

SURVEYING

VOLUME 2

FOURTH EDITION

ABOUT THE AUTHOR



S K Duggal is currently Professor in the Department of Civil Engineering at Motilal Nehru National Institute of Technology (MNNIT) Allahabad. He obtained his BE, ME and PhD degrees from Allahabad University. He has an experience of more than 36 years of teaching at both undergraduate and postgraduate levels. Prof. Duggal has published numerous research papers in various international and national journals/conferences and is also the author of *Design of Steel Structures*, 3e and *Limit State Design of Steel Structures*, 1e published by McGraw Hill Education (India) Private Limited.

SURVEYING

VOLUME 2

FOURTH EDITION

S K Duggal

*Professor
Department of Civil Engineering
Motilal Nehru National Institute of Technology
Allahabad*



McGraw Hill Education (India) Private Limited

NEW DELHI

McGraw Hill Education Offices

New Delhi New York St Louis San Francisco Auckland Bogotá Caracas
Kuala Lumpur Lisbon London Madrid Mexico City Milan Montreal
San Juan Santiago Singapore Sydney Tokyo Toronto



McGraw Hill Education (India) Private Limited

Published by McGraw Hill Education (India) Private Limited
P-24, Green Park Extension, New Delhi 110 016

Surveying, Volume 2, 4e

Copyright © 2013, 2009, 2004, 1996, by McGraw Hill Education (India) Private Limited.
No part of this publication may be reproduced or distributed in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise or stored in a database or retrieval system without the prior written permission of the publishers. The program listing (if any) may be entered, stored and executed in a computer system, but they may not be reproduced for publication.

This edition can be exported from India only by the publishers,
McGraw Hill Education (India) Private Limited.

ISBN (13): 978-1-25-902983-7

ISBN (10): 1-25-902983-2

Vice President and Managing Director: *Ajay Shukla*

Head—Higher Education Publishing and Marketing: *Vibha Mahajan*

Publishing Manager—SEM & Tech. Ed: *Shalini Jha*

Sr Editorial Researcher: *Harsha Singh*

Manager—Production Systems: *Satinder S Baveja*

Sr Production Manager: *P L Pandita*

Copy Editor: *Preyoshi Kundu*

Assistant General Manager—Higher Education (Marketing): *Vijay Sarathi*

Product Specialist—SEM & Tech Ed: *Sachin Tripathi*

Sr Graphic Designer (Cover): *Meenu Raghav*

General Manager—Production: *Rajender P Ghansela*

Production Manager—*Reji Kumar*

Information contained in this work has been obtained by McGraw Hill Education (India), from sources believed to be reliable. However, neither McGraw Hill Education (India) nor its authors guarantee the accuracy or completeness of any information published herein, and neither McGraw Hill Education (India) nor its authors shall be responsible for any errors, omissions, or damages arising out of use of this information. This work is published with the understanding that McGraw Hill Education (India) and its authors are supplying information but are not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought.

Typeset at Text-o-Graphics, B-1/56, Aravali Apartment, Sector-34, Noida 201 301, and printed at Magic International Pvt. Ltd., Plot No 26E, Sector 31 (Industrial) Site IV, Greater Noida 201306

Cover Printer: Magic International

RYZLCRZORAABL

*Dedicated to
my family
who inspired me to write this book*

CONTENTS

<i>Preface</i>	xv
<i>Visual Walkthrough</i>	xvi
1. Trigonometric Levelling	1
<i>Introduction</i>	1
1.1 Correction for Curvature and Refraction	2
1.2 Axis Signal Correction	3
1.3 Difference of Elevation of Two Stations by Single Observation	6
1.4 Difference of Elevation of Two Stations by Reciprocal Observations	11
1.5 Determination of Coefficient of Refraction	13
<i>Summary</i>	17
<i>Exercises</i>	17
<i>Objective-type Questions</i>	19
<i>Answers to Objective-type Questions</i>	20
2. Triangulation	21
<i>Introduction</i>	21
2.1 Triangulation Figures or Systems	22
2.2 System of Framework	30
2.3 Classification	31
2.4 Station Marks, Signals and Towers	32
2.5 Reconnaissance	37
2.6 Base Line Measurement	44
2.7 Base Measurement by Rigid Bars	44
2.8 Base Measurement by Flexible Apparatus	46
2.9 Tape Corrections	47
2.10 Radar Ranging	59
2.11 Problems in Base Measurement	62
2.12 Measurement of Angles	64
2.13 Satellite Station and Reduction to Centre	65
2.14 Field Checks in Triangulation	69
2.15 Trilateration	70
2.16 Flare Triangulation	71
2.17 Floating Triangulation	71
<i>Summary</i>	72
<i>Exercises</i>	73
<i>Objective-type Questions</i>	76
<i>Answers to Objective-type Questions</i>	80

3. Theory of Errors and Triangulation Adjustment	81
<i>Introduction</i>	81
3.1 Definitions	82
3.2 Laws of Weight	83
3.3 Laws of Accidental Errors	85
3.4 Principle of Least Squares	94
3.5 Distribution of Error to the Field Measurement	99
3.6 Normal Equations	100
3.7 Method of Correlates	103
3.8 Triangulation Adjustment	105
3.9 Angle Adjustment	105
3.10 Station Adjustment	107
3.11 Figure Adjustment	108
3.12 Adjustment of a Triangle	109
3.13 Spherical Excess	112
3.14 Adjustment of a Geodetic Quadrilateral	112
3.15 Adjustment of Quadrilateral with Central Station	115
3.16 Method of Equal Shifts	119
3.17 Method of Least Squares by Matrices	122
<i>Summary</i>	128
<i>Exercises</i>	128
<i>Objective-type Questions</i>	131
<i>Answers to Objective-type Questions</i>	133
4. Hydrography	134
<i>Introduction</i>	134
4.1 Tides	135
4.2 Mean Sea Level as a Datum Plane	136
4.3 Tide Gauges	137
4.4 Soundings	138
4.5 Sounding Equipment	139
4.6 Locating Sounding	141
4.7 Reduction of Soundings	147
4.8 Plotting the Soundings	147
4.9 Stream Measurements	149
<i>Summary</i>	160
<i>Exercises</i>	161
<i>Objective-type Questions</i>	162
<i>Answers to Objective-type Questions</i>	163
5. Aerial Survey	164
<i>Introduction</i>	164
5.1 Terrestrial Stereophotography	166
5.2 Aerial Photograph	169
5.3 Definitions	170
5.4 Scale of Vertical Photograph	172
5.5 Scale of Tilted Photograph	174

5.6	Displacements and Errors in Aerial Photogrammetry	178
5.7	Procedure of Aerial Survey	186
5.8	Ground Control	186
5.9	Flight Planning	186
5.10	Photo-Interpretation	196
5.11	Basic Photo-Interpretation Equipment	198
5.12	Parallax	201
5.13	Measurement of Parallax	206
5.14	Working up Data from the Photographs	209
5.15	Plotting the Details	209
5.16	Rectification	211
5.17	Fiducial Marks	212
5.18	Photogrammetric Products	212
5.19	Comparison of Products of Aerial Photography	215
5.20	Errors	216
5.21	Mistakes	216
	<i>Summary</i>	216
	<i>Exercises</i>	217
	<i>Objective-type Questions</i>	219
	<i>Answers to Objective-type Questions</i>	224

6. Remote Sensing

225

	<i>Introduction</i>	225
6.1	Remote Sensing Process	226
6.2	Electromagnetic Energy	228
6.3	Electromagnetic Spectrum	229
6.4	Interaction of Electromagnetic Energy with Matter	231
6.5	Effect of Atmosphere on Electromagnetic Radiation	232
6.6	Energy Interaction with Earth Surface Features	234
6.7	Remote-Sensing Sensor Systems	236
6.8	Platforms	239
6.9	Ideal and Real Remote-Sensing Systems	240
6.10	Data Acquisition and Interpretation	242
6.11	Resolution Concept in Remote Sensing	246
6.12	Applications of Remote Sensing	249
6.13	Land Use/Land Cover Analysis	251
6.14	Methodology for Land Use/Land Cover Mapping	256
6.15	Remote Sensing in India	258
6.16	Satellite Data Products	264
	<i>Summary</i>	264
	<i>Exercises</i>	265
	<i>Objective-type Questions</i>	266
	<i>Answers to Objective-type Questions</i>	267

7. Digital Image Processing and Analysis

268

	<i>Introduction</i>	268
7.1	Image Rectification and Restoration	270

7.2	Image Enhancement	280
7.3	Contrast Enhancement	282
7.4	Filtering	287
7.5	Image Transformation	290
7.6	Image Classification	291
7.7	Selective Dip Software	293
	<i>Solved Examples</i>	293
	<i>Summary</i>	297
	<i>Exercises</i>	298
	<i>Objective-type Questions</i>	300
	<i>Answers to Objective-type Questions</i>	300
8.	Geographic Information System (GIS)	301
	<i>Introduction</i>	301
8.1	Subsystems of GIS	303
8.2	Hardware of GIS	305
8.3	Data for GIS	306
8.4	Representation of Features	307
8.5	Data Structure for GIS	309
8.6	Vector Data Structure	309
8.7	Raster Data Structure	311
8.8	Vector vs Raster Data Structures	313
8.9	Data Format Conversions	315
8.10	Capabilities/Functionalities of GIS	317
8.11	Neighbourhood Functions	319
8.12	Map Overlay Analysis	322
8.13	Data Quality	327
8.14	Sources of Errors in GIS	328
8.15	Applications of GIS	332
8.16	Selective GIS Softwares	334
8.17	Remote Sensing and GIS	339
	<i>Summary</i>	340
	<i>Exercises</i>	341
	<i>Objective-type Questions</i>	341
	<i>Answers to Objective-type Questions</i>	342
9.	Global Positioning System (GPS)	343
	<i>Introduction</i>	343
9.1	GLONASS System	345
9.2	Galileo System	346
9.3	GPS Overview	347
9.4	Space Segment (Satellite Constellation)	347
9.5	Operational Control Segment (OCS)	349
9.6	User Equipment Segment	350
9.7	Principle of Position Determination via Satellite Generated Ranging Signals	351
9.8	Determining Satellite-to-User Range	353

9.9	Calculation of User Position	355
9.10	GPS System Time	358
9.11	Carrier Phase Measurement Technique	358
9.12	GPS Error Sources	359
9.13	Differential GPS	363
9.14	GPS Surveying Techniques	366
9.15	Indian Coordinate System for Using GPS	369
9.16	Uses and Applications of GPS	369
9.17	GPS Accuracy	373
	<i>Summary</i>	374
	<i>Exercises</i>	375
	<i>Objective-type Questions</i>	376
	<i>Answers to Objective-type Questions</i>	377

10. Field Astronomy

378

	<i>Introduction</i>	378
10.1	The Earth	379
10.2	Celestial Sphere	380
10.3	Solar System	381
10.4	Definitions	384
10.5	Position of a Celestial Body	390
10.6	Relationship between Coordinates	392
10.7	Spherical Trigonometry and Spherical Triangle	394
10.8	Astronomical Triangle	397
10.9	Napier's Rules of Circular Parts	399
10.10	Time	409
10.11	Conversion of Time	413
10.12	Astronomical Corrections	423
10.13	Determination of Time	427
10.14	Determination of True Meridian	429
10.15	Determination of Azimuth	431
10.16	Determination of Latitude	441
10.17	Determination of Longitude	445
10.18	Nautical Sextant	446
10.19	Theodolite Attachments	451
10.20	Zenith Telescope	456
	<i>Summary</i>	459
	<i>Exercises</i>	460
	<i>Objective-type Questions</i>	462
	<i>Answers to Objective-type Questions</i>	466

Appendices

467

I.	Conversion Tables	467
II.	Digital Terrain Model	468
III.	Physical Laws Governing EMR	470
IV.	Topology	472
V.	Pseudo-random Codes	473

- VI. Orbits of Planets 475
- VII. Physical Characteristics of the Sun, Moon and the Planets 476
- VIII. Physical Characteristics of the Satellites of the Planets 477
- IX. Astronomical Constants 479

Index**481**

PREFACE

About the Book

This volume is in continuation of Volume 1. It deals with the advanced topics of plane and geodetic surveying. As per the international standards, SI units have been used throughout the book. However, some examples in MKS units have also been added since MKS units are still in use. A large number of solved examples have been given to support the text. At the end of each chapter, sufficient objective questions have been given to test the understanding and grasp of the topic.

In this edition, the text has been thoroughly revised, updated, made more comprehensive and informative. Those familiar with the previous editions of the book will note that many changes have been made that update and improve the coverage of the previous editions. This was due to the numerous suggestions made by teachers and professionals who have used the previous edition of the book. Besides adding a chapter on digital image processing, many organisational changes have been made as well. Other changes made in this edition include new or updated material and increased coverage in the chapters on triangulation, aerial surveying, remote sensing, GIS, and GPS.

New to the Edition

- New Chapter on Digital Image Processing and Analysis
- Topical inclusions include
 - Remote Sensing Process
 - Line Scanners and Image Sensors
 - Remote Sensing Satellites in India
 - Glonass System
 - Galileo System
 - Carrier Phase Measurement Techniques and Indian Coordinate System
- Electromagnetic Distance Measurement shifted to Vol. 1
- Chapter-end summary added

Salient Features of this Book

- In-depth coverage of *Remote Sensing*, *GPS* and *GIS*
- Dedicated chapter on *Digital Image Processing and Analysis*
- Points to Remember given in the form of *Notes* at the end of the topics wherever necessary
- Provides exposure to recent developments in classification of surveys due to rapid changes in instrumentation
- Pedagogical features include
 - *Exercises 184*

- *Solved Examples 114*
- *MCQs 220*
- *Figures 228*

Chapter Organisation

In **Chapter 1**, trigonometric levelling as applied to geodetic surveying is discussed. In **Chapter 2**, triangulation has been discussed. Special emphasis has been laid on the electronic distance measurement method. **Chapter 3** deals with the theory of errors and triangulation adjustments. **Chapter 4** discusses hydrography. **Chapter 5** is devoted to aerial survey. **Chapter 6** deals with remote sensing. **Chapter 7** is devoted to digital image processing. **Chapters 8 and 9** deal with GIS and GPS respectively. **Chapter 10** discusses the quantities dealt with in astronomical observations. It is hoped that the arrangement of this chapter and the inclusion of typical solved examples will facilitate an understanding of the fundamental principles of astronomy.

The electronic measurement technique of distance measurement has been shifted to Vol. 1. It was felt that it goes better with the techniques of measurement of horizontal distances. The theory behind laws of accidental errors is included for better understanding of the underlying concepts and expressions presented in the chapter of theory of errors. The satellite Doppler shift technique for positioning hydrographic vessels has been included in Chapter 4. The section of photographic products in Chapter 5 has been revised and enhanced. New or revised coverage as well as new illustrations and figures have been included to explain the different approaches of remote sensing, GIS and GPS. This will make students have a better grip on these methods and techniques and with greater confidence.

In addition to serving as a textbook for the advanced surveying course, this new edition will also serve as a strong base for the professionals and beginners learning the foundation courses of remote sensing, GIS, and GPS. Another goal of this revision has been to broaden the scope of the book. As a result of this and the other changes affected, the fourth edition is more versatile than its predecessors.

Web Supplements

The following web supplements can be accessed at <http://www.mhhe.com/duggal/s4vol2>

For Instructors

- *PowerPoint slides*
- *Solution Manual*

For Students

- *Interactive Objective Questions*

Acknowledgements

I am thankful and would like to express my gratitude to the faculty and students of MNNIT Allahabad who have helped in the preparation of this new edition. The continuous support of Dr. Varun Singh, Assistant Professor in Civil Engineering Department, who helped in identifying the mistakes and made critical comments

about the revised text is gratefully acknowledged. I wish to thank the editorial team at McGraw Hill Education (India) for their sincere efforts in expediting the publication of this book.

I am thankful to the following reviewers who took out time to review the book:

Archana Bohra Gupta	<i>Mugneeram Bangur Memorial Engineering College, Jodhpur, Rajasthan</i>
Sanjeev Verma	<i>Technocrats Institute of Technology Excellence, Bhopal, Madhya Pradesh</i>
Sanjeev Kumar Suman	<i>National Institute of Technology, Patna, Bihar</i>
Arup Bhattacharya	<i>Jorhat Engineering College, Jorhat, Assam</i>
A V Pendse	<i>Vishwakarma Institute of Technology, Pune, Maharashtra</i>
P G Agnihotri	<i>Sardar Vallabhbhai National Institute of Technology, Surat, Gujarat</i>
G Bhaskar	<i>Institute of Road and Transport Technology, Erode, Tamil Nadu</i>
MS Ganesh Prasad	<i>National Institute of Engineering, Mysore, Karnataka</i>
Vazeer Mahmmood	<i>College of Engineering, Andhra University, Andhra Pradesh</i>
Pratik Parmar	<i>Vidhyapeeth Institute of Technology, Bhopal, Madhya Pradesh</i>

The critical comments and support of Prof. A K Sachan and Dr. N R Rawal, Assistant Professor in Civil Engineering Department of MNNIT Allahabad are gratefully acknowledged. Thanks are due to Prof. R D Gupta for his suggestions to make this volume versatile.

Feedback

In spite of all precautions, errors may have crept in. I will be highly obliged if such errors are brought to his notice. Readers may write to me at shashikantduggal@gmail.com.

S K DUGGAL

Publisher's Note

Do you have a feature request or a suggestion? We are always open to new ideas (the best ideas come from you!). You may send your comments to tmh.civilfeedback@gmail.com (kindly mention the title and author name in the subject line).

Piracy-related issue may also be reported.

VISUAL WALKTHROUGH

1

TRIGONOMETRIC LEVELLING

Introduction

Geodetic surveying consists of precisely measuring positions on the earth's surface, of a system of widely separated points. The positions are determined both relatively, in terms of the length and azimuths of the lines joining them, and absolutely, in terms of the coordinates—latitude, longitude, and elevation above mean sea level. These points (of observation) form the control points (stations) to which topographic, hydrographic, engineering and other surveys may be referred so as to localise the error between the geodetic stations.

In geodetic surveys, the horizontal control is established by precise traversing or triangulation and the latter will be discussed in Chapter 2. To establish vertical control there are two methods: precise levelling and trigonometric levelling. Precise levelling has already been discussed in Surveying Vol. I, Chapter 6. It is similar to spirit levelling but is carried out with greater refinement of instruments and techniques.

Trigonometrical levelling may be defined as the process of determining the relative elevations of different stations from observed vertical angles and known distances—either horizontal or geodetic lengths at mean sea level. When the horizontal distance between the points of interest is small, either the corrections for curvature and refraction are neglected or are applied linearly to the calculated

INTRODUCTION

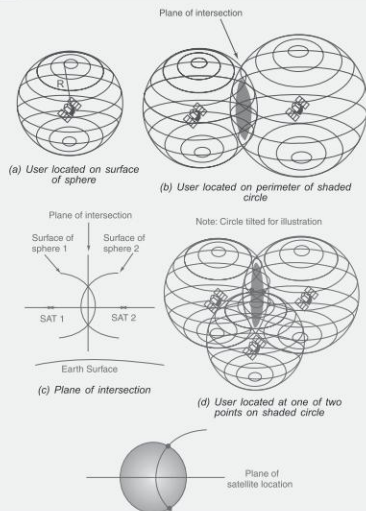
Each chapter begins with an Introduction giving an overview of the chapter as well as its scope.

FIGURES

Figures have been used to describe the instrument details for visualisation in simplified ways.

352

Surveying



UP-TO-DATE INFORMATION

Latest development of electronics-based equipments and electronic techniques of measurements are described.

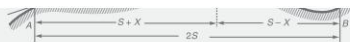


Fig. 2.14

Example 2.4 An observer standing on the deck of a ship just sees a lighthouse top with his eye at a height of 9 m. The top of the lighthouse is 64 m above m.s. Find the distance of the observer from the lighthouse.

Solution Refer to Fig. 2.15.

Let, $AO = D_1$ and $BO = D_2$.

This is a problem of curvature and refraction.

We know that combined correction $= 0.0673D^2$

Hence, $9 = 0.0673 AO^2$

or $AO = \sqrt{\frac{9 \times 100}{0.0673}} = 11.56 \text{ km}$

and $BO = \sqrt{\frac{64}{0.0673}} = 30.83 \text{ km}$

Hence, the distance between the observer and the lighthouse is
 $= 11.56 + 30.83 = 42.39 \text{ km}$

OBJECTIVE-TYPE QUESTIONS

Objective-type questions are helpful in the preparation for various competitive examinations.

