trial and error:
put $L_d = 45.00 \ \Rightarrow L_d = 40.95$
put $L_d = 45.95 \ \Rightarrow L_d = 40.50$
put $L_d = 45.50 \ \Rightarrow L_d = 40.50$
Minimum length of the dowel bar is $L_d + \delta \ = \ 40.5 + 2.0 \ = \ 42.5\ cm$, So, provide 45 cm long and 2.5 cm ϕ.
Therefore $L_d = 45 - 2 = 43\ cm$.
Step 2: Find the load transfer capacity of single dowel bar

$$\begin{aligned} P_s = &\quad 0.785 \times 2.5^2 \times 1000 &= 4906\ kg \\ P_f = &\quad \tfrac{2\times 2.5^3 \times 1400}{43.0+8.8\times 2} &= 722\ kg \\ P_b = &\quad \tfrac{100\times 2.5\times 43.0^2}{12.5\ (43.0+1.5\times 2)} &= 804\ kg \end{aligned}$$

Therefore, the required load transfer capacity

$$\max\left\{\frac{0.4 \times 5000}{4906}, \frac{0.4 \times 5000}{722}, \frac{0.4 \times 5000}{804}\right\}$$
$$\max\{0.41, 2.77, 2.487\} = 2.77$$

Step-3 : Find the required spacing: Effective distance of load transfer $= 1.8\ l \ = \ 1.8 \times 80 \ = \ 144\ cm$. Assuming 35 cm spacing,

Actual capacity is

$$1 + \frac{144 - 35}{144} + \frac{144 - 70}{144} + \frac{144 - 105}{144} + \frac{144 - 140}{144}$$
$$= 2.57 \ < \ 2.77 \quad \text{(the required capacity)}$$

Therefore assume 30 cm spacing and now the actual capacity is

$$1 + \frac{144 - 30}{144} + \frac{144 - 60}{144} + \frac{144 - 90}{144} + \frac{144 - 120}{144}$$
$$= 2.92 \ > \ 2.77 \quad \text{(the required capacity)}$$

Therefore provide 2.5 cm ϕ mild steel dowel bars of length 45 cm @ 30 cm center to center.

29.5.4 Tie bars

In contrast to dowel bars, tie bars are not load transfer devices, but serve as a means to tie two slabs. Hence tie bars must be deformed or hooked and must be firmly anchored into the concrete to function properly. They are smaller than dowel bars and placed at large intervals. They are provided across longitudinal joints.

Step 1 Diameter and spacing: The diameter and the spacing is first found out by equating the total sub-grade friction tot he total tensile stress for a unit length (one meter). Hence the area of steel per one meter in cm^2 is given by:

$$A_s \times S_s = b \times h \times W \times f$$
$$A_s = \frac{bhWf}{100 S_s} \tag{29.16}$$

where, b is the width of the pavement panel in m, h is the depth of the pavement in cm, W is the unit weight of the concrete (assume 2400 kg/cm^2), f is the coefficient of friction (assume 1.5), and S_s is the allowable working tensile stress in steel (assume 1750 kg/cm^2). Assume 0.8 to 1.5 cm ϕ bars for the design.

Step 2 Length of the tie bar: Length of the tie bar is twice the length needed to develop bond stress equal to the working tensile stress and is given by:

$$L_t = \frac{d\ S_s}{2\ S_b} \tag{29.17}$$

where, d is the diameter of the bar, S_s is the allowable tensile stress in kg/cm^2, and S_b is the allowable bond stress and can be assumed for plain and deformed bars respectively as 17.5 and 24.6 kg/cm^2.

Example

A cement concrete pavement of thickness 18 cm, has two lanes of 7.2 m with a joint. Design the tie bars.

(Solution:)
Given h=18 cm, b=7.2/2=3.6m, $S_s = 1700\ kg/cm^2$ $W = 2400\ kg/cm^2$ $f = 1.5$ $S_b = 24.6\ kg/cm^2$.

Step 1: diameter and spacing: Get A_s from

$$A_s = \frac{3.6 \times 18 \times 2400 \times 1.5}{100 \times 1750} = 1.33\ cm^2/m$$

Assume $\phi = 1\ cm$, $\Rightarrow$ $A = 0.785\ cm^2$. Therefore spacing is $\frac{100 \times 0.785}{1.33} = 59\ cm$, say 55 cm

Step 2: Length of the bar: Get L_t from

$$L_t = \frac{1 \times 1750}{2\ 246} \ = \ 36.0\ cm$$

[Ans] Use 1 cm ϕ tie bars of length of 36 cmi @ 55 cm c/c

29.6 Summary

Design of rigid pavements is based on Westergaard's analysis, where modulus of subgrade reaction, radius of relative stiffness, radius of wheel load distribution are used. For critical design, a combination of load stress, frictional stress and warping stress is considered. Different types of joints are required like expansion and contraction joints. Their design is also dealt with.

29.7 Problems

1. Design size and spacing of dowel bars at an expansion joint of concrete pavement of thickness 20 cm. Given the radius of relative stiffness of 90 cm. design wheel load 4000 kg. Load capacity of the dowel system is 40 percent of design wheel load. Joint width is 3.0 cm and the permissible stress in shear, bending and bearing stress in dowel bars are 1000,1500 and 100 kg/cm^2 respectively.

2. Design the length and spacing of tie bars given that the pavement thickness is 20cm and width of the road is 7m with one longitudinal joint. The unit weight of concrete is 2400 kg/m^3, the coefficient of friction is 1.5, allowable working tensile stress in steel is 1750 kg/cm^2, and bond stress of deformed bars is 24.6 kg/cm^2.

29.8 Solutions

1. Given, $P = 4000\ kg$, $l = 90\ cm$, $h = 20\ cm$, $\delta = 3\ cm$, $F_s = 1000\ kg/cm^2$, $F_f = 1500\ kg/cm^2$ and $F_b = 100\ kg/cm^2$; and assume $d = 2.5\ cm$ diameter.

   Step-1: length of the dowel bar L_d,

$$\begin{aligned} L_d &= 5 \times 2.5 \sqrt{\frac{1500}{100} \frac{(L_d + 1.5 \times 3)}{(L_d + 8.8 \times 3)}} \\ &= 12.5 \times \sqrt{15 \frac{(L_d + 4.5)}{(L_d + 26.4)}} \end{aligned}$$

   Solving for L_d by trial and error, it is =39.5cm Minimum length of the dowel bar is $L_d + \delta = 39.5 + 3.0 = 42.5\ cm$, So, provide 45 cm long and 2.5 cm ϕ. Therefore $L_d = 45 - 3 = 42\ cm$.

   Step 2: Find the load transfer capacity of single dowel bar

$$\begin{aligned} P_s = & \quad 0.785 \times 2.5^2 \times 1000 &= 4906.25\ kg \\ P_f = & \quad \tfrac{2 \times 2.5^3 \times 1500}{42.0 + 8.8 \times 3} &= 685.307\ kg \\ P_b = & \quad \tfrac{100 \times 2.5 \times 42.0^2}{12.5\ (42.0 + 1.5 \times 3)} &= 758.71\ kg \end{aligned}$$

   Therefore, the required load transfer capacity (refer equation)

$$\max\left\{\frac{0.4 \times 4000}{4906.25}, \frac{0.4 \times 4000}{685.307}, \frac{0.4 \times 4000}{758.71}\right\}$$
$$\max\{0.326, 2.335, 2.10\} = 2.335$$

   Step-3 : Find the required spacing: Effective distance of load transfer = $1.8 \times l = 1.8 \times 90 = 162\ cm$. Assuming 35 cm spacing,

   Actual capacity is

$$1 + \frac{162 - 35}{162} + \frac{162 - 70}{162} + \frac{162 - 105}{162} + \frac{162 - 140}{162}$$
$$= 2.83$$

Assuming 40cm spacing, capacity is,

$$1 + \frac{162 - 40}{162} + \frac{162 - 80}{162} + \frac{162 - 120}{162} + \frac{162 - 160}{162} = 2.52$$

So we should consider 2.52¿2.335 as it is greater and more near to other value. Therefore provide 2.5 cm ϕ mild steel dowel bars of length 45 cm @ 40 cm center to center.

2. 2.Given h=20 cm, b=7/2=3.5m, $S_s = 1750\ kg/cm^2$ $W = 2400\ kg/cm^2$ $f = 1.5$ $S_b = 24.6\ kg/cm^2$.

   Step 1: diameter and spacing:

   $$A_s = \frac{3.5 \times 20 \times 2400 \times 1.5}{100 \times 1750} = 1.44\ cm^2/m$$

   Assume $\phi = 1\ cm,\ \Rightarrow\ A = 0.785\ cm^2$. Therefore spacing is $\frac{100 \times 0.785}{1.44} = 54.57\ cm$, say 55 cm

   Step 2: Length of the bar:

   $$L_t = \frac{1 \times 1750}{2 \times 24.6} = 36.0\ cm$$

   [Ans] Use 1 cm ϕ tie bars of length of 36 cm @ 55 cm c/c

Chapter 30

Fundamental parameters of traffic flow

30.1 Overview

Traffic engineering pertains to the analysis of the behavior of traffic and to design the facilities for a smooth, safe and economical operation of traffic. Traffic flow, like the flow of water, has several parameters associated with it. The traffic stream parameters provide information regarding the nature of traffic flow, which helps the analyst in detecting any variation in flow characteristicis. Understanding traffic behavior requires a thorough knowledge of traffic stream parameters and their mutual relationships. In this chapter the basic concepts of traffic flow is presented.

30.2 Traffic stream parameters

The traffic stream includes a combination of driver and vehicle behavior. The driver or human behavior being non-uniform, traffic stream is also non-uniform in nature. It is influenced not only by the individual characteristics of both vehicle and human but also by the way a group of such units interacts with each other. Thus a flow of traffic through a street of defined characteristics will vary both by location and time corresponding to the changes in the human behavior.

The traffic engineer, but for the purpose of planning and design, assumes that these changes are within certain ranges which can be predicted. For example, if the maximum permissible speed of a highway is 60 kmph, the whole traffic stream can be assumed to move on an average speed of 40 kmph rather than 100 or 20 kmph.

Thus the traffic stream itself is having some parameters on which the characteristics can be predicted. The parameters can be mainly classified as : measurements of quantity, which includes density and flow of traffic and measurements of quality which includes speed. The traffic stream parameters can be macroscopic which characterizes the traffic as a whole or microscopic which studies the behavior of individual vehicle in the stream with respect to each other.

As far as the macroscopic characteristics are concerned, they can be grouped as measurement of quantity or quality as described above, i.e. flow, density, and speed. While the microscopic characteristics include the measures of separation, i.e. the headway or separation between vehicles which can be either time or space headway. The fundamental stream characteristics are speed, flow, and density and are discussed below.

30.3 Speed

Speed is considered as a quality measurement of travel as the drivers and passengers will be concerned more about the speed of the journey than the design aspects of the traffic. It is defined as the rate of motion in distance per unit of time. Mathematically speed or velocity v is given by,

$$v = \frac{d}{t} \tag{30.1}$$

where, v is the speed of the vehicle in m/s, d is distance traveled in m in time t seconds. Speed of different vehicles will vary with respect to time and space. To represent these variation, several types of speed can be defined. Important among them are spot speed, running speed, journey speed, time mean speed and space mean speed. These are discussed below.

30.3.1 Spot Speed

Spot speed is the instantaneous speed of a vehicle at a specified location. Spot speed can be used to design the geometry of road like horizontal and vertical curves, super elevation etc. Location and size of signs, design of signals, safe speed, and speed zone determination, require the spot speed data. Accident analysis, road maintenance, and congestion are the modern fields of traffic engineer, which uses spot speed data as the basic input. Spot speed can be measured using an enoscope, pressure contact tubes or direct timing procedure or radar speedometer or by time-lapse photographic methods. It can be determined by speeds extracted from video images by recording the distance traveling by all vehicles between a particular pair of frames.

30.3.2 Running speed

Running speed is the average speed maintained over a particular course while the vehicle is moving and is found by dividing the length of the course by the time duration the vehicle was in motion. i.e. this speed doesn't consider the time during which the vehicle is brought to a stop, or has to wait till it has a clear road ahead. The running speed will always be more than or equal to the journey speed, as delays are not considered in calculating the running speed

30.3.3 Journey speed

Journey speed is the effective speed of the vehicle on a journey between two points and is the distance between the two points divided by the total time taken for the vehicle to complete the journey including any stopped time. If the journey speed is less than running speed, it indicates that the journey follows a stop-go condition with enforced acceleration and deceleration. The spot speed here may vary from zero to some maximum in excess of the running speed. A uniformity between journey and running speeds denotes comfortable travel conditions.

30.3.4 Time mean speed and space mean speed

Time mean speed is defined as the average speed of all the vehicles passing a point on a highway over some specified time period. Space mean speed is defined as the average speed of all the vehicles occupying a given section of a highway over some specified time period. Both mean speeds will always be different from each other except in the unlikely event that all vehicles are traveling at the same speed. Time mean speed is a point measurement while space mean speed is a measure relating to length of highway or lane, i.e. the mean speed of vehicles over a period of time at a point in space is time mean speed and the mean speed over a space at a given instant is the space mean speed.

30.4 Flow

There are practically two ways of counting the number of vehicles on a road. One is flow or volume, which is defined as the number of vehicles that pass a point on a highway or a given lane or direction of a highway during a specific time interval. The measurement is carried out by counting the number of vehicles, n_t, passing a particular point in one lane in a defined period t. Then the flow q expressed in vehicles/hour is given by

$$q = \frac{n_t}{t} \tag{30.2}$$

Flow is expressed in planning and design field taking a day as the measurement of time.

30.4.1 Variations of Volume

The variation of volume with time, i.e. month to month, day to day, hour to hour and within a hour is also as important as volume calculation. Volume variations can also be observed from season to season. Volume will be above average in a pleasant motoring month of summer, but will be more pronounced in rural than in urban area. But this is the most consistent of all the variations and affects the traffic stream characteristics the least.

Weekdays, Saturdays and Sundays will also face difference in pattern. But comparing day with day, patterns for routes of a similar nature often show a marked similarity, which is useful in enabling predictions to be made.

The most significant variation is from hour to hour. The peak hour observed during mornings and evenings of weekdays, which is usually 8 to 10 per cent of total daily flow or 2 to 3 times the average hourly volume. These trips are mainly the work trips, which are relatively stable with time and more or less constant from day to day.

30.4.2 Types of volume measurements

Since there is considerable variation in the volume of traffic, several types of measurements of volume are commonly adopted which will average these variations into a single volume count to be used in many design purposes.

1. **Average Annual Daily Traffic(AADT)** : The average 24-hour traffic volume at a given location over a full 365-day year, i.e. the total number of vehicles passing the site in a year divided by 365.

2. **Average Annual Weekday Traffic(AAWT)** : The average 24-hour traffic volume occurring on weekdays over a full year. It is computed by dividing the total weekday traffic volume for the year by 260.

3. **Average Daily Traffic(ADT)** : An average 24-hour traffic volume at a given location for some period of time less than a year. It may be measured for six months, a season, a month, a week, or as little as two days. An ADT is a valid number only for the period over which it was measured.

4. **Average Weekday Traffic(AWT)** : An average 24-hour traffic volume occurring on weekdays for some period of time less than one year, such as for a month or a season.

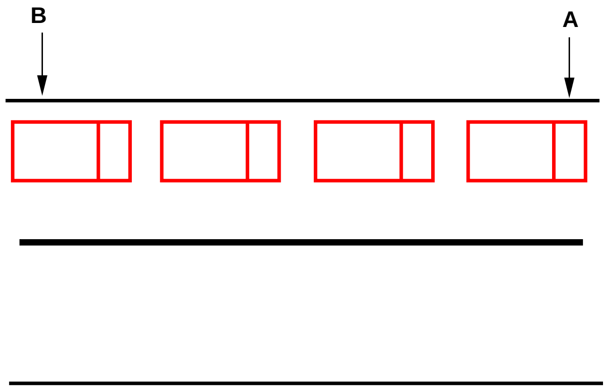


Figure 30:1: Illustration of density

The relationship between AAWT and AWT is analogous to that between AADT and ADT. Volume in general is measured using different ways like manual counting, detector/sensor counting, moving-car observer method, etc. Mainly the volume study establishes the importance of a particular route with respect to the other routes, the distribution of traffic on road, and the fluctuations in flow. All which eventually determines the design of a highway and the related facilities. Thus, volume is treated as the most important of all the parameters of traffic stream.

30.5 Density

Density is defined as the number of vehicles occupying a given length of highway or lane and is generally expressed as vehicles per km. One can photograph a length of road x, count the number of vehicles, n_x, in one lane of the road at that point of time and derive the density k as,

$$k = \frac{n_x}{x} \tag{30.3}$$

This is illustrated in figure 30:1. From the figure, the density is the number of vehicles between the point A and B divided by the distance between A and B. Density is also equally important as flow but from a different angle as it is the measure most directly related to traffic demand. Again it measures the proximity of vehicles in the stream which in turn affects the freedom to maneuver and comfortable driving.

30.6 Derived characteristics

From the fundamental traffic flow characteristics like flow, density, and speed, a few other parameters of traffic flow can be derived. Significant among them are the time headway, distance headway and travel time. They are discussed one by one below.

30.6.1 Time headway

The microscopic character related to volume is the time headway or simply headway. Time headway is defined as the time difference between any two successive vehicles when they cross a given point. Practically, it involves the measurement of time between the passage of one rear bumper and the next past a given point. If all headways h in time period, t, over which flow has been measured are added then,

$$\sum_{1}^{n_t} h_i = t \tag{30.4}$$

But the flow is defined as the number of vehicles n_t measured in time interval t, that is,

$$q = \frac{n_t}{t} = \frac{n_t}{\sum_1^{n_t} h_i} = \frac{1}{h_{av}} \tag{30.5}$$

where, h_{av} is the average headway. Thus average headway is the inverse of flow. Time headway is often referred to as simply the headway.

30.6.2 Distance headway

Another related parameter is the distance headway. It is defined as the distance between corresponding points of two successive vehicles at any given time. It involves the measurement from a photograph, the distance from rear bumper of lead vehicle to rear bumper of following vehicle at a point of time. If all th space headways in distance x over which the density has been measured are added,

$$\sum_{1}^{n_x} s_i = x \tag{30.6}$$

But the density (k) is the number of vehicles n_x at a distance of x, that is

$$k = \frac{n_x}{x} = \frac{n_x}{\sum_1^{n_x} s_i} = \frac{1}{s_{av}} \tag{30.7}$$

Where, s_{av} is average distance headway. The average distance headway is the inverse of density and is sometimes called as spacing.

30.6.3 Travel time

Travel time is defined as the time taken to complete a journey. As the speed increases, travel time required to reach the destination also decreases and viceversa. Thus travel time is inversely proportional to the speed. However, in practice, the speed of a vehicle fluctuates over time and the travel time represents an average measure.

30.7 Time-space diagram

Time space diagram is a convenient tool in understanding the movement of vehicles. It shows the trajectory of vehicles in the form of a two dimensional plot. Time space diagram can be plotted for a single vehicle as well as multiple vehicles. They are discussed below.

30.7.1 Single vehicle

Taking one vehicle at a time, analysis can be carried out on the position of the vehicle with respect to time. This analysis will generate a graph which gives the relation of its position on a road stretch relative to time.

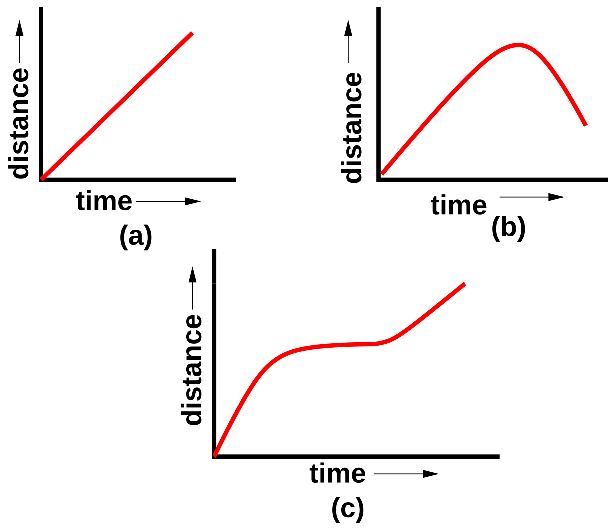


Figure 30:2: Time space diagram for a single vehicle

This plot thus will be between distance x and time t and x will be a functions the position of the vehicle for every t along the road stretch. This graphical representation of $x(t)$ in a (t, x) plane is a curve which is called as a trajectory. The trajectory provide an intuitive, clear, and complete summary of vehicular motion in one dimension.

In figure 30:2(a), the the distance x goes on increasing with respect to the origin as time progresses. The vehicle is moving at a smooth condition along the road way. In figure 30:2(b), the vehicle at first moves with a smooth pace after reaching a position reverses its direction of movement. In figure 30:2(c), the vehicle in between becomes stationary and maintains the same position.

From the figure, steeply increasing section of $x(t)$ denote a rapidly advancing vehicle and horizontal portions of $x(t)$ denote a stopped vehicle while shallow sections show a slow-moving vehicle. A straight line denotes constant speed motion and curving sections denote accelerated motion; and if the curve is concave downwards it denotes acceleration. But a curve which is convex upwards denotes deceleration.

30.7.2 Miultiple Vehicles

Time-space diagram can also be used to determine the fundamental parameters of traffic flow like speed, density and volume. It can also be used to find the derived characteristics like space headway and time headway. Figure 30:3 shows the time-space diagram for a set of vehicles traveling at constant speed. Density, by definition is the number of vehicles per unit length. From the figure, an observer looking into the stream can count 4 vehicles passing the stretch of road between x_1 and x_2 at time t. Hence, the density is given as

$$k = \frac{4 \ vehicles}{x_2 - x_1} \tag{30.8}$$

We can also find volume from this time-space diagram. As per the definition, volume is the number of vehicles counted for a particular interval of time. From the figure 30:3 we can see that 6 vehicles are present between the time t_1 and t_2. Therefore, the volume q is given as

$$q = \frac{3 \ vehicles}{t_2 - t_1} \tag{30.9}$$

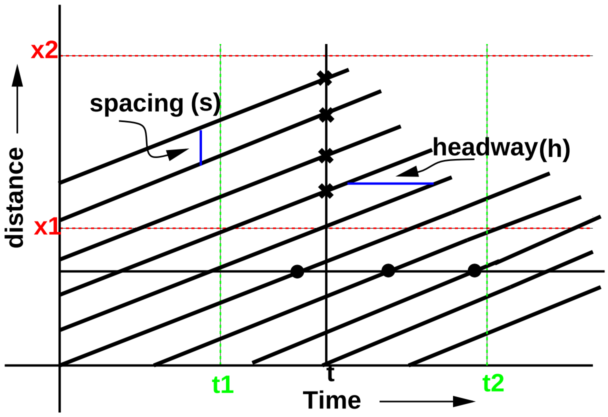


Figure 30:3: Time space diagram for many vehicles

Again the averages taken at a specific location (i.e., time ranging over an interval) are called time means and those taken at an instant over a space interval are termed as space means.

Another related definition which can be given based on the time-space diagram is the headway. Space headway is defined as the distance between corresponding points of two successive vehicles at any given time. Thus, the vertical gap between any two consecutive lines represents space headway. The reciprocal of density otherwise gives the space headway between vehicles at that time.

Similarly, time headway is defined as the time difference between any two successive vehicles when they cross a given point. Thus, the horizontal gap between the vehicles represented by the lines gives the time headway. The reciprocal of flow gives the average time headway between vehicles at that point.

30.8 Summary

Speed, flow and density are the basic parameters of traffic flow. Different measures of speed are used in traffic flow analysis like spot speed, time mean speed, space mean speed etc. Time-space diagram also can be used for determining these parameters. Speed and flow of the traffic stream can be computed using moving observer method.

30.9 Problems

1. The instantaneous speed of a vehicle at a specified location is called
   (a) Spot speed
   (b) Journey speed
   (c) Running speed
   (d) Time mean speed
2. Which of the following is not a derived characteristic?
   (a) Time headway

(b) Distance headway

(c) Travel time

(d) Density

30.10 Solutions

1. The instantaneous speed of a vehicle at a specified location is called
   (a) Spot speed$\surd$
   (b) Journey speed
   (c) Running speed
   (d) Time mean speed
2. Which of the following is not a derived characteristic?
   (a) Time headway
   (b) Distance headway
   (c) Travel time
   (d) Density$\surd$

Chapter 31

Fundamental relations of traffic flow

31.1 Overview

Speed is one of the basic parameters of traffic flow and time mean speed and space mean speed are the two representations of speed. Time mean speed and space mean speed and the relationship between them will be discussed in detail in this chapter. The relationship between the fundamental parameters of traffic flow will also be derived. In addition, this relationship can be represented in graphical form resulting in the fundamental diagrams of traffic flow.

31.2 Time mean speed (v_t)

As noted earlier, time mean speed is the average of all vehicles passing a point over a duration of time. It is the simple average of spot speed. Time mean speed v_t is given by,

$$v_t = \frac{1}{n}\sum_{i=1}^{n} v_i, \tag{31.1}$$

where v is the spot speed of i^{th} vehicle, and n is the number of observations. In many speed studies, speeds are represented in the form of frequency table. Then the time mean speed is given by,

$$v_t = \frac{\sum_{i=1}^{n} q_i v_i}{\sum_{i=1}^{n} q_i}, \tag{31.2}$$

where q_i is the number of vehicles having speed v_i, and n is the number of such speed categories.

31.3 Space mean speed (v_s)

The space mean speed also averages the spot speed, but spatial weightage is given instead of temporal. This is derived as below. Consider unit length of a road, and let v_i is the spot speed of i_{th} vehicle. Let t_i is the time the vehicle takes to complete unit distance and is given by $\frac{1}{v_i}$. If there are n such vehicles, then the average travel time t_s is given by,

$$t_s = \frac{\Sigma t_i}{n} = \frac{1}{n}\Sigma\frac{1}{v_i} \tag{31.3}$$

If t_{av} is the average travel time, then average speed $v_s = \frac{1}{t_s}$. Therefore from the above equation,

$$v_s = \frac{n}{\sum_{i=1}^{n} \frac{1}{v_i}} \tag{31.4}$$

No.	speed range	average speed (v_i)	volume of flow (q_i)	$v_i q_i$	$\frac{q_i}{v_i}$
1	2-5	3.5	1	3.5	2.29
2	6-9	7.5	4	30.0	0.54
3	10-13	11.5	0	0	0
4	14-17	15.5	7	108.5	0.45
	total		12	142	3.28

This is simply the harmonic mean of the spot speed. If the spot speeds are expressed as a frequency table, then,

$$v_s = \frac{\sum_{i=1}^{n} q_i}{\sum_{i=1}^{n} \frac{q_i}{v_i}} \tag{31.5}$$

where q_i vehicle will have v_i speed and n_i is the number of such observations.

Example 1

If the spot speeds are 50, 40, 60,54 and 45, then find the time mean speed and space mean speed.