7.1 IMAGE RECTIFICATION AND RESTORATION

This operation is often termed as *preprocessing* since it normally precedes further manipulation and analysis of the image data to extract specific information. Distortions or degradations in the image stem from the image acquisition process. Therefore, the aim of image rectification and restoration is to correct the image data and is accomplished by offsetting problems with the band data by recalculating the DN values. The factors that affect are the *digital image acquisition type*—digital camera, along-track scanner, across-track scanner; *platform*—airborne vs. satellite; and *total field of view*. The preprocessing operation attempts to correct or compensate the remotely sensed data for errors such as *radiometric errors*—removal of sensor or atmospheric noise to represent ground conditions more accurately, i.e., to improve image fidelity by correcting for data loss, removing haze; and then *geometric distortions*—converting data to ground coordinates by removal of distortions from sensor geometry, enabling mapping relative to data layers. Once the errors are removed, the data is said to be *restored* to its original condition; the correct values are never known. Further, the attempt to correct the data may itself introduce errors. However, all these corrections might not be required in every case and will depend upon the nature of information to be extracted from the data.

7.1.1 Radiometric Corrections

When the image data contains errors in the measured brightness values of the pixels, it limits the ability of the analyst to visually interpret or quantitatively process and analyse the images. The potential sources of errors are: periodic drift or malfunctioning of a detector; electronic interference between sensor components;

EXAMPLES

Examples help readers to relate and apply the chapter content to real-life problems.

whose top and bottom was 0.62 mm. Determine the height of the cliff, the base on the photograph measured was 50 mm. [Ans. 28.52 m]

5.25 Explain why the heights of buildings, trees, etc., appear to be exaggerated when a pair of vertical air photographs is viewed stereoscopically.

5.26 What is meant by stereophotography? Two distant objects, at the left and right of photoedditone line of sight, make an angle of 47° on the instrument station. In the positive print, they are seen to be 70 mm left and 60 mm right of the photo centre. Calculate the focal length of the camera lens.

[Ans. $f = 149.35 \text{ mm}$]

Objective-type Questions

- 5.1** A spatial view can be obtained with naked eyes for a distance up to
 - (a) 240 m
 - (b) 640 m
 - (c) 1140 m
 - (d) 1540 m
- 5.2** Which of the following is an appropriate scale of an air photograph?
 - (a) 1 cm = 10 m
 - (b) 1 cm = 100 m
 - (c) 1 cm = 500 m
 - (d) 1 cm = 1000 m
- 5.3** The focal length of an aerial camera is 12.5 cm and the exposure station was 2500 m above the datum. The scale of the photograph is
 - (a) 1 cm = 200 m
 - (b) 1 cm = 31 250 m
 - (c) 1:2000
 - (d) 1:3125
- 5.4** An aircraft is flying at a ground speed of 180 km/h. The focal length of the camera is 150 mm. The longitudinal coverage of the ground is 900 m. Which one is the exposure interval.
 - (a) 18 s
 - (b) 50 s
 - (c) 750 s
 - (d) 270 s
- 5.5** If the distance of the image of a point A from the isocentre, on a tilted photograph, is 15.6 mm and the same distance is equal to 14.9 mm on a vertical photograph, tilt distortion is
 - (a) 2.9 mm
 - (b) 0.007 mm
 - (c) 0.7 mm
 - (d) 0.29 mm
- 5.6** A vertical photograph was taken at an altitude of 1500 m above M.S.L. If the focal length of the camera is 20 m, the scale of photograph for a terrain lying at an elevation of 500 m is
 - (a) 1:50
 - (b) 1:100
 - (c) 1:1000
 - (d) 1:25
- 5.7** Two consecutive photographs were taken with an overlap of 55% and the print size was 22.5 cm $\times$ 22.5 cm for both the photographs. All the conditions were same and the scale of the photograph was 1 cm = 200 m. The length of the air base is

FIELD PROBLEMS

Applications of the principles are given to solve field problems.

9.12 GPS ERROR SOURCES

If the world was like a laboratory with perfect laboratory conditions, basic GPS would be a lot more accurate. Unfortunately, it is not so, and has plenty of opportunities for a radio-based system that spans the entire planet to get fouled up. Inaccuracies in GPS signals come from the variety of sources (Fig. 9.7), like satellite clocks, imperfect orbits and especially from the signal's trip through the earth's atmosphere. Since these inaccuracies are variable, it is hard to predict what they will be in order to correct for them. Although, these errors are small, but to get the kind of accuracies some critical positioning jobs require, all the errors, no matter how minor, are to be minimised. What is needed is a way to measure the actual errors as they happen. The error sources can be classified into three groups, namely satellite-related errors, propagation-medium related errors, and

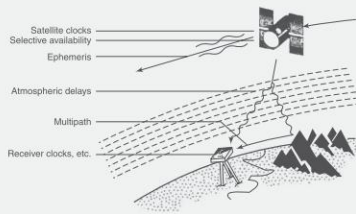


Fig. 9.7 Error sources

Appendix I

CONVERSION TABLES

1 mile	= 1.6093 km
	= 8 furlongs
	= 1760 yards
	= 5280 ft
	= 63360 in.
1 furlong	= 0.2012 km
	= 220 yards

APPENDICES

Appendices contain supplementary matter helpful in enriching the text.

QUANTIFICATION

Quantification of instrumental errors are explained with illustrative figures.

Let angle A be measured three times by the repetition method

$\angle A$	Weight
48°20'8"	3
48°20'10"	2
48°20'12"	3

Sum of the weights = 3 + 2 + 3 = 8

$$\begin{aligned}\text{Weighted arithmetic mean} &= 48^\circ 20' + \frac{1}{8} \times (8'' \times 3 + 10'' \times 2 + 12'' \times 3) \\ &= 48^\circ 20' 10''\end{aligned}$$

Weight of the weighted arithmetic mean 48°20'10" is 8.

3. If two or more quantities are added algebraically, the weight of the result is equal to the reciprocal of the sum of the reciprocals of the individual weights.

For example,

Angle	Weight
$\alpha = 46^\circ 30' 30''$	3
$\beta = 24^\circ 42' 20''$	2

$$\text{Weight of } \alpha + \beta (71^\circ 12' 50'') = \frac{1}{\frac{1}{3} + \frac{1}{2}} = \frac{6}{5}$$

1

TRIGONOMETRIC LEVELLING

Introduction

Geodetic surveying consists of precisely measuring positions on the earth's surface, of a system of widely separated points. The positions are determined both relatively, in terms of the length and azimuths of the lines joining them, and absolutely, in terms of the coordinates—latitude, longitude, and elevation above mean sea level. These points (of observation) form the control points (stations) to which topographic, hydrographic, engineering and other surveys may be referred so as to localise the error between the geodetic stations.

In geodetic surveys, the horizontal control is established by precise traversing or triangulation and the latter will be discussed in Chapter 2. To establish vertical control there are two methods: precise levelling and trigonometric levelling. Precise levelling has already been discussed in Surveying Vol. I, Chapter 6. It is similar to spirit levelling but is carried out with greater refinement of instruments and techniques.

Trigonometrical levelling may be defined as the process of determining the relative elevations of different stations from observed vertical angles and known distances—either horizontal or geodetic lengths at mean sea level. When the horizontal distance between the points of interest is small, either the corrections for curvature and refraction are neglected or are applied linearly to the calculated difference in elevation. This is applicable for plane surveying and this method of trigonometrical levelling is discussed in Surveying Vol. 1, Chapter 6. However, when distances are large, the instrument station is located at the centre of the area to observe the vertical angles outwards from this point. In this case, it is usual to apply corrections for curvature and refraction in angular measures directly to the observed vertical angles and the procedure is described in the present chapter.

The measurement of vertical angles between triangulation stations is usually undertaken at the same time as that of horizontal angles. The concurrent observation of the vertical angles alongwith the horizontal angles may prove to be economical, and for that reason it may be convenient to deal with the determination of height by the observation of vertical angles, rather than by precise levelling. However, the method of trigonometrical levelling is less accurate than precise (spirit) levelling in flat terrains. In mountainous regions and rugged terrains, the method is of great value, and the results are comparatively more accurate.

Depending upon the number of stations that can be occupied by the theodolite, there are two methods of determining the difference in elevations of two stations. When the observation is made from one of the stations, it is known as the *single observation method or direct method*, whereas in the other case when observations are made from both the stations, it is known as the *reciprocal observation method*. The latter one is preferred since the errors due to curvature and uncertain refraction are eliminated. Since the observations in the direct method are influenced by irregularities in the coefficient of refraction, it is desirable to make the measurement during the time of minimum refraction effect, i.e., early morning or otherwise, wherever practicable, the reciprocal method of observation should be preferred.

1.1 CORRECTION FOR CURVATURE AND REFRACTION

Let A and B be two stations and it is required to determine the true difference in their elevations (Fig. 1.1). The level lines through A and B are AA_1 and BB_1 , respectively. AA_1 and BB_1 are the horizontal lines through A and B , respectively. H is the true difference of elevation between A and B . D is the spheroidal distance AB and d is the corresponding horizontal distance AB . θ is the angle subtended by the distance AB at the centre of the earth, γ the angle of refraction, and m the coefficient of refraction. The angle of refraction at the two stations are assumed to be equal. C is the angle of curvature and R the radius of the earth. An observer at A sights along AB_2 instead of AB . α is the angle of elevation at A and β is the angle of depression at B . Therefore, the actual observed vertical angle at A is B_2AA_1 while the true angle is BAA_1 . Similarly at B , the actual observed vertical angle is A_2BB_1 while the true angle is ABB_1 . The corrected observed angles for refraction are therefore $(\alpha - \gamma)$ and $(\beta + \gamma)$. It is a usual practise to express the angle γ in terms of a fraction of angle θ , as $\gamma = m\theta$.

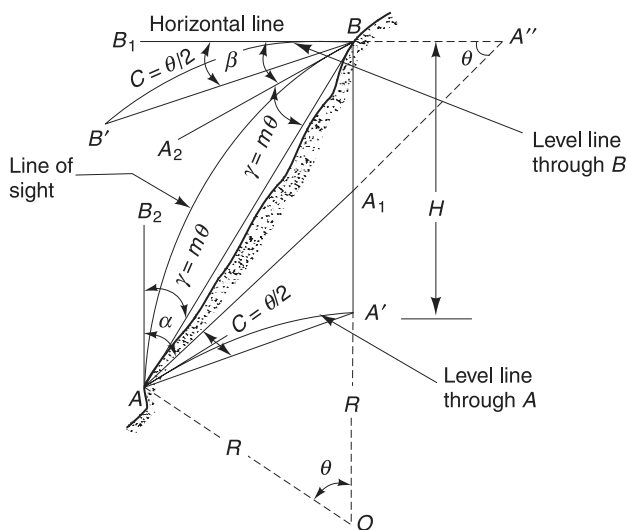


Fig. 1.1 Curvature and refraction

$$\text{or} \quad \gamma = \frac{mD}{R \sin 1''} \quad (1.1)$$

$$\text{Now,} \quad \angle B_1BA = \angle BAA'' + \angle BA''A$$

$$\text{or} \quad \beta + \gamma = (\alpha - \gamma) + \theta$$

$$\text{or} \quad \gamma = \frac{\alpha - \beta + \theta}{2} \quad (1.2)$$

When both the angles are angles of depression, then

$$\gamma = \frac{\theta - \alpha - \beta}{2} \quad (1.3)$$

Angular correction for curvature, $C = \angle A_1AA'$. Hence,

$$C = \frac{\theta}{2} \quad (1.4)$$

$$\text{or} \quad C = \frac{D}{2R \sin 1''} \quad (1.5)$$

$$\begin{aligned} \text{Total correction} &= -\frac{mD}{R \sin 1''} + \frac{D}{2R \sin 1''} \\ &= \frac{D}{2R \sin 1''} (1 - 2m) \end{aligned} \quad (1.6)$$

Note The combined correction is additive to an angle of elevation and subtractive to an angle of depression.

1.2 AXIS SIGNAL CORRECTION

Signals are used at stations to observe vertical angles. The height of signals used is usually different at various stations so as to ensure intervisibility. Further, the height of instrument axis and that of signal at the station sighted are never the same. This results in error in the observed vertical angles and they need to be corrected. The correction to be applied is known as *axis signal correction*, or *eye and object correction*, or *instrument and signal correction*.

Let A and B be two stations. The vertical angle is to be observed from A to B . If the height of instrument at A is equal to the height of signal at B , the observed angle will be the true vertical angle. But, since the two heights h_1 and S_2 (Fig. 1.2) will not be same, a correction has to be applied.

Let h_1 be the height of instrument at A and S_2 be the height of signal at B . α is the observed vertical angle and α_1 is the corrected vertical angle for the axis signal. Let D be the horizontal distance AB (Fig. 1.2) equal to the spheroidal distance AA' (Fig. 1.3). Axis signal correction is δ_1 ($\angle BAE$ in Fig. 1.3). In triangle ABO ,

$$\angle BAO = \angle A_1AO + \angle BAA_1 = 90^\circ + \alpha$$

$$\angle AOB = \theta$$

$$\angle ABO = 180^\circ - [(90^\circ + \alpha) + \theta] = 90^\circ - (\alpha + \theta)$$

$$\angle EBF = 90^\circ - [90^\circ - (\alpha + \theta)] = \alpha + \theta$$

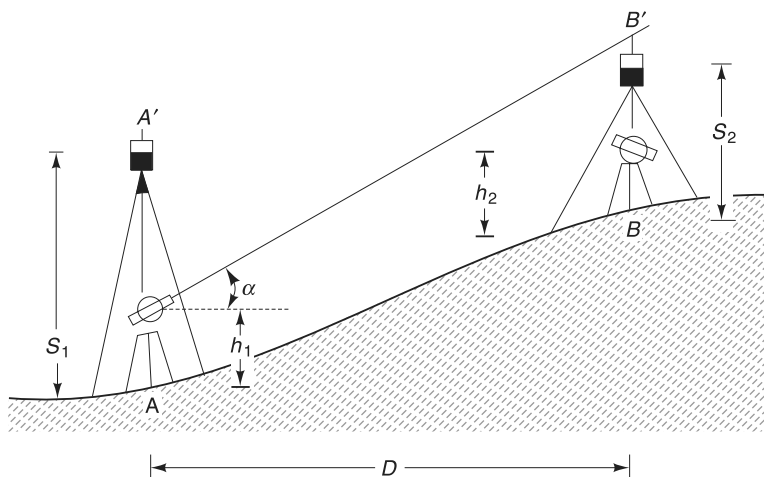


Fig. 1.2

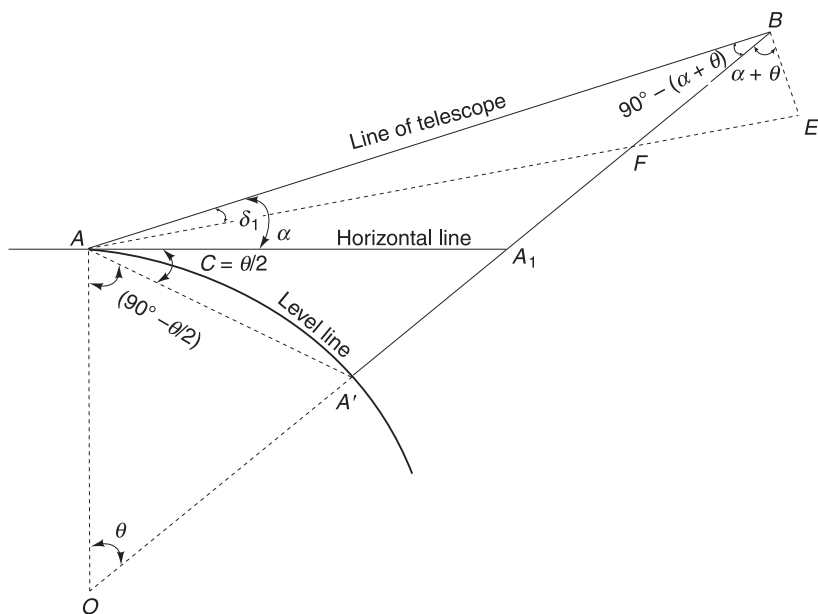


Fig. 1.3 Axis signal correction

The angle δ_1 is usually quite small and therefore angle $BEF = 90^\circ$

$$BE = BF \cos(\alpha + \theta) = (S_2 - h_1) \cos(\alpha + \theta)$$

In triangle $AA'B$

$$\angle BAA' = \alpha + \frac{\theta}{2}; \quad \angle ABA' = 90^\circ - (\alpha + \theta)$$

$$\angle AA'B = 180^\circ - \left[(90^\circ - (\alpha + \theta)) + \left(\alpha + \frac{\theta}{2} \right) \right] = 90^\circ + \frac{\theta}{2}$$

By sine rule,

$$\frac{AB}{\sin \angle AA'B} = \frac{AA'}{\sin \angle ABA'}$$

or
$$\frac{AB}{\sin \left(90^\circ + \frac{\theta}{2} \right)} = \frac{D}{\sin [90^\circ - (\alpha + \theta)]}$$

or
$$AB = D \frac{\sin \left(90^\circ + \frac{\theta}{2} \right)}{\sin [90^\circ - (\alpha + \theta)]} = D \frac{\cos \frac{\theta}{2}}{\cos(\alpha + \theta)} \quad (1.7)$$

From triangle ABE

$$\tan \delta_1 = \frac{BE}{AB} = \frac{(S_2 - h_1) \cos(\alpha + \theta)}{D \left[\frac{\cos \frac{\theta}{2}}{\cos(\alpha + \theta)} \right]}$$

or
$$\tan \delta_1 = \frac{(S_2 - h_1) \cos^2(\alpha + \theta)}{D \cos \frac{\theta}{2}} \quad (1.8)$$

Usually θ is very small as compared to α , hence the equation can be reduced to

$$\tan \delta_1 = \frac{(S_2 - h_1) \cos^2 \alpha}{D} \quad (1.9)$$

If the vertical angle is small, the equation reduces to

$$\delta_1 = \frac{S_2 - h_1}{D \sin 1''} \quad (1.10)$$

Similarly, if observations are made from B towards A with β as the observed vertical angle and β_1 the corrected vertical angle for axis signal correction, it can be shown that

$$\tan \delta_2 = \frac{(S_1 - h_2) \cos^2 \beta}{D}$$

Note The axis signal correction is negative for an angle of elevation and positive for an angle of depression.

Example 1.1 A vertical angle of elevation was observed from a station P as $2^\circ 32' 25''$. Determine its true value if the height of instrument at P is 1.2 m and height of signal at the other station Q is 5.2 m. The two stations P and Q are 5200 m apart. Take the value of $R \sin 1''$ as 30.88 m. The coefficient of refraction may be assumed to be 0.07. Find also the true value of the angle observed if it was an angle of depression.

Solution

Case I:

$$\begin{aligned}\alpha &= 2^\circ 32' 25'' \\ h_1 &= 1.2 \text{ m} \\ S_2 &= 5.2 \text{ m} \\ D &= 5200 \text{ m} \\ m &= 0.07\end{aligned}$$

Axis signal correction, $\delta = \frac{S_2 - h_1}{D \sin 1''}$

$$\begin{aligned}&= \frac{5.2 - 1.2}{5200} \times 206265 \\ &= 158.66'' = 2' 38.66''(-)\end{aligned}$$

Central angle, $\theta = \frac{D}{R \sin 1''}$

$$\begin{aligned}&= \frac{5200}{30.88} \\ &= 168.39'' = 2' 48.39''\end{aligned}$$

Curvature correction, $C = \frac{\theta}{2} = \frac{168.39}{2}$

$$= 84.19'' = 1' 24.19''(+)$$

Refraction correction, $\gamma = m\theta = 0.07 \times 168.39'' = 11.78''(-)$

$\therefore$ corrected vertical angle, $\alpha_1 = 2^\circ 32' 25'' - 2' 38.66'' + 1' 24.19'' - 11.78''$

$$= 2^\circ 30' 58.75''.$$

Case II:

Observed vertical angle $= -2^\circ 32' 25''$

Axis signal correction $= 2' 38.66'' (+)$

Curvature correction $= 1' 24.19''(-)$

Refraction correction $= 11.78''(+)$

The correct vertical angle $= -2^\circ 32' 25'' + 2' 38.66'' - 1' 24.19'' + 11.78''$

$$= -2^\circ 30' 58.7''$$

1.3 DIFFERENCE OF ELEVATION OF TWO STATIONS BY SINGLE OBSERVATION

When it is not possible to occupy both the stations, one of them being inaccessible, this method of determining the difference in elevation between the two stations is used. The vertical angle observed is corrected for curvature and refraction. Since the coefficient of refraction varies with temperature, this method does not yield accurate results. The effect of refraction is opposite to that of curvature tending to increase or decrease the vertical angle according to whether the angle is of elevation or depression.