Solution Time mean speed v_t is the average of spot speed. Therefore, $v_t = \frac{\Sigma v_i}{n} = \frac{50+40+60+54+45}{5} = \frac{249}{5} = 49.8$ Space mean speed is the harmonic mean of spot speed.Therefore, $v_s = \frac{n}{\Sigma \frac{1}{v_i}} = \frac{5}{\frac{1}{50}+\frac{1}{40}+\frac{1}{60}+\frac{1}{54}+\frac{1}{45}} = \frac{5}{0.12} = 48.82$

Example 2

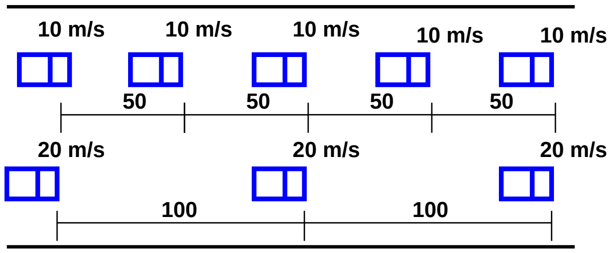
The results of a speed study is given in the form of a frequency distribution table. Find the time mean speed and space mean speed.

speed range	frequency
2-5	1
6-9	4
10-13	0
14-17	7

Solution The time mean speed and space mean speed can be found out from the frequency table given below. First, the average speed is computed, which is the mean of the speed range. For example, for the first speed range, average speed, $v_i = \frac{2+5}{2} = 3.5$ seconds. The volume of flow q_i for that speed range is same as the frequency. The terms $v_i.q_i$ and $\frac{q_i}{v_i}$ are also tabulated, and their summations in the last row. Time mean speed can be computed as, $v_t = \frac{\Sigma q_i v_i}{\Sigma q_i} = \frac{142}{12} = 11.83$ Similarly, space mean speed can be computed as, $v_s = \frac{\Sigma q_i}{\Sigma \frac{q_i}{v_i}} = \frac{12}{3.28} = 3.65$

31.4 Illustration of mean speeds

Inorder to understand the concept of time mean speed and space mean speed, following illustration will help. Let there be a road stretch having two sets of vehicle as in figure 31:1. The first vehicle is traveling at 10m/s with 50 m spacing, and the second set at 20m/s with 100 m spacing. Therefore, the headway of the slow vehicle



$h_s = 50/20 = 5sec$ $n_s = 60/5 = 12$ $k_s = 1000/50 = 20$

$h_f = 100/20 = 5sec$ $n_f = 60/5 = 12$ $k_f = 1000/100 = 10$

Figure 31:1: Illustration of relation between time mean speed and space mean speed

h_s will be 50 m divided by 10 m/s which is 5 sec. Therefore, the number of slow moving vehicles observed at A in one hour n_s will be $60/5 = 12$ vehicles. The density K is the number of vehicles in 1 km, and is the inverse of spacing. Therefore, $K_s = 1000/50 = 20$ vehicles/km. Therefore, by definition, time mean speed v_t is given by $v_t = \frac{12\times10+12\times20}{24} = 15$ m/s. Similarly, by definition, space mean speed is the mean of vehicle speeds over time. Therefore, $v_s = \frac{20\times10+10\times20}{30} = 13.3$ m/s This is same as the harmonic mean of spot speeds obtained at location A; ie $v_s = \frac{24}{12\times\frac{1}{10}+12\times\frac{1}{20}} = 13.3$ m/s. It may be noted that since harmonic mean is always lower than the arithmetic mean, and also as observed , space mean speed is always lower than the time mean speed. In other words, space mean speed weights slower vehicles more heavily as they occupy the road stretch for longer duration of time. For this reason, in many fundamental traffic equations, space mean speed is preferred over time mean speed.

31.5 Relation between time mean speed and space mean speed

The relation between time mean speed and space mean speed can be derived as below. Consider a stream of vehicles with a set of substream flow q_1,q_2, $\ldots q_i$, $\ldots q_n$ having speed v_1,v_2, $\ldots v_i$, $\ldots v_n$. The fundamental relation between flow(q), density(k) and mean speed v_s is,

$$q = k \times v_s \tag{31.6}$$

Therefore for any substream q_i, the following relationship will be valid.

$$q_i = k_i \times v_i \tag{31.7}$$

The summation of all substream flows will give the total flow q.

$$\Sigma q_i = q \tag{31.8}$$

Similarly the summation of all substream density will give the total density k.

$$\Sigma k_i = k \tag{31.9}$$

Let f_i denote the proportion of substream density k_i to the total density k,

$$f_i = \frac{k_i}{k} \tag{31.10}$$

Space mean speed averages the speed over space. Therefore, if k_i vehicles has v_i speed, then space mean speed is given by,

$$v_s = \frac{\Sigma k_i v_i}{k} \tag{31.11}$$

Time mean speed averages the speed over time.Therefore,

$$v_t = \frac{\Sigma q_i v_i}{q} \tag{31.12}$$

Substituting in 31.7 v_t can be written as,

$$v_t = \frac{\Sigma k_i {v_i}^2}{q} \tag{31.13}$$

Rewriting the above equation and substituting 31.11, and then substituting 31.6, we get,

$$\begin{aligned} v_t &= k\Sigma \frac{k_i}{k} v_i^2 \\ &= \frac{k\Sigma f_i {v_i}^2}{q} \\ &= \frac{\Sigma f_i {v_i}^2}{v_s} \end{aligned}$$

By adding and subtracting v_s and doing algebraic manipulations, v_t can be written as,

$$v_t = \frac{\Sigma f_i (v_s + (v_i - v_s))^2}{v_s} \tag{31.14}$$

$$= \frac{\Sigma f_i (v_s)^2 + (v_i - v_s)^2 + 2.v_s.(v_i - v_s)}{v_s} \tag{31.15}$$

$$= \frac{\Sigma f_i {v_s}^2}{v_s} + \frac{\Sigma f_i (v_i - v_s)^2}{v_s} + \frac{2.v_s.\Sigma f_i (v_i - v_s)}{v_s} \tag{31.16}$$

The third term of the equation will be zero because $\Sigma f_i(v_i - v_s)$ will be zero, since v_s is the mean speed of v_i. The numerator of the second term gives the standard deviation of v_i. Σf_i by definition is 1.Therefore,

$$= v_s \Sigma f_i + \frac{\sigma^2}{v_s} + 0 \tag{31.17}$$

$$v_t = v_s + \frac{\sigma^2}{v_s} \tag{31.18}$$

Hence, time mean speed is space mean speed plus standard deviation of the spot speed divided by the space mean speed. Time mean speed will be always greater than space mean speed since standard deviation cannot be negative. If all the speed of the vehicles are the same, then spot speed, time mean speed and space mean speed will also be same.

31.6 Fundamental relations of traffic flow

The relationship between the fundamental variables of traffic flow, namely speed, volume, and density is called the fundamental relations of traffic flow. This can be derived by a simple concept. Let there be a road with length v km, and assume all the vehicles are moving with v km/hr.(Fig 31:2). Let the number of vehicles

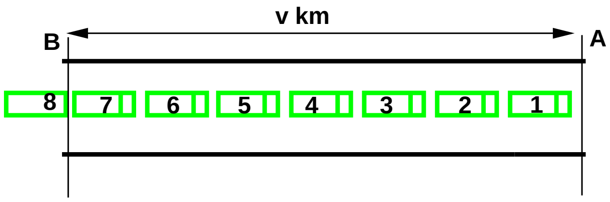


Figure 31:2: Illustration of relation between fundamental parameters of traffic flow

counted by an observer at A for one hour be n_1. By definition, the number of vehicles counted in one hour is flow(q). Therefore,

$$n_1 = q \tag{31.19}$$

Similarly, by definition, density is the number of vehicles in unit distance. Therefore number of vehicles n_2 in a road stretch of distance v_1 will be density $\times$ distance.Therefore,

$$n_2 = k \times v \tag{31.20}$$

Since all the vehicles have speed v, the number of vehicles counted in 1 hour and the number of vehicles in the stretch of distance v will also be same.(ie $n_1 = n_2$). Therefore,

$$q = k \times v \tag{31.21}$$

This is the fundamental equation of traffic flow. Please note that, v in the above equation refers to the space mean speed.

31.7 Fundamental diagrams of traffic flow

The relation between flow and density, density and speed, speed and flow, can be represented with the help of some curves. They are referred to as the fundamental diagrams of traffic flow. They will be explained in detail one by one below.

31.7.1 Flow-density curve

The flow and density varies with time and location. The relation between the density and the corresponding flow on a given stretch of road is referred to as one of the fundamental diagram of traffic flow. Some characteristics of an ideal flow-density relationship is listed below:

1. When the density is zero, flow will also be zero,since there is no vehicles on the road.

2. When the number of vehicles gradually increases the density as well as flow increases.

3. When more and more vehicles are added, it reaches a situation where vehicles can't move. This is referred to as the jam density or the maximum density. At jam density, flow will be zero because the vehicles are not moving.

4. There will be some density between zero density and jam density, when the flow is maximum. The relationship is normally represented by a parabolic curve as shown in figure 31:3

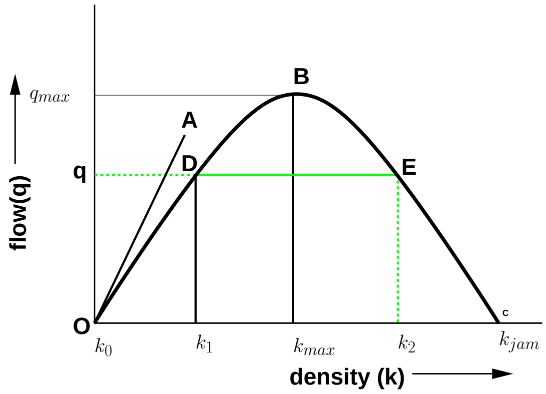


Figure 31:3: Flow density curve

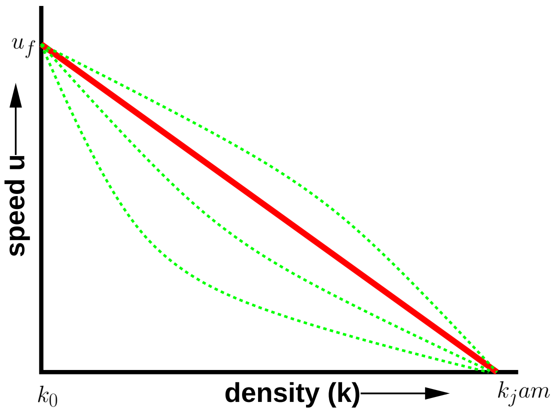


Figure 31:4: Speed-density diagram

The point O refers to the case with zero density and zero flow. The point B refers to the maximum flow and the corresponding density is k_{max}. The point C refers to the maximum density k_{jam} and the corresponding flow is zero. OA is the tangent drawn to the parabola at O, and the slope of the line OA gives the mean free flow speed, ie the speed with which a vehicle can travel when there is no flow. It can also be noted that points D and E correspond to same flow but has two different densities. Further, the slope of the line OD gives the mean speed at density k_1 and slope of the line OE will give mean speed at density k_2. Clearly the speed at density k_1 will be higher since there are less number of vehicles on the road.

31.7.2 Speed-density diagram

Similar to the flow-density relationship, speed will be maximum, referred to as the free flow speed, and when the density is maximum, the speed will be zero. The most simple assumption is that this variation of speed with density is linear as shown by the solid line in figure 31:4. Corresponding to the zero density, vehicles will be flowing with their desire speed, or free flow speed. When the density is jam density, the speed of the vehicles becomes zero. It is also possible to have non-linear relationships as shown by the dotted lines. These will be discussed later.

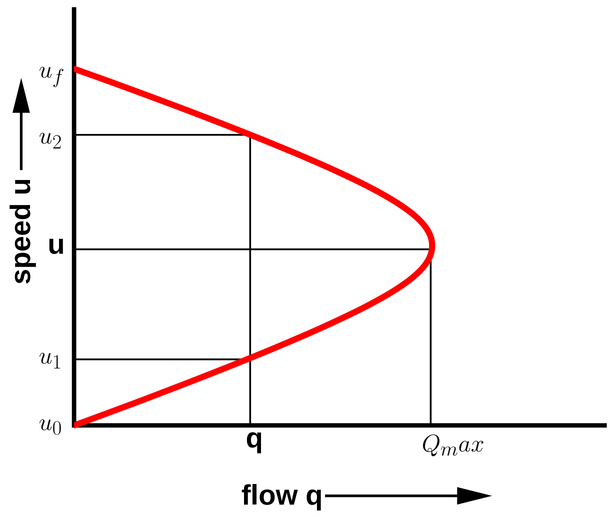


Figure 31:5: Speed-flow diagram

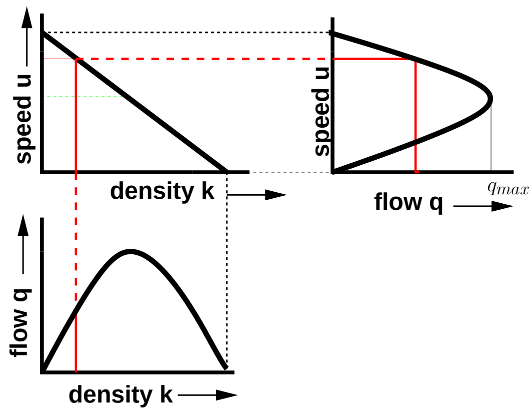


Figure 31:6: Fundamental diagram of traffic flow

31.7.3 Speed flow relation

The relationship between the speed and flow can be postulated as follows. The flow is zero either because there is no vehicles or there are too many vehicles so that they cannot move. At maximum flow, the speed will be in between zero and free flow speed. This relationship is shown in figure 31:5. The maximum flow q_{max} occurs at speed u. It is possible to have two different speeds for a given flow.

31.7.4 Combined diagrams

The diagrams shown in the relationship between speed-flow, speed-density, and flow-density are called the fundamental diagrams of traffic flow. These are as shown in figure 31:6

31.8 Summary

Time mean speed and space mean speed are two important measures of speed. It is possible to have a relation between them and was derived in this chapter. Also, time mean speed will be always greater than or equal to space mean speed. The fundamental diagrams of traffic flow are vital tools which enables analysis of fundamental relationships. There are three diagrams - speed-density, speed-flow and flow-density. They can be together combined in a single diagram as discussed in the last section of the chapter.

31.9 Problems

1. Space mean speed is
   - (a) the harmonic mean of spot speeds
   - (b) the sum of spot speeds
   - (c) the arithmetic mean of spot speeds
   - (d) the sum of journey speeds
2. Which among the following is the fundamental equation of traffic flow?
   - (a) $q = \frac{k}{v}$
   - (b) $q = k \times v$
   - (c) $v = q \times k$
   - (d) $q = k^2 \times v$

31.10 Solutions

1. Space mean speed is
   - (a) the harmonic mean of spot speeds$\surd$
   - (b) the sum of spot speeds
   - (c) the arithmetic mean of spot speeds
   - (d) the sum of journey speeds
2. Which among the following is the fundamental equation of traffic flow?
   - (a) $q = \frac{k}{v}$
   - (b) $q = k \times v\surd$
   - (c) $v = q \times k$
   - (d) $q = k^2 \times v$

Chapter 32

Traffic data collection

32.1 Overview

Unlike many other disciplines of the engineering, the situations that are interesting to a traffic engineer cannot be reproduced in a laboratory. Even if road and vehicles could be set up in large laboratories, it is impossible to simulate the behavior of drivers in the laboratory. Therefore, traffic stream characteristics need to be collected only from the field. There are several methods of data collection depending on the need of the study and some important ones are described in this chapter.

32.2 Data requirements

The most important traffic characteristics to be collected from the field includes sped, travel time, flow and density. Some cases, spacing and headway are directly measured. In addition, the occupancy, ie percentage of time a point on the road is occupied by vehicles is also of interest. The measurement procedures can be classified based on the geographical extent of the survey into five categories: (a) measurement at point on the road, (b) measurement over a short section of the road (less than 500 metres) (c) measurement over a length of the road (more than about 500 metres) (d) wide area samples obtained from number of locations, and (e) the use of an observer moving in the traffic stream. In each category, numerous data collection are there. However, important and basic methods will be discussed.

32.2.1 Measurements at a point

The most important point measurement is the vehicle volume count. Data can be collected manually or automatically. In manual method, the observer will stand at the point of interest and count the vehicles with the help of hand tallies. Normally, data will be collected for short interval of 5 minutes or 15 minutes etc. and for each types of vehicles like cars, two wheelers, three wheelers, LCV, HCV, multi axle trucks, nonmotorised traffic like bullock cart, hand cart etc. From the flow data, flow and headway can be derived.

Modern methods include the use of inductive loop detector, video camera, and many other technologies. These methods helps to collect accurate information for long duration. In video cameras, data is collected from the field and is then analyzed in the lab for obtaining results. Radars and microwave detectors are used to obtain the speed of a vehicle at a point. Since no length is involved, density cannot be obtained by measuring at a point.

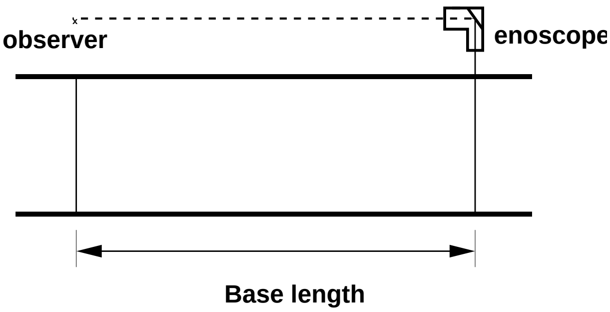


Figure 32:1: Illustration of measurement over short section using enoscope

32.2.2 Measurements over short section

The main objective of this study is to find the spot speed of vehicles. Manual methods include the use of enoscope. In this method a base length of about 30-90 metres is marked on the road. Enoscope is placed at one end and observer will stand at the other end. He could see the vehicle passing the farther end through enoscope and starts the stop watch. Then he stops the stop watch when the vehicle passes in front of him. The working of the enoscope is shown in figure 32:1.

An alternative method is to use pressure contact tube which gives a pressure signal when vehicle moves at either end. Another most widely used method is inductive loop detector which works on the principle of magnetic inductance. Road will be cut and a small magnetic loop is placed. When the metallic content in the vehicle passes over it, a signal will be generated and the count of the vehicle can be found automatically. The advantage of this detector is that the counts can be obtained throughout the life time of the road. However, chances of errors are possible because noise signals may be generated due to heavy vehicle passing adjacent lanes. When dual loops are used and if the spacing between them is known then speed also can be calculated in addition to the vehicle cost.

32.2.3 Measurements over long section

This is normally used to obtain variations in speed over a stretch of road. Usually the stretch will be having a length more than 500 metres. We can also get density. Most traditional method uses aerial photography. From a single frame, density can be measured, but not speed or volumes. In time lapse photography, several frames are available. If several frames are obtained over short time intervals, speeds can be measured from the distance covered between the two frames and time interval between them.

32.2.4 Moving observer method for stream measurement

Determination of any of the two parameters of the traffic flow will provide the third one by the equation $q = u.k$. Moving observer method is the most commonly used method to get the relationship between the fundamental stream characteristics. In this method, the observer moves in the traffic stream unlike all other previous methods.

Consider a stream of vehicles moving in the north bound direction. Two different cases of motion can be considered. The first case considers the traffic stream to be moving and the observer to be stationary. If n_o is

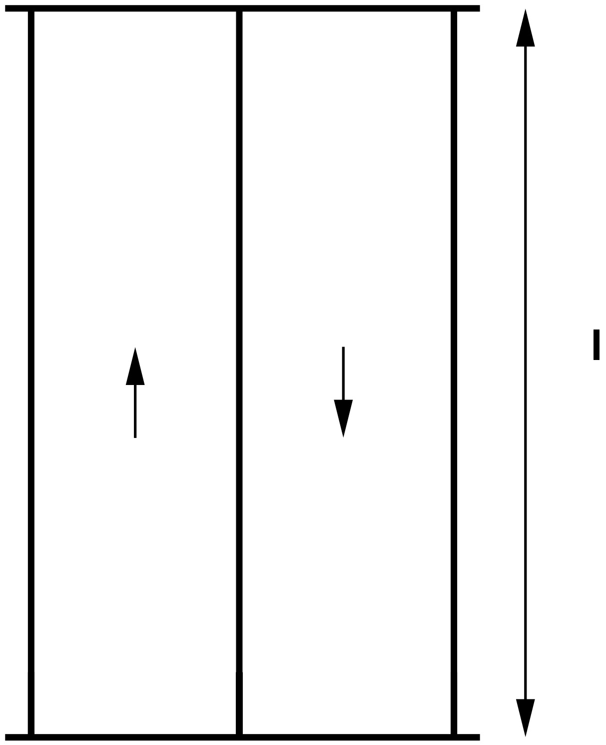


Figure 32:2: Illustration of moving observer method

the number of vehicles overtaking the observer during a period, t, then flow q is $\frac{n_0}{t}$, or

$$n_0 = q \times t \tag{32.1}$$

The second case assumes that the stream is stationary and the observer moves with speed v_o. If n_p is the number of vehicles overtaken by observer over a length l, then by definition, density k is $\frac{n_p}{l}$, or

$$n_p = k \times l \tag{32.2}$$

or

$$n_p = k.v_o.t \tag{32.3}$$

where v_0 is the speed of the observer and t is the time taken for the observer to cover the road stretch. Now consider the case when the observer is moving within the stream. In that case m_o vehicles will overtake the observer and m_p vehicles will be overtaken by the observer in the test vehicle. Let the difference m is given by m_0 - m_p, then from equation 32.5 and equation 32.7,

$$m = q.t - k.v_o.t \tag{32.4}$$

This equation is the basic equation of moving observer method, which relates q, k to the counts m, t and v_o that can be obtained from the test. However, we have two unknowns, q and k, but only one equation. For generating another equation, the test vehicle is run twice once with the traffic stream and another one against traffic stream, i.e.

$$m_w = q.t_w + k.v_w$$

$$\begin{aligned} &= q.t_w + k.l \\ m_a &= qt_a - k.v_a.t_a \\ &= q.t_a - k.l \end{aligned}$$

where, a, w denotes against and with traffic flow. It may be noted that the sign of equation 32.5 is negative, because test vehicle moving in the opposite direction can be considered as a case when the test vehicle is moving in the stream with negative velocity. Further, in this case, all the vehicles will be overtaking, since it is moving with negative speed. In other words, when the test vehicle moves in the opposite direction, the observer simply counts the number of vehicles in the opposite direction. Adding equation 32.5 and 32.5, we will get the first parameter of the stream, namely the flow(q) as:

$$q = \frac{m_w + m_a}{t_w + t_a} \tag{32.5}$$

Now calculating space mean speed from equation 32.5,

$$\begin{aligned} \frac{m_w}{t_w} &= q - k.v_w.t \\ &= q - \frac{q}{v} v_w \\ &= q - \frac{q}{v}\left[\frac{l}{t_w}\right] \\ &= q(1 - \frac{l}{v} \times \frac{1}{t_w}) \\ &= q(1 - \frac{t_{avg}}{t_w}) \end{aligned}$$

If v_s is the mean stream speed, then average travel time is given by $t_{avg} = \frac{l}{v_s}$. Therefore,

$$\begin{aligned} \frac{m_w}{q} &= t_w(1 - \frac{t_{avg}}{t_w}) = t_w - t_{avg} \\ t_{avg} &= t_w - \frac{m_w}{q} = \frac{l}{v}, \end{aligned}$$

Rewriting the above equation, we get the second parameter of the traffic flow, namely the mean speed v_s and can be written as,

$$v_s = \frac{l}{t_w - \frac{m_w}{q}} \tag{32.6}$$

Thus two parameters of the stream can be determined. Knowing the two parameters the third parameter of traffic flow density (k) can be found out as

$$k = \frac{q}{v_s} \tag{32.7}$$

For increase accuracy and reliability, the test is performed a number of times and the average results are to be taken.

Example 1

The length of a road stretch used for conducting the moving observer test is 0.5 km and the speed with which the test vehicle moved is 20 km/hr. Given that the number of vehicles encountered in the stream while the test

Sample no.	m_a	m_o	m_p	$m(m_o - m_p)$	t_a	t_w	$q = \frac{m_a+m_w}{t_a+t_w}$	$u = \frac{l}{t_w - \frac{m_a}{q}}$	$k = \frac{q}{v}$
1	107	10	74	-64	0.025	0.025	860	5.03	171
2	113	25	41	-16	0.025	0.025	1940	15.04	129
3	30	15	5	10	0.025	0.025	800	40	20
4	79	18	9	9	0.025	0.025	1760	25.14	70

vehicle was moving against the traffic stream is 107, number of vehicles that had overtaken the test vehicle is 10, and the number of vehicles overtaken by the test vehicle is 74, find the flow, density and average speed of the stream.

Solution Time taken by the test vehicle to reach the other end of the stream while it is moving along with the traffic is $t_w = \frac{0.5}{20} = 0.025$ hr Time taken by the observer to reach the other end of the stream while it is moving against the traffic is $t_a = t_w = 0.025$ hr Flow is given by equation, q = $\frac{107+(10-74)}{0.025+0.025} = 860$ veh/hr Stream speed v_s can be found out from equation$v_s = \frac{0.5}{0.025 - \frac{10.74}{860}} = 5$ km/hr Density can be found out from equation as k = $\frac{860}{5} = 172$veh/km

Example 2

The data from four moving observer test methods are shown in the table. Column 1 gives the sample number, column 2 gives the number of vehicles moving against the stream, column 3 gives the number of vehicles that had overtaken the test vehicle, and last column gives the number of vehicles overtaken by the test vehicle. Find the three fundamental stream parameters for each set of data. Also plot the fundamental diagrams of traffic flow.

Sample no.	1	2	3
1	107	10	74
2	113	25	41
3	30	15	5
4	79	18	9

Solution From the calculated values of flow, density and speed, the three fundamental diagrams can be plotted as shown in figure 32:3.

32.3 Summary

Traffic engineering studies differ from other studies in the fact that they require extensive data from the field which cannot be exactly created in any laboratory. Speed data are collected from measurements at a point or over a short section or over an area. Traffic flow data are collected at a point. Moving observer method is one in which both speed and traffic flow data are obtained by a single experiment.

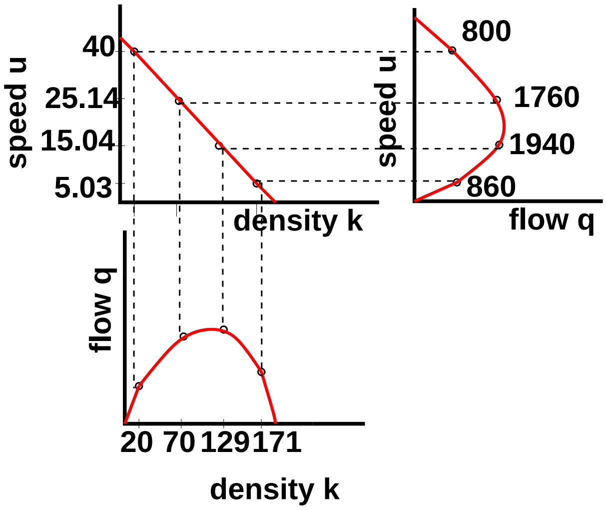


Figure 32:3: Fundamental diagrams of traffic flow

32.4 Problems

1. In the moving observer experiment, if the density is k, speed of the observer is v_o , length of the test stretch is l, t is the time taken by the observer to cover the road stretch, the number of vehicles overtaken by the observer n_p is given by,

   (a) $n_p = k.t$

   (b) $n_p = k.l$

   (c) $n_p = \frac{k}{v_o.t}$

   (d) $n_p = k.v_o.t$

2. If the length of the road stretch taken for conducting moving observer experiment is 0.4 km, time taken by the observer to move with the traffic is 5 seconds, number of vehicles moving with the test vehicle in the same direction is 10, flow is 10 veh/sec, find the mean speed.

   (a) 50 m/s

   (b) 100 m/s

   (c) 150 m/s

   (d) 200 m/s

32.5 Solutions

1. In the moving observer experiment, if the density is k, speed of the observer is v_o , length of the test stretch is l, t is the time taken by the observer to cover the road stretch, the number of vehicles overtaken by the observer n_p is given by,

   (a) $n_p = k.t$

(b) $n_p = k.l$

(c) $n_p = \frac{k}{v_o.t}$

(d) $n_p = k.v_o.t\surd$

2. If the length of the road stretch taken for conducting moving observer experiment is 0.4 km, time taken by the observer to move with the traffic is 5 seconds, number of vehicles moving with the test vehicle in the same direction is 10, flow is 10 veh/sec, find the mean speed.

   (a) 50 m/s

   (b) 100 m/s$\surd$

   (c) 150 m/s

   (d) 200 m/s

   Solution: Given that l=0.4 km, t_w=5seconds, m_w=10,q=10 veh/sec, substituting in equation,$v_s = \frac{l}{t_w - \frac{m_w}{q}}$=100m/s.

Chapter 33

Traffic stream models

33.1 Overview

To figure out the exact relationship between the traffic parameters, a great deal of research has been done over the past several decades. The results of these researches yielded many mathematical models. Some important models among them will be discussed in this chapter.

33.2 Greenshield's macroscopic stream model

Macroscopic stream models represent how the behaviour of one parameter of traffic flow changes with respect to another. Most important among them is the relation between speed and density. The first and most simple relation between them is proposed by Greenshield. Greenshield assumed a linear speed-density relationship as illustrated in figure 33:1 to derive the model. The equation for this relationship is shown below.

$$v = v_f - \left[\frac{v_f}{k_j}\right].k \tag{33.1}$$

where v is the mean speed at density k, v_f is the free speed and k_j is the jam density. This equation (33.1) is often referred to as the Greenshields' model. It indicates that when density becomes zero, speed approaches free flow speed (ie. $v \rightarrow v_f$ when $k \rightarrow 0$).

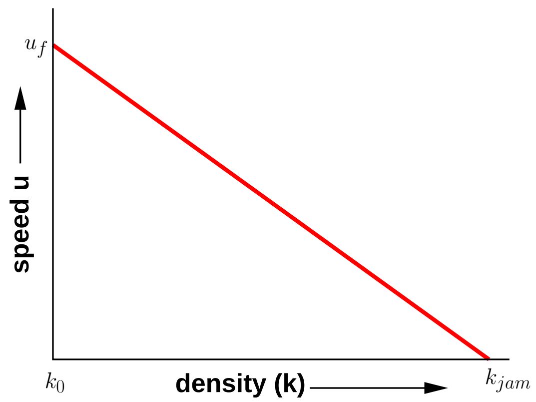


Figure 33:1: Relation between speed and density

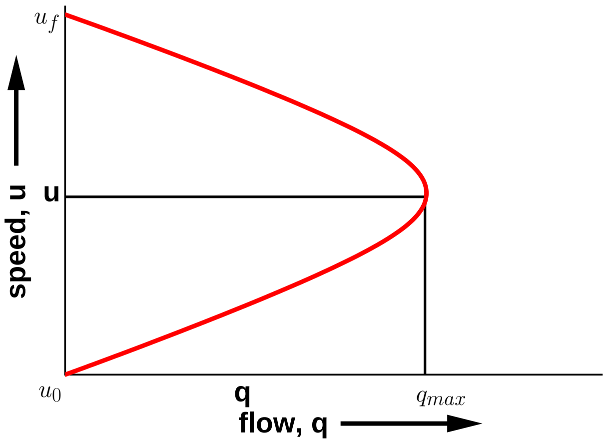


Figure 33:2: Relation between speed and flow

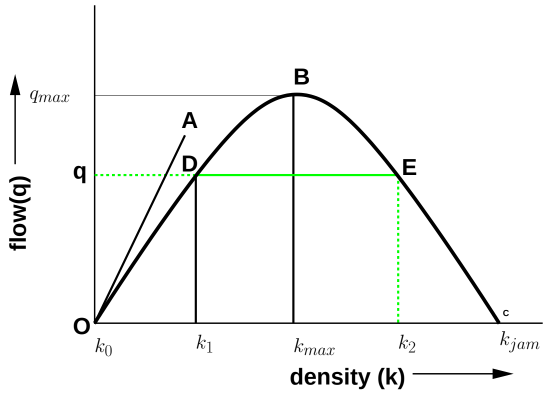


Figure 33:3: Relation between flow and density

Once the relation between speed and flow is established, the relation with flow can be derived. This relation between flow and density is parabolic in shape and is shown in figure 33:3. Also, we know that

$$q = k.v \tag{33.2}$$

Now substituting equation 33.1 in equation 33.2, we get

$$q = v_f.k - \left[\frac{v_f}{k_j}\right] k^2 \tag{33.3}$$

Similarly we can find the relation between speed and flow. For this, put $k = \frac{q}{v}$ in equation 33.1 and solving, we get

$$q = k_j.v - \left[\frac{k_j}{v_f}\right] v^2 \tag{33.4}$$

This relationship is again parabolic and is shown in figure 33:2. Once the relationship between the fundamental variables of traffic flow is established, the boundary conditions can be derived. The boundary conditions that are of interest are jam density, freeflow speed, and maximum flow. To find density at maximum flow, differentiate equation 33.3 with respect to k and equate it to zero. ie.,

$$\begin{aligned}\frac{dq}{dk} &= 0\\ v_f - \frac{v_f}{k_j}.2k &= 0\\ k &= \frac{k_j}{2}\end{aligned}$$

Denoting the density corresponding to maximum flow as k_0,

$$k_0 = \frac{k_j}{2} \tag{33.5}$$

Therefore, density corresponding to maximum flow is half the jam density Once we get k_0, we can derive for maximum flow, q_{max}. Substituting equation 33.5 in equation 33.3

$$\begin{aligned}q_{max} &= v_f.\frac{k_j}{2} - \frac{v_f}{k_j}.\left[\frac{k_j}{2}\right]^2\\ &= v_f.\frac{k_j}{2} - v_f.\frac{k_j}{4}\\ &= \frac{v_f.k_j}{4}\end{aligned}$$

Thus the maximum flow is one fourth the product of free flow and jam density. Finally to get the speed at maximum flow, v_0, substitute equation 33.5 in equation 33.1 and solving we get,

$$v_0 = v_f - \frac{v_f}{k_j}.\frac{k_j}{2}$$

$$v_0 = \frac{v_f}{2} \tag{33.6}$$

Therefore, speed at maximum flow is half of the free speed.

33.3 Calibration of Greenshield's model

Inorder to use this model for any traffic stream, one should get the boundary values, especially free flow speed (v_f) and jam density (k_j). This has to be obtained by field survey and this is called calibration process. Although it is difficult to determine exact free flow speed and jam density directly from the field, approximate values can be obtained from a number of speed and density observations and then fitting a linear equation between them. Let the linear equation be $y = a + bx$ such that y is density k and x denotes the speed v. Using linear regression method, coefficients a and b can be solved as,

$$b = \frac{n.\sum_{i=1}^{n} xy - \sum_{i=1}^{n} x.\sum_{i=1}^{n} y}{n.\sum_{i=1}^{n} x^2 - \sum_{i=1}^{n} x^2} \tag{33.7}$$

$$a = \bar{y} - b\bar{x} \tag{33.8}$$

Alternate method of solving for b is,

$$b = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{33.9}$$

where x_i and y_i are the samples, n is the number of samples, and $\bar{x}$ and $\bar{y}$ are the mean of x_i and y_i respectively.

Problem

For the following data on speed and density, determine the parameters of the Greenshields' model. Also find the maximum flow and density corresponding to a speed of 30 km/hr.

k	v
171	5
129	15
20	40
70	25

Solution Denoting y = v and x = k, solve for a and b using equation 33.8 and equation 33.9. The solution is tabulated as shown below.

x(k)	y(v)	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x}^2)$
171	5	73.5	-16.3	-1198.1	5402.3
129	15	31.5	-6.3	-198.5	992.3
20	40	-77.5	18.7	-1449.3	6006.3
70	25	-27.5	3.7	-101.8	756.3
390	85			-2947.7	13157.2

$\bar{x} = \frac{\Sigma x}{n} = \frac{390}{4} = 97.5$, $\bar{y} = \frac{\Sigma y}{n} = \frac{85}{4} = 21.3$. From equation 33.9, b = $\frac{-2947.7}{13157.2}$ = -0.2 $a = y - b\bar{x} = 21.3 + 0.2{\times}97.5 = 40.8$ So the linear regression equation will be,

$$v = 40.8 - 0.2k \tag{33.10}$$

Here $v_f = 40.8$ and $\frac{v_f}{k_j} = 0.2$ This implies, $k_j = \frac{40.8}{0.2} = 204$ veh/km The basic parameters of Greenshield's model are free flow speed and jam density and they are obtained as 40.8 kmph and 204 veh/km respectively. To find maximum flow, use equation 33.6, i.e., $q_{max} = \frac{40.8\times 204}{4} = 2080.8$ veh/hr Density corresponding to the speed 30 km/hr can be found out by substituting $v = 30$ in equation 33.10. i.e, 30 = 40.8 - 0.2 $\times$ k Therefore, k = $\frac{40.8-30}{0.2}$ = 54 veh/km

33.4 Other macroscopic stream models

In Greenshield's model, linear relationship between speed and density was assumed. But in field we can hardly find such a relationship between speed and density. Therefore, the validity of Greenshields' model was questioned and many other models came up. Prominent among them are Greenberg's logarithmic model, Underwood's exponential model, Pipe's generalized model, and multiregime models. These are briefly discussed below.

33.4.1 Greenberg's logarithmic model

Greenberg assumed a logarithmic relation between speed and density. He proposed,

$$v = v_0 \ln \frac{k_j}{k} \tag{33.11}$$

This model has gained very good popularity because this model can be derived analytically. (This derivation is beyond the scope of this notes). However, main drawbacks of this model is that as density tends to zero, speed tends to infinity. This shows the inability of the model to predict the speeds at lower densities.

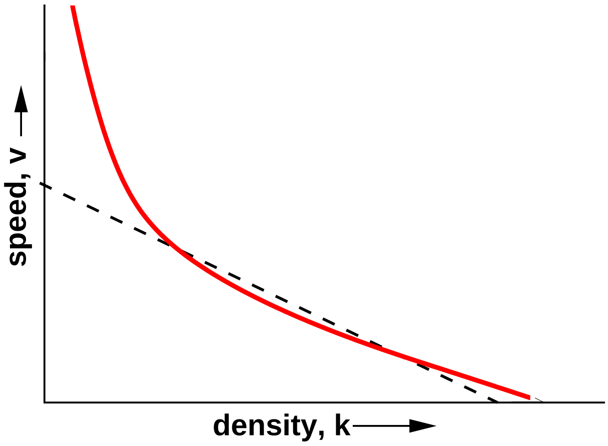


Figure 33:4: Greenberg's logarithmic model

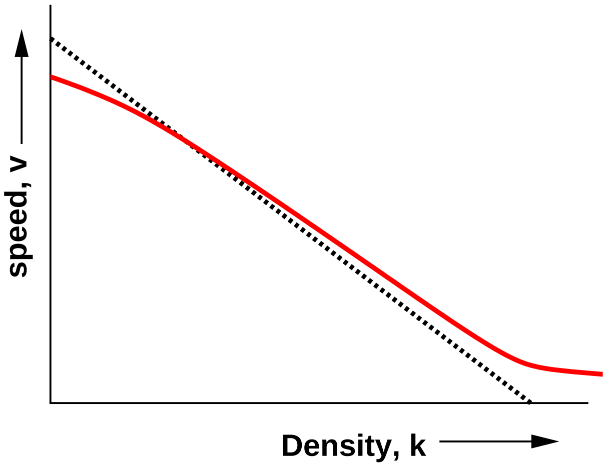


Figure 33:5: Underwood's exponential model

33.4.2 Underwood's exponential model

Trying to overcome the limitation of Greenberg's model, Underwood put forward an exponential model as shown below.

$$v = v_f . e^{\frac{-k}{k_0}} \tag{33.12}$$

where v_fThe model can be graphically expressed as in figure 33:5. is the free flow speedm and k_o is the optimum density, i.e. the densty corresponding to the maximum flow. In this model, speed becomes zero only when density reaches infinity which is the drawback of this model. Hence this cannot be used for predicting speeds at high densities.

Chapter 34

Microscopic traffic flow modeling

34.1 Overview

Macroscopic modeling looks at traffic flow from a global perspective, whereas microscopic modeling, as the term suggests, gives attention to the details of traffic flow and the interactions taking place within it. This chapter gives an overview of microscopic approach to modeling traffic and then elaborates on the various concepts associated with it.

A microscopic model of traffic flow attempts to analyze the flow of traffic by modeling driver-driver and driver-road interactions within a traffic stream which respectively analyzes the interaction between a driver and another driver on road and of a single driver on the different features of a road. Many studies and researches were carried out on driver's behavior in different situations like a case when he meets a static obstacle or when he meets a dynamic obstacle. Several studies are made on modeling driver behavior in another following car and such studies are often referred to as car following theories of vehicular traffic.

34.2 Notation

Longitudinal spacing of vehicles are of particular importance from the points of view of safety, capacity and level of service. The longitudinal space occupied by a vehicle depend on the physical dimensions of the vehicles as well as the gaps between vehicles. For measuring this longitudinal space, two microscopic measures are used- distance headway and distance gap. Distance headway is defined as the distance from a selected point (usually front bumper) on the lead vehicle to the corresponding point on the following vehicles. Hence, it includes the length of the lead vehicle and the gap length between the lead and the following vehicles. Before going in to the details, various notations used in car-following models are discussed here with the help of figure 34:1. The leader vehicle is denoted as n and the following vehicle as $(n + 1)$. Two characteristics at an instant t are of importance; location and speed. Location and speed of the lead vehicle at time instant t are represented by x_n^t and v_n^t respectively. Similarly, the location and speed of the follower are denoted by x_{n+1}^t and v_{n+1}^t respectively. The following vehicle is assumed to accelerate at time $t + \Delta T$ and not at t, where ΔT is the interval of time required for a driver to react to a changing situation. The gap between the leader and the follower vehicle is therefore $x_n^t - x_{n+1}^t$.

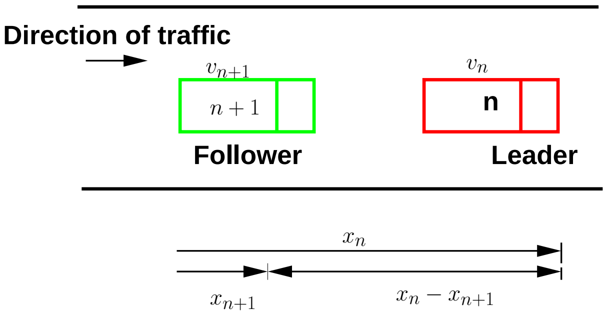


Figure 34:1: Notation for car following model

34.3 Car following models

Car following theories describe how one vehicle follows another vehicle in an uninterrupted flow. Various models were formulated to represent how a driver reacts to the changes in the relative positions of the vehicle ahead. Models like Pipes, Forbes, General Motors and Optimal velocity model are worth discussing.

34.3.1 Pipe's model

The basic assumption of this model is "A good rule for following another vehicle at a safe distance is to allow yourself at least the length of a car between your vehicle and the vehicle ahead for every ten miles per hour of speed at which you are traveling" According to Pipe's car-following model, the minimum safe distance headway increases linearly with speed. A disadvantage of this model is that at low speeds, the minimum headways proposed by the theory are considerably less than the corresponding field measurements.

34.3.2 Forbes' model

In this model, the reaction time needed for the following vehicle to perceive the need to decelerate and apply the brakes is considered. That is, the time gap between the rear of the leader and the front of the follower should always be equal to or greater than the reaction time. Therefore, the minimum time headway is equal to the reaction time (minimum time gap) and the time required for the lead vehicle to traverse a distance equivalent to its length. A disadvantage of this model is that, similar to Pipe's model, there is a wide difference in the minimum distance headway at low and high speeds.

34.3.3 General Motors' model

The General Motors' model is the most popular of the car-following theories because of the following reasons:

1. Agreement with field data; the simulation models developed based on General motors' car following models shows good correlation to the field data.

2. Mathematical relation to macroscopic model; Greenberg's logarithmic model for speed-density relationship can be derived from General motors car following model.

In car following models, the motion of individual vehicle is governed by an equation, which is analogous to the Newton's Laws of motion. In Newtonian mechanics, acceleration can be regarded as the response of the particle to *stimulus* it receives in the form of force which includes both the external force as well as those arising from the interaction with all other particles in the system. This model is the widely used and will be discussed in detail later.

34.3.4 Optimal velocity model

The concept of this model is that each driver tries to achieve an optimal velocity based on the distance to the preceding vehicle and the speed difference between the vehicles. This was an alternative possibility explored recently in car-following models. The formulation is based on the assumption that the desired speed $v_{n_{desired}}$ depends on the distance headway of the nth vehicle. i.e.$v^t_{n_{desired}} = v^{opt}(\Delta x^t_n)$ where v_{opt} is the optimal velocity function which is a function of the instantaneous distance headway Δx^t_n. Therefore a^t_n is given by

$$a^t_n = [1/\tau][V^{opt}(\Delta x^t_n) - v^t_n] \tag{34.1}$$

where $\frac{1}{\tau}$ is called as sensitivity coefficient. In short, the driving strategy of n^{th} vehicle is that, it tries to maintain a safe speed which inturn depends on the relative position, rather than relative speed.

34.4 General motor's car following model

34.4.1 Basic Philosophy

The basic philosophy of car following model is from Newtonian mechanics, where the acceleration may be regarded as the response of a matter to the stimulus it receives in the form of the force it receives from the interaction with other particles in the system. Hence, the basic philosophy of car-following theories can be summarized by the following equation

$$[\text{Response}]_n \ \alpha \ [\text{Stimulus}]_n \tag{34.2}$$

for the nth vehicle (n=1, 2, ...). Each driver can respond to the surrounding traffic conditions only by accelerating or decelerating the vehicle. As mentioned earlier, different theories on car-following have arisen because of the difference in views regarding the nature of the stimulus. The stimulus may be composed of the speed of the vehicle, relative speeds, distance headway etc, and hence, it is not a single variable, but a function and can be represented as,

$$a^t_n = f_{sti}(v_n, \Delta x_n, \Delta v_n) \tag{34.3}$$

where f_{sti} is the stimulus function that depends on the speed of the current vehicle, relative position and speed with the front vehicle.

34.4.2 Follow-the-leader model

The car following model proposed by General motors is based on follow-the leader concept. This is based on two assumptions; (a) higher the speed of the vehicle, higher will be the spacing between the vehicles and (b) to avoid collision, driver must maintain a safe distance with the vehicle ahead.

Let Δx^t_{n+1} is the gap available for $(n+1)^{th}$ vehicle, and let Δx_{safe} is the safe distance, v^t_{n+1} and v^t_n are the velocities, the gap required is given by,

$$\Delta x^t_{n+1} = \Delta x_{safe} + \tau v^t_{n+1} \tag{34.4}$$

where τ is a sensitivity coefficient. The above equation can be written as

$$x_n - x_{n+1}^t = \Delta x_{safe} + \tau v_{n+1}^t \tag{34.5}$$

Differentiating the above equation with respect to time, we get

$$v_n^t - v_{n+1}^t = \tau . a_{n+1}^t$$
$$a_{n+1}^t = \frac{1}{\tau}[v_n^t - v_{n+1}^t]$$

General Motors has proposed various forms of sensitivity coefficient term resulting in five generations of models. The most general model has the form,

$$a_{n+1}^t = \left[\frac{\alpha_{l,m}(v_{n+1}^t)^m}{(x_n^t - x_{n+1}^t)^l}\right]\left[v_n^t - v_{n+1}^t\right] \tag{34.6}$$

where l is a distance headway exponent and can take values from +4 to -1, m is a speed exponent and can take values from -2 to +2, and α is a sensitivity coefficient. These parameters are to be calibrated using field data. This equation is the core of traffic simulation models.

In computer, implementation of the simulation models, three things need to be remembered:

1. A driver will react to the change in speed of the front vehicle after a time gap called the reaction time during which the follower perceives the change in speed and react to it.
2. The vehicle position, speed and acceleration will be updated at certain time intervals depending on the accuracy required. Lower the time interval, higher the accuracy.
3. Vehicle position and speed is governed by Newton's laws of motion, and the acceleration is governed by the car following model.

Therefore, the governing equations of a traffic flow can be developed as below. Let ΔT is the reaction time, and Δt is the updation time, the governing equations can be written as,

$$v_n^t = v_n^{t-\Delta t} + a_n^{t-\Delta t} \times \Delta t \tag{34.7}$$
$$x_n^t = x_n^{t-\Delta t} + v_n^{t-\Delta t} \times \Delta t + \frac{1}{2} a_n^{t-\Delta t} \Delta t^2 \tag{34.8}$$
$$a_{n+1}^t = [\frac{\alpha_{l,m}(v_{n+1}^{t-\Delta T})}{(x_n^{t-\Delta T} - x_{n+1}^{t-\Delta T})}][v_n^{t-\Delta T} - v_{n+1}^{t-\Delta T}] \tag{34.9}$$

The equation 34.7 is a simulation version of the Newton's simple law of motion $v = u + at$ and equation 34.8 is the simulation version of the Newton's another equation $s = ut + \frac{1}{2}at^2$. The acceleration of the follower vehicle depends upon the relative velocity of the leader and the follower vehicle, sensitivity coefficient and the gap between the vehicles.

Problem

Let a leader vehicle is moving with zero acceleration for two seconds from time zero. Then he accelerates by 1 m/s^2 for 2 seconds, then decelerates by $1m/s^2$for 2 seconds. The initial speed is 16 m/s and initial location is 28 m from datum. A vehicle is following this vehicle with initial speed 16 m/s, and position zero. Simulate the behavior of the following vehicle using General Motors' Car following model (acceleration, speed and position) for 7.5 seconds. Assume the parameters l=1, m=0 , sensitivity coefficient $(\alpha_{l,m}) = 13$, reaction time as 1 second and scan interval as 0.5 seconds.

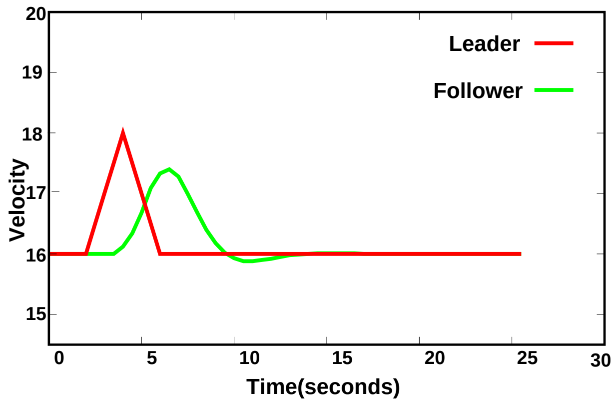


Figure 34:2: Velocity vz Time

Solution The first column shows the time in seconds. Column 2, 3, and 4 shows the acceleration, velocity and distance of the leader vehicle. Column 5,6, and 7 shows the acceleration, velocity and distance of the follower vehicle. Column 8 gives the difference in velocities between the leader and follower vehicle denoted as dv. Column 9 gives the difference in displacement between the leader and follower vehicle denoted as dx. Note that the values are assumed to be the state at the beginning of that time interval. At time t=0, leader vehicle has a velocity of 16 m/s and located at a distance of 28 m from a datum. The follower vehicle is also having the same velocity of 16 m/s and located at the datum. Since the velocity is same for both, $dv = 0$. At time t = 0, the leader vehicle is having acceleration zero, and hence has the same speed. The location of the leader vehicle can be found out from equation as, x = 28+16×0.5 = 36 m. Similarly, the follower vehicle is not accelerating and is maintaining the same speed. The location of the follower vehicle is, x = 0+16×0.5 = 8 m. Therefore, dx = 36-8 =28m. These steps are repeated till t = 1.5 seconds. At time t = 2 seconds, leader vehicle accelerates at the rate of 1 m/s^2 and continues to accelerate for 2 seconds. After that it decelerates for a period of two seconds. At t= 2.5 seconds, velocity of leader vehicle changes to 16.5 m/s. Thus dv becomes 0.5 m/s at 2.5 seconds. dx also changes since the position of leader changes. Since the reaction time is 1 second, the follower will react to the leader's change in acceleration at 2.0 seconds only after 3 seconds. Therefore, at t=3.5 seconds, the follower responds to the leaders change in acceleration given by equation i.e., a = $\frac{13\times0.5}{28.23}$ = 0.23 m/s^2. That is the current acceleration of the follower vehicle depends on dv and reaction time Δ of 1 second. The follower will change the speed at the next time interval. i.e., at time t = 4 seconds. The speed of the follower vehicle at t = 4 seconds is given by equation as v= 16+0.231×0.5 = 16.12 The location of the follower vehicle at t = 4 seconds is given by equation as x = 56+16×0.5+$\frac{1}{2}$×0.231×0.5^2 = 64.03 These steps are followed for all the cells of the table.

The earliest car-following models considered the difference in speeds between the leader and the follower as the stimulus. It was assumed that every driver tends to move with the same speed as that of the corresponding leading vehicle so that

$$a_n^t = \frac{1}{\tau}(v_n^{t+1} - v_{n+1}^t) \tag{34.10}$$

where τ is a parameter that sets the time scale of the model and $\frac{1}{\tau}$ can be considered as a measure of the sensitivity of the driver. According to such models, the driving strategy is to follow the leader and, therefore, such car-following models are collectively referred to as the follow the leader model. Efforts to develop this stimulus function led to five generations of car-following models, and the most general model is expressed

Table 34:1: Car-following example

t (1)	$a(t)$ (2)	$v(t)$ (3)	$x(t)$ (4)	$a(t)$ (5)	$v(t)$ (6)	$x(t)$ (7)	dv (8)	dx (9)
t	$a(t)$	$v(t)$	$x(t)$	$a(t)$	$v(t)$	$x(t)$	dv	dx
0.00	0.00	16.00	28.00	0.00	16.00	0.00	0.00	28.00
0.50	0.00	16.00	36.00	0.00	16.00	8.00	0.00	28.00
1.00	0.00	16.00	44.00	0.00	16.00	16.00	0.00	28.00
1.50	0.00	16.00	52.00	0.00	16.00	24.00	0.00	28.00
2.00	1.00	16.00	60.00	0.00	16.00	32.00	0.00	28.00
2.50	1.00	16.50	68.13	0.00	16.00	40.00	0.50	28.13
3.00	1.00	17.00	76.50	0.00	16.00	48.00	1.00	28.50
3.50	1.00	17.50	85.13	0.23	16.00	56.00	1.50	29.13
4.00	-1.00	18.00	94.00	0.46	16.12	64.03	1.88	29.97
4.50	-1.00	17.50	102.88	0.67	16.34	72.14	1.16	30.73
5.00	-1.00	17.00	111.50	0.82	16.68	80.40	0.32	31.10
5.50	-1.00	16.50	119.88	0.49	17.09	88.84	-0.59	31.03
6.00	0.00	16.00	128.00	0.13	17.33	97.45	-1.33	30.55
6.50	0.00	16.00	136.00	-0.25	17.40	106.13	-1.40	29.87
7.00	0.00	16.00	144.00	-0.57	17.28	114.80	-1.28	29.20
7.50	0.00	16.00	152.00	-0.61	16.99	123.36	-0.99	28.64
8.00	0.00	16.00	160.00	-0.57	16.69	131.78	-0.69	28.22
8.50	0.00	16.00	168.00	-0.45	16.40	140.06	-0.40	27.94
9.00	0.00	16.00	176.00	-0.32	16.18	148.20	-0.18	27.80
9.50	0.00	16.00	184.00	-0.19	16.02	156.25	-0.02	27.75
10.00	0.00	16.00	192.00	-0.08	15.93	164.24	0.07	27.76
10.50	0.00	16.00	200.00	-0.01	15.88	172.19	0.12	27.81
11.00	0.00	16.00	208.00	0.03	15.88	180.13	0.12	27.87
11.50	0.00	16.00	216.00	0.05	15.90	188.08	0.10	27.92
12.00	0.00	16.00	224.00	0.06	15.92	196.03	0.08	27.97
12.50	0.00	16.00	232.00	0.05	15.95	204.00	0.05	28.00
13.00	0.00	16.00	240.00	0.04	15.98	211.98	0.02	28.02
13.50	0.00	16.00	248.00	0.02	15.99	219.98	0.01	28.02
14.00	0.00	16.00	256.00	0.01	16.00	227.98	0.00	28.02
14.50	0.00	16.00	264.00	0.00	16.01	235.98	-0.01	28.02
15.00	0.00	16.00	272.00	0.00	16.01	243.98	-0.01	28.02
15.50	0.00	16.00	280.00	0.00	16.01	251.99	-0.01	28.01
16.00	0.00	16.00	288.00	-0.01	16.01	260.00	-0.01	28.00
16.50	0.00	16.00	296.00	0.00	16.01	268.00	-0.01	28.00
17.00	0.00	16.00	304.00	0.00	16.00	276.00	0.00	28.00
17.50	0.00	16.00	312.00	0.00	16.00	284.00	0.00	28.00
18.00	0.00	16.00	320.00	0.00	16.00	292.00	0.00	28.00
18.50	0.00	16.00	328.00	0.00	16.00	300.00	0.00	28.00
19.00	0.00	16.00	336.00	0.00	16.00	308.00	0.00	28.00
19.50	0.00	16.00	344.00	0.00	16.00	316.00	0.00	28.00
20.00	0.00	16.00	352.00	0.00	16.00	324.00	0.00	28.00
20.50	0.00	16.00	360.00	0.00	16.00	332.00	0.00	28.00

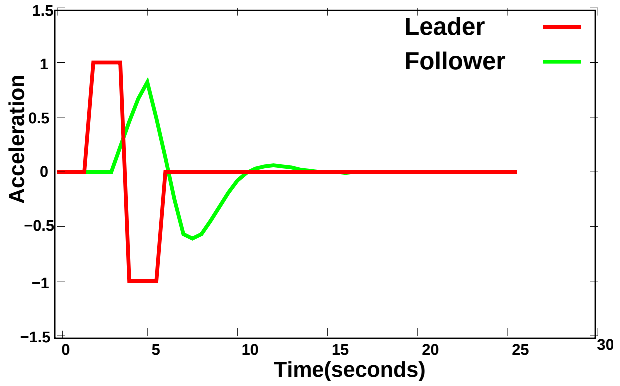


Figure 34:3: Acceleration vz Time

mathematically as follows.

$$a_{n+1}^{t+\Delta T} = \frac{\alpha_{l,m}\,[v_{n+1}^{t-\Delta T}]^m}{[x_n^{t-\Delta T} - x_{n+1}^{t-\Delta T}]^l}(v_n^{t-\Delta T} - v_{n+1}^{t-\Delta T}) \tag{34.11}$$

where l is a distance headway exponent and can take values from +4 to -1, m is a speed exponent and can take values from -2 to +2, and α is a sensitivity coefficient. These parameters are to be calibrated using field data.