1.3.1 Approximate Expressions

In Fig. 1.4, let A and B be the two stations and it is required to determine the difference of elevation between them. It may be computed by observing the vertical angle from any one of these stations, say A , and if the spheroidal distance D between these stations is known. The observed vertical angle is α (angle of the elevation), the angle of refraction is γ and the angle of curvature is C . Assuming $AA'B$ as a plane right angled triangle, an approximate expression for height difference can be developed. Difference in height of A and $B = A'B = H = D \tan \phi$

where

$$\phi = \alpha + C - \gamma$$

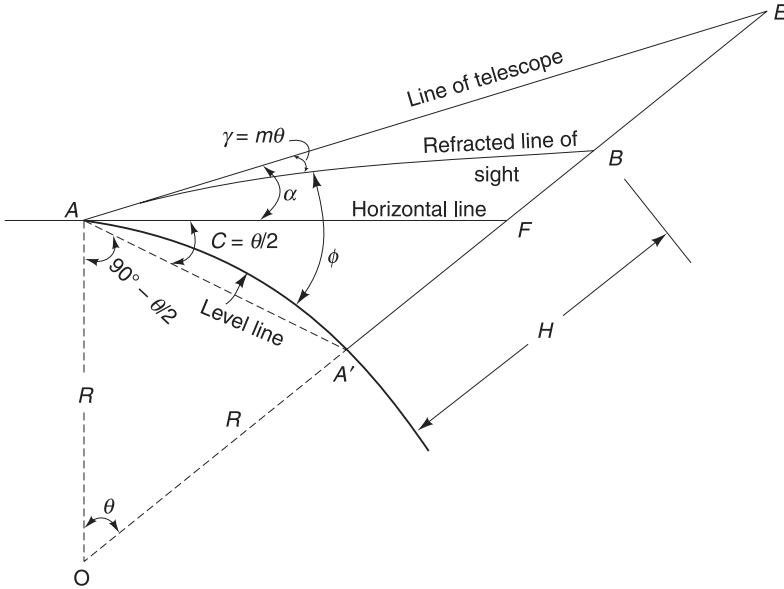


Fig. 1.4 Single observation method (angle of elevation)

$$\therefore H = D \tan [\alpha + C - \gamma] \quad (1.11)$$

$$\text{or} \quad H = D \tan [\alpha + (\theta/2 - m\theta)]$$

$$\text{or} \quad H = D \tan \left[\alpha + (1-2m) \frac{D}{2R \sin 1''} \right] \quad (1.12)$$

In case if the vertical angle (β) observed is the angle of depression, as at station B (Fig. 1.5) the approximate expression for difference in elevation, will be the difference in heights of B and A .

$$B'A = H = D \tan \phi$$

where

$$\phi = \beta - C + \gamma = \beta - (C - \gamma)$$

or

$$H = D \tan [\beta - (C - \gamma)]$$

or

$$H = D \tan \left[\beta - \left(\frac{\theta}{2} - m\theta \right) \right]$$

or

$$H = D \tan \left[\beta - (1-2m) \frac{D}{2R \sin 1''} \right] \quad (1.13)$$

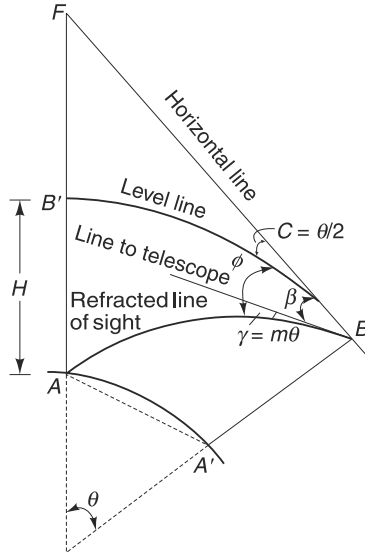


Fig. 1.5 Single observation method (angle of depression)

1.3.2 Exact Expressions

In Fig. 1.4, α is the observed vertical angle of elevation, D is the horizontal distance AA' and let it be equal to the spheroidal distance AA' . Let α_1 be the angle corrected for axis signal

$$\angle BAA' = \alpha + C - \gamma = \alpha + \frac{\theta}{2} - m\theta$$

$$\angle AA'B = \left(90^\circ - \frac{\theta}{2} \right) + \theta = 90^\circ + \frac{\theta}{2}$$

$$\angle ABA' = 180^\circ - \left(\alpha + \frac{\theta}{2} - m\theta \right) - \left(90^\circ + \frac{\theta}{2} \right) = 90^\circ - (\alpha + \theta - m\theta)$$

Applying sine rule

$$\frac{BA'}{\sin BAA'} = \frac{AA'}{\sin ABA'}$$

$$\text{or} \quad \frac{H}{\sin \left(\alpha + \frac{\theta}{2} - m\theta \right)} = \frac{D}{\sin [90^\circ - (\alpha + \theta - m\theta)]}$$

$$\text{or} \quad H = D \frac{\sin(\alpha + \theta/2 - m\theta)}{\cos(\alpha + \theta - m\theta)}$$

We know that,

$$D = R\theta$$

$$\begin{aligned}
 \text{or} \quad \theta &= \frac{D}{R} = \frac{D}{R \sin 1''} \\
 \therefore H &= D \frac{\sin \left(\alpha + \frac{D}{2R \sin 1''} - \frac{mD}{R \sin 1''} \right)}{\cos \left(\alpha + \frac{D}{R \sin 1''} - \frac{mD}{R \sin 1''} \right)} \\
 \text{or} \quad H &= D \frac{\sin \left(\alpha + (1-2m) \frac{D}{2R \sin 1''} \right)}{\cos \left(\alpha + (1-m) \frac{D}{R \sin 1''} \right)} \quad (1.14)
 \end{aligned}$$

Equation (1.14) can be modified by replacing α by α_1 .

In Fig. 1.5, β is the observed vertical angle of depression, β_1 is the angle corrected for axis signal

$$\begin{aligned}
 \angle ABB' &= \beta - C + \gamma \\
 &= \beta - \frac{\theta}{2} + m\theta \\
 \angle BB'A &= (90^\circ - \theta) + \frac{\theta}{2} \\
 &= 90^\circ - \frac{\theta}{2} \\
 \angle B'AB &= 180^\circ - (90^\circ - \theta/2) - (\beta - \theta/2 + m\theta) \\
 &= 90^\circ - (\beta - \theta + m\theta)
 \end{aligned}$$

In triangle BAB'

$$\begin{aligned}
 \frac{AB'}{\sin ABB'} &= \frac{BB'}{\sin B'AB} \\
 \text{or} \quad \frac{H}{\sin \left(\beta - \frac{\theta}{2} + m\theta \right)} &= \frac{D}{\sin [90^\circ - (\beta - \theta + m\theta)]} \\
 \text{or} \quad H &= D \frac{\sin(\beta - \theta/2 + m\theta)}{\cos(\beta - \theta + m\theta)} \\
 \text{or} \quad H &= D \frac{\sin \left[\beta - (1-2m) \frac{D}{2R \sin 1''} \right]}{\cos \left[\beta - (1-m) \frac{D}{R \sin 1''} \right]} \quad (1.15)
 \end{aligned}$$

Equation (1.15) can be modified by replacing β by β_1 .

Example 1.2 Two triangulation stations A and B are 3200.65 m apart. Find the difference of elevation of the two stations for the following data:

Angle of depression at B to A	$= 2^\circ 18' 16''$
Height of signal at A	$= 4.23 \text{ m}$
Height of instrument at B	$= 1.24 \text{ m}$
Coefficient of refraction at B	$= 0.07$
$R \sin 1''$	$= 30.88 \text{ m}$
R.L. of B	$= 242.6 \text{ m}$

Solution

$$\begin{aligned}\text{Axis signal correction } \delta &= \frac{S_2 - h_1}{D \sin 1''} \\ &= \frac{4.23 - 1.24}{3200.65} \times 206\,265 \\ &= 192.69'' = 3' 12.69'' (+)\end{aligned}$$

$$\begin{aligned}\text{Vertical angle corrected for axis signal, } \beta_1 &= 2^\circ 18' 16'' + 3' 12.69'' \\ &= 2^\circ 21' 28.69''\end{aligned}$$

$$\begin{aligned}\text{Central angle, } \theta &= \frac{D}{R \sin 1''} \\ &= \frac{3200.65}{30.88} = 103.64'' = 1' 43.64''\end{aligned}$$

$$\text{Curvature correction, } C = \frac{\theta}{2} = \frac{103.64''}{2} = 51.82''$$

$$\text{Refraction correction, } \gamma = m\theta = 0.07 \times 103.64 = 7.25''$$

$$\begin{aligned}H &= D \frac{\sin(\beta_1 + m\theta - \theta/2)}{\cos(\beta_1 + m\theta - \theta)} \\ &= \frac{3200.65 \times \sin(2^\circ 21' 28.69'' + 7.25'' - 51.82'')}{\cos(2^\circ 21' 28.69'' + 7.25'' - 1' 43.64'')} \\ &= 131.12 \text{ m} \\ \therefore \text{ R.L. of } A &= 242.6 - 131.12 \\ &= 111.48 \text{ m}\end{aligned}$$

Example 1.3 Two triangulation stations A and B are 2800 m apart. Observations were made for vertical angle of elevation from A to B and the mean angle observed was $1^\circ 28' 32''$. The height of the instrument was 1.38 m and the signal was 2.46 m high. If the reduced level of station A was 125 m and the coefficient of refraction was 0.07, calculate the reduced level of B . The radius of the earth is 6372 km.

$$\text{Solution} \quad \text{Axis signal correction} = \frac{2.46 - 1.38}{2800} \times 206\,265 = 1' 19.56''$$

$$\begin{aligned}\text{Vertical angle corrected for axis signal, } \alpha_1 &= 1^\circ 28' 32'' - 1' 19.56'' \\ &= 1^\circ 27' 12.44''\end{aligned}$$

$$\text{Central angle, } \theta = \frac{D}{R \sin 1''}$$

$$= \frac{2800}{6372 \times 10^3} \times 206265 = 90.637''$$

$$\text{Correction for curvature, } C = \frac{\theta}{2} = \frac{90.637''}{2} = 45.318''$$

$$\text{Correction for refraction, } \gamma = m\theta = 0.07 \times 90.637'' = 6.34''$$

$$\begin{aligned} H &= D \frac{\sin(\alpha_1 - m\theta + \theta/2)}{\cos(\alpha_1 - m\theta + \theta)} \\ &= \frac{2800 \times \sin(1^\circ 27' 12.44'' - 6.34'' + 45.318'')}{\cos(1^\circ 27' 12.44'' - 6.34'' + 90.637'')} \\ &= 71.578 \text{ m} \end{aligned}$$

$\therefore$ the reduced level of $B = 125 + 71.578 = 196.578 \text{ m}$

1.4 DIFFERENCE OF ELEVATION OF TWO STATIONS BY RECIPROCAL OBSERVATIONS

In this method of finding out the difference of elevation of two stations, the observations are made simultaneously from both the stations to eliminate the effect of refraction completely. As far as possible, the observations are made at the time of minimum refraction, i.e., between 10 a.m. and 3 p.m. The mean of the vertical angles is obtained and used for calculating the difference in elevation.

In Fig. 1.6, A and B are the two stations and the difference of elevation of these is required. α and β are the vertical angles observed simultaneously at the two stations. The angle of refraction is γ and the angle of curvature is C at the two stations. Let the corrected vertical angles for signal be α_1 and β_1 .

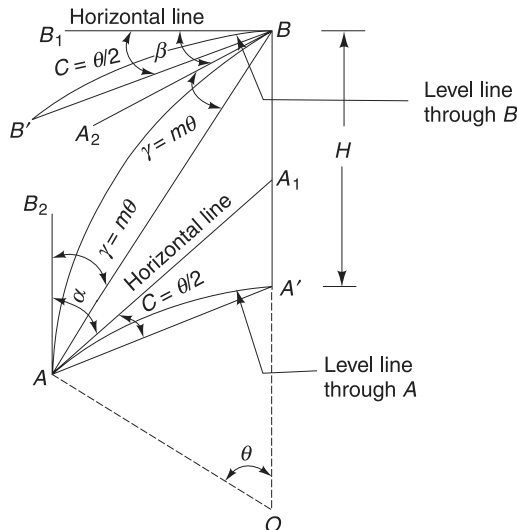


Fig. 1.6 Difference in elevation by reciprocal observations

The mean vertical angle = $\frac{\alpha_1 + \beta_1}{2}$ (1.16)

$$\angle BAA' = \left[\alpha + \left(\frac{\theta}{2} - m\theta \right) \right]$$

The sign of correction being positive for angle of elevation.

$$\angle ABB' = \left[\beta - \left(\frac{\theta}{2} - m\theta \right) \right]$$

The sign of correction being negative for angle of depression.

Since the chords AA' and BB' are parallel, $\angle BAA' = \angle ABB'$

$$\therefore \alpha + \left(\frac{\theta}{2} - m\theta \right) = \beta - \left(\frac{\theta}{2} - m\theta \right) = \frac{\alpha + \beta}{2} \quad (1.17)$$

In triangle $AA'B$

$$AA' = D \text{ (the spheroidal distance)}$$

$$\angle BAA' = \alpha + \frac{\theta}{2} - m\theta$$

$$\angle ABA' = 90^\circ - (\beta + m\theta)$$

By sine rule

$$\frac{BA'}{\sin BAA'} = \frac{AA'}{\sin ABA'}$$

$$\frac{H}{\sin \left(\alpha + \frac{\theta}{2} - m\theta \right)} = \frac{D}{\sin [90^\circ - (\beta + m\theta)]}$$

or
$$H = D \frac{\sin \left(\alpha + \frac{\theta}{2} - m\theta \right)}{\cos(\beta + m\theta)} \quad (1.18)$$

But from Eq. (1.17)

$$\beta + m\theta = \frac{\alpha + \beta}{2} + \frac{\theta}{2}$$

and
$$\alpha + \frac{\theta}{2} - m\theta = \frac{\alpha + \beta}{2}$$

$$\therefore H = D \frac{\left(\sin \frac{\alpha + \beta}{2} \right)}{\cos \left(\frac{\alpha + \beta}{2} + \frac{\theta}{2} \right)} \quad (1.19)$$

For small distance, $\theta/2$ being small may be neglected and hence,

$$H = D \frac{\sin \left(\frac{\alpha + \beta}{2} \right)}{\cos \left(\frac{\alpha + \beta}{2} \right)} = D \tan \left(\frac{\alpha + \beta}{2} \right) \quad (1.20)$$

Equation (1.20) can be modified by replacing α and β by α_1 and β_1 , the angles corrected for axis signal.

1.5 DETERMINATION OF COEFFICIENT OF REFRACTION

The coefficient of refraction m is a measure of the curvature of the line of sight and is the ratio of the radius of curvature of the earth R , to the radius of curvature of the line of sight R_s .

$$m = \frac{R}{R_s}$$

To determine the value of m , several reciprocal observations are made and a mean of all the values of coefficient of refraction is obtained.

In Fig. 1.6,

$$C = \frac{\theta}{2} \quad (1.21)$$

$$\text{and} \quad \gamma = m\theta \quad (1.22)$$

$$\text{Also,} \quad \alpha + (C - \gamma) = \beta - (C - \gamma)$$

$$\text{or} \quad 2\gamma = \alpha - \beta + 2C$$

$$\text{or} \quad \gamma = \frac{1}{2}(\alpha - \beta + 2C)$$

$$\text{or} \quad \gamma = \frac{1}{2}\left(\alpha - \beta + 2\frac{\theta}{2}\right)$$

$$\text{or} \quad \gamma = \frac{1}{2}(\theta + \alpha - \beta) \quad (1.23)$$

From Eqs. (1.22) and (1.23)

$$m\theta = \frac{1}{2}(\theta + \alpha - \beta)$$

$$\text{or} \quad m = \frac{\theta + \alpha - \beta}{2\theta} \quad (1.24)$$

Note The usual value of m obtained by this method is 0.07. It should also be noted that while defining the refraction correction linearly, it is assumed to be 1/7th of the curvature correction $\left(\frac{1}{7} \frac{D^2}{2R} = 0.14 \frac{D^2}{2R} = 2m \frac{D^2}{2R}\right)$. Therefore, in such a case $2m = 0.14$, i.e., m has a value of 0.07.

Example 1.4 Two stations A and B are 1700 m apart. The observations recorded were as follows.

	Station A	Station B	
Height of instrument	1.39 m	1.46 m	
Height of signal	2.2 m	2.00 m	$R \sin 1'' = 30.88 \text{ m}$
Vertical angle	$+1^\circ 08' 05''$	$-1^\circ 06' 10''$	

Calculate the difference of level between A and B , the coefficient of refraction and the refraction correction.

Solution

Vertical angle from A to B , $\alpha = 1^\circ 08' 05''$

$$\text{Axis signal correction to } \alpha = \frac{2 - 1.39}{1700} \times 206\,265 = 74.012'' \text{ (-)}$$

$$\therefore \text{vertical angle corrected for axis signal, } \alpha_1 = 1^\circ 08' 05'' - 74.012'' \\ = 1^\circ 06' 50.98''$$

Vertical angle from B to A , $\beta = -1^\circ 06' 10''$

$$\text{Axis signal correction to } \beta = \frac{2.2 - 1.46}{1700} \times 206\,265 = 89.78''$$

$$\text{Vertical angle corrected for axis signal, } \beta_1 = 1^\circ 06' 10'' + 89.78'' \\ = 1^\circ 07' 39.78''$$

$$\text{Central angle, } \theta = \frac{D}{R \sin 1''} = \frac{1700}{30.88} = 55.05''$$

$$H = D \frac{\sin\left(\frac{\alpha_1 + \beta_1}{2}\right)}{\cos\left[\left(\frac{\alpha_1 + \beta_1}{2}\right) + \frac{\theta}{2}\right]} \\ = 1700 \times \frac{\sin\left(\frac{(1^\circ 06' 50.98'' + 1^\circ 07' 39.78'')}{2}\right)}{\cos\left[\left(\frac{1^\circ 06' 50.98'' + 1^\circ 07' 39.78''}{2}\right) + \frac{55.05''}{2}\right]} \\ = 33.263 \text{ m}$$

$$\text{Refraction correction, } \gamma = \frac{1}{2}(\theta + \alpha_1 - \beta_1) \\ = \frac{1}{2}(55.05'' + 1^\circ 06' 50.98'' - 1^\circ 07' 39.78'') = 3.125''$$

$$\text{Coefficient of refraction } m = \frac{\gamma}{\theta/2} = \frac{3.125}{55.05/2} = 0.113$$

Example 1.5 Two points A and B are 9 km apart with respective reduced levels as 104.5 m and 290.5 m. The signal height at A is 1.50 m and that at B is 3.10 m. The instrument height at both the stations was 1.50 m. If 30.88 m on the earth's surface subtends $1''$ of arc at the earth's centre and the effect of refraction is $1/7$ th of that of the curvature, determine the observed angles from A to B and B to A .

Solution

$$H = 290.5 - 104.5 = 186 \text{ m}$$

We know that, $H = D \tan \phi$

$$\text{or} \quad \tan \phi = \phi = \frac{H}{D} \times 206\,265$$

$$\text{where} \quad \phi = \alpha + (C - \gamma) = \beta - (C - \gamma)$$

$$\text{Now,} \quad \phi = \frac{186}{9000} \times 206\,265 = 1^\circ 11' 2.81''$$

$$\text{Central angle, } \theta = \frac{9000}{30.88} = 291.45''$$

$$\text{Curvature correction, } C = \frac{\theta}{2} = \frac{291.45}{2} = 145.725''$$

$$\text{Refraction correction, } \gamma = \frac{1}{7} C = \frac{145.725}{7} = 20.817''$$

$$\alpha = \phi - (C - \gamma) = 1^\circ 11' 2.81'' - (145.725'' - 20.817'') \\ = 1^\circ 9' 0.72''$$

$$\text{and} \quad \beta = \phi + (C - \gamma) = 1^\circ 11' 2.81'' + (145.725'' - 20.817'') \\ = 1^\circ 13' 4.9''$$

Since the instrument height at A is 1.5 m and signal height at B is 3.1 m

$$\text{Axis signal correction} = \frac{3.1 - 1.5}{9000} \times 206\,265 = 36.67''$$

To obtain the true angle, the correction is subtracted from the observed angle of elevation. But in this example, the angle (α) computed is the true reciprocal angle. Thus, the correction has to be added to obtain the observed angle.

$$\therefore \text{observed angle of elevation} = 1^\circ 9' 0.72'' + 36.67'' \\ = 1^\circ 9' 37.39''$$

Since, the instrument height at B and the signal height at A both are equal (1.5 m), no axis signal correction will be required. Therefore, the observed angle from B to A is $1^\circ 13' 4.9''$.