34.5 Simulation Models

Simulation modeling is an increasingly popular and effective tool for analyzing a wide variety of dynamical problems which are difficult to be studied by other means. Usually, these processes are characterized by the interaction of many system components or entities.

34.5.1 Applications of simulation

Traffic simulations models can meet a wide range of requirements:

1. Evaluation of alternative treatments
2. Testing new designs
3. As an element of the design process
4. Embed in other tools
5. Training personnel
6. Safety Analysis

34.5.2 Need for simulation models

Simulation models are required in the following conditions

1. Mathematical treatment of a problem is infeasible or inadequate due to its temporal or spatial scale

2. The accuracy or applicability of the results of a mathematical formulation is doubtful, because of the assumptions underlying (e.g., a linear program) or an heuristic procedure (e.g., those in the Highway Capacity Manual)
3. The mathematical formulation represents the dynamic traffic/control environment as a simpler quasi steady state system.
4. There is a need to view vehicle animation displays to gain an understanding of how the system is behaving
5. Training personnel
6. Congested conditions persist over a significant time.

34.5.3 Classification of Simulation Model

Simulation models are classified based on many factors like

1. Continuity
   (a) Continuous model
   (b) Discrete model
2. Level of detail
   (a) Macroscopic models
   (b) Mesoscopic models
   (c) Microscopic models
3. Based on Processes
   (a) Deterministic
   (b) Stochastic

34.6 Summary

Microscopic traffic flow modeling focuses on the minute aspects of traffic stream like vehicle to vehicle interaction and individual vehicle behavior. They help to analyze very small changes in the traffic stream over time and space. Car following model is one such model where in the stimulus-response concept is employed. Optimal models and simulation models were briefly discussed.

34.7 Problems

1. The minimum safe distance headway increases linearly with speed. Which model follows this assumption?
   (a) Forbe's model
   (b) Pipe's model
   (c) General motor's model

(d) Optimal velocity model

2. The most popular of the car following models is

   (a) Forbe's model
   (b) Pipe's model
   (c) General motor's model
   (d) Optimal velocity model

34.8 Solutions

1. The minimum safe distance headway increases linearly with speed. Which model follows this assumption?

   (a) Forbe's model
   (b) Pipe's model$\surd$
   (c) General motor's model
   (d) Optimal velocity model

2. The most popular of the car following models is

   (a) Forbe's model
   (b) Pipe's model
   (c) General motor's model$\surd$
   (d) Optimal velocity model

Chapter 35

Capacity and Level of service

35.1 Overview

Capacity and Level of service are two related terms. Capacity analysis tries to give a clear understanding of how much traffic a given transportation facility can accommodate. Level of service tries to answer how good is the present traffic situation on a given facility. Thus it gives a qualitative measure of traffic, where as capacity analysis gives a quantitative measure of a facility. Capacity and level of service varies with the type of facility, prevailing traffic and road conditions etc. These concepts are discussed in this chapter.

35.2 Capacity

Capacity is defined as the maximum number of vehicles, passengers, or the like, per unit time, which can be accommodated under given conditions with a reasonable expectation of occurrence. Some of the observations that are found from this definition can be now discussed. Capacity is independent of the demand. It speaks about the physical amount of vehicles and passengers a road can afford. It does not depend on the total number of vehicles demanding service. On the other hand, it depends on traffic conditions, geometric design of the road etc. For example, a curved road has lesser capacity compared to a straight road. Capacity is expressed in terms of units of some specific thing (car, people, etc.), so it also does depend on the traffic composition. In addition, the capacity analysis depends on the environmental conditions too. Capacity is a probabilistic measure and it varies with respect to time and position. Hence it is not always possible to completely derive analytically the capacity. In most cases it is obtained, through field observations.

35.3 Level of service

A term closely related to capacity and often confused with it is service volume. When capacity gives a quantitative measure of traffic, level of service or LOS tries to give a qualitative measure. A service volume is the maximum number of vehicles, passengers, or the like, which can be accommodated by a given facility or system under given conditions at a given level of service.

For a given road or facility, capacity could be constant. But actual flow will be different for different days and different times in a day itself. The intention of LOS is to relate the traffic service quality to a given flow rate of traffic. It is a term that designates a range of operating conditions on a particular type of facility. Highway capacity manual (HCM) developed by the transportation research board of USA provides some procedure to determine level of service. It divides the quality of traffic into six levels ranging form level A to level F. Level

A represents the best quality of traffic where the driver has the freedom to drive with free flow speed and level F represents the worst quality of traffic. Level of service is defined based on the measure of effectiveness or (MOE). Typically three parameters are used under this and they are speed and travel time, density, and delay. One of the important measures of service quality is the amount of time spent in travel. Therefore, speed and travel time are considered to be more effective in defining LOS of a facility. Density gives the proximity of other vehicles in the stream. Since it affects the ability of drivers to maneuver in the traffic stream, it is also used to describe LOS. Delay is a term that describes excess or unexpected time spent in travel. Many specific delay measures are defined and used as MOE's in the highway capacity manual.

35.4 Types of facilities

Most important classification of transportation facilities from the engineering perspective is based on the continuity of flow, that is *uninterrupted flow* and *interrupted flow*. Uninterrupted flow is the flow of traffic in which there is no obstructions to the movement of vehicles along the road. Freeway is one example for this type of facility. In a freeway, when a vehicle enters a freeway, there is no need for the vehicle to stop anywhere till it leaves the freeway. There are three sections in a freeway - basic unit, weaving section and ramps(on/off). Vehicles will be entering the freeway through ramps. Ramps used for entering the freeway is called on-ramps and those used for exiting the freeway are called off-ramps. Freeways generally have 4, 6, or 8 lane alignments. Multi lanes also provide uninterrupted flow. HCM defines the levels of service of freeway sections based on density, as described in tables 35:1and 35:2.

Table 35:1: LOS for a basic freeway segment

LOS	K (veh/km/lane)	FFS (Km/hr)	v/c
A	0-7	120	0.35
B	7-11	120	0.55
C	11-16	114	0.77
D	16-22	99	0.92
E	22-28	85	1.0
F	> 28	< 85	> 1.0

In many roads, there will be signalized as well as unsignalized intersections. Uninterrupted flow is possible in sections of rural and suburban multilane highways between signalized intersections where signal spacing is sufficient to allow for uninterrupted flow. Two lane highways also provide uninterrupted flow facilities.

Interrupted flow refers to the condition when the traffic flow on the road is obstructed due to some reasons. This is experienced in signalized intersections, unsignalized intersections, arterials etc. At signalized intersections, there will be some kind of active control and the vehicle will have to stop or sometimes to reduce its speed and the flow of traffic is interrupted. Thus the capacity is defined in terms of control delay ie sec/veh. Arterials are roads of long stretches with many intersections in between and obviously there will be interruption to the flow of traffic. Here, the capacity is expressed in terms of average travel speed. Some other facilities are facilities for pedestrians, bicycles, bus-transit, rail-transit etc. Example for pedestrian facility is a provision of subway exclusively for the use of pedestrians. Here, the capacity may be expressed in terms of number of

Table 35:2: LOS for an intersection

LOS	Control Delay sec/veh(signalised)	Delay sec/veh (unsignalised)
A	≤ 10	≤ 10
B	10-20	10-15
C	20-35	15-25
D	35-55	25-35
E	55-80	35-50
F	> 80	> 50

passengers. In bus transit system, the buses has to stop at the bus bays and also it has to share the road with the other vehicles. Hence the capacity will be affected by the control characteristics and the traffic conditions prevailing in the road. Since trains have exclusive right of way, the capacity is strictly governed by the control characteristics. It has two types of capacities - line capacity and station capacity. Line capacity is based on the number of tracks available between two stations. Station capacity refers to the facilities available in the platform of the station,and other facilities.

For uninterrupted flow of traffic, measure of effectiveness (MOE) is density in freeways. Speed also becomes important in two-lane highways and multilane highways. In the case of interrupted flow, MOE is delay. The delay of travel time becomes an important factor in calculating the capacity.

35.5 Highway capacity

Highway capacity is defined by the Highway Capacity Manual as the maximum hourly rate at which persons or vehicles can be reasonably expected to traverse a point or a uniform segment of a lane or roadway during a given time period under prevailing roadway, traffic and control conditions. The highway capacity depends on certain conditions as listed below;

1. **Traffic conditions**: It refers to the traffic composition in the road such as the mix of cars, trucks, buses etc in the stream. It also include peaking characteristics, proportions of turning movements at intersections and the like.

2. **Road way characteristics**: This points out to the geometric characteristics of the road. These include lane width, shoulder width, lane configuration, horizontal alignment and vertical alignment.

3. **Control conditions**: This primarily applies to surface facilities and often refer to the signals at intersections etc.

Again capacity can be defined for a *point* or *uniform* section. Capacity is estimated for segments having uniform conditions. Points where these conditions change represent the boundaries where separate analysis may be required. Capacity is the maximum flow rate that a facility can afford. This maximum flow rate is taken for the worst 15 minutes of the peak hours while finding out the capacity. Capacity is measured as a reasonably expected value and not the maximum flow rate ever observed in the facility. This is because the

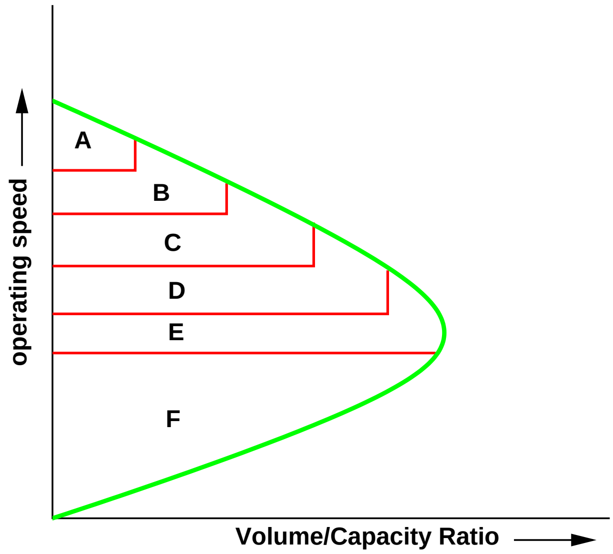


Figure 35:1: Level of service A to F

measured capacity at a single location will show significant variation from day to day. Further, local driving habits also produce variations in the observed capacity.

35.6 Factors affecting level of service

The level of service can be derived from a road under different operating characteristics and traffic volumes. The factors affecting level of service (LOS) can be listed as follows:

1. Speed and travel time
2. Traffic interruptions/restrictions
3. Freedom to travel with desired speed
4. Driver comfort and convenience
5. Operating cost.

Highway Capacity Manual(HCM) used travel speed and volume by capacity ratio (v/c ratio) to distinguish between various levels of service. The value of v/c ratio can vary between 0 and 1. Depending upon the travel speed and v/c ratio, HCM has defined six levels of service, level A to level F based on a graph between operating speed and v/c ratio as shown in the figure 35:1. Level of service A represents the zone of free flow. Here the traffic volume will be less, traffic will be experiencing free flow also. The drivers will be having the complete freedom to choose their desired speed. Even at maximum density, for this LOS the average spacing between vehicles is 167 m. Lane changes within the traffic stream, as well as merging and diverging movements, are made relatively easy. The effect of minor incidents and point breakdowns are easily aborted at this level. Level of service B represents zone of reasonably free flow. Free flow speeds are still maintained at this level of service. The drivers freedom to choose their desired speed is only slightly restricted. The lowest average spacing between vehicles is about 100 m. The effects of small incidents and point breakdowns are still easily contained.

At level of service C, the presence of other vehicles begins to restrict the maneuverability within the traffic stream. Average speeds remain at or near the free flow speed level, but significant increase in driver vigilance is required at this level. Minimum average spacing between the vehicles is in the range of 67 m. Queues may be expected to form behind any significant blockage. At level of service D, the average speeds begin to decline with increasing flows. Freedom to maneuver within the traffic stream is noticeably restricted. At this level, density deteriorates more quickly with flow. The spacing between the vehicles is about 50 m. As the traffic stream has little space to absorb disruptions, minor incidents can lead to queuing of vehicles. Level of service E define operation at capacity. At this level, the stream reaches it's maximum density limit. There will be no usable gaps in the stream and even slight disruptions will cause a breakdown, with queues forming rapidly behind the disruption. Maneuvering within the traffic stream becomes extremely difficult. Level of service F describes conditions in a queue that has formed behind a point of breakdown or disruption. As vehicles shuffle through the queue, there may be periods when they move quickly, and others when they are stopped completely. Thus this level of service is used to describe the point of breakdown as well, eventhough operations downstream of such a breakdown may appear good. Level of service F represents the region of forced flow, having low speed, and complete breakdown of the system .

35.7 Summary

Capacity and level of service are two important terms applied to traffic operation and are given suitable definitions by highway capacity manual. Capacity represents the ability of the system to handle traffic whereas level of service looks at the system from the drivers perspective. The fundamental diagrams of traffic flow can be used in the representation of level of service. Level of service ranges from level A to F, representing the free flow conditions and F representing the worst traffic conditions like less speed, high density etc.

35.8 Problems

1. The parameter which is not used under measure of effectiveness
   (a) Density
   (b) Delay
   (c) Speed
   (d) Flow
2. How many levels of service are defined by HCM?
   (a) Three
   (b) Six
   (c) Five
   (d) Four

35.9 Solutions

1. The parameter which is not used under measure of effectiveness

(a) Density

(b) Delay

(c) Speed

(d) Flow$\surd$

2. How many levels of service are defined by HCM?

(a) Three

(b) Six$\surd$

(c) Five

(d) Four

Chapter 36

Traffic signs

36.1 Overview

Traffic control device is the medium used for communicating between traffic engineer and road users. Unlike other modes of transportation, there is no control on the drivers using the road. Here traffic control devices comes to the help of the traffic engineer. The major types of traffic control devices used are- traffic signs, road markings , traffic signals and parking control. This chapter discusses traffic control signs. Different types of traffic signs are regulatory signs, warning signs and informatory signs.

36.2 Requirements of traffic control devices

1. **The control device should fulfill a need** : Each device must have a specific purpose for the safe and efficient operation of traffic flow. The superfluous devices should not be used.

2. **It should command attention from the road users**: This affects the design of signs. For commanding attention, proper visibility should be there. Also the sign should be distinctive and clear. The sign should be placed in such a way that the driver requires no extra effort to see the sign.

3. **It should convey a clear, simple meaning**: Clarity and simplicity of message is essential for the driver to properly understand the meaning in short time. The use of color, shape and legend as codes becomes important in this regard. The legend should be kept short and simple so that even a less educated driver could understand the message in less time.

4. **Road users must respect the signs**: Respect is commanded only when the drivers are conditioned to expect that all devices carry meaningful and important messages. Overuse, misuse and confusing messages of devices tends the drivers to ignore them.

5. **The control device should provide adequate time for proper response from the road users**: This is again related to the design aspect of traffic control devices. The sign boards should be placed at a distance such that the driver could see it and gets sufficient time to respond to the situation. For example, the STOP sign which is always placed at the stop line of the intersection should be visible for atleast one safe stopping sight distance away from the stop line.

36.3 Communication tools

A number of mechanisms are used by the traffic engineer to communicate with the road user. These mechanisms recognize certain human limitations, particularly eyesight. Messages are conveyed through the following elements.

1. **Color**: It is the first and most easily noticed characteristics of a device. Usage of different colors for different signs are important. The most commonly used colors are red, green, yellow, black, blue, and brown . These are used to code certain devices and to reinforce specific messages. Consistent use of colors helps the drivers to identify the presence of sign board ahead.

2. **Shape** : It is the second element discerned by the driver next to the color of the device. The categories of shapes normally used are circular, triangular, rectangular, and diamond shape. Two exceptional shapes used in traffic signs are octagonal shape for STOP sign and use of inverted triangle for GIVE WAY (YIELD) sign. Diamond shape signs are not generally used in India.

3. **Legend** : This is the last element of a device that the drive comprehends. This is an important aspect in the case of traffic signs. For the easy understanding by the driver, the legend should be short, simple and specific so that it does not divert the attention of the driver. Symbols are normally used as legends so that even a person unable to read the language will be able to understand that. There is no need of it in the case of traffic signals and road markings.

4. **Pattern**: It is normally used in the application of road markings, complementing traffic signs. Generally solid, double solid and dotted lines are used. Each pattern conveys different type of meaning. The frequent and consistent use of pattern to convey information is recommended so that the drivers get accustomed to the different types of markings and can instantly recognize them.

36.4 Types of traffic signs

There are several hundreds of traffic signs available covering wide variety of traffic situations. They can be classified into three main categories.

1. **Regulatory signs**: These signs require the driver to obey the signs for the safety of other road users.

2. **Warning signs**:These signs are for the safety of oneself who is driving and advice the drivers to obey these signs.

3. **Informative signs**: These signs provide information to the driver about the facilities available ahead, and the route and distance to reach the specific destinations

In addition special type of traffic sign namely *work zone signs* are also available. These type of signs are used to give warning to the road users when some construction work is going on the road. They are placed only for short duration and will be removed soon after the work is over and when the road is brought back to its normal condition. The first three signs will be discussed in detail below.

36.4.1 Regulatory signs

These signs are also called mandatory signs because it is mandatory that the drivers must obey these signs. If the driver fails to obey them, the control agency has the right to take legal action against the driver. These signs are primarily meant for the safety of other road users. These signs have generally black legend on a white background. They are circular in shape with red borders. The regulatory signs can be further classified into :

1. **Right of way series**: These include two unique signs that assign the right of way to the selected approaches of an intersection. They are the STOP sign and GIVE WAY sign For example, when one minor road and major road meets at an intersection, preference should be given to the vehicles passing through the major road. Hence the give way sign board will be placed on the minor road to inform the driver on the minor road that he should give way for the vehicles on the major road. In case two major roads are meeting, then the traffic engineer decides based on the traffic on which approach the sign board has to be placed. Stop sign is another example of regulatory signs that comes in right of way series which requires the driver to stop the vehicle at the stop line.

2. **Speed series**: Number of speed signs may be used to limit the speed of the vehicle on the road. They include typical speed limit signs, truck speed, minimum speed signs etc. Speed limit signs are placed to limit the speed of the vehicle to a particular speed for many reasons. Separate truck speed limits are applied on high speed roadways where heavy commercial vehicles must be limited to slower speeds than passenger cars for safety reasons. Minimum speed limits are applied on high speed roads like expressways, freeways etc. where safety is again a predominant reason. Very slow vehicles may present hazard to themselves and other vehicles also.

3. **Movement series**: They contain a number of signs that affect specific vehicle maneuvers. These include turn signs, alignment signs, exclusion signs, one way signs etc. Turn signs include turn prohibitions and lane use control signs. Lane use signs make use of arrows to specify the movements which all vehicles in the lane must take. Turn signs are used to safely accommodate turns in unsignalized intersections.

4. **Parking series**: They include parking signs which indicate not only parking prohibitions or restrictions, but also indicate places where parking is permitted, the type of vehicle to be parked, duration for parking etc.

5. **Pedestrian series**: They include both legend and symbol signs. These signs are meant for the safety of pedestrians and include signs indicating pedestrian only roads, pedestrian crossing sites etc.

6. **Miscellaneous**: Wide variety of signs that are included in this category are: a "KEEP OF MEDIAN" sign, signs indicating road closures, signs restricting vehicles carrying hazardous cargo or substances, signs indicating vehicle weight limitations etc.

Some examples of the regulatory signs are shown in figure 36:1. They include a stop sign, give way sign, signs for no entry, sign indicating prohibition for right turn, vehicle width limit sign, speed limit sign etc.

36.4.2 Warning signs

Warning signs or cautionary signs give information to the driver about the impending road condition. They advice the driver to obey the rules. These signs are meant for the own safety of drivers. They call for extra vigilance from the part of drivers. The color convention used for this type of signs is that the legend will be black

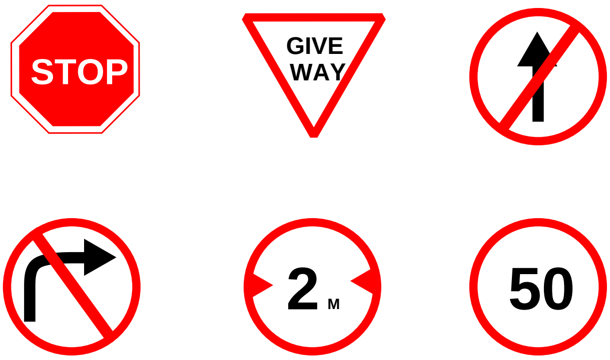


Figure 36:1: Examples of regulatory signs (stop sign, give way sign, signs for no entry, sign indicating prohibition for right turn, vehicle width limit sign, speed limit sign)

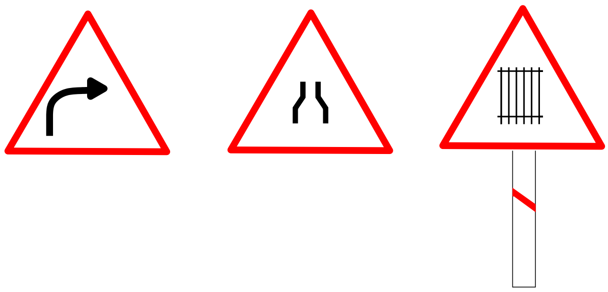


Figure 36:2: Examples of cautionary signs (right hand curve sign board, signs for narrow road, sign indicating railway track ahead)

in color with a white background. The shape used is upward triangular or diamond shape with red borders. Some of the examples for this type of signs are given in fig 36:2 and includes right hand curve sign board, signs for narrow road, sign indicating railway track ahead etc.

36.4.3 Informative signs

Informative signs also called guide signs, are provided to assist the drivers to reach their desired destinations. These are predominantly meant for the drivers who are unfamiliar to the place. The guide signs are redundant for the users who are accustomed to the location.

Some of the examples for these type of signs are route markers, destination signs, mile posts, service information, recreational and cultural interest area signing etc. Route markers are used to identify numbered highways. They have designs that are distinctive and unique. They are written black letters on yellow background. Destination signs are used to indicate the direction to the critical destination points, and to mark important intersections. Distance in kilometers are sometimes marked to the right side of the destination. They are, in general, rectangular with the long dimension in the horizontal direction. They are color coded as white letters with green background.

Mile posts are provided to inform the driver about the progress along a route to reach his destination. Service guide signs give information to the driver regarding various services such as food, fuel, medical assistance etc. They are written with white letters on blue background. Information on historic, recreational and other cultural

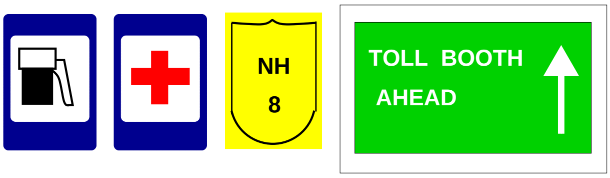


Figure 36:3: Examples of informative signs (route markers, destination signs, mile posts, service centre information etc)

area is given on white letters with brown background. In the figure 36:3 we can see some examples for informative signs which include route markers, destination signs, mile posts, service centre information etc..

36.5 Summary

Traffic signs are means for exercising control on or passing information to the road users. They may be regulatory, warning, or informative. Among the design aspects of the signs, the size, shape, color and location matters. Some of the signs along with examples were discussed in this chapter. A few web sites discussing on traffic signs are giben below:

www.aptransport.org/html/signs.htm, www.indiacar.com/infobank/Traffic-signs.htm.

36.6 Problems

1. Regulatory signs are also called
   (a) Mandatory signs
   (b) Cautionary signs
   (c) Informative signs
   (d) Warning signs
2. Stop sign comes under
   (a) Regulatory signs
   (b) Cautionary signs
   (c) Informative signs
   (d) none of these

36.7 Solutions

1. Regulatory signs are also called
   (a) Mandatory signs$\surd$
   (b) Cautionary signs
   (c) Informative signs

   (d) Warning signs

2. Stop sign comes under

   (a) Regulatory signs$\surd$
   (b) Cautionary signs
   (c) Informative signs
   (d) none of these

Chapter 37

Road markings

37.1 Overview

The essential purpose of road markings is to guide and control traffic on a highway. They supplement the function of traffic signs. The markings serve as a psychological barrier and signify the delineation of traffic path and its lateral clearance from traffic hazards for the safe movement of traffic. Hence they are very important to ensure the safe, smooth and harmonious flow of traffic. Various types of road markings like longitudinal markings, transverse markings, object markings and special markings to warn the driver about the hazardous locations in the road etc. will be discussed in detail in this chapter.

37.2 Classification of road markings

The road markings are defined as lines, patterns, words or other devices, except signs, set into applied or attached to the carriageway or kerbs or to objects within or adjacent to the carriageway, for controlling, warning, guiding and informing the users. The road markings are classified as longitudinal markings, transverse markings, object markings, word messages, marking for parkings, marking at hazardous locations etc.

37.3 Longitudinal markings

Longitudinal markings are placed along the direction of traffic on the roadway surface, for the purpose of indicating to the driver, his proper position on the roadway. Some of the guiding principles in longitudinal markings are also discussed below.

Longitudinal markings are provided for separating traffic flow in the same direction and the predominant color used is white. Yellow color is used to separate the traffic flow in opposite direction and also to separate the pavement edges. The lines can be either broken, solid or double solid. Broken lines are permissive in character and allows crossing with discretion, if traffic situation permits. Solid lines are restrictive in character and does not allow crossing except for entry or exit from a side road or premises or to avoid a stationary obstruction. Double solid lines indicate severity in restrictions and should not be crossed except in case of emergency. There can also be a combination of solid and broken lines. In such a case, a solid line may be crossed with discretion, if the broken line of the combination is nearer to the direction of travel. Vehicles from the opposite directions are not permitted to cross the line. Different types of longitudinal markings are centre line, traffic lanes, no passing zone, warning lines, border or edge lines, bus lane markings, cycle lane markings.

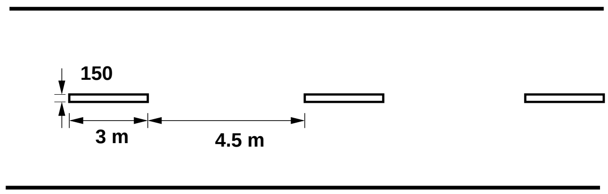


Figure 37:1: Centre line marking for a two lane road

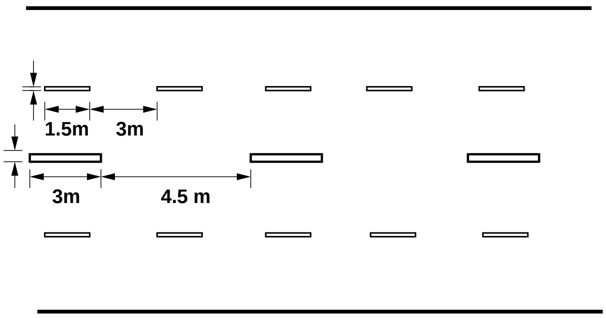


Figure 37:2: Centre line and lane marking for a four lane road

37.3.1 Centre line

Centre line separates the opposing streams of traffic and facilitates their movements. Usually no centre line is provided for roads having width less than 5 m and for roads having more than four lanes. The centre line may be marked with either single broken line, single solid line, double broken line, or double solid line depending upon the road and traffic requirements. On urban roads with less than four lanes, the centre line may be single broken line segments of 3 m long and 150 mm wide. The broken lines are placed with 4.5 m gaps (figure 37:1). On curves and near intersections, gap shall be reduced to 3 metres. On undivided urban roads with at least two traffic lanes in each direction, the centre line marking may be a single solid line of 150 mm wide as in figure 37:2, or double solid line of 100 mm wide separated by a space of 100 mm as shown in figure 37:3. The centre barrier line marking for four lane road is shown in figure 37:4.

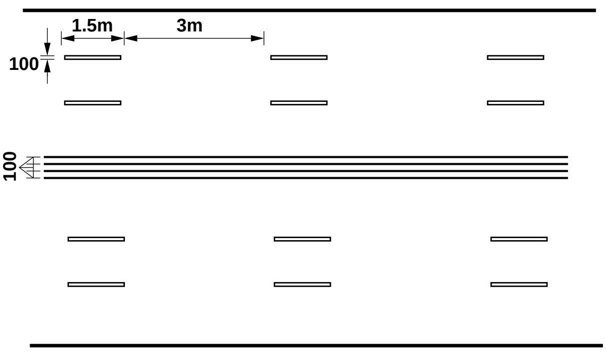


Figure 37:3: Double solid line for a two lane road

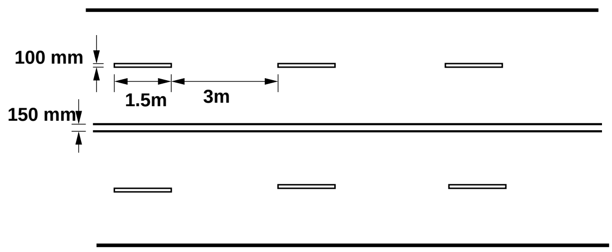


Figure 37:4: Centre barrier line marking for four lane road

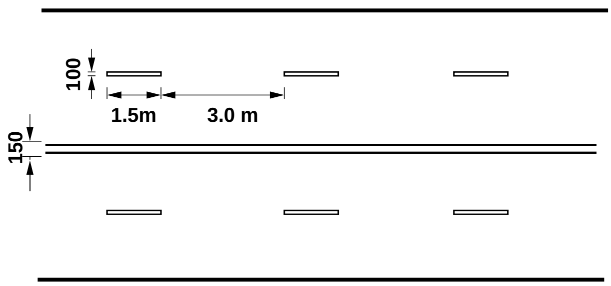


Figure 37:5: Lane marking for a four lane road with solid barrier line

37.3.2 Traffic lane lines

The subdivision of wide carriageways into separate lanes on either side of the carriage way helps the driver to go straight and also curbs the meandering tendency of the driver. At intersections, these traffic lane lines will eliminate confusion and facilitates turning movements. Thus traffic lane markings help in increasing the capacity of the road in addition ensuring more safety. The traffic lane lines are normally single broken lines of 100 mm width. Some examples are shown in figure 37:5 and figure 37:6.

37.3.3 No passing zones

No passing zones are established on summit curves, horizontal curves, and on two lane and three lane highways where overtaking maneuvers are prohibited because of low sight distance. It may be marked by a solid yellow line along the centre or a double yellow line. In the case of a double yellow line, the left hand element may be a solid barrier line, the right hand may be a either a broken line or a solid line . These solid lines are also called barrier lines. When a solid line is to the right of the broken line, the passing restriction shall apply only to the

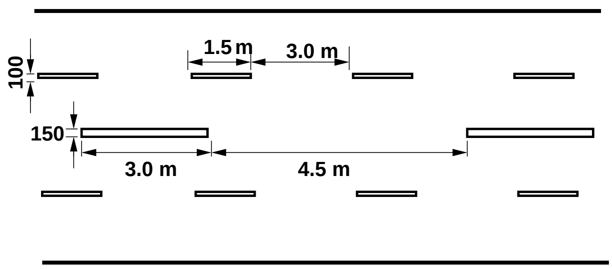


Figure 37:6: Traffic lane marking for a four lane road with broken centre line

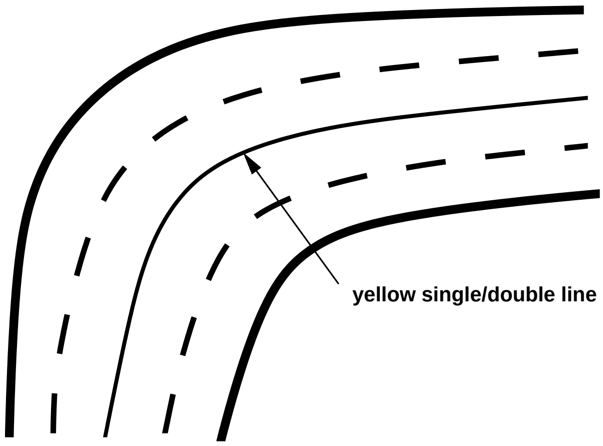


Figure 37:7: Barrier line marking for a four lane road

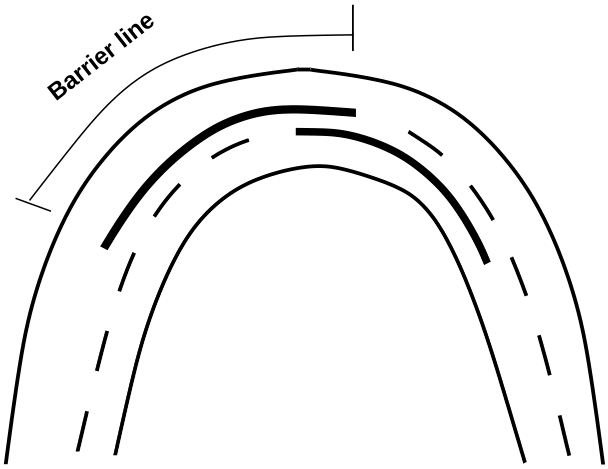


Figure 37:8: No passing zone marking at horizontal curves

opposing traffic. Some typical examples are shown in figure 37:7 and figure 37:8. In the latter case, the no passing zone is staggered for each direction.

37.3.4 Warning lines

Warning lines warn the drivers about the obstruction approaches. They are marked on horizontal and vertical curves where the visibility is greater than prohibitory criteria specified for no overtaking zones. They are broken lines with 6 m length and 3 m gap. A minimum of seven line segments should be provided. A typical example is shown in figure 37:9

37.3.5 Edge lines

Edge lines indicate edges of rural roads which have no kerbs to delineate the limits upto which the driver can safely venture. They should be at least 150 mm from the actual edge of the pavement. They are painted in yellow or white.

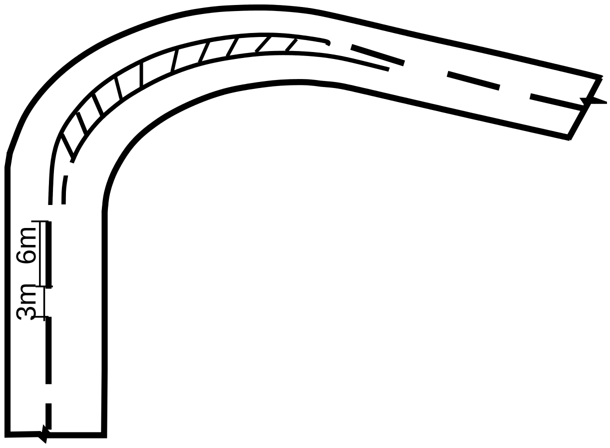


Figure 37:9: Warning line marking for a two lane road

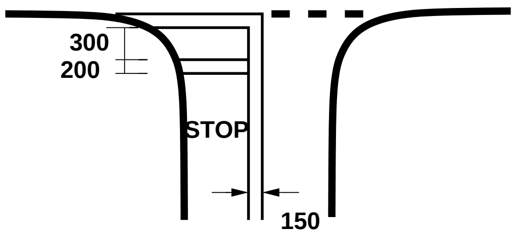


Figure 37:10: Stop line marking near an intersection

All the lines should be preferably light reflective, so that they will be visible during night also. Improved night visibility may also be obtained by the use of minute glass beads embedded in the pavement marking materials to produce a retroreflective surface.

37.4 Transverse markings

Transverse markings are marked across the direction of traffic. They are marked at intersections etc. The site conditions play a very important role. The type of road marking for a particular intersection depends on several variables such as speed characteristics of traffic, availability of space etc. Stop line markings, markings for pedestrian crossing, direction arrows, etc. are some of the markings on approaches to intersections.

37.4.1 Stop line

Stop line indicates the position beyond which the vehicles should not proceed when required to stop by control devices like signals or by traffic police. They should be placed either parallel to the intersecting roadway or at right angles to the direction of approaching vehicles. An example for a stop line marking is shown in figure 37:10.

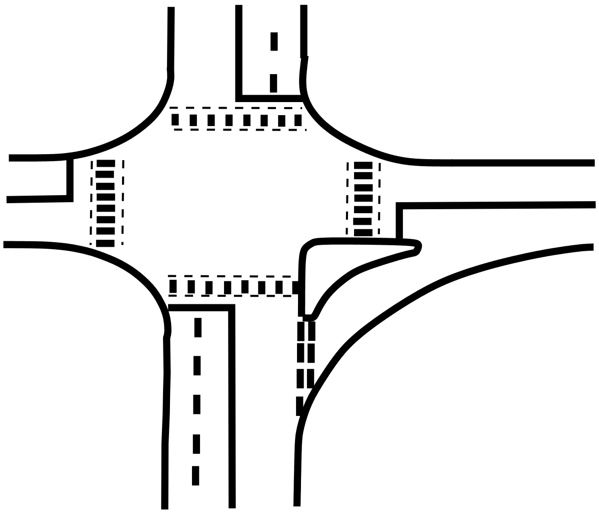


Figure 37:11: Pedestrian marking near an intersection

37.4.2 Pedestrian crossings

Pedestrian crossings are provided at places where the conflict between vehicular and pedestrian traffic is severe. The site should be selected that there is less inconvenience to the pedestrians and also the vehicles are not interrupted too much. At intersections, the pedestrian crossings should be preceded by a stop line at a distance of 2 to 3m for unsignalized intersections and at a distance of one metre for signalized intersections. Most commonly used pattern for pedestrian crossing is Zebra crossing consisting of equally spaced white strips of 500 mm wide. A typical example of an intersection illustrating pedestrian crossings is shown in figure 37:11.

37.4.3 Directional arrows

In addition to the warning lines on approaching lanes, directional arrows should be used to guide the drivers in advance over the correct lane to be taken while approaching busy intersections. Because of the low angle at which the markings are viewed by the drivers, the arrows should be elongated in the direction of traffic for adequate visibility. The dimensions of these arrows are also very important. A typical example of a directional arrow is shown in figure 37:12.

37.5 Object marking

Physical obstructions in a carriageway like traffic island or obstructions near carriageway like signal posts, pier etc. cause serious hazard to the flow of traffic and should be adequately marked. They may be marked on the objects adjacent to the carriageway.

37.5.1 Objects within the carriageway

The obstructions within the carriageway such as traffic islands, raised medians, etc. may be marked by not less than five alternate black and yellow stripes. The stripes should slope forward at an angle of 45° with respect to

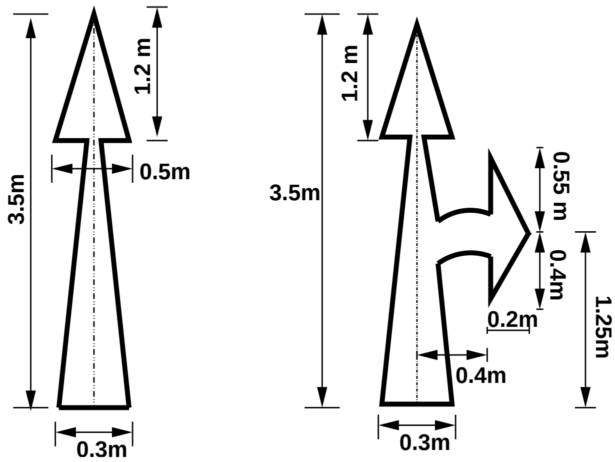


Figure 37:12: Directional arrow marking

the direction of traffic. These stripes shall be uniform and should not be less than 100 m wide so as to provide sufficient visibility.

37.5.2 Objects adjacent to carriageway

Sometimes objects adjacent to the carriageway may pose some obstructions to the flow of traffic. Objects such as subway piers and abutments, culvert head walls etc. are some examples for such obstructions. They should be marked with alternate black and white stripes at a forward angle of 45° with respect to the direction of traffic. Poles close to the carriageway should be painted in alternate black and white up to a height of 1.25 m above the road level. Other objects such as guard stones, drums, guard rails etc. where chances of vehicles hitting them are only when vehicle runs off the carriageway should be painted in solid white. Kerbs of all islands located in the line of traffic flow shall be painted with either alternating black and white stripes of 500 mm wide or chequered black and white stripes of same width. The object marking for central pier and side walls of an underpass is illustrated in figure 37:13.

37.6 Word messages

Information to guide, regulate, or warn the road user may also be conveyed by inscription of word message on road surface. Characters for word messages are usually capital letters. The legends should be as brief as possible and shall not consist of more than three words for any message. Word messages require more and important time to read and comprehend than other road markings. Therefore, only few and important ones are usually adopted. Some of the examples of word messages are STOP, SLOW, SCHOOL, RIGHT TUN ONLY etc. The character of a road message is also elongated so that driver looking at the road surface at a low angle can also read them easily. The dimensioning of a typical alphabet is shown in figure 37:14.

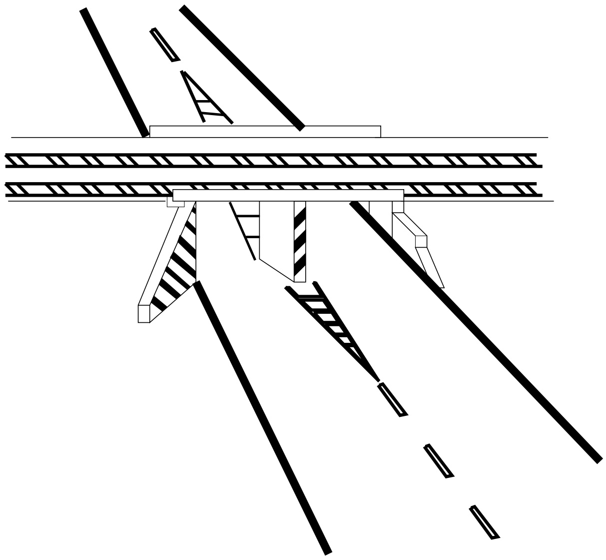


Figure 37:13: Marking for objects adjacent to the road way

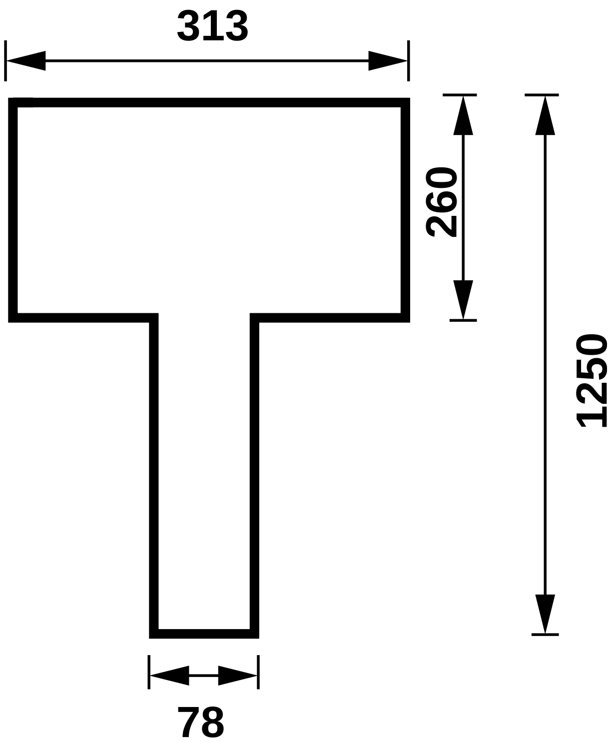


Figure 37:14: Typical dimension of the character T used in road marking

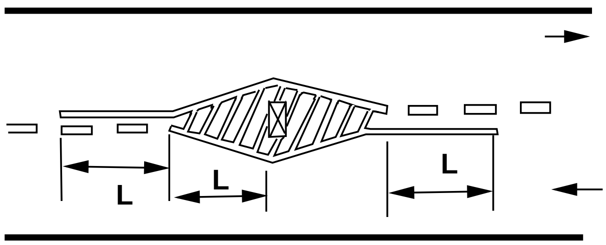


Figure 37:15: Approach marking for obstructions on the road way

37.7 Parking

The marking of the parking space limits on urban roads promotes more efficient use of the parking spaces and tends to prevent encroachment on places like bus stops, fire hydrant zones etc. where parking is undesirable. Such parking space limitations should be indicated with markings that are solid white lines 100 mm wide. Words TAXI, CARS, SCOOTERS etc. may also be written if the parking area is specific for any particular type of vehicle. To indicate parking restriction, kerb or carriage way marking of continuous yellow line 100 mm wide covering the top of kerb or carriageway close to it may be used.

37.8 Hazardous location

Wherever there is a change in the width of the road, or any hazardous location in the road, the driver should be warned about this situation with the help of suitable road markings. Road markings showing the width transition in the carriageway should be of 100 mm width. Converging lines shall be 150 mm wide and shall have a taper length of not less than twenty times the off-set distance. Typical carriageway markings showing transition from wider to narrower sections and vice-versa is shown in figure 37:15. In the figure, the driver is warned about the position of the pier through proper road markings.

37.9 Summary

Road markings are aids to control traffic by exercising psychological control over the road users. They are made use of in delineating the carriage way as well as marking obstructions, to ensure safe driving. They also assist safe pedestrian crossing. Longitudinal markings which are provided along the length of the road and its various classifications were discussed. Transverse markings are provided along the width of the road. Road markings also contain word messages, but since it is time consuming to understand compared to other markings there are only very few of them. Markings are also used to warn the driver about the hazardous locations ahead. Thus road markings ensure smooth flow of traffic providing safety also to the road users. The following web link give further insight in to the road markings: mutcd.fhwa.dot.gov/pdfs/200311/pdf-index.htm.

37.10 Problems

1. Broken lines

   (a) allows crossing with discretion

(b) does not allow crossing except for entry or exit from a side road

(c) allows crossing only in case of extreme emergency

(d) are not at all used as road markings.

2. Stop line comes under

   (a) Longitudinal markings

   (b) Object markings

   (c) Transverse markings

   (d) None of these

37.11 Solutions

1. Broken lines

   (a) allows crossing with discretion√

   (b) does not allow crossing except for entry or exit from a side road

   (c) allows crossing only in case of extreme emergency

   (d) are not at all used as road markings.

2. Stop line comes under

   (a) Longitudinal markings

   (b) Object markings

   (c) Transverse markings√

   (d) None of these

Chapter 38

Parking

38.1 Overview

Parking is one of the major problems that is created by the increasing road traffic. It is an impact of transport development. The availability of less space in urban areas has increased the demand for parking space especially in areas like Central business district. This affects the mode choice also. This has a great economical impact.

38.2 Parking studies

Before taking any measures for the betterment of conditions, data regarding availability of parking space, extent of its usage and parking demand is essential. It is also required to estimate the parking fares also. Parking surveys are intended to provide all these information. Since the duration of parking varies with different vehicles, several statistics are used to access the parking need.

38.2.1 Parking statistics

Parking accumulation: It is defined as the number of vehicles parked at a given instant of time. Normally this is expressed by accumulation curve. Accumulation curve is the graph obtained by plotting the number of bays occupied with respect to time.

Parking volume: Parking volume is the total number of vehicles parked at a given duration of time. This does not account for repetition of vehicles. The actual volume of vehicles entered in the area is recorded.

Parking load : Parking load gives the area under the accumulation curve. It can also be obtained by simply multiplying the number of vehicles occupying the parking area at each time interval with the time interval. It is expressed as vehicle hours.

Average parking duration: It is the ratio of total vehicle hours to the number of vehicles parked.

$parkingduration = \frac{parkingload}{parkingvolume}$

Parking turnover: It is the ratio of number of vehicles parked in a duration to the number of parking bays available.

$parkingturnover = \frac{parkingvolume}{No.ofbaysavailable}$

This can be expressed as number of vehicles per bay per time duration.

Parking index: Parking index is also called occupancy or efficiency. It is defined as the ratio of number of bays occupied in a time duration to the total space available. It gives an aggregate measure of how effectively the parking space is utilized. Parking index can be found out as follows

$$\text{parking index} = \frac{\text{parking load}}{\text{parking capacity}} \times 100 \tag{38.1}$$

To illustrate the various measures, consider a small example in figure 38:1, which shows the duration for which each of the bays are occupied(shaded portion). Now the accumulation graph can be plotted by simply noting the number of bays occupied at time interval of 15, 30, 45 etc. minutes ias shown in the figure.

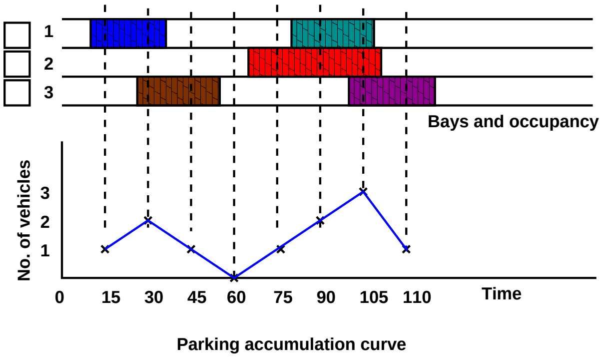


Figure 38:1: Parking bays and accumulation curve

The various measures are calculated as shown below:
Parking volume= 5 vehicles.
Parking load $= (1+2+1+0+1+2+3+1)\frac{15}{60} = \frac{11\times15}{60} = 2.75$ veh hour.
Average parking duration $= \frac{2.75\ veh\ hours}{5veh} = 33$ minutes.
Parking turnover $= \frac{5\ veh/2\ hours}{3bays} = 0.83$ veh/hr/bay.
Parking index $= \frac{2.75\ veh\ hour}{3\times2\ veh\ hours} \times 100 = 45.83\%$

38.3 Parking surveys

Parking surveys are conducted to collect the above said parking statistics. The most common parking surveys conducted are in-out survey, fixed period sampling and license plate method of survey.

1. **In-out survey**: In this survey, the occupancy count in the selected parking lot is taken at the beginning. Then the number of vehicles that enter the parking lot for a particular time interval is counted. The number of vehicles that leave the parking lot is also taken. The final occupancy in the parking lot is also taken. Here the labor required is very less. Only one person may be enough. But we wont get any data regarding the time duration for which a particular vehicle used that parking lot. Parking duration and turn over is not obtained. Hence we cannot estimate the parking fare from this survey.

2. **Fixed period sampling**: This is almost similar to in-out survey. All vehicles are counted at the beginning of the survey. Then after a fixed time interval that may vary between 15 minutes to i hour, the count is again taken. Here there are chances of missing the number of vehicles that were parked for a short duration.

3. **License plate method of survey**: This results in the most accurate and realistic data. In this case of survey, every parking stall is monitored at a continuous interval of 15 minutes or so and the license plate number is noted down. This will give the data regarding the duration for which a particular vehicle was using the parking bay. This will help in calculating the fare because fare is estimated based on the duration for which the vehicle was parked. If the time interval is shorter, then there are less chances of missing short-term parkers. But this method is very labor intensive.

38.4 Ill effects of parking

Parking has some ill-effects like congestion, accidents, pollution, obstruction to fire-fighting operations etc.

Congestion: Parking takes considerable street space leading to the lowering of the road capacity. Hence, speed will be reduced, journey time and delay will also subsequently increase. The operational cost of the vehicle increases leading to great economical loss to the community.

Accidents: Careless maneuvering of parking and unparking leads to accidents which are referred to as parking accidents. Common type of parking accidents occur while driving out a car from the parking area, careless opening of the doors of parked cars, and while bringing in the vehicle to the parking lot for parking.

Environmental pollution: They also cause pollution to the environment because stopping and starting of vehicles while parking and unparking results in noise and fumes. They also affect the aesthetic beauty of the buildings because cars parked at every available space creates a feeling that building rises from a plinth of cars.

Obstruction to fire fighting operations: Parked vehicles may obstruct the movement of firefighting vehicles. Sometimes they block access to hydrants and access to buildings.

38.5 Parking requirements

There are some minimum parking requirements for different types of building. For residential plot area less than 300 sq.m require only community parking space. For residential plot area from 500 to 1000 sq.m, minimum one-fourth of the open area should be reserved for parking. Offices may require atleast one space for every 70 sq.m as parking area. One parking space is enough for 10 seats in a restaurant where as theatres and cinema halls need to keep only 1 parking space for 20 seats. Thus, the parking requirements are different for different land use zones.

38.6 On street parking

On street parking means the vehicles are parked on the sides of the street itself. This will be usually controlled by government agencies itself. Common types of on-street parking are as listed below. This classification is based on the angle in which the vehicles are parked with respect to the road alignment. As per IRC the standard dimensions of a car is taken as 5$\times$ 2.5 metres and that for a truck is 3.75$\times$ 7.5 metres.

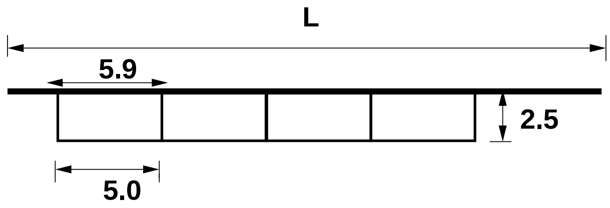


Figure 38:2: Illustration of parallel parking

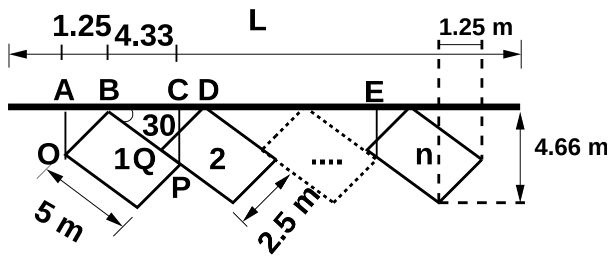


Figure 38:3: Illustration of 30° parking

Parallel parking: The vehicles are parked along the length of the road. Here there is no backward movement involved while parking or unparking the vehicle. Hence, it is the most safest parking from the accident perspective. However, it consumes the maximum curb length and therefore only a minimum number of vehicles can be parked for a given kerb length. This method of parking produces least obstruction to the on-going traffic on the road since least road width is used. Parallel parking of cars is shown in figure 38:2. The length available to park N number of vehicles, L = $\frac{N}{5.9}$

30° parking: In thirty degree parking, the vehicles are parked at 30° with respect to the road alignment. In this case, more vehicles can be parked compared to parallel parking. Also there is better maneuverability. Delay caused to the traffic is also minimum in this type of parking. An example is shown in figure 38:3. From the figure,

$$
\begin{aligned}
AB &= OBsin30^\circ = && 1.25, \\
BC &= OPcos30^\circ = && 4.33, \\
BD &= DQcos60^\circ = && 5, \\
CD &= BD - BC = 5 - 4.33 = && 0.67, \\
AB + BC &= 1.25 + 4.33 = && 5.58
\end{aligned}
$$

For N vehicles, L = AC + (N-1)CE =5.58+(N-1)5 =0.58+5N

45° parking: As the angle of parking increases, more number of vehicles can be parked. Hence compared to parallel parking and thirty degree parking, more number of vehicles can be accommodated in this type of parking. From figure 38:4, length of parking space available for parking N number of vehicles in a given kerb is L = 3.54 N+1.77

60° parking: The vehicles are parked at 60° to the direction of road. More number of vehicles can be accommodated in this parking type. From the figure 38:5, length available for parking N vehicles =2.89N+2.16.

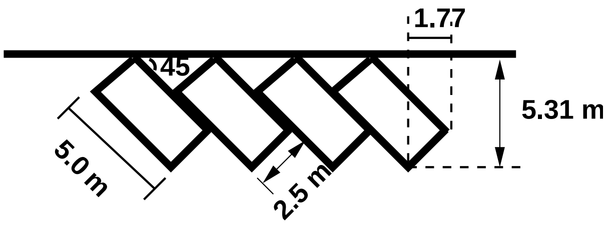


Figure 38:4: Illustration of 45° parking

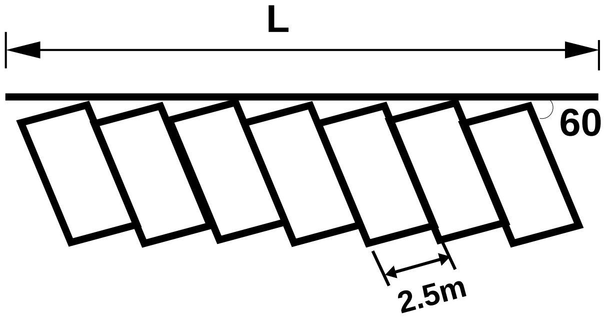


Figure 38:5: Illustration of 60° parking

Right angle parking: In right angle parking or 90° parking, the vehicles are parked perpendicular to the direction of the road. Although it consumes maximum width kerb length required is very little. In this type of parking, the vehicles need complex maneuvering and this may cause severe accidents. This arrangement causes obstruction to the road traffic particularly if the road width is less. However, it can accommodate maximum number of vehicles for a given kerb length. An example is shown in figure 38:6. Length available for parking N number of vehicles is L = 2.5N.

38.7 Off street parking

In many urban centres, some areas are exclusively allotted for parking which will be at some distance away from the main stream of traffic. Such a parking is referred to as off-street parking. They may be operated by either public agencies or private firms. A typical layout of an off-street parking is shown in figure 38:7.