Example 1.6 It is required to determine the elevation of a station O . Observations were made to three stations A , B and C already fixed and of known elevations. The following data was recorded:

Instrument station	Station observed	Height of instrument (m)	Distance (m)	Height of signal (m)	Vertical angle
	A		3600	5.6	$1^\circ 1' 20''$
O	B	1.50	4700	4.1	$-53' 00''$
	C		5000	4.9	$-34' 10''$

The reduced level of A , B and C were 294, 159.5 and 181 m, respectively. Take $m = 0.07$ and $R \sin 1'' = 30.88$ m.

Solution *Axis signal correction*Distance *OA*

$$\delta_1 = \frac{5.6 - 1.50}{3600} \times 206\,265 = 234.91'' (-)$$

Distance *OB*

$$\delta_2 = \frac{4.1 - 1.50}{4700} \times 206\,265 = 114.10'' (+)$$

Distance *OC*

$$\delta_3 = \frac{4.9 - 1.50}{5000} \times 206\,265 = 140.26'' (+)$$

Combined correction for curvature and refraction

$$C_r = \frac{D}{2R \sin 1''} (1 - 2m)$$

Distance *OA*

$$C_{r_1} = \frac{3600}{2 \times 30.88} \times (1 - 2 \times 0.07) = 50.13''$$

Distance *OB*

$$C_{r_2} = \frac{4700}{2 \times 30.88} \times (1 - 2 \times 0.07) = 65.446''$$

Distance *OD*

$$C_{r_3} = \frac{5000}{2 \times 30.88} \times (1 - 2 \times 0.07) = 69.624''$$

*Corrected vertical angles*From *O* to *A*

$$\alpha_1 = 1^\circ 1' 20'' - 234.91'' + 50.13'' = 58' 15.12''$$

From *O* to *B*

$$\alpha_2 = 53' 00'' + 114.10'' - 65.446'' = 53' 48.654'' (-)$$

From *O* to *C*

$$\alpha_3 = 34' 10'' + 140.26'' - 69.624'' = 35' 20.936'' (-)$$

Difference of levels of stations

$$H = D \tan \alpha$$

$$H_1 = 3600 \tan 58' 15.12'' = +61.007 \text{ m}$$

$$H_2 = 4700 \tan 53' 48.654'' = -73.575 \text{ m}$$

$$H_3 = 5000 \tan 35' 20.936'' = -51.414 \text{ m}$$

R.L. of *O*

$$\text{In first case} = 294 - 61.007 = 232.993 \text{ m}$$

$$\text{In second case} = 159.5 + 73.575 = 233.075 \text{ m}$$

$$\text{In third case} = 181.0 + 51.414 = 232.414 \text{ m}$$

$$\therefore \text{Mean elevation of } O = 232.827 \text{ m.}$$

Summary

For establishing vertical control in geodetic surveying either precise levelling or trigonometric levelling is used; the former type of levelling has been discussed in Surveying Vol. 1. In situations such as for points on hills that rise abruptly above the surrounding plain (i.e., with steep approach), the ordinary spirit levelling is impractical; the equality of back sight distance and fore sight cannot be maintained. Trigonometric levelling under such conditions becomes the most important part of geodetic surveying. Moreover, in geodetic surveying, for finding elevations level instrument has to be used after the necessary angular measurements are made with theodolite which is more time consuming. Trigonometrical levelling, on the other hand, has the advantage of observing vertical angles at the same time as the horizontal angles are measured with the same instrument saving time. However, trigonometric levelling with all the necessary precautions will not result in more accurate observations than that made even with ordinary spirit levelling.

Procedure to find the corrections due to difference between heights of signals the axis signal correction and that due to curvature and refraction are elaborated. The chief source of error arises from uncertainty regarding the amount of refraction. It varies with the rate at which temperature changes with height, as well as it varies throughout the day. Further the line of sight is assumed to be a circular arc and the coefficient of refraction is assumed to be independent of elevation or inclination, which is not true. Moreover, the coefficient of refraction is assumed to be minimum average value of the area, and if possible should be determined by simultaneous observations of reciprocal angles.

To determine difference of elevation between two widely spaced points two methods are in use. One of these methods consists in observing vertical angles from each station to the other. This method eliminates the effect of uncertainty regarding the value of coefficient of refraction. In the other method, the difference of elevation is determined from the vertical angle measured at one of the station only and the knowledge of the coefficient of refraction is required. Both the methods have been described in details.

Exercises

- 1.1 Define the coefficient of refraction. Explain how its value can be obtained by simultaneous reciprocal observations.
- 1.2 Describe the difference between the techniques of reciprocal levelling and reciprocal trigonometrical levelling, and discuss the conditions in which each is most effectively used.
- 1.3 Derive an expression for the difference of level between two points A and B a distance D apart, with the vertical angle as the angle of elevation from A to B . The height of the instrument at A and that of the signal at B are equal.
- 1.4 It is required to find the difference of elevation between two stations A and B 9290 m apart. The angle of elevation from A to B was observed with a theodolite as $2^{\circ}06'18''$ and the height of the instrument was 1.25 m. The height of signal at station B was 3.96 m. Take the value of $R \sin 1''$ as 30.88 and coefficient of refraction as 0.07. Also, find the R.L. of station B if the R.L. of station A is 300 m.
[Ans. 344.59 m, 644.59 m]
- 1.5 The distance between two stations A and B is 6370 m. The station B was 200 m above the station A . Calculate the angles observed from A and B . Assume that the instrument and signal heights to be equal and the effect of refraction as $1/7$ th of that of the curvature. The radius of earth is 6370 km.
[Ans. $1^{\circ}46'27.73''$, $1^{\circ}49'24.53''$]

- 1.6** Two stations *A* and *B* are 16.44 km apart. The following data were recorded:
 Instrument at *A*, angle of depression to *B* = $3'42''$
 Instrument at *B*, angle of depression to *A* = $2'04''$
 Height of instrument at *A* = 1.42 m
 Height of instrument at *B* = 1.42 m
 Height of signal at *A* = 5.53 m
 Height of signal at *B* = 5.53 m
 $R \sin 1''$ = 30.88 m
 Find the difference in level between *A* and *B* and the coefficient of refraction at the time of observation.

[Ans. 3.91 m, 0.0784]

- 1.7** The distance between two stations *A* and *B* was 3489.96 m. Determine the reduced level of the station *B*, if the R.L. of *A* was 950.75 m. The following observations were recorded:

	Station <i>A</i>	Station <i>B</i>
Height of instrument	1.433 m	1.463 m
Height of signal	4.572 m	3.962 m
Vertical angle	$+1^{\circ}52'4''$ (to <i>B</i>)	$-1^{\circ}48'20''$ (to <i>A</i>)
Take the value of $R \sin 1''$ as 30.88.		

[Ans. 1062.932 m]

- 1.8** The following data refer to the elevations of the ground stations of a triangle *ABC* in a trigonometrical survey. Find the closing error and the R.L.s of *B* and *C* if the R.L. of *A* is 1600 m.

(a) Vertical angle from <i>A</i> to <i>B</i> = $+1^{\circ}20'20''$	Weight 2	Height of instrument at <i>A</i> = 1.50 m Height of signal at <i>A</i> = 4.92 m Height of instrument at <i>B</i> = 1.47 m Height of signal at <i>B</i> = 4.44 m Distance <i>AB</i> = 4777.8 m
Vertical angle from <i>B</i> to <i>A</i> = $-1^{\circ}12'24''$		
(b) Vertical angle from <i>B</i> to <i>C</i> = $-49'24''$	Weight 3	Height of instrument at <i>B</i> = 1.44 m Height of signal at <i>B</i> = 4.71 m Height of instrument at <i>C</i> = 1.41 m Height of signal at <i>C</i> = 5.52 m Distance <i>BC</i> = 4068.2 m
Vertical angle from <i>C</i> to <i>B</i> = $+55'12''$		
(c) Vertical angle from <i>C</i> to <i>A</i> = $-47'12''$	Weight 1	Height of instrument at <i>A</i> = 1.44 m Height of signal at <i>C</i> = 3.96 m Distance <i>AC</i> = 3187.5 m

[Ans. 2410 m, 1706.803, 1644.79 m]

- 1.9** Two stations *A* and *B* were at a distance 1800.50 m apart. Reciprocal observations were made to determine the difference of level between them. The following data was recorded:

Height of instrument at	<i>A</i> = 1.463 m
Height of instrument at	<i>B</i> = 1.457 m
Height of signal at	<i>A</i> = 1.647 m
Height of signal at	<i>B</i> = 1.762 m
Vertical angle from	<i>A</i> to <i>B</i> = $+1^{\circ}42'2''$
Vertical angle from	<i>B</i> to <i>A</i> = $-1^{\circ}41'46''$

Calculate also the height of B above A , using the vertical angle from A only. Assume the coefficient of refraction as 0.07. Prove the formula used.

[Ans. 53.32 m, 53.356 m]

Objective-type Questions

- 1.1 The process of determining the elevations of stations from vertical angles and geodetic lengths at mean sea level is known as
 - (a) levelling
 - (b) trigonometric levelling
 - (c) triangulation
 - (d) hypsometry
- 1.2 Trigonometric levelling by reciprocal observations
 - (a) eliminates error due to uncertain refraction.
 - (b) is essentially done in the early morning.
 - (c) proves to be more accurate than spirit levelling.
 - (d) All are correct
- 1.3 In trigonometric levelling, the correction due to refraction and curvature are determined
 - (a) in linear measure
 - (b) in angular measure
 - (c) Either (a) or (b)
 - (d) graphically
- 1.4 Which one of the following is a correction to be applied in trigonometric levelling?
 - (a) Correction for dip
 - (b) Correction for semi-diameter of sun
 - (c) Axis signal correction
 - (d) Parallax correction
- 1.5 In trigonometric levelling the correction is subtracted to the observed vertical angle if it is for
 - (a) axis signal
 - (b) combined refraction and curvature
 - (c) Both (a) and (b)
 - (d) None of the above
- 1.6 In trigonometric leveling, axis-signal correction for an angle of elevation is given by

$$(a) + \frac{S_1 - h_2}{D} \cos^2 \alpha$$

$$(b) - \frac{S_1 - h_2}{D} \cos^2 \alpha$$

$$(c) - \frac{S_2 - h_1}{D} \cos^2 \alpha$$

$$(d) + \frac{S_2 - h_1}{D} \cos^2 \alpha$$

- 1.7 For reciprocal observations in trigonometric levelling, the height difference between two points will be

$$(a) D \frac{\sin(\beta_1 + \alpha_1)}{\cos\left(\frac{\beta_1 + \alpha_1}{2} - \frac{\theta}{2}\right)}$$

$$(b) D \frac{\sin(\beta_1 + \alpha_1)}{\cos\left(\frac{\beta_1 + \alpha_1}{2} + \frac{\theta}{2}\right)}$$

$$(a) D \frac{\sin\left(\frac{\beta_1 + \alpha_1}{2}\right)}{\cos\left(\frac{\beta_1 + \alpha_1}{2} + \frac{\theta}{2}\right)}$$

$$(b) D \frac{\sin\left(\frac{\beta_1 + \alpha_1}{2}\right)}{\cos\left(\frac{\beta_1 + \alpha_1}{2} - \frac{\theta}{2}\right)}$$

Answers to Objective-type Questions

1.1 (b)

1.2 (a)

1.3 (b)

1.4 (c)

1.5 (b)

1.6 (b)

1.7 (c)

2

TRIANGULATION

Introduction

Every survey, from mapping a continent to a small plot of land, depends upon a carefully measured framework which is thereafter treated as free from error. Subsequently, the details are filled in the framework by less elaborate methods. The fixation of a framework for a survey is known as *horizontal control*. It can, in general, be established either by traversing or triangulation or by a combination of both. Some of the other methods to provide horizontal control are trilateration and electromagnetic distance measurement (E.D.M.) traversing. The E.D.M. traversing is rapidly replacing triangulation, since the former is as accurate and far more economical. However, trilateration has not gained popularity, probably due to the lack of easy checks and more tedious computations.

Horizontal control usually consists of a combination of triangulation and traverse. For most of the surveys of small extent, e.g., plane surveys, where direct linear measurements are impossible, triangulation is most suited. It is also suitable when long sights are taken. In triangulation, a number of lines of sight are required at each station. When the sights are long, the stations are elevated by building towers. In case the distances are short, the expense of towers offsets any saving and traverse becomes economical. Therefore, the character of the physical conditions encountered rather than the accuracy desired, governs the choice of the method of horizontal control. Triangulation is most suited for hilly areas, whereas traversing is suitable for flat areas.

Triangulation is preferred for hills and undulating areas, since it is easy to establish stations at reasonable distances apart, with intervisibility. In plane and crowded areas, it is not suitable as the intervisibility of stations is affected. The difficulty is overcome by building towers which is quite expensive. The main disadvantage of triangulation is the accumulation of error in the lengths and directions of lines, since both of them, for successive lines, depend upon the computations for those of the preceding line, which necessitates a number of check bases.

The direct object of triangulation is not to provide a complete plan showing details and topographical features, but simply to locate a number of isolated points (stations) over the surface of the country. When these points have been located,

the topographical and other details are surveyed by means of chain, traverse, plane table, stadia, total station or other surveys.

In triangulation, the entire area to be surveyed is covered with a framework of triangles. If the length and direction of one side, and all three angles of a triangle are measured precisely, the lengths and directions of the remaining two sides of the triangle can be computed. The length of the first line, which is measured precisely is known as *base line*. The other two computed sides are used as new base lines for two other triangles interconnected with the first triangle. By extending this process, i.e., the measurement of the further interconnected triangles and using the computed sides, a chain or network of triangles can be spread over the entire area. The apex of the triangles so located with a relatively greater accuracy provide horizontal control of the survey. Thus, *triangulation may be defined as a system of multiplying ground controls on the earth's surface*. As a check, the length of one of the sides of the last triangle is also measured and compared with the computed one; this side is known as check base. This method of surveying was first introduced by a Dutchman named Snelli. This approach of classical triangulation in which the base lines and check base lines are measured by invar tapes in catenaries and connected into the triangulation by angular extension procedures is now obsolete. The more modern approach would be to measure the base line with EDM equipment and to include many more check lines in the network, to afford greater control of *scale error* (the difference between the measured and computed check base).

The field work of a triangulation is carried out in the following well-defined operations:

1. Reconnaissance
2. Station preparation
3. Base line measurement
4. Measurement of angles

Besides field work, triangulation consists of the specifications, the design of stations and signals, and the reduction and adjustment of the observations. The description which follows includes specifications, signals and station marks, field work, some geodetic practices, and office work.

Applications

Triangulation may be used for

1. establishing accurately located control points for plane and geodetic surveys of large areas,
2. establishing accurately located control points in connection with aerial surveying and
3. accurate location of engineering projects such as centre lines, terminal points and shafts for long tunnels; and centre lines and abutments for long span bridges.

2.1 TRIANGULATION FIGURES OR SYSTEMS

Triangulation figures may be defined as a system consisting of triangulation stations connected by a chain of triangles. The complete figure is called *triangulation figure* or *triangulation system*. The most common type of figures used in a triangulation

system are triangles, quadrilaterals and polygons. All of these figures should fulfill the rigid geometric conditions given as follows:

1. The sum of the interior angles should be $(2n - 4) \times 90^\circ$, where n is the number of sides of the figure. The average number of seconds by which the sum of angles deviates from 180° , plus the required spherical excess is known as *triangular misclosure*.
2. If all the angles are measured at a station, their sum should be 360° .
3. The length of sides calculated through more than one routes should agree.

It is impossible to fulfill all the geometrical conditions, owing to the errors, until the field measurements have been adjusted.

2.1.1 Triangle

A chain of triangles is very rapid and economical when a narrow strip of terrain is to be surveyed, e.g., highway, river, valley, etc. Angles less than 30° or more than 120° are not permitted; for well-conditioned triangles, angles should not be less than 30° or more than 120° .

Advantages Simple, rapid and economical.

Disadvantages Since it is used to survey a long narrow strip, a number of base lines must be introduced frequently to reduce the accumulation of errors. Therefore, a single chain of triangles is never permitted in high order triangulation. The number of rigid geometrical conditions to be fulfilled in the *figures adjustment* is small; so this system is least accurate.

Note It is always advisable to use a double row of single triangles.

2.1.2 Quadrilaterals

Quadrilaterals afford an excellent system since the various combinations of sides and angles can be used to compute the lengths of the required sides, and checks can be made frequently. The best quadrilateral is a square. A quadrilateral with both diagonals having no station at their intersection is usually employed. These are best suited for hilly areas.

Advantage Most accurate system as the number of checks are more.

2.1.3 Polygons

When areas that are very wide in proportion to their lengths are to be surveyed, then pentagonal or hexagonal figures may be economical. These may or may not have a central station. However, the progress of work is slow because of more settings of the instrument.

Advantages Such a system provides the desired checks on the computations and usually better triangles can be selected from such a system than those provided by quadrilaterals.

In Fig. 2.1, the double dark lines are bases and the dotted lines are the 'check base lines'. If any error is made in the measurement of a base line, it will result in a scale error throughout the network. The *scale error* as already defined is the

difference between the measured and computed check base. Therefore, to control this error, check base lines should be measured frequently.

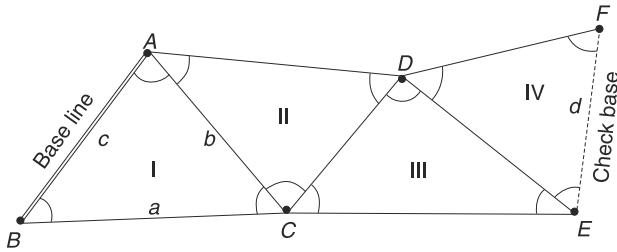


Fig. 2.1(a) Chain of simple triangles

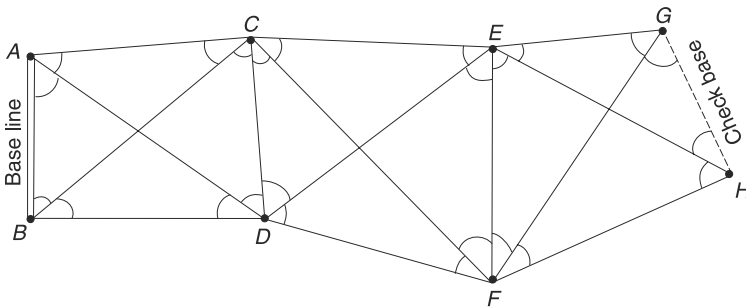


Fig. 2.1(b) Braced quadrilaterals

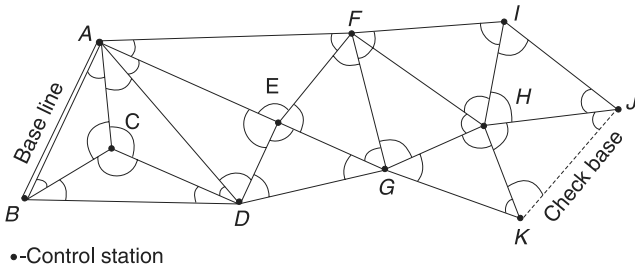


Fig. 2.1(c) Polygons with central points

2.1.4 Choice of Figure

Of the three figures discussed above, a chain of single triangles is the most simple. It requires the measurement of fewer angles as compared to quadrilaterals and polygons. The only check available for a chain of triangles is the sum of the angles of each triangle considered by itself. This necessitates providing many check bases to acquire high precision which, in turn, offsets the economy achieved by adopting this system. Therefore, when relatively less precise work is required, this system is most useful. Quadrilaterals are best for a narrow chain whereas polygons for a wide chain system.

Criteria for the selection of figures The figure should be such that the computations can be done through two independent routes with at least one route having well-conditioned triangles. The selected figure should yield maximum progress with minimum field work and effort. All the lines of the figure should be of comparable length and very long lines should be avoided. A complex figure should not involve more than 12 conditions.

2.1.5 Strength of Figure

In a system of triangulation, the angles and the base line are subjected to the errors of measurement and since the sines change more rapidly for small angles than for large ones, it is evident that the percent error in a computed side of a triangle will be larger if the side is opposite a small angle than if it is opposite a large one. Therefore, in choosing the stations for a system of triangulation, it is a matter of great importance that the angles shall be of such size that the requisite degree of accuracy is maintained in the computed results. Also, in a system of triangles which is more complex than a chain of simple triangles, there will be a variety of paths through the system whereby the length of any given side may be computed, and it is important to choose that route which will yield the highest accuracy.

Strength of figure, thus, may be defined as *a figure which gives the least error in the calculated length of the last line in the system due to the shape of triangle and the composition of the figures.*

Distance Angles In computing a given side in any triangle, there are always two *distance angles*, namely, the angle opposite the known side and the angle opposite the unknown side. In triangle ABC (Fig. 2.2), if c is the known side and b is the side to be computed, then C and B will be the distance angles. This implies that in a given triangle there is always one angle which has no effect on computing a given side. Hence, this non-effective angle can be very small and still the accuracy of the computed length will in nowise be affected. Therefore, the stations should be so selected that the non-effective angle is kept quite low and the distance angles are kept large for accurate computations of the sides.