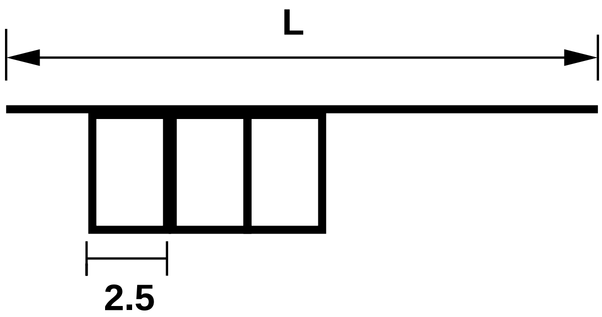


Figure 38:6: Illustration of 90° parking

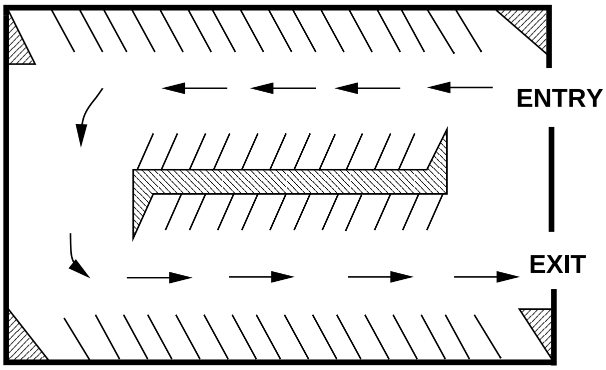


Figure 38:7: Illustration of off-street parking

Example 1

From an in-out survey conducted for a parking area consisting of 40 bays, the initial count was found to be 25. Table gives the result of the survey. The number of vehicles coming in and out of the parking lot for a time interval of 5 minutes is as shown in the table 38:1. Find the accumulation, total parking load, average occupancy and efficiency of the parking lot.

Table 38:1: In-out survey data

Time	In	Out
5	3	2
10	2	4
15	4	2
20	5	4
25	7	3
30	8	2
35	2	7
40	4	2
45	6	4
50	4	1
55	3	3
60	2	5

Solution The solution is shown in table 38:2

- Accumulation can be found out as initial count plus number of vehicles that entered the parking lot till that time minus the number of vehicles that just exited for that particular time interval. For the first time interval of 5 minutes, accumulation can be found out as 25+3-2 = 26. It is being tabulated in column 4.
- Occupancy or parking index is given by equation For the first time interval of five minutes, $Parking\ index =$

Table 38:2: In-out parking survey solution

Time (1)	In (2)	Out (3)	Accumulation (4)	Occupancy (5)	Parking load (6)
5	3	2	26	65	130
10	2	4	24	60	120
15	4	2	26	65	130
20	5	4	27	67.5	135
25	7	3	31	77.5	155
30	8	2	37	92.5	185
35	2	7	32	80	160
40	4	2	34	85	170
45	6	4	36	90	180
50	4	1	39	97.5	195
55	3	3	39	97.5	195
60	2	5	36	90	180
Total					1735

$\frac{26}{40} \times 100 = 65\%$. The occupancy for the remaining time slot is similarly calculated and is tabulated in column 5.

Average occupancy is the average of the occupancy values for each time interval. Thus it is the average of all values given in column 5 and the value is 80.63%.

- Parking load is tabulated in column 6. It is obtained by multiplying accumulation with the time interval. For the first time interval, parking load = 26 $\times$ 5 = 130 vehicle minutes.
- Total parking load is the summation of all the values in column 5 which is equal to 1935 vehicle minutes or 32.25 vehicle hours

Example 2

The parking survey data collected from a parking lot by license plate method is s shown in the table 38:3 below. Find the average occupancy, average turn over, parking load, parking capacity and efficiency of the parking lot.

Solution See the following table for solution 38:4. Columns 1 to 5 is the input data. The parking status in every bay is coded first. If a vehicle occupies that bay for that time interval, then it has a code 1. This is shown in columns 6, 7, 8 and 9 of the table corresponding to the time intervals 15, 30, 45 and 60 seconds.

- Turn over is computed as the number of vehicles present in that bay for that particular hour. For the first bay, it is counted as 3. Similarly, for the second bay, one vehicle is present throughout that hour and hence turnout is 1 itself. This is being tabulated in column 10 of the table. Average turn over = $\frac{\text{Sum of turn-over}}{\text{Total number of bays}} = 2.25$

Table 38:3: Licence plate parking survey data

Bay	Time			
	0-15	15-30	30-45	45-60
1	1456	9813	-	5678
2	1945	1945	1945	1945
3	3473	5463	5463	5463
4	3741	3741	9758	4825
5	1884	1884	-	7594
6	-	7357	-	7893
7	-	4895	4895	4895
8	8932	8932	8932	-
9	7653	7653	8998	4821
10	7321	-	2789	2789
11	1213	1213	3212	4778
12	5678	6678	7778	8888

- Accumulation for a time interval is the total of number of vehicles in the bays 1 to 12 for that time interval. Accumulation for first time interval of 15 minutes = 1+1+1+1+1+0+0+1+1+1+1+1 = 10
- Parking volume = Sum of the turn over in all the bays = 27 vehicles
- Average duration is the average time for which the parking lot was used by the vehicles. It can be calculated as sum of the accumulation for each time interval $\times$ time interval divided by the parking volume $=\frac{(10+11+9+11)\times 15}{27} = 22.78$ minutes/vehicle.
- Occupancy for that time interval is accumulation in that particular interval divided by total number of bays. For first time interval of 15 minutes, occupancy = $(10\times 100)/12 = 83\%$ Average occupancy is found out as the average of total number of vehicles occupying the bay for each time interval. It is expressed in percentage. Average occupancy = $\frac{0.83+0.92+0.75+0.92}{4} \times 100 = 85.42\%$.
- Parking capacity = number of bays $\times$ number of hours = $12\times 1 = 12$ vehicle hours
- Parking load = total number of vehicles accumulated at the end of each time interval $\times$ time = $\frac{(10+11+9+11)\times 15}{60}$ = 10.25 vehicle hours
- Efficiency = $\frac{\text{Parking load}}{\text{Total number of bays}} = \frac{10.25}{12} = 85.42\%$.

38.8 Summary

Providing suitable parking spaces is a challenge for traffic engineers and planners in the scenario of ever increasing vehicle population. It is essential to conduct traffic surveys in order to design the facilities or plan the fares. Different types of parking layout, surveys and statistics were discussed in this chapter.

Table 38:4: Licence plate parking survey solution

Bay	Time				Time				
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	15	30	45	60	15	30	45	60	Turn over
1	1456	9813	-	5678	1	1	0	1	3
2	1945	1945	1945	1945	1	1	1	1	1
3	3473	5463	5463	5463	1	1	1	1	2
4	3741	3741	9758	4825	1	1	1	1	3
5	1884	1884	-	7594	1	1	0	1	2
6	-	7357	-	7893	0	1	0	1	2
7	-	4895	4895	4895	0	1	1	1	1
8	8932	8932	8932	-	1	1	1	0	1
9	7653	7653	8998	4821	1	1	1	1	3
10	7321	-	2789	2789	1	0	1	1	2
11	1213	1213	3212	4778	1	1	1	1	3
12	5678	6678	7778	8888	1	1	1	1	4
	Accumulation				10	11	9	11	
	Occupancy				0.83	0.92	0.75	0.92	2.25

38.9 Problems

1. The parking survey data collected from a parking lot by license plate method is shown in table 38:5 below. Find the average occupancy, average turnover, parking load, parking capacity and efficiency of parking lot.

Table 38:5: Licence plate: problem

Bay	Time			
	0-15	15-30	30-45	45-60
1	1501	1501	4021	-
2	1255	1255	1255	1255
3	3215	3215	3215	3215
4	-	-	3100	3100
5	1623	1623	1623	-
6	2204	2204	-	-

Solution Refer table 38:6. Column 1 to 5 is the input data. The parking status in every bay is coded first. If a vehicle occupies that bay for that time interval, then it has a code 1. This is shown in columns 6, 7, 8 and 9 of the tables corresponding to the time intervals 15,30,45 and 60 seconds.

Table 38:6: License Plate Problem: Solution

Bay	Time				Time				
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	15	30	45	60	15	30	45	60	Turn over
1	1501	1501	4021	-	1	1	1	0	2
2	1255	1255	1255	1255	1	1	1	1	1
3	3215	3215	3215	3215	1	1	1	1	1
4	-	-	3100	3100	0	0	1	1	1
5	1623	1623	1623	-	1	1	1	0	1
6	2204	2204	-	-	1	1	0	0	1
	Accumulation				5	5	5	3	
	Occupancy				0.83	0.83	0.83	0.5	

- Turn over is computed as the number of vehicles present in that bay for that particular hour. For the first bay, it is counted as 2. Similarly, for the second bay, one vehicle is present throughout that hour and hence turnout is 1 itself This is being tabulated in column 10 of the table. Total turn over in all the bays or parking volume= 2+1+1+1+1+1 = 7 vehicles Average turn over = $\frac{\text{Sum of turn−over}}{\text{Total number of bays}} = \frac{7}{6}$=1.17
- Average duration is the average time for which the parking lot was used by the vehicles. It can be calculated as sum of the accumulation for each time interval $\times$ time interval divided by the parking volume $=\frac{(5+5+5+3)\times 15}{7} = 38.57$ minutes/vehicle.
- Average occupancy is found out as the average of total number of vehicles occupying the bay for each time interval. It is expressed in percentage. Average occupancy $= \frac{0.83+0.83+0.83+0.5}{4} \times 100 = 75\%$.
- Parking capacity = number of bays $\times$ number of hours = $6\times 1 = 6$ vehicle hours
- Parking load = total number of vehicles accumulated at the end of each time interval $\times$ time = $\frac{(5+5+5+3)\times 15}{60}$ = 4.5 vehicle hours
- Efficiency = $\frac{\text{Parking load}}{\text{Total number of bays}} = \frac{4.5}{6} = 75\%$.

Chapter 39

Traffic intersections

39.1 Overview

Intersection is an area shared by two or more roads. This area is designated for the vehicles to turn to different directions to reach their desired destinations. Its main function is to guide vehicles to their respective directions. Traffic intersections are complex locations on any highway. This is because vehicles moving in different direction wan to occupy same space at the same time. In addition, the pedestrians also seek same space for crossing. Drivers have to make split second decision at an intersection by considering his route, intersection geometry, speed and direction of other vehicles etc. A small error in judgment can cause severe accidents. It also causes delay and it depends on type, geometry, and type of control. Overall traffic flow depends on the performance of the intersections. It also affects the capacity of the road. Therefore, both from the accident perspective and the capacity perspective, the study of intersections very important for the traffic engineers especially in the case of urban scenario.

39.2 Conflicts at an intersection

Conflicts at an intersection are different for different types of intersection. Consider a typical four-legged intersection as shown in figure. The number of conflicts for competing through movements are 4, while competing right turn and through movements are 8. The conflicts between right turn traffics are 4, and between left turn and merging traffic is 4. The conflicts created by pedestrians will be 8 taking into account all the four approaches. Diverging traffic also produces about 4 conflicts. Therefore, a typical four legged intersection has about 32 different types of conflicts. This is shown in figure 39:1.

The essence of the intersection control is to resolve these conflicts at the intersection for the safe and efficient movement of both vehicular traffic and pedestrians. Two methods of intersection controls are there: time sharing and space sharing. The type of intersection control that has to be adopted depends on the traffic volume, road geometry, cost involved, importance of the road etc.

39.3 Levels of intersection control

The control of an intersection can be exercised at different levels. They can be either passive control, semi control, or active control. In passive control, there is no explicit control on the driver . In semi control, some amount of control on the driver is there from the traffic agency. Active control means the movement of the traffic is fully controlled by the traffic agency and the drivers cannot simply maneuver the intersection according

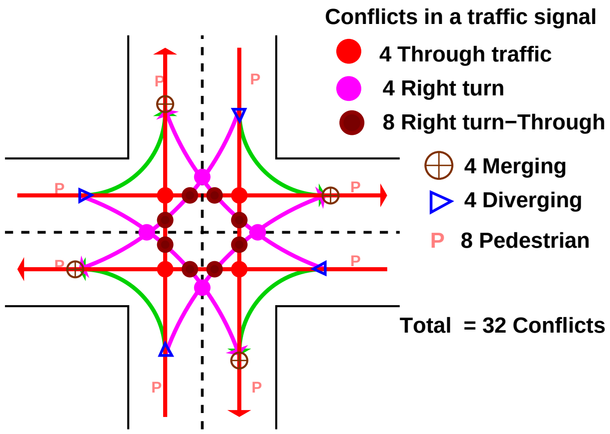


Figure 39:1: Conflicts at an intersection

to his choice.

39.3.1 Passive control

When the volume of traffic is less, no explicit control is required. Here the road users are required to obey the basic rules of the road. Passive control like traffic signs, road markings etc. are used to complement the intersection control. Some of the intersection control that are classified under passive control are as follows:

1. **No control** If the traffic coming to an intersection is low, then by applying the basic rules of the road like driver on the left side of the road must yield and that through movements will have priority than turning movements. The driver is expected to obey these basic rules of the road.

2. **Traffic signs**: With the help of warning signs, guide signs etc. it is able to provide some level of control at an intersection. Give way control, two-way stop control, and all-way stop control are some examples. The GIVE WAY control requires the driver in the minor road to slow down to a minimum speed and allow the vehicle on the major road to proceed. Two way stop control requires the vehicle drivers on the minor streets should see that the conflicts are avoided. Finally an all-way stop control is usually used when it is difficult to differentiate between the major and minor roads in an intersection. In such a case, STOP sign is placed on all the approaches to the intersection and the driver on all the approaches are required to stop the vehicle. The vehicle at the right side will get priority over the left approach. The traffic control at 'at-grade' intersection may be uncontrolled in cases of low traffic. Here the road users are required to obey the basic rules of the road. Passive control like traffic signs, road markings etc. are used to complement the intersection control.

3. **Traffic signs plus marking**: In addition to the traffic signs, road markings also complement the traffic control at intersections. Some of the examples include stop line marking, yield lines, arrow marking etc.

39.3.2 Semi control

In semi control or partial control, the drivers are gently guided to avoid conflicts. Channelization and traffic rotaries are two examples of this.

1. **Channelization**: The traffic is separated to flow through definite paths by raising a portion of the road in the middle usually called as islands distinguished by road markings. The conflicts in traffic movements are reduced to a great extent in such a case. In channelized intersections, as the name suggests, the traffic is directed to flow through different channels and this physical separation is made possible with the help of some barriers in the road like traffic islands, road markings etc.

2. **Traffic rotaries**: It is a form of intersection control in which the traffic is made to flow along one direction around a traffic island. The essential principle of this control is to convert all the severe conflicts like through and right turn conflicts into milder conflicts like merging, weaving and diverging. It is a form of 'at-grade' intersection laid out for the movement of traffic such that no through conflicts are there. Free-left turn is permitted where as through traffic and right-turn traffic is forced to move around the central island in a clock-wise direction in an orderly manner. Merging, weaving and diverging operations reduces the conflicting movements at the rotary.

39.3.3 Active control

Active control implies that the road user will be forced to follow the path suggested by the traffic control agencies. He cannot maneuver according to his wish. Traffic signals and grade separated intersections come under this classification.

1. **Traffic signals**: Control using traffic signal is based on time sharing approach. At a given time, with the help of appropriate signals, certain traffic movements are restricted where as certain other movements are permitted to pass through the intersection. Two or more phases may be provided depending upon the traffic conditions of the intersection. When the vehicles traversing the intersection is very large, then the control is done with the help of signals. The phases provided for the signal may be two or more. If more than two phases are provided, then it is called multiphase signal.

   The signals can operate in several modes. Most common are fixed time signals and vehicle actuated signals. In fixed time signals, the cycle time, phases and interval of each signal is fixed. Each cycle of the signal will be exactly like another. But they cannot cater to the needs of the fluctuating traffic. On the other hand, vehicle actuated signals can respond to dynamic traffic situations. Vehicle detectors will be placed on the streets approaching the intersection and the detector will sense the presence of the vehicle and pass the information to a controller. The controller then sets the cycle time and adjusts the phase lengths according to the prevailing traffic conditions.

2. **Grade separated intersections**: The intersections are of two types. They are at-grade intersections and grade-separated intersections. In at-grade intersections, all roadways join or cross at the same vertical level. Grade separated intersections allows the traffic to cross at different vertical levels. Sometimes the topography itself may be helpful in constructing such intersections. Otherwise, the initial construction cost required will be very high. Therefore, they are usually constructed on high speed facilities like expressways, freeways etc. These type of intersection increases the road capacity because vehicles can flow with high speed and accident potential is also reduced due to vertical separation of traffic.

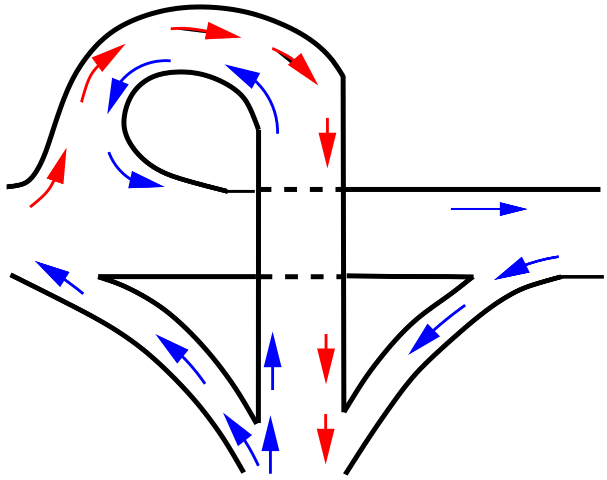


Figure 39:2: Trumpet interchange

39.4 Grade separated intersections

As we discussed earlier, grade-separated intersections are provided to separate the traffic in the vertical grade. But the traffic need not be those pertaining to road only. When a railway line crosses a road, then also grade separators are used. Different types of grade-separators are flyovers and interchange. Flyovers itself are subdivided into overpass and underpass. When two roads cross at a point, if the road having major traffic is elevated to a higher grade for further movement of traffic, then such structures are called overpass. Otherwise, if the major road is depressed to a lower level to cross another by means of an under bridge or tunnel, it is called under-pass.

Interchange is a system where traffic between two or more roadways flows at different levels in the grade separated junctions. Common types of interchange include trumpet interchange, diamond interchange , and cloverleaf interchange.

1. **Trumpet interchange**: Trumpet interchange is a popular form of three leg interchange. If one of the legs of the interchange meets a highway at some angle but does not cross it, then the interchange is called trumpet interchange. A typical layout of trumpet interchange is shown in figure 39:2.

2. **Diamond interchange**: Diamond interchange is a popular form of four-leg interchange found in the urban locations where major and minor roads crosses. The important feature of this interchange is that it can be designed even if the major road is relatively narrow. A typical layout of diamond interchange is shown in figure 39:3.

3. **Clover leaf interchange**: It is also a four leg interchange and is used when two highways of high volume and speed intersect each other with considerable turning movements. The main advantage of cloverleaf intersection is that it provides complete separation of traffic. In addition, high speed at intersections can be achieved. However, the disadvantage is that large area of land is required. Therefore, cloverleaf interchanges are provided mainly in rural areas. A typical layout of this type of interchange is shown in figure 39:4.

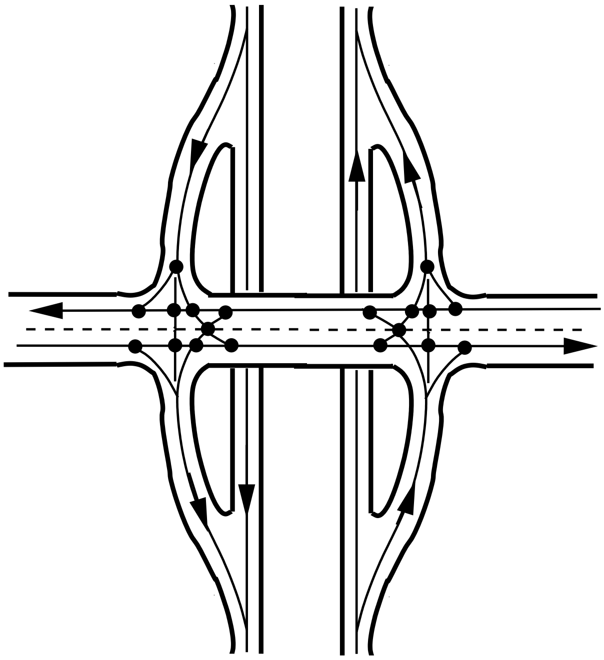


Figure 39:3: Diamond interchange

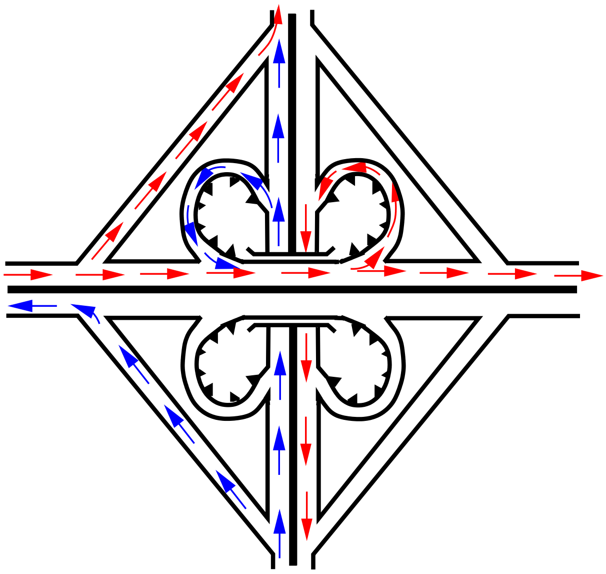


Figure 39:4: Cloverleaf interchange

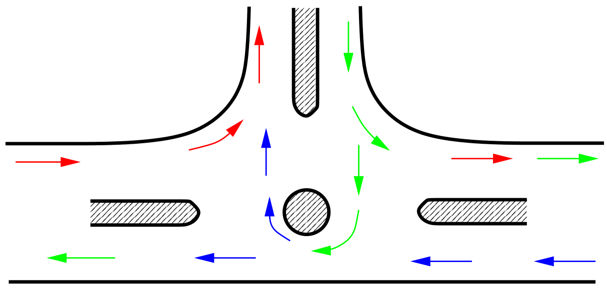


Figure 39:5: Channelization of traffic through a three-legged intersection

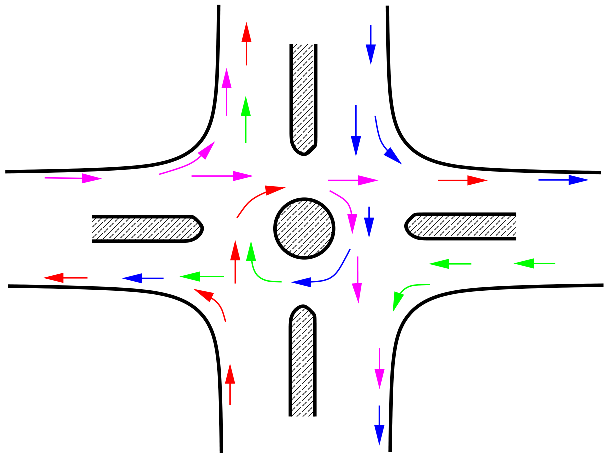


Figure 39:6: Channelization of traffic through a four-legged intersection

39.5 Channelized intersection

Vehicles approaching an intersection are directed to definite paths by islands, marking etc. and this method of control is called channelization. Channelized intersection provides more safety and efficiency. It reduces the number of possible conflicts by reducing the area of conflicts available in the carriageway. If no channelizing is provided the driver will have less tendency to reduce the speed while entering the intersection from the carriageway. The presence of traffic islands, markings etc. forces the driver to reduce the speed and becomes more cautious while maneuvering the intersection. A channelizing island also serves as a refuge for pedestrians and makes pedestrian crossing safer. Channelization of traffic through a three-legged intersection (refer figure 39:5) and a four-legged intersection (refer figure 39:6) is shown in the figure.

39.6 Summary

Traffic intersections are problem spots on any highway, which contribute to a large share of accidents. For safe operation, these locations should be kept under some level of control depending upon the traffic quantity and behavior. Based on this, intersections and interchanges are constructed, the different types of which were discussed in the chapter.

39.7 Problems

1. The GIVE WAY control

   (a) requires the driver in the minor road to slow down to a minimum speed and allow the vehicle on the major road to proceed.

   (b) requires the driver in the major road to slow down to a minimum speed and allow the vehicle on the minor road to proceed.

   (c) requires the drivers on both minor and major roads to stop.

   (d) is similar to one way control.

2. Traffic signal is an example of

   (a) Passive control

   (b) No control

   (c) Active control

   (d) none of these

39.8 Solutions

1. The GIVE WAY control

   (a) requires the driver in the minor road to slow down to a minimum speed and allow the vehicle on the major road to proceed.$\surd$

   (b) requires the driver in the major road to slow down to a minimum speed and allow the vehicle on the minor road to proceed.

   (c) requires the drivers on both minor and major roads to stop.

   (d) is similar to one way control.

2. Traffic signal is an example of

   (a) Passive control

   (b) No control

   (c) Active control$\surd$

   (d) none of these

Chapter 40

Traffic rotaries

40.1 Overview

Rotary intersections or round abouts are special form of at-grade intersections laid out for the movement of traffic in one direction around a central traffic island. Essentially all the major conflicts at an intersection namely the collision between through and right-turn movements are converted into milder conflicts namely merging and diverging. The vehicles entering the rotary are gently forced to move in a clockwise direction in orderly fashion. They then weave out of the rotary to the desired direction. The benefits, design principles, capacity of rotary etc. will be discussed in this chapter.

40.2 Advantages and disadvantages of rotary

The key advantages of a rotary intersection are listed below:

1. Traffic flow is regulated to only one direction of movement, thus eliminating severe conflicts between crossing movements.
2. All the vehicles entering the rotary are gently forced to reduce the speed and continue to move at slower speed. Thus, none of the vehicles need to be stopped,unlike in a signalized intersection.
3. Because of lower speed of negotiation and elimination of severe conflicts, accidents and their severity are much less in rotaries.
4. Rotaries are self governing and do not need practically any control by police or traffic signals.
5. They are ideally suited for moderate traffic, especially with irregular geometry, or intersections with more than three or four approaches.

Although rotaries offer some distinct advantages, there are few specific limitations for rotaries which are listed below.

1. All the vehicles are forced to slow down and negotiate the intersection. Therefore, the cumulative delay will be much higher than channelized intersection.
2. Even when there is relatively low traffic, the vehicles are forced to reduce their speed.
3. Rotaries require large area of relatively flat land making them costly at urban areas.
4. The vehicles do not usually stop at a rotary. They accelerate and exit the rotary at relatively high speed. Therefore, they are not suitable when there is high pedestrian movements.

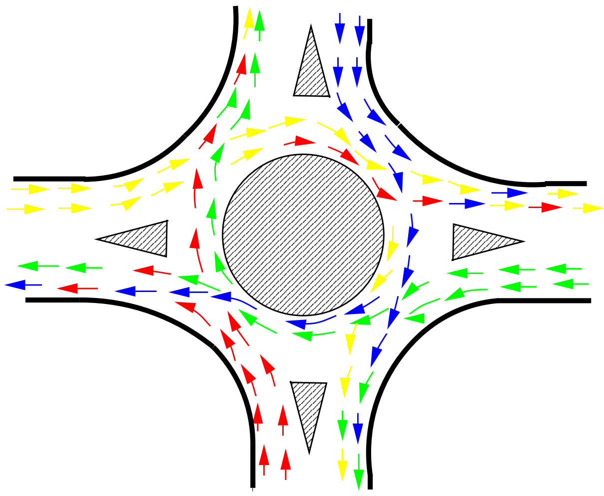


Figure 40:1: Traffic operations in a rotary

40.3 Guidelines for the selection of rotaries

Because of the above limitation, rotaries are not suitable for every location. There are few guidelines that help in deciding the suitability of a rotary. They are listed below.

1. Rotaries are suitable when the traffic entering from all the four approaches are relatively equal.
2. A total volume of about 3000 vehicles per hour can be considered as the upper limiting case and a volume of 500 vehicles per hour is the lower limit.
3. A rotary is very beneficial when the proportion of the right-turn traffic is very high; typically if it is more than 30 percent.
4. Rotaries are suitable when there are more than four approaches or if there is no separate lanes available for right-turn traffic. Rotaries are ideally suited if the intersection geometry is complex.

40.4 Traffic operations in a rotary

As noted earlier, the traffic operations at a rotary are three; diverging, merging and weaving. All the other conflicts are converted into these three less severe conflicts.

1. **Diverging**: It is a traffic operation when the vehicles moving in one direction is separated into different streams according to their destinations.
2. **Merging**: Merging is the opposite of diverging. Merging is referred to as the process of joining the traffic coming from different approaches and going to a common destination into a single stream.
3. **Weaving**: Weaving is the combined movement of both merging and diverging movements in the same direction.

These movements are shown in figure 40:1. It can be observed that movements from each direction split into three; left, straight, and right turn.