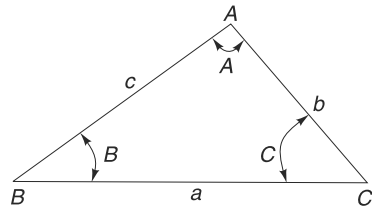


Fig. 2.2

Criterion of Strength It is often necessary that some criterion of strength be applied to the triangulation system so that it shall yield satisfactory results in the most economical manner. There are two reasons for this. Firstly, in establishing a triangulation system in the field, there may be alternative stations which may be used and it is desirable to have a quantitative measure of the relative strength of the two or more possible systems of figures. Secondly, in order to insure a desired degree of precision throughout the systems, definite limitations are imposed upon the shapes of the triangles of which it is composed.

If it is assumed that all the angles are measured with the same degree of precision, then the strength of a figure as compared with another depends upon various factors such as the number of observed angles or directions, the number

of geometric conditions, and the proportions of each triangle, i.e., the sizes of the distance angles used in the computations. Strength of a triangulation figure can be evaluated by the method developed by the US Coast and Geodetic survey. The method is based on an expression for the square of probable error, L^2 , that would occur in the sixth place of the logarithm of any side.

$$\begin{aligned} L^2 &= \frac{4}{3} d^2 \frac{D-C}{D} \sum (\delta A^2 + \delta A \delta B + \delta^2 B) \\ &= \frac{4}{3} d^2 R \end{aligned}$$

and

$$\begin{aligned} R &= \frac{D-C}{D} \sum (\delta A^2 + \delta A \delta B + \delta B^2) \\ &= \frac{D-C}{D} \sum d \end{aligned}$$

where R = strength of figure,

D = number of directions observed, excluding those for the known side of a given figure,

δA = the tabular difference for 1" in the logarithmic sine of distance angle A in the sixth decimal place,

δB = the tabular difference for 1" in the logarithmic sine of distance angle B in the sixth decimal place.

$d = \delta A^2 + \delta A \delta B + \delta B^2$. The values of d , the probable error of an observed direction (in seconds), are listed in Table 2.1.

C = the number of geometric conditions

= angle conditions + side conditions

= $(n' - s' + 1) + (n - 2s + 3)$

where

n = number of lines,

s = total number of stations,

n' = number of lines observed in both directions (lines observed in one direction only are usually shown by broken lines), and

s' = number of stations occupied.

Distance angle A = angle opposite to known side.

Distance angle B = angle opposite to the side which is to be computed.

Note The relative strength of figure is computed in terms of R . The stronger figures will have lower values of R . Hence, the different courses of routes are compared and the one which has the lowest value of R , is chosen.

Example 2.1 In a triangle ABC , angles A , B and C were observed as 70° , 50° and 60° . The line AC was used as a base of known length. Calculate the strength of the figure.

Table 2.1 Table for determining relative strength of triangulation figures (Values of $\delta A^2 + \delta\delta A\delta B + \delta\delta B^2$)

	10°	12°	14°	16°	18°	20°	22°	24°	26°	28°	30°	35°	40°	45°	50°	55°	60°	65°	70°	75°	80°	85°	90°
10°	428	359																					
12°	359	295	253																				
14°	315	253	214	187																			
16°	284	225	187	162	143																		
18°	262	204	168	143	126	113																	
20°	245	189	153	130	113	100	91																
22°	232	177	142	119	103	91	81	74															
24°	221	167	134	111	95	83	74	67	61														
26°	213	160	126	104	89	77	68	61	56	51													
28°	206	153	120	99	83	72	63	57	51	47	43												
30°	199	148	115	94	79	68	59	53	48	43	40	33											
35°	188	137	106	85	71	60	52	46	41	37	33	27	23										
40°	179	129	99	79	65	54	47	41	36	32	29	23	19	16									
45°	172	124	93	74	60	50	43	37	32	28	25	20	16	13	11								
50°	167	119	89	70	57	47	39	34	29	26	23	18	14	11	9	8							
55°	162	115	86	67	54	44	37	32	27	24	21	16	12	10	8	7	5						
60°	159	112	83	64	51	42	35	30	25	22	19	14	11	9	7	5	4	4					
65°	155	109	80	62	49	40	33	28	24	21	18	13	10	7	6	5	4	3	2				
70°	152	106	78	60	48	38	32	27	23	19	17	12	9	7	5	4	3	2	2	1			
75°	150	104	76	58	46	37	30	25	21	18	16	11	8	6	4	3	2	2	1	1	1		
80°	147	102	74	57	45	36	29	24	20	17	15	10	7	5	4	3	2	1	1	1	0	0	
85°	145	100	73	55	43	34	28	23	19	16	14	10	7	5	3	2	2	1	1	0	0	0	0

(Contd.)

Solution

$$C = (n' - s' + 1) + (n - 2s + 3)$$

$$n = 3, s = 3, n' = 3, s' = 3$$

$$D = 6 - 2 = 4$$

$$C = (3 - 3 + 1) + (3 - 6 + 3) = 1$$

For angle $A = 70^\circ$ and $B = 50^\circ$, $d = +5$

$$\text{Strength of figure, } R = \frac{D-C}{D} \sum d = \left(\frac{4-1}{4} \right) \times 5 = 3.75$$

Example 2.2 In Fig. 2.3, the side AC was known and BD was computed; the angles measured have been shown in the figure. Calculate the strength.

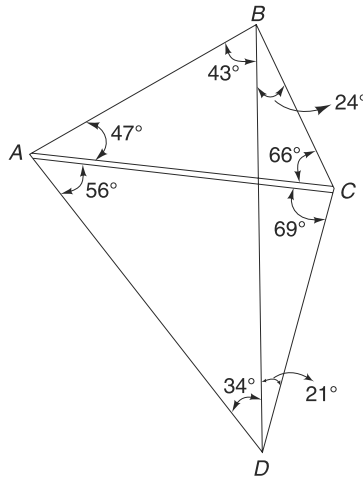


Fig. 2.3

Solution There can be several routes to compute the strength. The one which gives the minimum strength will be opted.

$$n = 6, n' = 6, s = 4, s' = 4$$

$$D = 6 \times 2 - 2 = 10$$

$$C = (6 - 4 + 1) + (6 - 2 \times 4 + 3) = 4$$

$$R = \frac{D-C}{D} = \frac{10-4}{10} = 0.6$$

Route 1:

Using triangle ACD containing measured base AC and triangle ADB containing BD to be computed.

For triangle ACD , the distance angles are 55° and 69° ; and $d = 4$.

For triangle ADB , the distance angles are 43° and 103° ; and $d = 5$.

$$\therefore \sum d = 4 + 5 = 9$$

$$\text{and } R_1 = 0.6 \times 9 = 5.4$$

Route 2:

Using triangles ACD and DCB .

For triangle ACD , the distance angles are 56° and 55° ; and $d = 7$.

For triangle DCB , the distance angles are 24° and 135° ; and $d = 17$.

$$\therefore \quad \Sigma d = 7 + 17 = 24$$

$$\text{and} \quad R_2 = 0.6 \times 24 = 14.4$$

Route 3:

Using triangles ACB and ABD .

For triangle ACB , the distance angles are 66° and 67° ; and $d = 3$.

For triangle ABD , the distance angles are 34° and 103° ; and $d = 8$.

$$\therefore \quad \Sigma d = 3 + 8 = 11$$

$$\text{and} \quad R_3 = 0.6 \times 11 = 6.6$$

Route 4:

Using triangles ACB and BCD .

For triangle ACB , the distance angles are 67° and 47° ; and $d = 7$.

For triangle BCD , the distance angles are 21° and 135° ; and $d = 26$.

$$\therefore \quad \Sigma d = 7 + 26 = 33$$

$$\text{and} \quad R_4 = 0.6 \times 33 = 19.8$$

Hence, the best route for calculating BD is the first one.

2.2 SYSTEM OF FRAMEWORK

The purpose of a system of framework is to establish a system of horizontal control points. By the use of these control points, different parts of the survey can be tied into the triangulation framework. For example, a traverse may begin and end at different control points.

2.2.1 Grid Iron System

The primary network consists of two chains of triangles, one roughly in the direction of the meridian and the other roughly perpendicular to the meridian (Fig. 2.4 (a)). The areas enclosed are then surveyed by forming a network of small triangles of second and third order triangulations. This is used on very extensive surveys. This triangulation system is adopted in India.

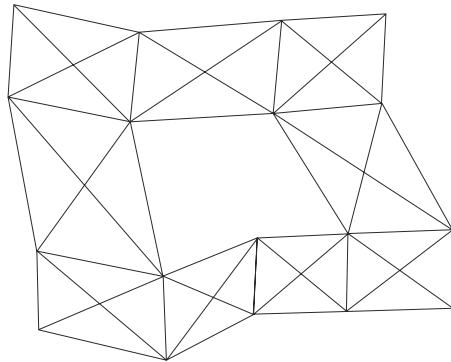


Fig. 2.4(a) Grid iron system

2.2.2 Central System

The area to be surveyed is covered by a chain of triangles extending outwards in all directions from the initial base line (Fig. 2.4(b)). This system is adopted in the U.K.

A suitable system of triangles should be selected to provide an appropriate strength of figure. It must be possible to carry out the computation by two separate routes, by well-conditioned triangles. The objective should be to cover the given

area on the ground with a minimum number of stations, and at the same time to achieve the required accuracy. The number of stations can be reduced by using a chain of quadrilaterals (parallelograms) by increasing the length relative to the width.

2.3 CLASSIFICATION

The classification of a triangulation system is based upon the degree of accuracy required, the extent of the area to be surveyed, length of the base, length of the sides, and triangular misclosure.

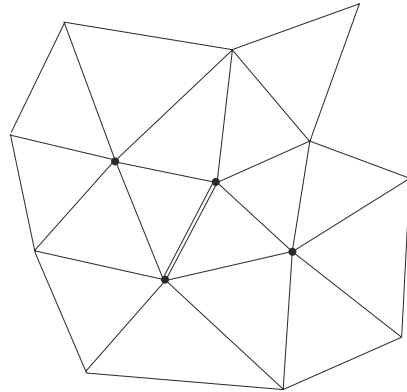


Fig. 2.4(b) Central system

2.3.1 Primary or First-Order Triangulation

A first order triangulation is the highest-order triangulation and is employed for very large areas, for example, for the earth's figure, for obtaining the most precise control in mapping, and for small-scale mapping.

It consists of forming large, well-conditioned triangles. Precise instruments are used for observations and every possible refinement is exercised.

2.3.2 Secondary or Second-Order Triangulation

A second-order triangulation is employed for running a second series of triangles by fixing points at close intervals inside the primary series of triangles. It consists of forming small, well-conditioned triangles with less precise instruments.

2.3.3 Tertiary or Third-Order Triangulation

A third-order triangulation is employed for running in a third series of triangles, by fixing points inside the secondary triangles at short intervals to furnish horizontal control for details on a topographic survey. The triangles are of the smallest size in comparison with the other two orders of triangulation.

Figure 2.5 illustrates a scheme of triangulation network wherein one of the first-order polygons is shown. The rows of network are built along meridians and parallels intersecting to form first-order polygons. The areas inside these polygons are filled with a continuous net of second-order triangles which are further filled with the third-order nets.

Note A topographic triangulation is usually of limited extent, and is not necessarily based upon a primary system. It would generally be quite independent and be of the third order of precision. In such a case it is necessary to measure at least two bases followed by a tie-in of this net to a higher-order triangulation. However, for geodetic triangulation intended for the measurement of an arc of a meridian, or for the determination of other data concerning the shape of the earth as a whole demands a very precise system of triangles.

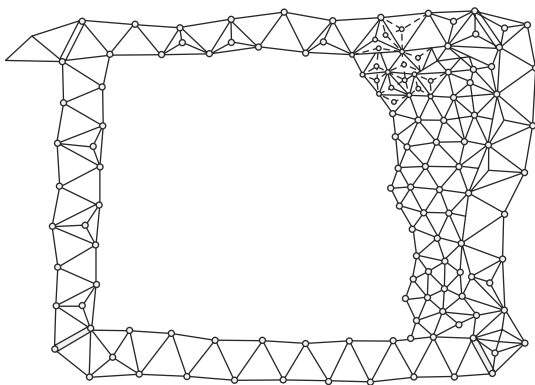


Fig. 2.5 A typical scheme of triangulation network

2.3.4 Precision of Triangulation Networks

The recommendations regarding the precision to be obtained during triangulation are given in Table 2.2.

Table 2.2 Precision in triangulation

S. No.	Requirement	Triangulation of		
		Ist order	IIInd order	IIIrd order
1	Length of sides of triangles	30–160 km	8–70 km	1.5–10 km
2	Length of base line	5–20 km	2–5 km	0.5–3 km
3	Average triangular closure	1"	3"	6"
4	Maximum triangular closure	3"	8"	12"
5	Instrument	Wild or Zeiss theodolite	Wild theodolite	Any theodolite with least count of 5"
6	Limiting strength of figure	25	40	50
7	Sets of observations with theodolite	16	8	3

Note On Indian triangulation, the average length of a side in the hilly country was 48 km, while in the plains the average length was about 17.5 km only. In the secondary triangulation chains in the Himalayas, sights as long as 320 km were obtained.

2.4 STATION MARKS, SIGNALS AND TOWERS

The object of station marks is to provide a surface mark with a permanent mark buried below, and a good target, on which to observe, accurately centred over it. It should also include a complete description to enable the station to be recovered even after many years. Often when the stations must be elevated above the ground, a rigid support is required for the instrument and the signal. Sometimes it is

possible to utilise the existing tall structures, if they exist, with or without the addition of towers (scaffolds), but usually a tower has to be erected.

2.4.1 Station Marks

The station marks are bronze or copper markers cemented into rock or concrete in the form of a pedestal. The purpose is twofold. It ensures that the signal over the station mark is central when emplaced and to hold the instrument in a centred position while making the observations. An underground mark immediately below the reference station mark is established since the surface mark may get disturbed as it is on the earth. The mark may be a frustum of a cone or pyramid, or have the form of a post with an enlarged base.

Marking of the Station Stations are permanently marked so that they can be identified in the future. When the area is rocky, a hole 7.5–15 cm deep is drilled in the rock and a copper or iron bolt is fixed with cement concrete.

When the area is earth which can be excavated, two marks are used—one buried and the other on the surface. A stone with a small hole fitted with a brass screw is buried to a depth of 75 cm and is embedded in cement mortar. It is covered with earth and a similar marked stone is placed with its top surface level with the ground. A bush is erected at the station and a mound of stone is erected around the bush (Fig. 2.6). This arrangement is also called a *Cairn*. The buried mark is not referred to until it appears that the surface mark has been disturbed. A full description of each station is prepared with its number, locality, position of witness marks (with respect to some permanent objects), etc.

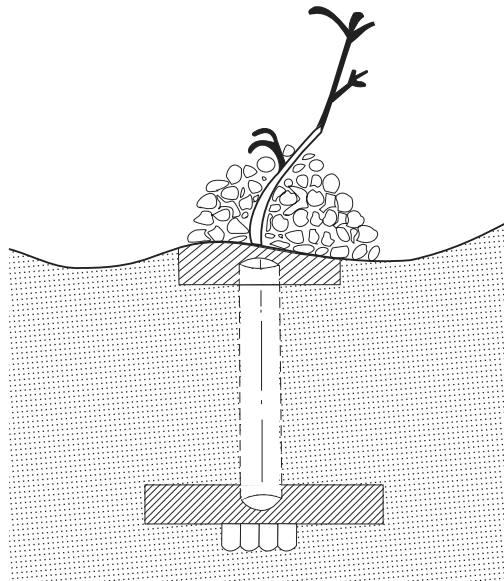


Fig. 2.6 Cairn

Description of the Station The reconnaissance party should prepare a well-defined description of the stations selected by them because of two reasons. Firstly,

the survey party responsible for conducting the observations will work from the report of the reconnaissance party. Therefore, the stations should have a clear and well-defined identification. Secondly, the life of a well-executed survey is of decades. Therefore, such references which are going to last for 50–100 years must be included in the report.

The description must include the approximate directions of the prominent topographical features. An aerial photographic identification, if possible, should be included. The method of approach and the path to be followed to the station should also be included in the report. This helps in assessing the difficulties in carrying up instruments and stores. The location of water and aggregates should also be noted. The camping sites and local food supplies, if properly recorded, prove to be of great value.

2.4.2 Signals

They are devices erected to define the exact position of an observed station. A signal is placed at each station so that lines of sight are established between triangulation stations. These may be opaque (e.g., a pole, or a three or four-legged type) or luminous.

Characteristics of a Good Signal A signal should fulfill the following requirements:

1. It should be of sufficient height, capable of being vertical, and accurately centred over the station mark. A general rule for the height of signal is $h = 13.3D$, h being the height of the signal in centimetres and D being the distance in kilometres.
2. It should be conspicuous, i.e., should be clearly visible.
3. It should be of suitable size for accurate bisection. A general rule for diameter of signal is $d(\text{cm}) = 1.3D$ to $1.9D$, D being the distance in kilometres.
4. It should be easy to erect in minimum time.
5. It should be symmetrical.
6. The shape should be such that the phase error is minimum.

Opaque Signals These are used for triangulation with less accuracy and for sights not exceeding 30 km. A pole is used when the angle of measurement at the station is not required and the sights are less than 6 km. The signal should subtend about $1\text{--}1.5''$ at the observer, corresponding to a width of 4.8–7.3 mm per km. But for short sights, considerations of rigidity govern the diameter of the pole and necessitate a greater angular width. On the other hand, a mast to subtend those angles at long distances would be inconveniently heavy for erection, and beyond 30 km, the width must be increased by nailing on rectangular or diamond shaped targets. The targets are fixed in pairs at right angles to each other so that these are visible from different directions.

On the stations, where angles are to be measured, a three-legged signal with a flag (Fig. 2.7(a)) or with a pole (Fig. 2.7(b)) should be used. It should be sufficiently high so as to keep provision for setting up the instrument. The three-legged signal should have its legs braced and firmly set into the ground. The pole

should be plumb and over the station, and is painted red and white for a clear vision. The bracing of the legs should be high enough to allow the instrument to be placed underneath and to permit free movement of the surveyor. Usually, the height of the signal is kept as 3.0 m. A three-legged signal is preferred over a four-legged one since in some cases it is difficult to set the latter and it may also obstruct a number of lines of sight.

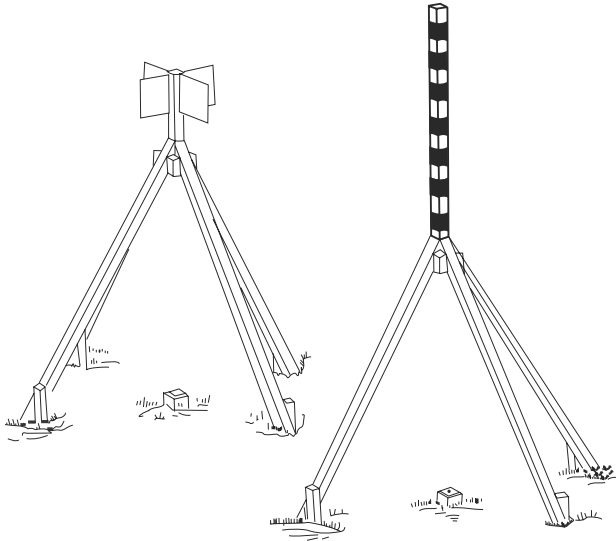


Fig. 2.7(a)

Signal with flag

Fig. 2.7(b)

Signal with pole

Luminous Signals For geodetic work, it is a general practise to make observations on luminous signals. These have the advantage of distinct and clear visibility even for long distances (more than 30 km), and enable the observations to be made both by day and by night. Examples of such signals are electric beacon lamps, heliographs and heliotropes. The latter two can be used only if the climate is favourable.

When the observations are to be made at night and if longer lines are to be ranged, either acetylene or electric lamps may be used. A beacon lamp (Fig. 2.8) is a type of electric lamp run on 6 V either by an accumulator or by dry cells. When properly adjusted, the beam of light has a spread of 5° . These are suitable up to a length of 80 km. For lines of over 80 km, the magnesium lamp with parabolic reflector is used.

Heliotropes or heliographs, generally referred to as 'helios', give light similar to bright stars over very long ranges. These are suitable in places where sunshine is expected for reasonably long periods. A heliotrope consists of plane mirrors and reflects the sun rays in the required direction; mirror of about 2.5 cm diameter

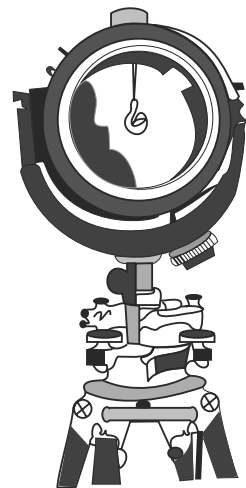


Fig. 2.8

Beacon lamp

covers a distance of about 16 km. A heliotrope may either be telescopic (Fig. 2.9) or defined by a sight vane and an aperture carrying cross-wires. It is very cheap and has no running and operating cost. But its use for first order triangulation is restricted since it can be used on sunny days only. The best time for observation with helios is towards sunset. The heliotrope reflects a continuous beam of light, whereas a heliograph reflects periodic beams.

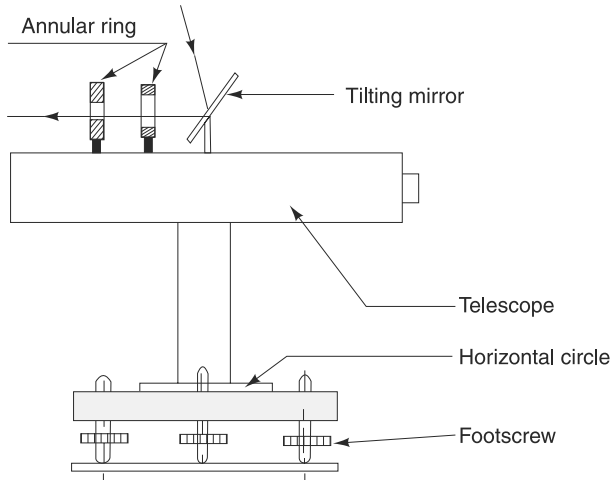


Fig. 2.9 *Telescopic heliotrope*

The opaque signals have an advantage over luminous signals on account of the fact that once the former have been erected they require little attention. On the other hand, though luminous signals increase the accuracy of measurement, they are, however, uneconomical and require an additional person to take care.

2.4.3 Phase of the Signal

When the observations are made in the sunlight on a signal of circular shape (Fig. 2.10), due to lateral illumination, some part of the signal is lighted up while the other part is in the shade. The observer has a tendency to bisect the centre of the illuminated part. This causes an angular error which is known as 'phase error'. The observed angles must, therefore, be reduced to that corresponding to the centre of the signal. If a single target is used and set normal to the line of sight during observation, the phase may be avoided.

The effect of phase is more common in cylindrical signals.

Let O be the centre of a signal at a distance D from the observer, and O' the false centre. Let r be the radius of the signal and A the observer's position. The phase error is given by

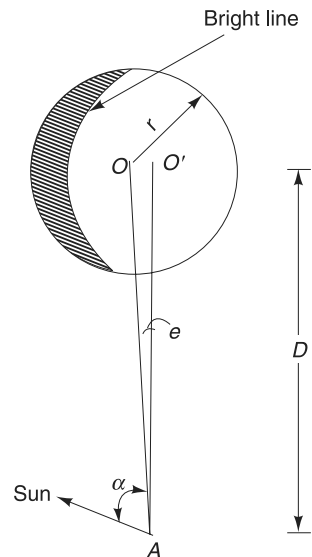


Fig. 2.10 *Phase error*

$$e = \frac{r \cos^2(\alpha/2)}{D \sin 1''}$$

$$= \frac{r \cos^2(\alpha/2)}{D} \times 206\,265 \text{ sec}$$

where α is the angle which the direction of the sun makes with AO . Sometimes, the observation is made on the bright line instead of on the centre of the signal. Then the phase error is

$$e = \frac{r \cos(\alpha/2)}{D} \times 206\,265 \text{ sec}$$

2.4.4 Towers

A tower (Fig. 2.11) is erected when it is desired to elevate the instrument or signal, above the instrument station, to obtain a line of sight clear of obstructions. These are built in duplicate, one to support the instrument and the other to support the observer and signal. Usually, these are made of steel sections which can be bolted together. These are known as 'Bilby towers'.

2.5.4 Scaffolds

These are occasionally used for first and second order triangulation across plains. These are portable and may be made with steel or timber. The height ranges between 10 and 15 m. The scaffolds are also made with a double structure similar to towers. This arrangement prevents the disturbances of the instrument due to movement of the observer.

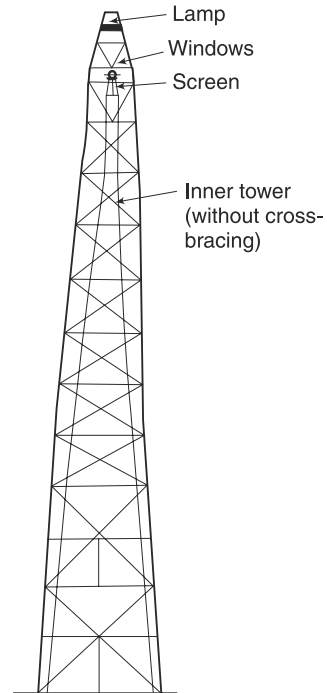


Fig. 2.11 Typical steel tower

Example 2.3 Observations were made on the centre of a 10 cm diameter signal, from the instrument at A. The distance AB is 6 km and the sun makes an angle of 50° with the line AB . Determine the phase error if the observation are made on the bright portion.

Solution

$$D = 6 \text{ km} = 6 \times 10^5 \text{ cm}$$

$$e = \frac{r \cos^2(\alpha/2)}{D} \times 206\,265 = \frac{5 \times \cos^2 25^\circ}{6 \times 10^5} \times 206\,265$$

$$= 1.41''$$

2.5 RECONNAISSANCE

The object of reconnaissance is to plan a system of triangulation in accordance with the specifications laid down for the type of triangulation. A thorough reconnaissance of the area contributes to the accuracy, geometrical strength,

simplicity and economy. It is impossible to realise all the requirements and a compromise may have to be made.

The most favourable conditions for triangulation are found in the hilly countries. It is often helpful to make sketches of the area from a highest point, noting all high points which may be suitable for use as stations. Also, the information regarding access to the stations, transport facilities, and company accommodation should be collected. Once the tentative locations of stations are ascertained, the intervisibility of these stations is checked. At least two locations of base lines are also selected. In the following subsections, some of the requirements for consideration are presented briefly.

2.5.1 Size of Triangles

The main factor influencing the size of a triangle in a framework is the purpose for which the survey is being conducted. For example, it should be small for a topographic survey, since the stations should be quite close so as to provide necessary control for traverses. Whereas, for a geodetic survey, the triangles are quite large having sides of several kilometres, with due allowance being made for curvature of the earth. Hence, in geodetic surveying, the practicable height of the tower limits the size of the triangles.

2.5.2 Shape of Triangles

The shape of a triangle should be such that any error in the measurement of an angle should affect the calculated sides the least. In triangulation, the sides are computed by using the sine rule and since the sines change more rapidly for small angles than for large ones, the percentage error in the computed side opposite to a small angle of a triangle will be larger as compared to a side opposite a large angle.

To compute the optimum angle, let A, B, C be the three angles of a triangle and a, b, c be the three opposite sides, respectively (Fig. 2.12). Let $\delta A, \delta B, \delta C$ be the error in measurement of the angles A, B, C , and $\delta a_1, \delta a_2, \delta a_3$ be the corresponding errors in the side a due to the errors in the measurement of the angles A, B and C .

By the sine formula, $a = c \frac{\sin A}{\sin C}$

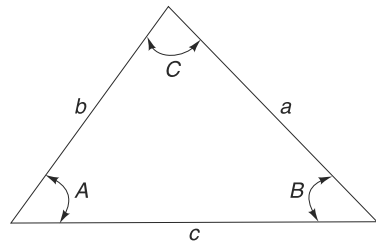


Fig. 2.12

Then, differentiating partially with respect to A

$$\delta a_1 = c \frac{\cos A \delta A}{\sin C}$$

$$\therefore \frac{\delta a_1}{a} = \frac{\cos A}{\sin A} \delta A = \cot A \delta A$$

and differentiating partially with respect to C

$$\delta a_2 = -c \frac{\sin A \cos C \delta C}{\sin^2 C}$$

$$\therefore \frac{\delta a_2}{a} = -\frac{\cos C}{\sin C} \delta C = -\cot C \delta C$$

If δA and δC , the probable errors in the angles, be $\pm \beta$, then the probable fractional error

$$\frac{\delta a}{a} = \sqrt{\left(\frac{\delta a_1}{a}\right)^2 + \left(\frac{\delta a_2}{a}\right)^2}$$

$$\text{or} \quad \frac{\delta a}{a} = \pm \beta \sqrt{\cot^2 A + \cot^2 C}$$

The probable fractional error will be minimum when $\cot^2 A + \cot^2 C$ is minimum.

Assuming the two angles A and B to be equal, then

$$C = 180^\circ - A - B = 180^\circ - 2A$$

$$\therefore \cot^2 2A + \cot^2 2A \text{ should be minimum.}$$

Differentiating with respect to A and equating to zero leads to

$$4 \cos^4 A + 2 \cos^2 A - 1 = 0$$

$$\text{or} \quad A = 56^\circ 14'$$

Therefore, the best shape of a triangle is an isosceles with base angles of $56^\circ 14'$. However, practically speaking, an equilateral triangle is the best.

2.5.3 Selection of Stations

The triangulation stations must be selected carefully. An experienced surveyor can save a lot of time and funds by keeping the following key points in mind:

1. Stations should be intervisible.
2. Stations should form well-conditioned triangles.
3. Stations should be easily accessible.
4. Stations should be so selected that the survey lines are not too long.
5. Stations should be so located that the cost of clearing, cutting and building towers is minimum.
6. To avoid trespassing, the stations should be selected such that the main lines are within the area to be surveyed.
7. Stations are so selected that the complete details of the area can be covered with them.

2.5.4 Intervisibility and Height of Stations

The most essential condition of triangulation is that the stations should be intervisible. When the distances are small, the intervisibility can be checked by the reconnaissance party directly, but if the distance between the stations is too large or the difference in elevation between the stations is less, then both the signals at the stations and the instrument are raised to overcome the effect of curvature and to achieve the desired intervisibility. The calculations of the height of a signal as well as that of an instrument depend upon the following factors:

1. *Distance between stations:* The distance of visible horizon above the datum is given by

$$h = \frac{D^2}{2R} (1 - 2m) = 0.06735D^2$$

where D = distance to the visible horizon (in km)
 h = height of the station above datum (in m)
 R = mean radius of earth
 m = coefficient of refraction
 = 0.07 for sights over land
 = 0.08 for sights over sea

2. *Relative elevation of stations:* If there is no obstruction due to intervening ground, then the required elevation of a signal at the other station can be calculated by using the preceding expression for h . Applying the same at station A (Fig. 2.13) leads to

$$R_1 = \frac{D_1^2}{2R} (1 - 2m)$$

Similarly, at station B (Fig. 2.11)

$$R_2 = \frac{D_2^2}{2R} (1 - 2m)$$

where R_1 = known elevation of A above datum
 R_2 = required elevation of B above datum
 D_1 = distance of station A from O
 D_2 = distance of station B from O
 D = total distance

If we know R_1 , then D_1 can be computed, and this determines D_2 since, $D_2 = D - D_1$; R_2 can then be evaluated by the stated formula. If the actual ground level at B is known, it can be ascertained whether the station B is to be raised, above the ground, or not.

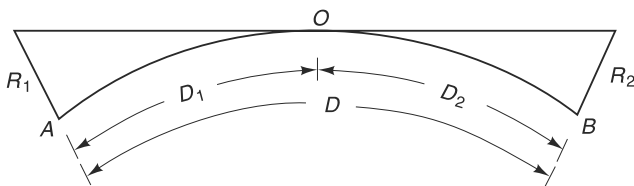


Fig. 2.13

Note While making such calculations, the line of sight is kept at least 3.0 m above the ground, as refraction is maximum near the ground. Also, the line of sight should not graze the ground.

3. *Profile of the intervening ground:* A profile of the intervening ground is plotted. This is done by plotting the peaks of the undulating ground. Then it is ensured that the proposed line of sight passes above the obstructions. This can be checked as discussed above or by Captain G.T. McCaw's solution.

In Fig. 2.14, let h_1 = height of station A above datum, h_2 = height of station B above datum, h = height of line of sight at the obstruction, $R = 2S$ = distance between stations A and B, and ξ = zenith distance between A and B. Then

$$h = \frac{1}{2}(h_2 + h_1) + \frac{1}{2}(h_2 - h_1) \frac{X}{S} - (S^2 - X^2) \operatorname{cosec}^2 \xi \frac{(1-2m)}{2R}$$

where $\operatorname{cosec}^2 \xi = 1 + \left(\frac{h_2 - h_1}{2S} \right)^2 \approx 1$ and $\frac{1-2m}{2R} = 0.06735$

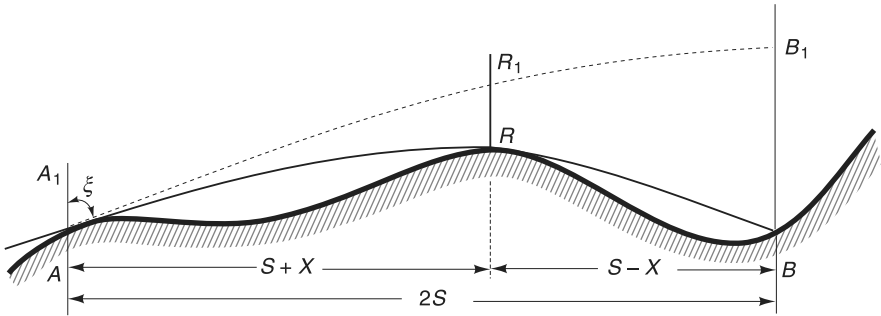


Fig. 2.14

Example 2.4 An observer standing on the deck of a ship just sees a lighthouse top with his eye at a height of 9 m. The top of the lighthouse is 64 m above m.s.l. Find the distance of the observer from the lighthouse.

Solution Refer to Fig. 2.15.

Let, $AO = D_1$ and $BO = D_2$.

This is a problem of curvature and refraction.

We know that combined correction = $0.0673D^2$

Hence, $9 = 0.0673 AO^2$

or $AO = \sqrt{\frac{9.00}{0.0673}} = 11.56 \text{ km}$

and $BO = \sqrt{\frac{64}{0.0673}} = 30.83 \text{ km}$

Hence, the distance between the observer and the lighthouse is

$$= 11.56 + 30.83 = 42.39 \text{ km}$$

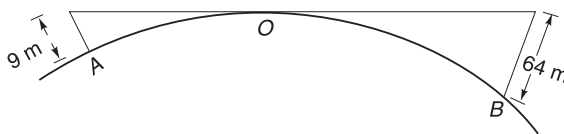


Fig. 2.15

Example 2.5 The elevations of two triangulation stations A and B , 120 km apart, are 200 m and 1000 m above m.s.l. The elevations of two peaks C and D on the profile between them are, respectively, 290 m and 540 m. The distance $AC = 50$ km and $AD = 80$ km. Design a suitable signal required at B , so that it is visible from the ground station A .

Solution Let the horizontal sight from A cut the horizon in e (Fig. 2.16). The distance Ae to the visible horizon,

$$D = \sqrt{\frac{h}{0.0673}} = \sqrt{\frac{200}{0.0673}} = 54.51 \text{ km}$$

We know that $AC = 50$ km, $AD = 80$ km and $AB = 120$ km. Thus,

$$ec = 54.51 - 50 = 4.51 \text{ km}$$

$$ed = 80 - 54.51 = 25.49 \text{ km}$$

$$eb = 120 - 54.51 = 65.49 \text{ km}$$

The corresponding heights are:

$$cc_1 = 0.0673 \times (4.51)^2 = 1.369 \text{ m}$$

$$dd_1 = 0.0673 \times (25.49)^2 = 43.52 \text{ m}$$

$$bb_1 = 0.0673 \times (65.49)^2 = 288.645 \text{ m}$$

To ascertain if the line of sight AB will clear the peaks C and D , consider triangles Ac_1c_2 and ABb_1 .

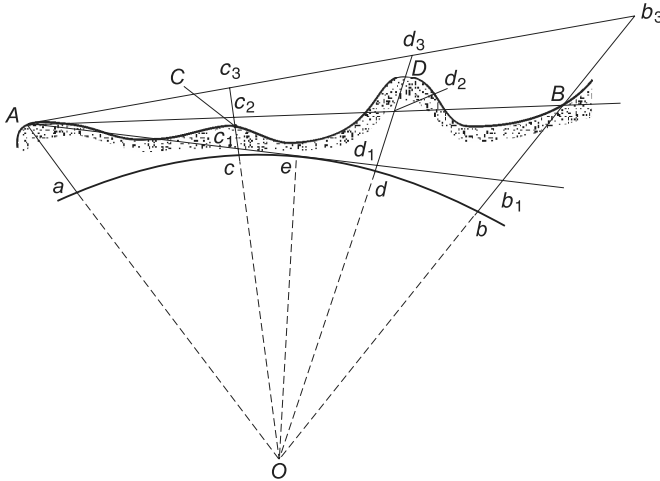


Fig. 2.16

$$\frac{c_1c_2}{b_1B} = \frac{Ac_1}{Ab_1} = \frac{50}{120}$$

and

$$\frac{d_1d_2}{b_1B} = \frac{Ad_1}{Ab_1} = \frac{80}{120}$$

$$b_1B = Bb - bb_1 = 1000 - 288.645 = 711.35 \text{ m}$$

Hence,
$$c_1c_2 = \frac{50}{120} \times 711.35 = 296.40 \text{ m}$$

$$d_1d_2 = \frac{80}{120} \times 711.35 = 474.233 \text{ m}$$

The elevation of the line of sight at $c = c_1c_2 + cc_1$

$$= 296.40 + 1.369 = 297.769 \text{ m}$$

The elevation of C is 290 m. Hence, the line of sight will clear the peak C by the distance $290 - 297.769 = -7.769 \text{ m}$.

The elevation of line of sight at $D = d_1d_2 + dd_1$

$$= 474.233 + 43.52$$

$$= 517.753$$

The elevation of D is 540 m. Hence, the line of sight will fail to clear the peak D by the distance $540 - 517.753 = 22.247 \text{ m}$.

The line of sight should be at least 3.0 m above ground level. Hence,

$$c_3c_2 = -7.769 + 3 = -4.769 \text{ m}$$

So, the line of sight clears the obstruction, and

$$d_3d_2 = 22.247 + 3 = 25.247 \text{ m}$$

Hence, to clear all obstructions, the height of the signal at B is

$$Bb_3 = \frac{AB}{AD} \times d_2d_3 = \frac{120}{80} \times 25.247 = 37.8705 \text{ m}$$

Solution by Captain McCaw's method

$$h = \frac{1}{2}(h_2 + h_1) + \frac{1}{2}(h_2 - h_1)\frac{X}{S} - (S^2 - X^2)\left(\frac{1-2m}{2R}\right)$$

where

$$h_1 = 200 \text{ m}$$

$$h_2 = 1000 \text{ m}$$

$$2S = 120 \text{ km, i.e., } S = 60 \text{ km}$$

$$(S + X) = 80 \text{ km or } X = 80 - 60 = 20 \text{ km}$$

$$\text{and } \frac{1-2m}{2R} = 0.06735$$

$$\begin{aligned} \therefore h &= \frac{1}{2}(1000 + 200) + \frac{1}{2}(1000 - 200) \times \frac{20}{60} - (60^2 - 20^2) \times 0.06735 \\ &= 600 + 133.33 - 215.52 \\ &= 517.81 \text{ m} \end{aligned}$$

Hence, line of sight fails to clear by $540 - 517.81 = 22.19 \text{ m}$.

The line of sight should be 3 m above ground. Therefore, the line of sight should be raised by $22.19 + 3 = 25.19 \text{ m}$.

Hence, height of signal at $B = (120/80) \times 25.19 = 37.785 \text{ m}$.

2.6 BASE LINE MEASUREMENT

In triangulation, the base line is the most important line and is measured very precisely. This is so because the accuracy of triangulation depends upon the accuracy of the sides of the triangles computed from the base line. In addition to being a supplement to angular measurement in triangulation, the base measurement may also be used as a base to determine the scale of triangulation. This may also be used as a substitute for triangulation and is known as *trilateration*.

2.6.1 Site Selection for Base Line

Following points should be borne in mind:

1. The ground should be firm and level.
2. The extremities of the base line should be intervisible at ground level.
3. It should be free from obstructions throughout its length in order to have a clear line of sight.
4. It should provide a system of well-conditioned triangles.
5. The site should be level as far as possible or uniformly sloping.
6. The base line should pass through the centre of the area.
7. It is preferable to choose a base as small as possible so as to measure it precisely.

2.6.2 Base Measuring Equipments

Following are the base measuring equipments.