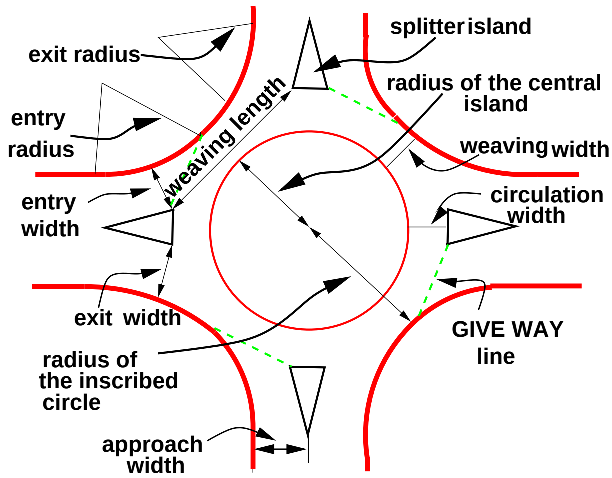


Figure 40:2: Design of a rotary

40.4.1 Design elements

The design elements include design speed, radius at entry, exit and the central island, weaving length and width, entry and exit widths. In addition the capacity of the rotary can also be determined by using some empirical formula. A typical rotary and the important design elements are shown in figure 40:2

40.4.2 Design speed

All the vehicles are required to reduce their speed at a rotary. Therefore, the design speed of a rotary will be much lower than the roads leading to it. Although it is possible to design roundabout without much speed reduction, the geometry may lead to very large size incurring huge cost of construction. The normal practice is to keep the design speed as 30 and 40 kmph for urban and rural areas respectively.

40.4.3 Entry, exit and island radius

The radius at the entry depends on various factors like design speed, super-elevation, and coefficient of friction. The entry to the rotary is not straight, but a small curvature is introduced. This will force the driver to reduce the speed. The entry radius of about 20 and 25 metres is ideal for an urban and rural design respectively.

The exit radius should be higher than the entry radius and the radius of the rotary island so that the vehicles will discharge from the rotary at a higher rate. A general practice is to keep the exit radius as 1.5 to 2 times the entry radius. However, if pedestrian movement is higher at the exit approach, then the exit radius could be set as same as that of the entry radius.

The radius of the central island is governed by the design speed, and the radius of the entry curve. The radius of the central island, in practice, is given a slightly higher radius so that the movement of the traffic already in the rotary will have priority. The radius of the central island which is about 1.3 times that of the entry curve is adequate for all practical purposes.

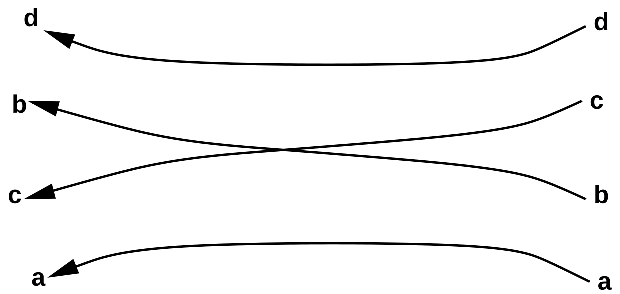


Figure 40:3: Weaving operation in a rotary

40.4.4 Width of the rotary

The entry width and exit width of the rotary is governed by the traffic entering and leaving the intersection and the width of the approaching road. The width of the carriageway at entry and exit will be lower than the width of the carriageway at the approaches to enable reduction of speed. IRC suggests that a two lane road of 7 m width should be kept as 7 m for urban roads and 6.5 m for rural roads. Further, a three lane road of 10.5 m is to be reduced to 7 m and 7.5 m respectively for urban and rural roads.

The width of the weaving section should be higher than the width at entry and exit. Normally this will be one lane more than the average entry and exit width. Thus weaving width is given as,

$$\mathrm{w_{weaving}} = \left(\frac{\mathrm{e_1} + \mathrm{e_2}}{2}\right) + 3.5\mathrm{m} \tag{40.1}$$

where e_1 is the width of the carriageway at the entry and e_2 is the carriageway width at exit.

Weaving length determines how smoothly the traffic can merge and diverge. It is decided based on many factors such as weaving width, proportion of weaving traffic to the non-weaving traffic etc. This can be best achieved by making the ratio of weaving length to the weaving width very high. A ratio of 4 is the minimum value suggested by IRC. Very large weaving length is also dangerous, as it may encourage over-speeding.

40.5 Capacity

The capacity of rotary is determined by the capacity of each weaving section. Transportation road research lab (TRL) proposed the following empirical formula to find the capacity of the weaving section.

$$Q_w = \frac{280w[1 + \frac{e}{w}][1 - \frac{p}{3}]}{1 + \frac{w}{l}} \tag{40.2}$$

where e is the average entry and exit width, i.e, $\frac{(e_1+e_2)}{2}$, w is the weaving width, l is the length of weaving, and p is the proportion of weaving traffic to the non-weaving traffic. Figure 40:3 shows four types of movements at a weaving section, a and d are the non-weaving traffic and b and c are the weaving traffic. Therefore,

$$p = \frac{b + c}{a + b + c + d} \tag{40.3}$$

This capacity formula is valid only if the following conditions are satisfied.

1. Weaving width at the rotary is in between 6 and 18 metres.

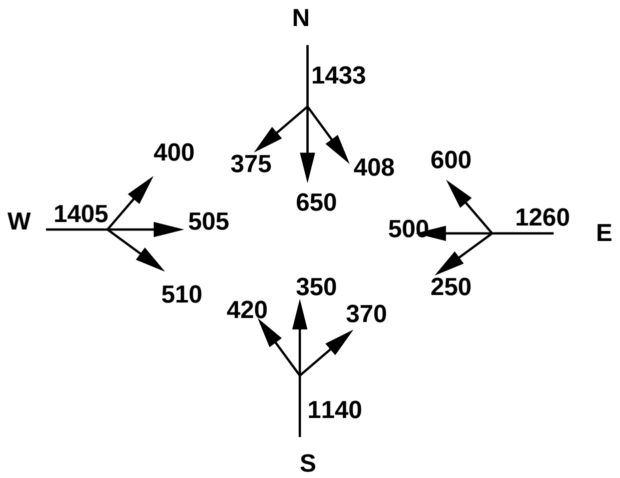


Figure 40:4: Traffic approaching the rotary

2. The ratio of average width of the carriage way at entry and exit to the weaving width is in the range of 0.4 to 1.

3. The ratio of weaving width to weaving length of the roundabout is in between 0.12 and 0.4.

4. The proportion of weaving traffic to non-weaving traffic in the rotary is in the range of 0.4 and 1.

5. The weaving length available at the intersection is in between 18 and 90 m.

Example

The width of a carriage way approaching an intersection is given as 15 m. The entry and exit width at the rotary is 10 m. The traffic approaching the intersection from the four sides is shown in the figure 40:4 below. Find the capacity of the rotary using the given data.

Solution

- The traffic from the four approaches negotiating through the roundabout is illustrated in figure 40:5.
- Weaving width is calculated as, w = $[\frac{e_1+e_2}{2}] + 3.5 = 13.5$ m
- Weaving length, l is calculated as = 4×w = 54 m
- The proportion of weaving traffic to the non-weaving traffic in all the four approaches is found out first.
- It is clear from equation,that the highest proportion of weaving traffic to non-weaving traffic will give the minimum capacity. Let the proportion of weaving traffic to the non-weaving traffic in West-North direction be denoted as p_{WN}, in North-East direction as p_{NE}, in the East-South direction as p_{ES}, and finally in the South-West direction as p_{SW}.
- The weaving traffic movements in the East-South direction is shown in figure 40:6. Then using equation,p_{ES} $=\frac{510+650+500+600}{510+650+500+600+250+375}=\frac{2260}{2885}=0.783$
$p_{WN} = \frac{505+510+350+600}{505+510+350+600+400+370}=\frac{1965}{2735}=0.718$

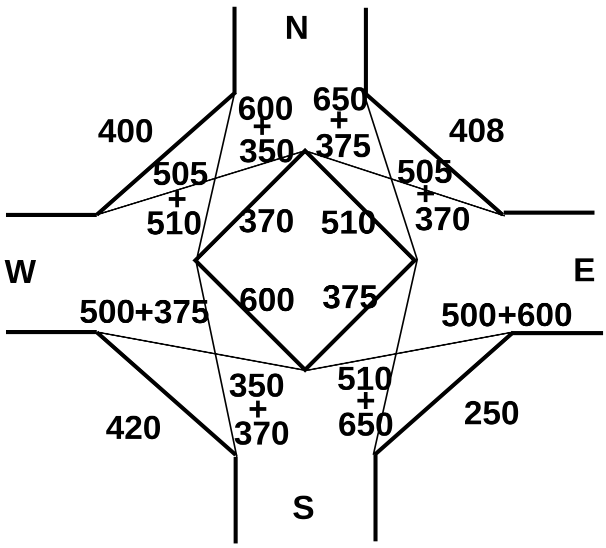


Figure 40:5: Traffic negotiating a rotary

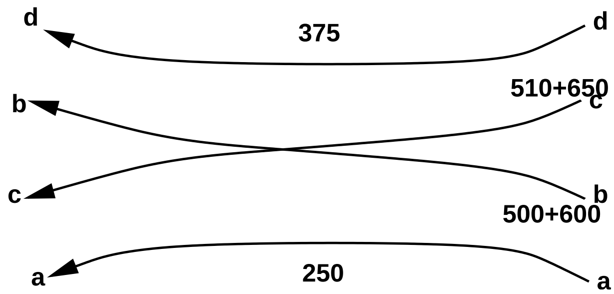


Figure 40:6: Traffic weaving in East-South direction

$$p_{NE} = \frac{650+375+505+370}{650+375+505+370+510+408} = \frac{1900}{2818} = 0.674$$
$$p_{SW} = \frac{350+370+500+375}{350+370+500+375+420+600} = \frac{1595}{2615} = 0.6099$$

- Thus the proportion of weaving traffic to non-weaving traffic is highest in the East-South direction.
- Therefore, the capacity of the rotary will be capacity of this weaving section. From equation,

$$Q_{ES} = \frac{280 \times 13.5[1 + \frac{10}{13.5}][1 - \frac{0.783}{3}]}{1 + \frac{13.5}{54}} = 2161.164 veh/hr. \tag{40.4}$$

40.6 Summary

Traffic rotaries reduce the complexity of crossing traffic by forcing them into weaving operations. The shape and size of the rotary are determined by the traffic volume and share of turning movements. Capacity assessment of a rotary is done by analyzing the section having the greatest proportion of weaving traffic. The analysis is done by using the formula given by TRL.

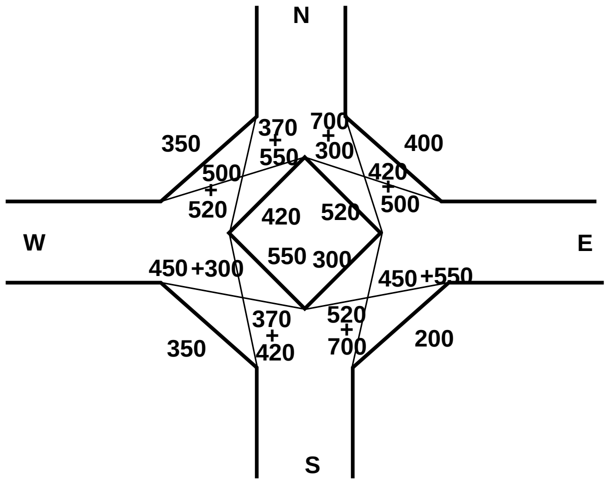


Figure 40:7: Traffic negotiating a rotary

40.7 Problems

1. The width of approaches for a rotary intersection is 12 m. The entry and exit width at the rotary is 10 m. Table below gives the traffic from the four approaches, traversing the intersection. Find the capacity of the rotary.

Approach	Left turn	Straight	Right turn
North	400	700	300
South	350	370	420
East	200	450	550
West	350	500	520

Solution

- The traffic from the four approaches negotiating through the roundabout is illustrated in figure 40:7.
- Weaving width is calculated as, w = $[\frac{e_1+e_2}{2}] + 3.5 = 13.5$ m
- Weaving length can be calculated as, l = 4×w = 54 m
- The proportion of weaving traffic to the non-weaving traffic in all the four approaches is found out first.
- It is clear from equation,that the highest proportion of weaving traffic to non-weaving traffic will give the minimum capacity. Let the proportion of weaving traffic to the non-weaving traffic in West-North direction be denoted as p_{WN}, in North-East direction as p_{NE}, in the East-South direction as p_{ES}, and finally in the South-West direction as p_{SW}. Then using equation,$p_{ES} = \frac{450+550+700+520}{200+450+550+700+520+300} = \frac{2220}{2720} = 0.816$
  $p_{WN} = \frac{370+550+500+520}{350+370+550+500+520+420} = \frac{1740}{2510} = 0.69$
  $p_{NE} = \frac{420+500+700+300}{520+400+420+500+700+300} = \frac{1920}{2840} = 0.676$
  $p_{SW} = \frac{450+300+370+420}{550+450+400+370+420+350} = \frac{1540}{2540} = 0.630$
- Thus the proportion of weaving traffic to non-weaving traffic is highest in the East-South direction.

- Therefore, the capacity of the rotary will be the capacity of this weaving section. From equation, $Q_{ES} = \frac{280 \times 13.5[1+\frac{10}{13.5}][1-\frac{0.816}{3}]}{1+\frac{13.5}{54}} = 380.56$veh/hr.

Chapter 41

Traffic signal design-I

41.1 Overview

The conflicts arising from movements of traffic in different directions is solved by time sharing of the principle. The advantages of traffic signal includes an orderly movement of traffic, an increased capacity of the intersection and requires only simple geometric design. However the disadvantages of the signalized intersection are it affects larger stopped delays, and the design requires complex considerations. Although the overall delay may be lesser than a rotary for a high volume, a user is more concerned about the stopped delay.

41.2 Definitions and notations

A number of definitions and notations need to be understood in signal design. They are discussed below:

- **Cycle**: A signal cycle is one complete rotation through all of the indications provided.
- **Cycle length**: Cycle length is the time in seconds that it takes a signal to complete one full cycle of indications. It indicates the time interval between the starting of of green for one approach till the next time the green starts. It is denoted by C.
- **Interval**: Thus it indicates the change from one stage to another. There are two types of intervals - change interval and clearance interval. *Change interval* is also called the yellow time indicates the interval between the green and red signal indications for an approach. *Clearance interval* is also called *all red* is included after each yellow interval indicating a period during which all signal faces show red and is used for clearing off the vehicles in the intersection.
- **Green interval**: It is the green indication for a particular movement or set of movements and is denoted by G_i. This is the actual duration the green light of a traffic signal is turned on.
- **Red interval**: It is the red indication for a particular movement or set of movements and is denoted by R_i. This is the actual duration the red light of a traffic signal is turned on.
- **Phase**: A phase is the green interval plus the change and clearance intervals that follow it. Thus, during green interval, non conflicting movements are assigned into each phase. It allows a set of movements to flow and safely halt the flow before the phase of another set of movements start.
- **Lost time**: It indicates the time during which the intersection is not effectively utilized for any movement. For example, when the signal for an approach turns from red to green, the driver of the vehicle which is

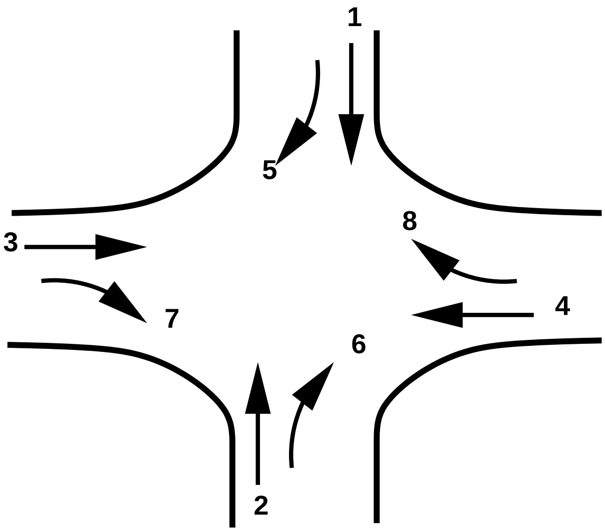


Figure 41:1: Four legged intersection

in the front of the queue, will take some time to perceive the signal (usually called as reaction time) and some time will be lost here before he moves.

41.3 Phase design

The signal design procedure involves six major steps. They include the (1) phase design, (2) determination of amber time and clearance time, (3) determination of cycle length, (4)apportioning of green time, (5) pedestrian crossing requirements, and (6) the performance evaluation of the above design. The objective of phase design is to separate the conflicting movements in an intersection into various phases, so that movements in a phase should have no conflicts. If all the movements are to be separated with no conflicts, then a large number of phases are required. In such a situation, the objective is to design phases with minimum conflicts or with less severe conflicts.

There is no precise methodology for the design of phases. This is often guided by the geometry of the intersection, flow pattern especially the turning movements, the relative magnitudes of flow. Therefore, a trial and error procedure is often adopted. However, phase design is very important because it affects the further design steps. Further, it is easier to change the cycle time and green time when flow pattern changes, where as a drastic change in the flow pattern may cause considerable confusion to the drivers. To illustrate various phase plan options, consider a four legged intersection with through traffic and right turns. Left turn is ignored. See figure 41:1. The first issue is to decide how many phases are required. It is possible to have two, three, four or even more number of phases.

41.3.1 Two phase signals

Two phase system is usually adopted if through traffic is significant compared to the turning movements. For example in figure 41:2, non-conflicting through traffic 3 and 4 are grouped in a single phase and non-conflicting through traffic 1 and 2 are grouped in the second phase. However, in the first phase flow 7 and 8 offer some conflicts and are called permitted right turns. Needless to say that such phasing is possible only if the turning

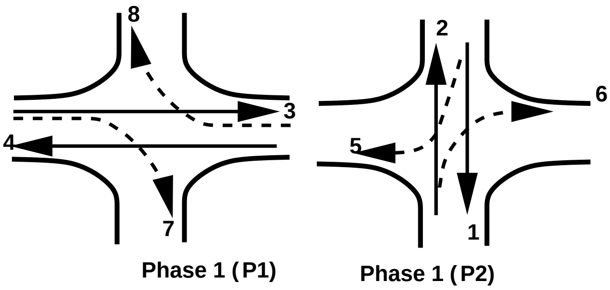


Figure 41:2: Two phase signal

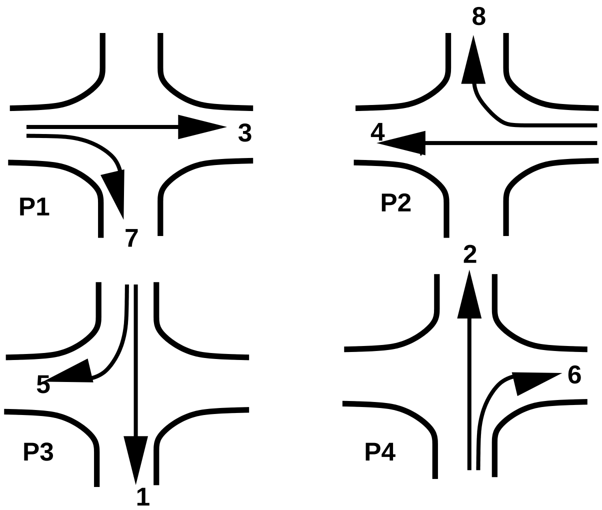


Figure 41:3: One way of providing four phase signals

movements are relatively low. On the other hand, if the turning movements are significant ,then a four phase system is usually adopted.

41.3.2 Four phase signals

There are at least three possible phasing options. For example, figure 41:3 shows the most simple and trivial phase plan. where, flow from each approach is put into a single phase avoiding all conflicts. This type of phase plan is ideally suited in urban areas where the turning movements are comparable with through movements and when through traffic and turning traffic need to share same lane. This phase plan could be very inefficient when turning movements are relatively low.

Figure 41:4 shows a second possible phase plan option where opposing through traffic are put into same phase. The non-conflicting right turn flows 7 and 8 are grouped into a third phase. Similarly flows 5 and 6 are grouped into fourth phase. This type of phasing is very efficient when the intersection geometry permits to have at least one lane for each movement, and the through traffic volume is significantly high. Figure 41:5 shows yet another phase plan. However, this is rarely used in practice.

There are five phase signals, six phase signals etc. They are normally provided if the intersection control is adaptive, that is, the signal phases and timing adapt to the real time traffic conditions.

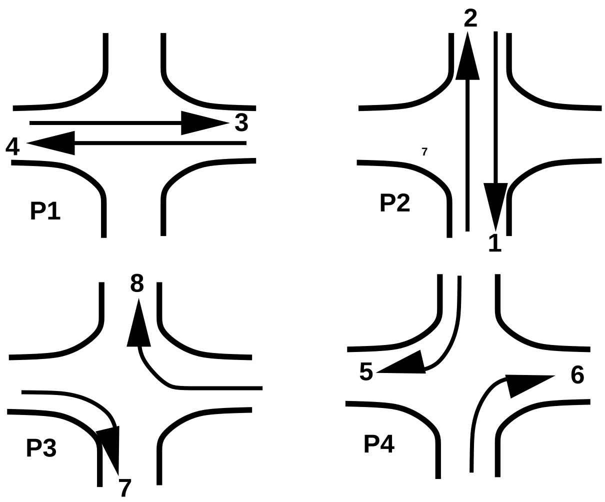


Figure 41:4: Second possible way of providing a four phase signal

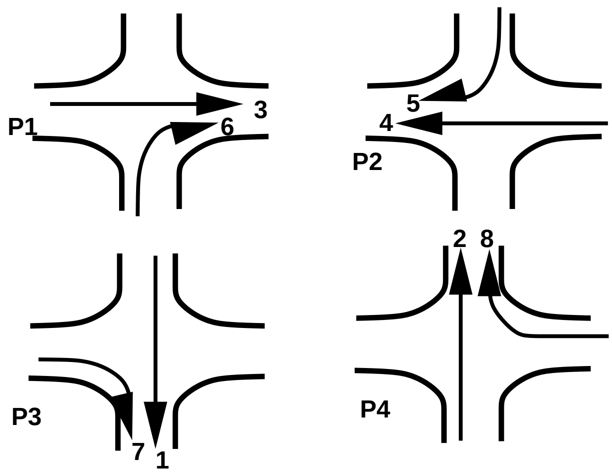


Figure 41:5: Third possible way of providing a four-phase signal

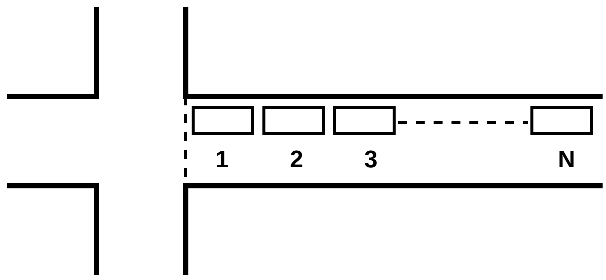


Figure 41:6: Group of vehicles at a signalized intersection waiting for green signal

41.4 Interval design

There are two intervals, namely the change interval and clearance interval, normally provided in a traffic signal. The change interval or yellow time is provided after green time for movement. The purpose is to warn a driver approaching the intersection during the end of a green time about the coming of a red signal. They normally have a value of 3 to 6 seconds.

The design consideration is that a driver approaching the intersection with design speed should be able to stop at the stop line of the intersection before the start of red time. Institute of transportation engineers (ITE) has recommended a methodology for computing the appropriate length of change interval which is as follows:

$$y = t + \frac{v_{85}}{2a + 19.6g} \tag{41.1}$$

where y is the length of yellow interval in seconds, t is the reaction time of the driver, v_{85} is the 85^{th} percentile speed of approaching vehicles in m/s, a is the deceleration rate of vehicles in m/s^2, g is the grade of approach expressed as a decimal. Change interval can also be approximately computed as $y = \frac{SSD}{v}$, where SSD is the stopping sight distance and v is the speed of the vehicle. The clearance interval is provided after yellow interval and as mentioned earlier, it is used to clear off the vehicles in the intersection. Clearance interval is optional in a signal design. It depends on the geometry of the intersection. If the intersection is small, then there is no need of clearance interval whereas for very large intersections, it may be provided.

41.5 Cycle time

Cycle time is the time taken by a signal to complete one full cycle of iterations. i.e. one complete rotation through all signal indications. It is denoted by C. The way in which the vehicles depart from an intersection when the green signal is initiated will be discussed now. Figure 41:6 illustrates a group of N vehicles at a signalized intersection, waiting for the green signal. As the signal is initiated, the time interval between two vehicles, referred as headway, crossing the curb line is noted. The first headway is the time interval between the initiation of the green signal and the instant vehicle crossing the curb line. The second headway is the time interval between the first and the second vehicle crossing the curb line. Successive headways are then plotted as in figure 41:7. The first headway will be relatively longer since it includes the reaction time of the driver and the time necessary to accelerate. The second headway will be comparatively lower because the second driver can overlap his/her reaction time with that of the first driver's. After few vehicles, the headway will become constant. This constant headway which characterizes all headways beginning with the fourth or fifth vehicle, is defined as the saturation headway, and is denoted as h. This is the headway that can be achieved by a stable

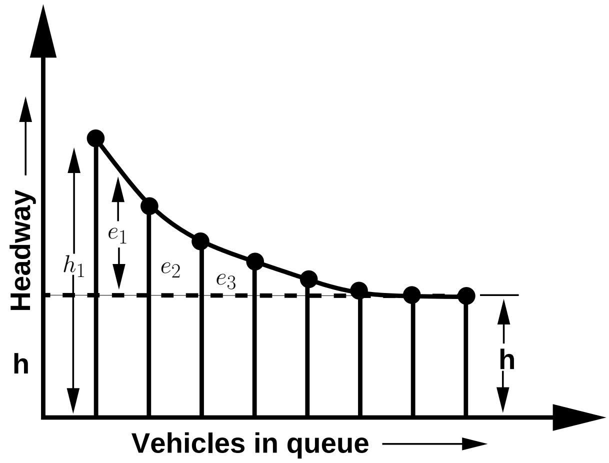


Figure 41:7: Headways departing signal

moving platoon of vehicles passing through a green indication. If every vehicles require h seconds of green time, and if the signal were always green, then s vehicles/per hour would pass the intersection. Therefore,

$$s = \frac{3600}{h} \tag{41.2}$$

where s is the saturation flow rate in vehicles per hour of green time per lane, h is the saturation headway in seconds. vehicles per hour of green time per lane. As noted earlier, the headway will be more than h particularly for the first few vehicles. The difference between the actual headway and h for the i^{th} vehicle and is denoted as e_i shown in figure 41:7. These differences for the first few vehicles can be added to get start up lost time, l_1 which is given by,

$$l_1 = \sum_{i=1}^{n} e_i \tag{41.3}$$

The green time required to clear N vehicles can be found out as,

$$T = l_1 + h.N \tag{41.4}$$

where T is the time required to clear N vehicles through signal, l_1 is the start-up lost time, and h is the saturation headway in seconds.

41.5.1 Effective green time

Effective green time is the actual time available for the vehicles to cross the intersection. It is the sum of actual green time (G_i) plus the yellow minus the applicable lost times. This lost time is the sum of start-up lost time (l_1) and clearance lost time (l_2) denoted as t_L. Thus effective green time can be written as,

$$g_i = G_i + Y_i - t_L \tag{41.5}$$

41.5.2 Lane capacity

The ratio of effective green time to the cycle length ($\frac{g_i}{C}$)is defined as green ratio. We know that saturation flow rate is the number of vehicles that can be moved in one lane in one hour assuming the signal to be green always.

Then the capacity of a lane can be computed as,

$$c_i = s_i \frac{g_i}{C} \tag{41.6}$$

where c_i is the capacity of lane in vehicle per hour, s_i is the saturation flow rate in vehicle per hour per lane, C is the cycle time in seconds.

Problem

Let the cycle time of an intersection is 60 seconds, the green time for a phase is 27 seconds, and the corresponding yellow time is 4 seconds. If the saturation headway is 2.4 seconds/vehicle, the start-up lost time is 2 seconds/phase, and the clearance lost time is 1 second/phase, find the capacity of the movement per lane?

Solution Total lost time, t_L = 2+1 = 3 seconds. From equation effective green time, g_i = 27+4-3 = 28 seconds. From equationsaturation flow rate, $s_i = \frac{3600}{h} = \frac{3600}{2.4}$ = 1500 veh/hr. Capacity of the given phase can be found out from equation,$C_i = 1500 \times \frac{28}{60}$ = 700 veh/hr/lane.

41.5.3 Critical lane

During any green signal phase, several lanes on one or more approaches are permitted to move. One of these will have the most intense traffic. Thus it requires more time than any other lane moving at the same time. If sufficient time is allocated for this lane, then all other lanes will also be well accommodated. There will be one and only one critical lane in each signal phase. The volume of this critical lane is called critical lane volume.