1. Rigid bars, e.g., contact apparatus and Colby apparatus.
2. Flexible apparatus, e.g., chain, wire and tape.
3. Electronic apparatus, e.g., radar and electromagnetic distance measurement (EDM).
4. Optical instruments, e.g., tacheometer.
5. Sound ranging.

Though all the ten bases of the Great Trigonometrical Survey of India were measured with the Colby apparatus, rigid bars are obsolete nowadays. On the other hand, flexible apparatus like chain or tape do not give very accurate results but for small bases, as in topographic triangulation, they provide sufficient accuracy. Radars and tellurometers give very good result but are expensive and involve skill. These are preferred for first and second order triangulation. Tacheometer is used for measuring small bases on undulating grounds, whereas sound ranging is used in hydrographic surveys.

2.7 BASE MEASUREMENT BY RIGID BARS

In the early 20th century, the triangulation work was done by rigid bars. With the introduction of wires and tapes, the rigid bars have become redundant (*obsolete*) and of historical interest only. But to keep the theoretical interest, general features of these bars are explained briefly.

The rigid bars are classed as *contact apparatus*—the ends of a predetermined number of bars are brought into contact for measuring the length in steps (e.g., Eimbeck Duplex apparatus); and *optical apparatus*—the effective lengths of the bars are engraved on them and the observations on it are made with a microscope (examples are Colby apparatus and Woodward Iced Bar apparatus).

A change in atmospheric temperature will alter the length of rigid base and to encounter this some provision in the apparatus had to be made. The Colby apparatus, used in India, is a compensating base bar in which length of the bar is maintained by a combination of two metals. The Eimbeck Duplex apparatus, used by US coast and Geodetic Survey, consists of two measuring bars and is an example of bimetallic non-compensating base bars. The Woodward Iced bar, used in Russia, is a monometallic base bar in which the temperature is kept constant at melting point of ice.

Colby Apparatus This apparatus was designed by Major General Colby. All the ten bases of Great Trigonometrical Survey of India were measured with the Colby apparatus. It consists of an iron and a brass bar (Fig. 2.17), each 10 ft $1\frac{1}{2}$ in. long, fixed together at the middle by means of two steel pins.

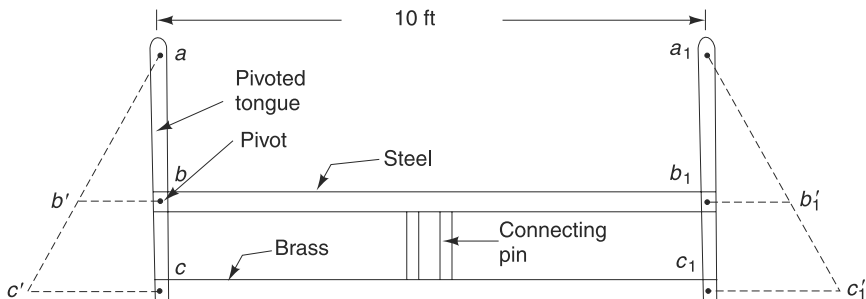


Fig. 2.17 Colby compensating bar

A flat steel tongue, about 6 inches long, is pivoted at each end of the bar. Each of the tongue carries one microscopic platinum dot a and a_1 making the distance aa_1 exactly 10 ft.

To secure compensation, the ratio ab/ac is made equal to the ratio of coefficients of linear expansion of iron and brass, i.e., $3/5$. Since the tongue is free to pivot, the position of the dot remains constant under the change of temperature. If due to a change of temperature, the length bb_1 changes to $b'b'_1$, say be x , then the length cc_1 will change to $c'c'_1$ by $\frac{5}{3}x$. Thus, the positions of the dots a and a_1 remains unchanged. The bar is held in a box at the middle of its length. A spirit level is placed on the bar, and is observed through a window in the top of the box.

For measuring the bases in India, five such bars were simultaneously used with a gap of 6 in. between the forward mark of one bar and the rear mark of the next bar by means of a framework. The framework was equipped with two microscopes with their cross-wires 6 in. apart. A small telescope, parallel to the microscopes,

is fixed at the middle of this bar for sighting reference marks on the ground. Each compensated microscope bar is supported on the box containing the main bar.

2.8 BASE MEASUREMENT BY FLEXIBLE APPARATUS

There are two methods, Wheeler's method and Jaderin's method, which use steel or invar types for the measurement of base lines. On the other hand, Hunter's short base method uses four chains, each 66 ft. long.

Wheeler's Method

Procedure

1. The line to be measured is cleared of the obstructions.
2. Stout poles 10 cm × 10 cm, also known as *posts*, are driven into the ground along the line at a distance of about 1 km apart. The pole alignment is done with a transit.
3. Then a series of such posts are driven between these posts at about tape length. These posts are called *marking posts* and project about 60 cm above the ground.
4. Some additional posts 2.5 cm × 5 cm are also driven in between the marking posts to support the suspended tape (Fig. 2.18). The post intervals range from 15 to 30 m.

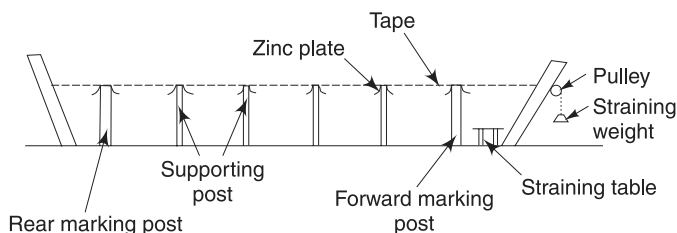


Fig. 2.18 *Wheeler's method of base measurement*

5. A strip of zinc or copper is nailed above each post to provide a flat surface.
6. The tape is stretched between the marking posts and is allowed to hang freely.
7. The rear end of the tape is connected to the 'straining pole' behind the 'rear marking post' and the forward end to a straining weight or spring balance attachment. A weighted device is usually preferred over a spring balance, since the former is easier to manipulate.
8. The rear end of the tape is adjusted to coincide with the mark on the zinc plate on the rear end.
9. A standard tension is applied by means of a straining weight and the tape is read on the zinc plate mark on the 'forward marking post'.
10. The temperature of the tape is determined by three thermometers, two at the ends and one placed at the middle of the tape.
11. The tape is then carried forward and the process is repeated.

Jaderin's Method This method was introduced by Jaderin and has become very popular due to the following reasons:

1. It is less expensive and the speed of measurement is high.
2. Due to longer lengths of the tape, the choice of the base sites become wider.

Procedure

1. The plumb bob is suspended on one of the two stations. On the other station, a theodolite is set up to provide a line of sight.
2. The tripods are placed at a spacing depending upon the length of the base line. The spacing between the tripods can be from 5 m to the tape length. If tripods are placed at the tape length, the weights are attached at the end of the tape to provide tension. Tension is measured with the help of a spring balance.
3. Another theodolite is placed somewhere in the field and is used to keep the tripod heads at the same level (Fig. 2.19). This avoids the correction for slope.

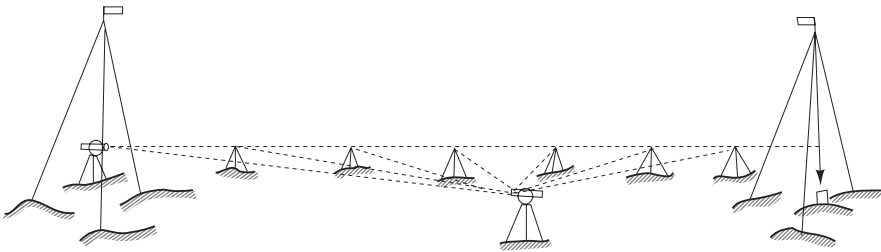


Fig. 2.19 Jaderin's method of base measurement

4. Total base length is obtained by adding the lengths between the tripods and the distances between the first and last tripod from the stations.
5. The exact length of the base line is obtained by applying tape corrections as required.

Hunter's Short Base Method It consists of a steel tape in four sections joined together by swivel joint. The tape has a length of 4 chains ($4 \times 66'$) and is rolled on a drum. The tape is supported at its two ends on three legged wooden posts, and in between it is supported by two tripods. A tension is applied on one end of the tape to produce a moment of 12 lbs. Targets are provided for sighting and reading angles. Corrections are directly found for the measured length from the tables. The instrument has been designed by the Department of Survey of India.

2.9 TAPE CORRECTIONS

The various sources of errors in taping and the corrections to be applied to the measured length have been discussed in Chapter 2 of Surveying Vol. 1. Some of the special conditions which may exist in the measurement of a base are presented

here and should be given due consideration while estimating the combined effect of corrections as applicable.

The tapes used for base measurement are either of invar or steel. The latter is more practicable since invar deforms very quickly if mishandled, but at the same time, an invar tape is more precise than a steel tape.

The temperature of the tape should be measured at least at three places. Two thermometers should be placed at the ends and one at the mid-length.

The effect of sag should be controlled by applying proper tension by a spring balance and weights. For all measurements the tape should, as far as possible, be supported at the same level to reduce the effect of sag as well as slope. Also, the supporting posts should be carefully aligned with a theodolite to reduce the errors due to non-alignment.

Finally, the effect of the second term of the formula for slope correction should also be investigated. Sometimes, for example, on steep slopes, the effect may be appreciable.

2.9.1 Modified Tape Corrections

Since the accuracy of the triangulation survey depends largely on the degree of precision with which the base has been measured, some modification or refinement may be required to account for the field conditions and the instruments used.

Correction for Slope

1. When the difference of elevation between the two ends is measured

Slope correction,
$$C_{sl} = -\frac{h^2}{2l}$$

where h = difference in elevation between the two ends of a line of length l measured on an uniform slope.

This slope correction is applicable only to measurements by rigid apparatus. For the measurements made with flexible apparatus, the deformation of catenary should be accounted for as the ends will not be at the same elevation. Henrici has developed the following formula for the measured difference of elevation.

Let L = nominal length of tape or wire

l = length along catenary between tripod marks

h = difference of elevation between measuring tripods

C_a = standardisation correction to nominal length

t = difference of mean temperature and standard temperature ($T_n - T_o$)

α = coefficient of thermal expansion

α_1 = algebraic sum of scale readings (when wires are used)

λ = apparent shortening due to sag $\left(C_{sa} = \frac{W^2 l}{24P^2} \right).$

X = required horizontal distance between tripod marks

X_0 = value of X when $h = 0$

Then $A = C_a + \alpha_1 + l \alpha t$

$$X_0 = l - \lambda = L + A$$

$$\begin{aligned}
 \text{and} \quad X &= X_0 \left(1 - \frac{h^2}{l^2} \right)^{\frac{1}{2}} \\
 &= (L + A) \left[1 - \frac{h^2}{(L + A + \lambda)^2} \right]^{\frac{1}{2}} \quad (2.1)
 \end{aligned}$$

Expanding Eq. (2.1), and retaining only the first powers of A and λ ,

$$X = L - B + A + C + D \quad (2.2)$$

where

$$\begin{aligned}
 B &= L \left(\frac{h^2}{2L^2} + \frac{h^4}{8L^4} + \frac{h^6}{16L^6} + \frac{5h^8}{128L^8} + \dots \right) \\
 C &= A \left(\frac{h^2}{2L^2} + \frac{3h^4}{5L^4} + \frac{5h^6}{16L^6} + \dots \right) \\
 D &= \lambda \left(\frac{h^2}{L^2} + \frac{h^4}{2L^4} + \frac{3h^6}{8L^6} + \dots \right)
 \end{aligned}$$

Note Equation (2.2) gives a result with an accuracy of 1 in 10^7 if $\frac{h}{L}$ is less than 0.1, A is less than $L/1000$ and the tension applied exceeds 20 times the weight of the tape.

2. When the slope is measured in terms of θ , the angle of elevation or depression

Slope correction, $C_{sl} = l (1 - \cos \theta)$

when the ends of the tape are not at the same level

Sag correction, $\lambda = -kl$

where $k = E^2 \cos^2 \theta (1 \pm E_1 \sin \theta)/24$ (2.3)

where $E_1 = W/P$

Note Use '+' when the tension (P) is applied at the upper end of the tape and '-' when the tension is applied at the lower end of the tape.

Correction for Pull (Tension) The correction is given by

$$C_p = \frac{P - P_o}{AE} L$$

where P is the pull applied during measurement, P_o is the standard pull, A is the area of cross section of the tape in cm^2 , E is the modulus of elasticity of the tape (Steel: $2.1 \times 10^5 \text{ N/mm}^2$; Invar: $1.54 \times 10^5 \text{ N/mm}^2$), and L is the measured length in metres.

Sometimes, a wire or tape is standardised in catenary, and if so, an alteration in tension will alter the sag correction. In this case, the sag correction for the tension used during standardisation should be added to the standardised chord length to give the equivalent length on the flat. The sag correction for the tension applied during measurement may then be subtracted from this result to give the new chord length.

Correction for temperature. The correction is given by

$$C_t = \alpha (T_m - T_0)L$$

where T_m is the mean temperature during measurement, T_0 is the temperature of standardisation, α is the coefficient of thermal expansion of the tape material (Steel: $0.0000035/^\circ\text{C}$; Invar: $0.000000122/^\circ\text{C}$), and L is the measured length in metres. The temperature measurement should be made at least with three thermometers (of least count 0.1°C).

Correction for Index Error of Spring Balance Usually, spring balance is used for measuring standard tension. A balance is adjusted to read correctly in the vertical position while standardising, whereas it indicates less tension than the true tension when used horizontally in the field. The difference known as index correction must be applied to the tension measured in the field.

$$\text{Index error for horizontal position} = \frac{W - I - I'}{2}$$

where I = index error when balance is vertical, i.e., the reading when suspended hook is downward, without load (taken as negative when less than zero)
 I' = reading with balance inverted and suspended from hook
 W = total weight of balance

Correction for Change of Gravity When a wire or tape measurement is made at a different elevation from that of the place of standardisation, allowance is made in geodetic work for the change in the value of acceleration due to gravity, g .

When tension is applied by means of suspended weights, the pull they exert and the weight of the wire are proportional to g . The form of the catenary remains constant, but the wire extends or shortens by the resulting increase or decrease from the standard tension. If the tension is measured by spring balance, it is unaffected by change of g , but the weight of the wire is altered and an increase of g will cause increase of sag.

The value of g over the earth can be found by using the Helmert formula.

$$g = g_0(1 + 0.005302 \sin^2 \phi)$$

where g = acceleration due to gravity at sea level in latitude ϕ
 g_0 = acceleration due to gravity at sea level at the equator

Correction for gravity,

$$C_g = -g \frac{2h}{R}, \text{ for height above mean sea level} \\ = +g \frac{3h}{4R}, \text{ for the mass between sea level and the station}$$

where h = elevation of station above mean sea level (in metre)
 R = mean radius of earth (in metre)

In the case of straining by weights, let g_1 , and g_2 be the respective values of gravity at the places of standardisation and of measurement. Then,

$$\frac{g_2 - g_1}{g_1} = \frac{\delta P}{P}$$

where δP = change produced in the nominal tension P .

Note The correction is positive for g_2 greater than g_1 .

Correction for Inclination of End Readings When the tape is suspended in catenary, the end scales will be inclined to the horizontal since the end of the tape itself is inclined to the horizontal. The inclination is given by

$$\tan \phi = \phi = \frac{W}{2P} \quad (\text{for small value of } \phi, \tan \phi = \phi).$$

Let Σr = algebraic sum of the end readings.

Then, the correction to be applied to these readings is

$$\begin{aligned} &= -\Sigma r(1 - \cos \phi) \\ &= -\Sigma r \left(1 - 1 + \frac{\phi^2}{2} + \dots \right) = -\Sigma r \frac{W^2}{8P^2} \\ &= -3\Sigma r \frac{C_{sa}}{l} \end{aligned}$$

The value Σr is usually very small and the correction is often neglected. However, this must be accounted for if the tapes are standardised in catenary.

Example 2.6 Show that for a flexible, uniform, inextensible tape of total weight $2W$, hanging freely between two supports at the same level under a tension T at each support, the horizontal distance between the supports is

$$\frac{H}{w} \log \left(\frac{T + W}{T - W} \right)$$

where H = horizontal tension at the centre of the tape, and
 w = weight per metre run of the tape.

Solution Take the origin of coordinates at the lowest point of the tape, with the X -axis horizontal and the Y -axis vertical. Consider a length s of tape measured from the origin. The tension at the origin is horizontal and equal to H . Let the tension at the other end of this length be P , and the inclination to the horizontal of the tangent at this point be α .

Resolving forces vertically and horizontally for this portion of the tape (Fig. 2.20 (a)).

$$P \sin \alpha = ws \quad (\text{i})$$

$$P \cos \alpha = H \quad (\text{ii})$$

Dividing Eq. (i) by Eq. (ii) gives

$$\tan \alpha = \frac{ws}{H}$$

Differentiating with regard to x gives

$$\sec^2 \alpha \times \frac{d\alpha}{dx} = \frac{w}{H} \times \frac{ds}{dx}$$

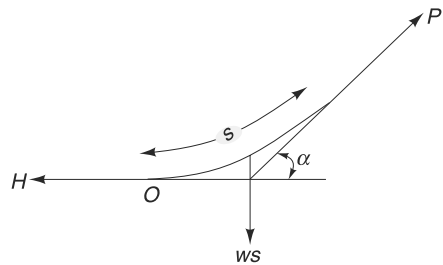


Fig. 2.20(a)

But from the elemental triangle (Fig. 2.22(b))

$$\begin{aligned} \frac{ds}{dx} &= \sec \alpha \\ \Rightarrow \sec^2 \alpha \times \frac{d\alpha}{dx} &= \frac{w}{H} \sec \alpha \\ \text{or } \sec \alpha \frac{d\alpha}{dx} &= \frac{w}{H} \end{aligned}$$

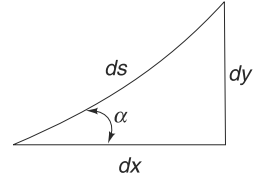


Fig. 2.20(b)

Then, if α' is the inclination of the tangent at the support and x' is half the distance between the supports (Fig. 2.20(c)).

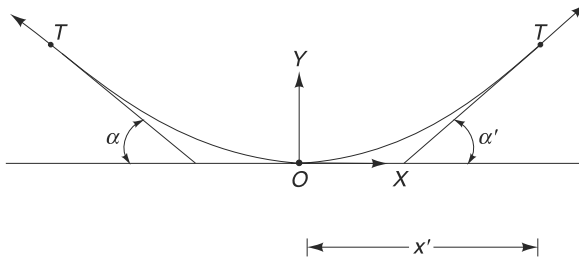


Fig. 2.20(c)

$$\int_0^{\alpha'} \sec \alpha \, d\alpha = \int_0^{\alpha'} \frac{w}{H} dx$$

$$[\log (\sec \alpha + \tan \alpha)]_0^{\alpha'} = \frac{w}{H} x'$$

Hence,
$$x' = \frac{H}{w} \log_e \left(\frac{\sec \alpha' + \tan \alpha'}{1+0} \right)$$

or
$$x' = \frac{H}{w} \log_e (\sec \alpha' + \tan \alpha') \quad (\text{iii})$$

Resolving vertically for one whole half of the tape;

$$T \sin \alpha' = W \quad \Rightarrow \quad \sin \alpha' = \frac{W}{T}$$

and since
$$\cos \alpha' = \sqrt{1 - \sin^2 \alpha'}$$

$$\cos \alpha' = \sqrt{\frac{T^2 - W^2}{T^2}} = \frac{\sqrt{T^2 - W^2}}{T}$$

$$\Rightarrow \sec \alpha' = \frac{T}{\sqrt{T^2 - W^2}}$$

Also
$$\tan \alpha' = \frac{\sin \alpha'}{\cos \alpha'} = \frac{W}{\sqrt{T^2 - W^2}}$$

Substituting for $\sec \alpha'$ and $\tan \alpha'$ in Eq. (iii) leads to

$$\begin{aligned} x' &= \frac{H}{w} \log_e \left(\frac{T}{\sqrt{T^2 - W^2}} + \frac{W}{\sqrt{T^2 - W^2}} \right) \\ &= \frac{H}{w} \log_e \left(\frac{T + W}{\sqrt{T - W} \sqrt{T + W}} \right) \\ &= \frac{H}{w} \log_e \sqrt{\frac{T + W}{T - W}} \\ &= \frac{1}{2} \times \frac{H}{w} \log_e \left(\frac{T + W}{T - W} \right) \end{aligned}$$

Then, total horizontal distance

$$= 2x' = \frac{H}{w} \log_e \frac{T + W}{T - W}$$

Example 2.7 In the process of measuring a base by Jadrine method, the tape was suspended between the two tripods A and B with the elevation of B higher than that of A . The distance between the successive tripods was kept equal to l . Derive an expression accounting for the effect of sag and slope correction in the base measurement.