41.6 Determination of cycle length

The cycle length or cycle time is the time taken for complete indication of signals in a cycle. Fixing the cycle length is one of the crucial steps involved in signal design.

If t_{Li} is the start-up lost time for a phase i, then the total start-up lost time per cycle, $L = \sum_{i=1}^{N} t_L$, where N is the number of phases. If start-up lost time is same for all phases, then the total start-up lost time is $L = Nt_L$. If C is the cycle length in seconds, then the number of cycles per hour $= \frac{3600}{C}$ The total lost time per hour is the number of cycles per hour times the lost time per cycle and is $= \frac{3600}{C}.L$ Substituting as $L = Nt_L$, total lost time per hour can be written as $= \frac{3600.N.t_L}{C}$ The total effective green time T_g available for the movement in a hour will be one hour minus the total lost time in an hour. Therefore,

$$T_g = 3600 - \frac{3600.N.t_L}{C} \tag{41.7}$$

$$= 3600\left[1 - \frac{N.t_L}{C}\right] \tag{41.8}$$

$$\tag{41.9}$$

Let the total number of critical lane volume that can be accommodated per hour is given by V_c, then $V_c = \frac{T_g}{h}$ Substituting for T_g, from equation 41.9 and s_i from the maximum sum of critical lane volumes that can be accommodated within the hour is given by,

$$= \frac{T_g}{h} \tag{41.10}$$

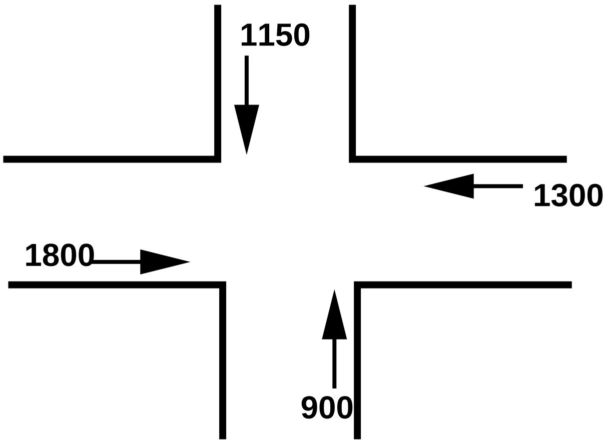


Figure 41:8: Traffic flow in the intersection

$$V_c = \frac{3600}{h}\left[1 - \frac{N.t_L}{C}\right] \tag{41.11}$$

$$= s_i\left[1 - \frac{N.t_L}{C}\right] \tag{41.12}$$

$$Therefore\ C = \frac{N.t_L}{1 - \frac{V_c}{s}} \tag{41.13}$$

(41.14)

The expression for C can be obtained by rewriting the above equation. The above equation is based on the assumption that there will be uniform flow of traffic in an hour. To account for the variation of volume in an hour, a factor called peak hour factor, (PHF) which is the ratio of hourly volume to the maximum flow rate, is introduced. Another ratio called v/c ratio indicating the quality of service is also included in the equation. Incorporating these two factors in the equation for cycle length, the final expression will be,

$$C = \frac{N.t_L}{1 - \frac{V_c}{s_i \times PHF \times \frac{v}{c}}} \tag{41.15}$$

Highway capacity manual (HCM) has given an equation for determining the cycle length which is a slight modification of the above equation. Accordingly, cycle time C is given by,

$$C = \frac{N.L.X_C}{X_C - \Sigma(\frac{V}{s})_i} \tag{41.16}$$

where N is the number of phases, L is the lost time per phase, $(\frac{V}{s})_i$ is the ratio of volume to saturation flow for phase i, X_C is the quality factor called critical $\frac{V}{C}$ratio where V is the volume and C is the capacity.

Problem

The traffic flow in an intersection is shown in the figure 41:8. Given start-up lost time is 3 seconds, saturation head way is 2.3 seconds, compute the cycle length for that intersection. Assume a two-phase signal.

Solution

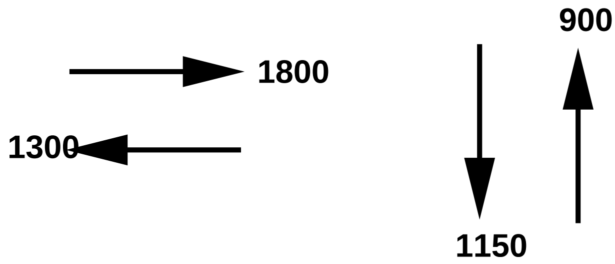


Figure 41:9: One way of providing phases

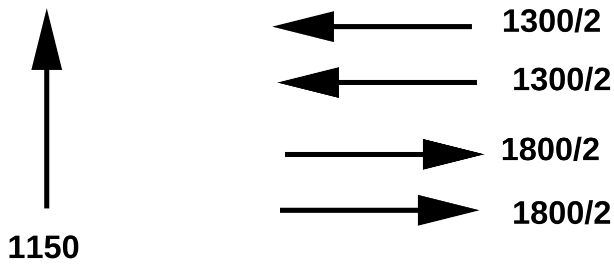


Figure 41:10: second way of providing phases

- If we assign two phases as shown below figure 41:9, then the critical volume for the first phase which is the maximum of the flows in that phase = 1150 vph. Similarly critical volume for the second phase = 1800 vph. Therefore, total critical volume for the two signal phases = 1150+1800 = 2950 vph.
- Saturation flow rate for the intersection can be found out from the equation as $s_i = \frac{3600}{2.3} = 1565.2$ vph. This means, that the intersection can handle only 1565.2 vph. However, the critical volume is 2950 vph . Hence the critical lane volume should be reduced and one simple option is to split the major traffic into two lanes. So the resulting phase plan is as shown in figure (41:10).
- Here we are dividing the lanes in East-West direction into two, the critical volume in the first phase is 1150 vph and in the second phase it is 900 vph. The total critical volume for the signal phases is 2050 vph which is again greater than the saturation flow rate and hence we have to again reduce the critical lane volumes.
- Assigning three lanes in East-West direction, as shown in figure 41:11, the critical volume in the first phase is 575 vph and that of the second phase is 600 vph, so that the total critical lane volume = 575+600 = 1175 vph which is lesser than 1565.2 vph.
- Now the cycle time for the signal phases can be computed from equation,$C = \frac{2\times3}{1-\frac{1175}{1565.2}} = 24$ seconds.

41.7 Summary

Traffic signal is an aid to control traffic at intersections where other control measures fail. The signals operate by providing right of way to a certain set of movements in a cyclic order. Depending on the requirements they can be either fixed or vehicle actuated and two or multivalued. The design procedure discussed in this chapter

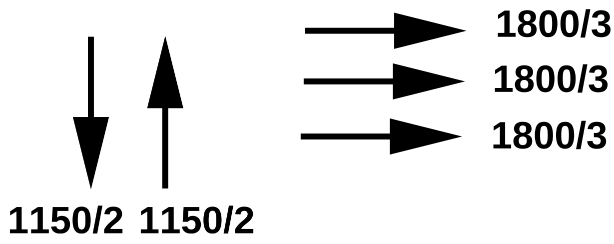


Figure 41:11: Third way of providing phases

include interval design, determination of cycle time, and computation of saturation flow making use of HCM guidelines.

41.8 Problems

1. Saturation flow rate can be computed as,
   (a) $\frac{3600}{h}$
   (b) $\frac{h}{3600}$
   (c) 3600×h
   (d) none of these
2. Lane capacity is
   (a) $c_i = s_i \times \frac{g_i}{C}$
   (b) $c_i = s_i \times g_i$
   (c) $c_i = \frac{s_i}{C}$
   (d) none of these

41.9 Solutions

1. Saturation flow rate can be computed as,
   (a) $\frac{3600}{h}\surd$
   (b) $\frac{h}{3600}$
   (c) 3600×h
   (d) none of these
2. Lane capacity is
   (a) $c_i = s_i \times \frac{g_i}{C}\surd$
   (b) $c_i = s_i \times g_i$
   (c) $c_i = \frac{s_i}{C}$
   (d) none of these

Chapter 42

Traffic signal design-II

42.1 Overview

In the previous chapter, a simple design of cycle time was discussed. Here we will discuss how the cycle time is divided in a phase. The performance evaluation of a signal is also discussed.

42.2 Green splitting

Green splitting or apportioning of green time is the proportioning of effective green time in each of the signal phase. The green splitting is given by,

$$g_i = \left[\frac{V_c i}{\sum_{i=1}^{n} V_{ci}}\right] \times t_g \tag{42.1}$$

where $V_c i$ is the critical lane volume and t_g is the total effective green time available in a cycle. This will be cycle time minus the total lost time for all the phases. Therefore,

$$T_g = C - n.t_L \tag{42.2}$$

where C is the cycle time in seconds, n is the number of phases, and t_L is the lost time per phase. If lost time is different for different phases, then cycle time can be computed as follows.

$$T_g = C - \sum_{i=1}^{n} t_{Li} \tag{42.3}$$

where t_{Li} is the lost time for phase i, n is the number of phases and C is the lost time in seconds. Actual greentime can be now found out as,

$$G_i = g_i - y_i + t_{Li} \tag{42.4}$$

where G_i is the actual green time, g_i is the effective green time available, y_i is the amber time, and L_i is the lost time for phase i.

Problem

The phase diagram with flow values of an intersection with two phases is shown in figure 42:1. The lost time and yellow time for the first phase is 2.5 and 3 seconds respectively. For the second phase the lost time and yellow time are 3.5 and 4 seconds respectively. If the cycle time is 120 seconds, find the green time allocated for the two phases.

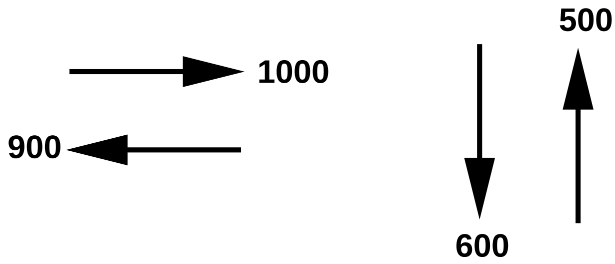


Figure 42:1: Phase diagram for an intersection

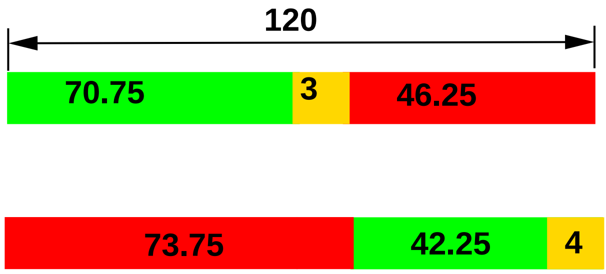


Figure 42:2: Timing diagram

Solution

- Critical lane volume for the first phase, V_{C1} = 1000 vph.
- Critical lane volume for the second phase, V_{C2} = 600 vph.
- The sum of the critical lane volumes, $V_C = V_{C1} + V_{C2}$ = 1000+600 = 1600 vph.
- Effective green time can be found out from equationas T_g=120-(2.5-3.5)= 114 seconds.
- Green time for the first phase, g_1 can be found out from equationas $g_1 = \frac{1000}{1600} \times 114 = 71.25$ seconds.
- Green time for the second phase, g_2 can be found out from equationas $g_2 = \frac{600}{1600} \times 114$= 42.75 seconds.
- Actual green time can be found out from equationThus actual green time for the first phase, G_1 = 71.25-3+2.5 = 70.75 seconds.
- Actual green time for the second phase, G_2 = 42.75-3+2.5 = 42.25 seconds.
- The phase diagram is as shown in figure 42:2.

42.3 Pedestrian crossing requirements

Pedestrian crossing requirements can be taken care by two ways; by suitable phase design or by providing an exclusive pedestrian phase. It is possible in some cases to allocate time for the pedestrians without providing an exclusive phase for them. For example, consider an intersection in which the traffic moves from north to south and also from east to west. If we are providing a phase which allows the traffic to flow only in north-south direction, then the pedestrians can cross in east-west direction and vice-versa. However in some cases, it may

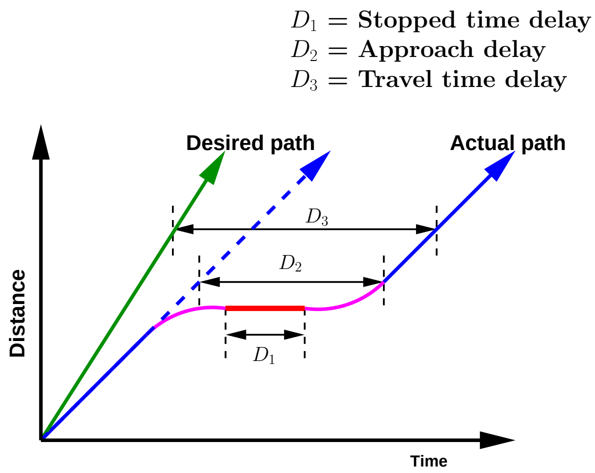


Figure 42:3: Illustration of delay measures

be necessary to provide an exclusive pedestrian phase. In such cases, the procedure involves computation of time duration of allocation of pedestrian phase. Green time for pedestrian crossing G_p can be found out by,

$$G_p = t_s + \frac{dx}{u_P} \tag{42.5}$$

where G_p is the minimum safe time required for the pedestrians to cross, often referred to as the "pedestrian green time", t_s is the start-up lost time, dx is the crossing distance in metres, and u_p is the walking speed of pedestrians which is about 15th percentile speed. The start-up lost time t_s can be assumed as 4.7 seconds and the walking speed can be assumed to be 1.2 m/s.

42.4 Performance measures

Performance measures are parameters used to evaluate the effectiveness of the design. There are many parameters involved to evaluate the effectiveness of the design and most common of these include delay, queuing, and stops. Delay is a measure that most directly relates the driver's experience. It describes the amount of time that is consumed while traversing the intersection. The figure 42:3 shows a plot of distance versus time for the progress of one vehicle. The desired path of the vehicle as well as the actual progress of the vehicle is shown. There are three types of delay as shown in the figure. They are stopped delay, approach delay and control delay. *Stopped time delay* includes only the time at which the vehicle is actually stopped waiting at the red signal. It starts when the vehicle reaches a full stop, and ends when the vehicle begins to accelerate. *Approach delay* includes the stopped time as well as the time lost due to acceleration and deceleration. It is measured as the time differential between the actual path of the vehicle, and path had there been green signal. *Control delay* is measured as the difference between the time taken for crossing the intersection and time taken to traverse the same section, had been no intersection. For a signalized intersection, it is measured at the stop-line as the vehicle enters the intersection. Among various types of delays, stopped delay is easy to derive and often used as a performance indicator and will be discussed.

Vehicles are not uniformly coming to an intersection. i.e., they are not approaching the intersection at constant time intervals. They come in a random manner. This makes the modeling of signalized intersection delay complex. Most simple of the delay models is Webster's delay model. It assumes that the vehicles are

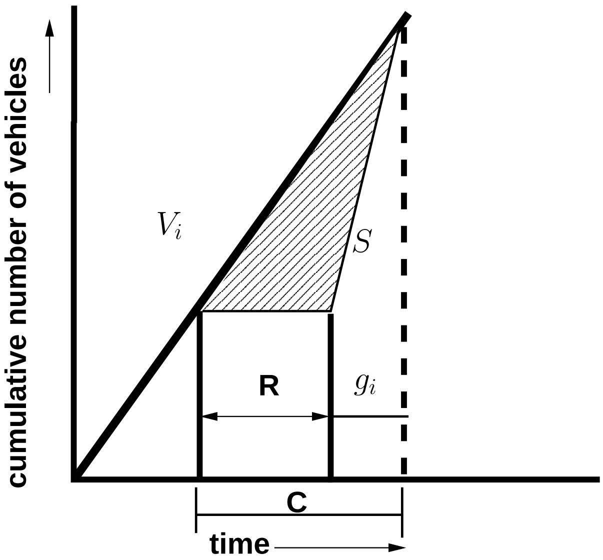


Figure 42:4: Graph between time and cumulative number of vehicles at an intersection

arriving at a uniform rate. Plotting a graph with time along the x-axis and cumulative vehicles along the y-axis we get a graph as shown in figure 42:4. The delay per cycle is shown as the area of the hatched portion in the figure. Webster derived an expression for delay per cycle based on this, which is as follows.

$$d_i = \frac{\frac{C}{2}[1 - \frac{g_i}{C}]^2}{1 - \frac{V_i}{S}} \tag{42.6}$$

where g_i is the effective green time, C is the cycle length, V_i is the critical flow for that phase, and S is the saturation flow.

Delay is the most frequently used parameter of effectiveness for intersections. Other measures like length of queue at any given time (Q_T) and number of stops are also useful. Length of queue is used to determine when a given intersection will impede the discharge from an adjacent upstream intersection. The number of stops made is an important input parameter in air quality models.

Problem

The traffic flow for a four-legged intersection is as shown in figure 42:5. Given that the lost time per phase is 2.4 seconds, saturation headway is 2.2 seconds, amber time is 3 seconds per phase, find the cycle length, green time and performance measure(delay per cycle). Assume critical v/c ratio as 0.9.

Solution

- The phase plan is as shown in figure 42:6. Sum of critical lane volumes is the sum of maximum lane volumes in each phase, ΣV_{Ci} = 433+417+233+215 = 1298 vph.
- Saturation flow rate, S_i from equation=$\frac{3600}{2.2}$ = 1637 vph. $\frac{V_c}{S_i} = \frac{433}{1637} + \frac{417}{1637} + \frac{233}{1637} + \frac{1298}{1637} = 0.793$.
- Cycle length can be found out from the equation as C=$\frac{4\times2.4\times0.9}{0.9-\frac{1298}{1637}}$ = 80.68 seconds $\approx$ 80 seconds.

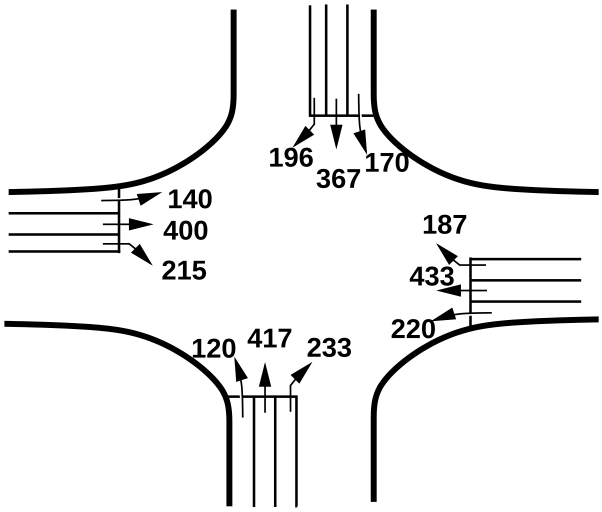


Figure 42:5: Traffic flow for a typical four-legged intersection

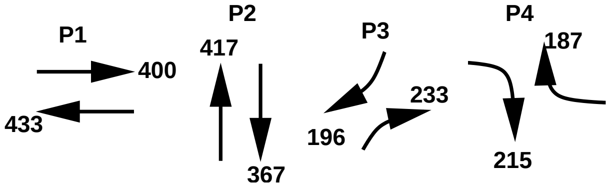


Figure 42:6: Phase plan

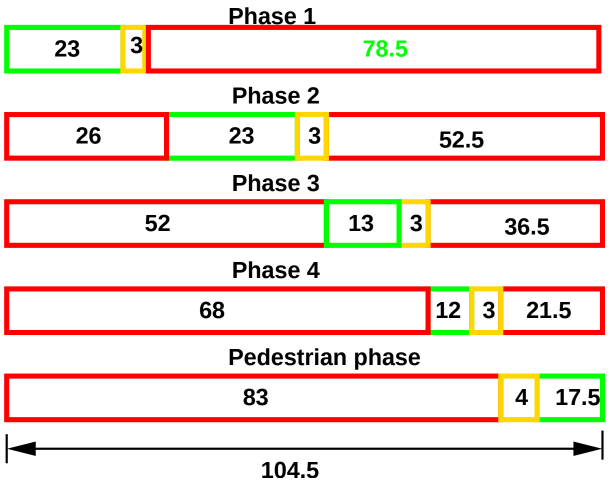


Figure 42:7: Timing diagram

- The effective green time can be found out as $G_i = \frac{V_{Ci}}{V_C} \times (C - L)$ = 80-(4×2.4)= 70.4 seconds, where L is the lost time for that phase = 4× 2.4.
- Green splitting for the phase 1 can be found out as $g_1 = 70.4 \times [\frac{483}{1298}] = 22.88$ seconds.
- Similarly green splitting for the phase 2,$g_2 = 70.4 \times [\frac{417}{1298}] = 22.02$ seconds.
- Similarly green splitting for the phase 3,$g_3 = 70.4 \times [\frac{233}{1298}] = 12.04$ seconds.
- Similarly green splitting for the phase 4,$g_4 = 70.4 \times [\frac{215}{1298}] = 11.66$ seconds.
- The actual green time for phase 1 from equationas G_1= 22.88-3+2.4 ≈ 23 seconds.
- Similarly actual green time for phase 2, G_2 = 22.02-3+2.4 ≈ 23 seconds.
- Similarly actual green time for phase 3, G_3 = 12.04-3+2.4 ≈ 13 seconds.
- Similarly actual green time for phase 4, G_4 = 11.66-3+2.4 ≈ 12 seconds.
- Pedestrian time can be found out from as $G_p = 4 + \frac{6\times3.5}{1.2} = 21.5$ seconds. The phase diagram is shown in figure 42:7. The actual cycle time will be the sum of actual green time plus amber time plus actual red time for any phase. Therefore, for phase 1, actual cycle time = 23+3+78.5 = 104.5 seconds.
- Delay at the intersection in the east-west direction can be found out from equationas

$$d_{EW} = \frac{\frac{104.5}{2}[1 - \frac{23-2.4+3}{104.5}]^2}{1 - \frac{433}{1637}} = 42.57 sec/cycle. \quad (42.7)$$

- Delay at the intersection in the west-east direction can be found out from equation,as

$$d_{WE} = \frac{\frac{104.5}{2}[1 - \frac{23-2.4+3}{104.5}]^2}{1 - \frac{400}{1637}} = 41.44 sec/cycle. \quad (42.8)$$

- Delay at the intersection in the north-south direction can be found out from equation,

$$d_{NS} = \frac{\frac{104.5}{2}[1-\frac{23-2.4+3}{104.5}]^2}{1-\frac{367}{1637}} = 40.36 sec/cycle. \tag{42.9}$$

- Delay at the intersection in the south-north direction can be found out from equation,

$$d_{SN} = \frac{\frac{104.5}{2}[1-\frac{23-2.4+3}{104.5}]^2}{1-\frac{417}{1637}} = 42.018 sec/cycle. \tag{42.10}$$

- Delay at the intersection in the south-east direction can be found out from equation,

$$d_{SE} = \frac{\frac{104.5}{2}[1-\frac{13-2.4+3}{104.5}]^2}{1-\frac{233}{1637}} = 46.096 sec/cycle. \tag{42.11}$$

- Delay at the intersection in the north-west direction can be found out from equation,

$$d_{NW} = \frac{\frac{104.5}{2}[1-\frac{13-2.4+3}{104.5}]^2}{1-\frac{196}{1637}} = 44.912 sec/cycle. \tag{42.12}$$

- Delay at the intersection in the west-south direction can be found out from equation,

$$d_{WS} = \frac{\frac{104.5}{2}[1-\frac{12-2.4+3}{104.5}]^2}{1-\frac{215}{1637}} = 46.52 sec/cycle. \tag{42.13}$$

- Delay at the intersection in the east-north direction can be found out from equation,

$$d_{EN} = \frac{\frac{104.5}{2}[1-\frac{12-2.4+3}{104.5}]^2}{1-\frac{187}{1637}} = 45.62 sec/cycle. \tag{42.14}$$

42.5 Summary

Green splitting is done by proportioning the green time among various phases according to the critical volume of the phase. Pedestrian phases are provided by considering the walking speed and start-up lost time. Like other facilities, signals are also assessed for performance, delay being th e important parameter used.

42.6 Problems

1. Table shows the traffic flow for a four-legged intersection. The lost time per phase is 2.4 seconds, saturation headway is 2.2 seconds, amber time is 3 seconds per phase. Find the cycle length, green time and performance measure. Assume critical volume to capacity ratio as 0.85. Draw the phasing and timing diagrams.

From	To	Flow(veh/hr)
North	South	750
East	West	650
West	East	500

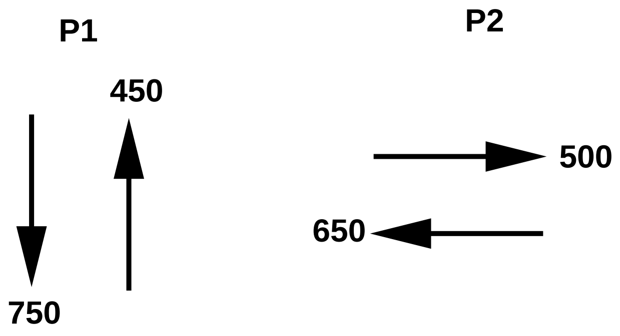


Figure 42:8: Phase diagram

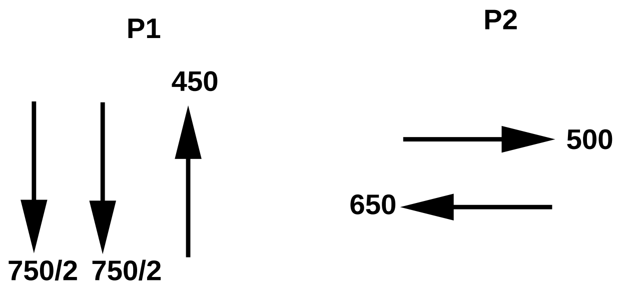


Figure 42:9: Phase diagram

Solution

- Given, saturation headway is 2.2 seconds, total lost time per phase (t_L) is 2.4 seconds, saturation flow = $\frac{3600}{2.2}$ = 1636.36 veh/hr. Phasing diagram can be assumed as in figure 42:9.
- Cycle time C can be found from equation= $\frac{2\times2.4\times0.85}{0.85-\frac{750+650}{1636.36}}$ as negative.
- Hence the traffic flowing from north to south can be allowed to flow into two lanes.
- Now cycle time can be find out as $\frac{2\times2.4\times0.85}{0.85-\frac{450+650}{1636.36}}$ = 22.95 or 23 seconds.
- The effective green time , $t_g = C - (N \times t_L) = 23 - (2 \times 2.4) = 18.2$ seconds.
- This green time can be split into two phases as, For phase 1, $g_1 = \frac{450\times18.2}{1100}$= 7.45 seconds. For phase 2, $g_2 = \frac{650\times18.2}{1100}$ = 10.75 seconds. Now actual green time ,$G_1 = g_1$ minus amber time plus lost time. Therefore, G_1 = 7.45-3+2.4 = 6.85 seconds. G_2 = 10.75-3+2.4 = 10.15 seconds.
- Timing diagram is as shown in figure 42:10
- Delay at the intersection in the east-west direction can be found out from equationas

$$d_{EW} = \frac{\frac{23}{2}[1-\frac{10.75-2.4+3}{23}]^2}{1-\frac{650}{1636.36}} = 4.892 sec/cycle. \tag{42.15}$$

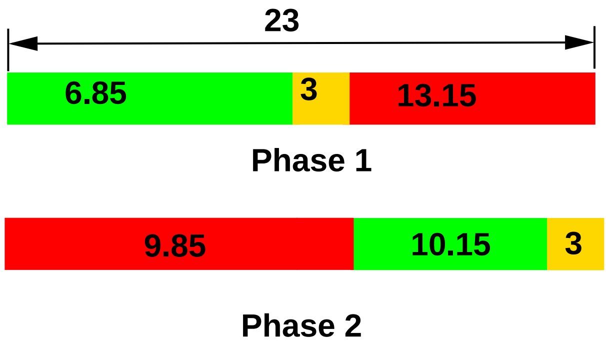


Figure 42:10: Timing diagram

- Delay at the intersection in the west-east direction can be found out from equation,as

$$d_{WE} = \frac{\frac{23}{2}[1 - \frac{10.75-2.4+3}{23}]^2}{1 - \frac{500}{1636.36}} = 4.248 sec/cycle. \tag{42.16}$$

- Delay at the intersection in the north-south direction can be found out from equation,

$$d_{NS} = \frac{\frac{23}{2}[1 - \frac{7.45-2.4+3}{23}]^2}{1 - \frac{750}{1636.36}} = 8.97 sec/cycle. \tag{42.17}$$

- Delay at the intersection in the south-north direction can be found out from equation,

$$d_{SN} = \frac{\frac{23}{2}[1 - \frac{7.45-2.4+3}{23}]^2}{1 - \frac{450}{1636.36}} = 6.703 sec/cycle. \tag{42.18}$$