Solution Refer to Fig. 2.21.

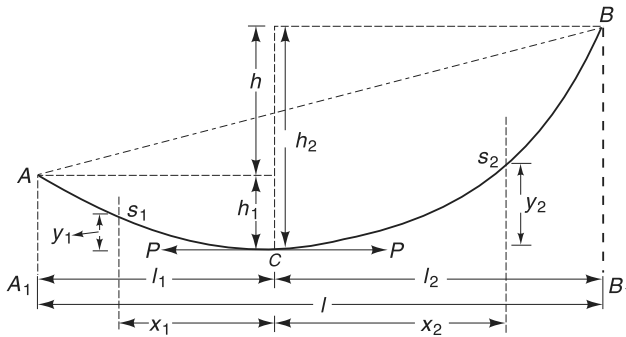


Fig. 2.21

Let

A and B = the positions of the tripods

C = the lowest point of the catenary

P = the tension at the lowest point C

l_1 and l_2 = the horizontal lengths A_1C and CB_1

s_1 and s_2 = the lengths CA and CB along the suspended tape

h_1 = difference of elevation between A and C

h_2 = difference of elevation between B and C

h = difference of elevation between B and A

w = weight per unit length of the tape

W = total weight of the tape

Approximating the catenary to be parabola and taking C the origin.

Equation of CA is $y_1 = k_1 x_1^2$ (i)
 at $x_1 = l_1, \quad y_1 = h_1;$

$$\Rightarrow k_1 = \frac{h_1}{l_1^2} \quad \therefore y_1 = \frac{h_1 x_1^2}{l_1^2} \quad \text{and slope, } \frac{dy_1}{dx_1} = \frac{2h_1 x_1}{l_1^2}$$

similarly for the CB portion, $y_2 = k_2 x_2^2$

$$k_2 = \frac{h_2}{l_2^2}$$

$$\therefore y_2 = \frac{h_2 x_2^2}{l_2^2} \quad \text{and slope, } \frac{dy_2}{dx_2} = \frac{2h_2 x_2}{l_2^2}$$

The length of curve

$$\begin{aligned} s = s_1 + s_2 &= \int_0^{l_1} \left[1 + \left(\frac{2h_1 x_1}{l_1^2} \right)^2 \right] dx_1 + \int_0^{l_2} \left[1 + \left(\frac{2h_2 x_2}{l_2^2} \right)^2 \right] dx_2 \\ s &= \left[l_1 + l_2 + \frac{2}{3} \left(\frac{h_1^2}{l_1} + \frac{h_2^2}{l_2} \right) \right] \\ s &= l + \frac{2}{3} \left(\frac{h_1^2}{l_1} + \frac{h_2^2}{l_2} \right) \end{aligned} \quad \text{(ii)}$$

For the equilibrium of CA

$$P \times h_1 = \frac{wl_1^2}{2}$$

$$\Rightarrow P = \frac{wl_1^2}{2h_1} \quad \text{(iii)}$$

$$\text{or } h_1 = \frac{wl_1^2}{2P} \quad \text{(iv)}$$

$$\text{and also } P \times h_2 = \frac{wl_2^2}{2}$$

$$\Rightarrow P = \frac{wl_2^2}{2h_2} \quad \text{(v)}$$

$$\text{or } h_2 = \frac{wl_2^2}{2P} \quad \text{(vi)}$$

From Eqs. (iii) and (v)

$$\frac{wl_1^2}{2h_1} = \frac{wl_2^2}{2h_2}$$

or
$$\frac{h_1}{l_1^2} = \frac{h_2}{l_2^2} \quad (\text{vii})$$

Substituting the values h_1 and h_2 in Eq. (ii)

$$s = l + \frac{2}{3} \left[\left(\frac{wl_1^2}{2P} \right)^2 \times \frac{1}{l_1} + \left(\frac{wl_2^2}{2P} \right) \times \frac{1}{l_2} \right]$$

or
$$s - l = \frac{2}{3} \frac{w^2}{P^2} (l_1^3 + l_2^3) \quad (\text{viii})$$

where

$s - l$ = total correction, C_{tl}

Substituting

$$l_1 = \frac{l}{2} - e, \text{ and } l_2 = \frac{l}{2} + e$$

$$\Rightarrow e = \frac{l_2 - l_1}{2}$$

From Eq. (viii), we get

$$\begin{aligned} C_{tl} &= \frac{1}{6} \frac{w^2}{P^2} \left[\left(\frac{l}{2} - e \right)^3 + \left(\frac{l}{2} + e \right)^3 \right] \\ &= \frac{w^2}{6P^2} \left[\frac{l^3}{4} + 3le^2 \right] \end{aligned}$$

Substituting the value of e ,

$$\begin{aligned} &= \frac{w^2 l^3}{24P^2} + \frac{w^2 l}{8P^2} (l_2 - l_1)^2 \\ &= \frac{w^2 l^3}{24P^2} + \frac{w^2 l^2}{8P^2} \frac{(l_2 - l_1)^2}{l} \\ &= \frac{w^2 l^3}{24P^2} + \frac{w^2}{8P^2} \frac{(l_2 + l_1)^2 (l_2 - l_1)^2}{l} \\ C_{tl} &= \frac{w^2 l^3}{24P^2} + \frac{w^2}{8P^2} \frac{(l_2^2 - l_1^2)^2}{l} \quad (\text{ix}) \end{aligned}$$

From Eq. (iii),
$$\frac{w}{P} = \frac{2h_1}{l_1^2}$$

From Eq. (vii),
$$\frac{l_2^2 - l_1^2}{l_1^2} = \frac{h_2 - h_1}{h_1}$$

Substituting these values in second part of Eq. (ix)

$$C_{tl} = \frac{w^2 l^3}{24P^2} + \frac{1}{8} \left(\frac{2h_1}{l_1^2} \right)^2 \left(\frac{h_2 - h_1}{h_1} \right)^2 \frac{l_1^4}{l}$$

$$\begin{aligned}
 &= \frac{w^2 l}{24P^2} + \frac{4h_1^2}{8l_1^4} \frac{(h_2 - h_1)^2}{h_1^2} \frac{l_1^4}{l} \\
 &= \frac{w^2 l}{24P^2} + \frac{(h_2 - h_1)^2}{2l} \\
 &= \frac{w^2 l}{24P^2} + \frac{h^2}{2l}
 \end{aligned}$$

Example 2.8 The three bays of a base line were measured by a steel tape and the following data were collected:

No.	Length (m)	Temperature (°C)	Pull (kg)	Difference of level supports (m)
1	30.084	12	7	0.3
2	29.973	13	7	0.7
3	25.233	17	5	0.7

The 30 m tape was standardised at 15°C under a pull of 4.5 kg. What are the true length of the bays? The tape weighs 1 kg. The unit weight of steel is 8300 kg/m³ and the coefficient of thermal expansion is 0.000011/°C and $E = 2.1 \times 10^4$ kg/mm².

Solution

(a) Length measured = 30.084 m

Length of tape, $L = 30$ m

Temperature = 12°C

Pull = 7 kg

Difference of level supports = 0.3 m

Standard temperature = 15°C

Standard pull = 4.5 kg

$E = 2.1 \times 10^4$ kg/mm²

$\alpha = 0.000011/^\circ\text{C}$

Unit weight of steel = 8300 kg/m³

Weight of tape, $W = 1$ kg

Correction for pull:

$$C_P = \frac{(P - P_0)L}{AE}$$

$$\text{Area of tape, } A = \frac{1}{30 \times 8300}$$

$$\Rightarrow A = 4.016 \times 10^{-6} \text{ m}^2$$

$$C_P = \frac{(7 - 4.5) \times 30}{4.016 \times 10^{-6} \times 2.1 \times 10^4 \times 10^6}$$

$$C_P = 8.893 \times 10^{-4} \text{ m} \quad (\text{i})$$

Correction for temperature:

$$\begin{aligned} C_T &= \alpha \Delta T L \\ &= 0.000011 \times (12 - 15) \times 30.084 \\ \Rightarrow C_T &= -9.927 \times 10^{-4} \text{ m} \end{aligned} \quad (\text{ii})$$

Correction for difference of level supports:

$$\begin{aligned} C_L &= -\frac{h^2}{2L} \\ C_L &= -\frac{(0.3)^2}{2 \times 30.084} = -1.5 \times 10^{-3} \text{ m} \end{aligned} \quad (\text{iii})$$

Correction for sag:

$$\begin{aligned} C_S &= \frac{W^2 L}{24n^2 p^2} \\ &= \frac{1^2 \times 30 \times 1^2}{24 \times 1 \times 7^2} \\ C_S &= -0.0255 \end{aligned} \quad (\text{iv})$$

$$\begin{aligned} \text{Total corrected length} &= 30.084 + (8.893 \times 10^{-4}) - (9.927 \times 10^{-4}) \\ &\quad - (1.5 \times 10^{-3}) - (0.0255) \\ &= 30.057 \text{ m} \end{aligned}$$

- (b) Length = 29.973 m
 Temperature = 13°C
 Pull = 7 kg
 Difference of level supports = 0.7 m

Correction for pull:

$$\begin{aligned} C_P &= \frac{(P - P_o)L}{AE} \\ &= \frac{(7 - 4.5) \times 30}{4.016 \times 10^{-6} \times 2.1 \times 10^4 \times 10^6} \\ C_P &= 8.893 \times 10^{-4} \text{ m} \end{aligned} \quad (\text{i})$$

Correction for temperature:

$$\begin{aligned} C_T &= \alpha(\Delta T)L \\ &= 0.000011 \times (13 - 15) \times 29.973 \\ &= -6.59 \times 10^{-4} \text{ m} \end{aligned} \quad (\text{ii})$$

Correction for difference of level supports:

$$\begin{aligned} C_L &= -\frac{h^2}{2L} \\ &= -\frac{(0.7)^2}{2 \times 29.973} = 8.174 \times 10^{-3} \end{aligned} \quad (\text{iii})$$

Correction for sag:

$$C_S = \frac{W^2 L}{24n^2 p^2}$$

$$= \frac{1^2 \times 30 \times 1^2}{24 \times 1 \times 7^2}$$

$$C_S = -0.0255 \quad (\text{iv})$$

$$\begin{aligned} \text{Total corrected length} &= 29.973 + (8,893 \times 10^{-4}) - (6.59 \times 10^{-4}) \\ &\quad - (8.174 \times 10^{-3}) - (0.0255) \\ &= 29.9395 \text{ m} \\ &= 29.94 \text{ m} \end{aligned}$$

(c) Length = 25.233 m

Temperature = 17°C

Pull = 5 kg

Difference of level supports = 0.7 m

Correction for pull:

$$C_P = \frac{(P - P_o)L}{AE}$$

$$= \frac{(5 - 4.5) \times 30}{4.016 \times 10^{-6} \times 2.1 \times 10^4 \times 10^6}$$

$$= 1.78 \times 10^{-4} \text{ m}$$

Correction for temperature:

$$C_T = \alpha \Delta T L$$

$$= 0.000011 \times (17 - 15) \times 25.233$$

$$= 5.5512 \times 10^{-4} \text{ m}$$

Correction for difference of level supports:

$$C_L = -\frac{h^2}{2L}$$

$$= -\frac{(0.7)^2}{2 \times 25.233}$$

$$= -9.71 \times 10^{-3} \text{ m}$$

Correction for sag:

$$C_S = \frac{W^2 L}{24n^2 p^2}$$

$$\Rightarrow C_S = \frac{1^2 \times 30}{24 \times 1 \times 5^2}$$

$$\Rightarrow C_S = -0.05 \text{ m}$$

$$\begin{aligned}\text{Total corrected length} &= 25.233 + (1.78 \times 10^{-4}) + (5.55 \times 10^{-4}) \\ &\quad - (9.71 \times 10^{-3}) - 0.05 \\ &= 25.1740 \text{ m}\end{aligned}$$

Example 2.9 Compute the value of the correction due to change of gravity in the case of a base line, the length of which is 7170 m. The 30 m tapes were standardised in latitude $54^\circ 54'$ at an elevation of 23 m, the tension being applied by 90 N weights. The same weights were used in the field measurement in latitude $57^\circ 42'$ and at an elevation of 7 m. The tapes stretch 0.15 mm per newton tension per tape length. Take the radius of earth as 6400 km.

Solution

$$\begin{aligned}\phi &= 54^\circ 54' = 54.9^\circ, g_0 = 9.81 \text{ m/s}^2 \\ g &= g_0 (1 + 0.005302 \sin^2 \phi) \\ g_2 &= g_0 [1 + 0.005302 \sin^2 54.9^\circ] = 9.8448 \text{ m/s}^2 \\ h &= 23 \text{ m}, R = 6400 \text{ km} = 6400 \times 10^3 \text{ m} \\ C_g &= -g \frac{2h}{R}, \text{ and also } +g \frac{3h}{4R} \\ C_g &= -g \frac{2h}{R} = -9.8448 \times \frac{2 \times 23}{6400 \times 10^3} \\ &= -7.0759 \times 10^{-5} \text{ m (above m.s.l.)}\end{aligned}$$

$$\text{Also} \quad C_g = g \frac{3h}{4R} = 2.6534 \times 10^{-5} \text{ m (below, m.s.l.)}$$

$$\begin{aligned}\text{Now,} \quad g_1 &= g_0 [1 + 0.005302 \sin^2 57.7^\circ] \quad [\phi = 57^\circ 42' = 57.7^\circ] \\ &= 9.84716 \text{ m/s}^2\end{aligned}$$

$$\delta P/P = \frac{g_2 - g_1}{g_1} = 2.38139 \times 10^{-4}$$

$$[\text{but } P = 90 \text{ N/m}^2]$$

$$\Rightarrow \quad \delta P = 2.38139 \times 10^{-4} \times 90 = 0.02143 \text{ N}$$

The, stretched length = 0.15 mm per N tension per tape length

$$= 0.15 \times 0.02 \times 3 \times 30 \times 10^3$$

$$= 96.4465 \times 10^{-3} \text{ m} = 0.096446 \text{ m}$$

$$L = 7170 - 7.0759 \times 10^{-5} - 0.096446 \quad (\text{above m.s.l.})$$

$$= 7169.903 \text{ m}$$

$$\text{and} \quad L = 7170 + 2.6534 \times 10^{-5} - 0.096446 = 7169.903 \text{ m (below m.s.l.)}$$

2.10 RADAR RANGING

The introduction of radar has led to *radar triangulation* in which the lengths of the sides of triangles are measured directly by radar. This form of triangulation can also be used to

1. establish fixed points at very considerable distances apart,
2. establish networks of ordinary triangulation, which are separated by very wide gaps, to be connected together,
3. fix the position of aircraft in aerial survey and

4. fix the position of vessel engaged on sounding work in hydrographic survey.

The limits of accuracy of a single measure of a distance by radar are about 17–20 m, and these values are, for all practical purposes, independent of the length of the line.

There are two forms of distance measurement by radar—non-cooperative and cooperative. *Cooperative radars* are more accurate and preferred for survey works. In this form, a special instrument receives the signals at the distant point and, after a short delay of known duration, re-transmits them on a different frequency to the sending station. In *non-cooperative radars*, the waves are reflected directly from the object and are used in aerial and hydrographic surveys.

Principle The principle of measurements by radar is based on the measurement of the time taken for a short, intense pulse of short-wave radio energy to travel from one point to another and then return to the sending station after reflection or re-transmission at the distant station. The distance d between the two stations is given by $vt/2$, where v is the velocity of the propagation of radio waves and t is the time interval between emission and reception of the radio waves.

Distance Measurement by Radar Distances between two ground stations can be measured directly by radar, provided the range is limited to the distance of visual observation (100 km). However, atmospheric conditions near the ground may introduce errors in the assumed value of the velocity of propagation of the radio pulses. An error of 1/5000 may be expected, which would be too great for ordinary geodetic purposes. However, long distances, say up to 600 km can be measured with an error of 1/30 000 with the aid of an aeroplane.

For this purpose, let it be required to measure a distance AB . An aeroplane flies across the line AB , approximately at right angles to it and near the centre. The receiving and transmitting apparatus are fixed at A and B , respectively. Let c , d , e and f (Fig. 2.22) be the positions of aeroplane as it flies. The distances measured with the aeroplane position c will be cA and cB . Similarly, a number of sets of distances dA and dB , eA and eB , and fA and fB are measured at close interval of times. The distance AB will be the sum of distances cA and cB . In this way, the sum of the distances from the aeroplane positions to the end points are obtained for a number of points on the flight line. The results are plotted as a curve against time and the minimum sum of all possible positions of the aeroplane. The minimum sum will be the distance AB . An accuracy of 1/50 000 may be achieved by this method of radar ranging.

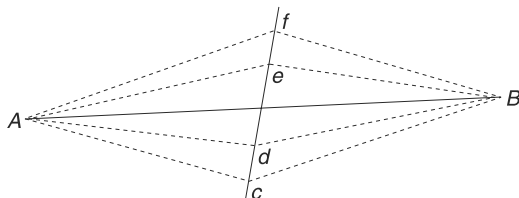


Fig. 2.22 Radar ranging

The radar ranging, as discussed above, can be used to fix control points over wide gaps. Let A and B be the two points fixed by ground survey on one side of

a wide gap, and C and D are to be fixed on the other side as shown in Fig. 2.23. The length and the azimuth of the side AB are known. AD , AC , BC and BD can be found by radar ranging. Since lengths of all the sides in each triangle are now known, the angles can be calculated and the points C and D can be fixed.

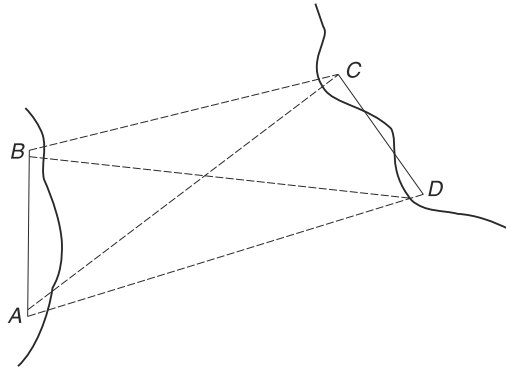


Fig. 2.23 Fixing control points by radar ranging

Variations of Velocity of Radio Waves The velocity of radio waves vary with variations in the value of coefficient of refraction of the atmosphere, which in turn is dependent on the values of atmospheric pressure, temperature and humidity. These values vary with height above mean sea level. The coefficient of refraction at a height H_1 metres above the sea level is given by

$$\mu_{H_1} = \mu_0 \left(1 - \frac{H_1}{4R} \right)$$

where μ_0 = index of refraction at sea level

R = mean radius of earth in metres

Then, if v is the mean velocity along the path,

$$v = v_0 \left(1 + \frac{H_2 + H_1}{8R} - \frac{D^2}{64R^2} \right)$$

where v_0 = velocity at sea level under normal atmospheric condition

D = approximate slant distance

H_2 = height of the aeroplane above sea level

H_1 = height of the ground station above sea level

Reduction of Observed Distance to Surface Distance The distances observed in radar ranging, such as cA , cB , etc. (Fig. 2.22) are the slant distances from the aeroplane position to the end points A and B . These should be corrected to bring them to horizontal lines at sea level.

Let A be the ground station, C the position of aircraft, O the centre of earth, and let the curved arc EF be the mean sea level as shown in Fig. 2.24. The measured length $AC = D$ and the required length is the arc $EF = S$. Let K be the chord length EF . Then,

$$K = R \left[\frac{[D - (H_2 - H_1)][D + (H_2 - H_1)]}{(R + H_1)(R + H_2)} \right]^{\frac{1}{2}}$$

and the arc distance

$$S = K + \frac{K^3}{24R^2} + \frac{3K^5}{640R^4} + \dots$$

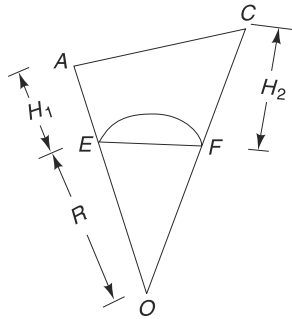


Fig. 2.24

Choice of Wavelength For radar works, it is necessary to use very short-length waves. Ultra-short waves of lengths upto say 1–2 m do not bent much by refraction and defraction and are the ideal choice. Another factor in the favour of short-wave transmission is that these give a sharper and narrower impression of the pulse on the recorder than the long waves do. This provides a more accurately defined commencement of the pulse front from which measurements can be taken. Hence, short waves give more accurate results. However, to circumvent electronic difficulties, wave lengths between 1 and 10 m length are used. Ultra-short waves of lengths less than a metre travel along an optical path and give the most accurate results.

2.11 PROBLEMS IN BASE MEASUREMENT

Some of the problems that are quite frequent in base measurement are extending, breaking and interpolating a base. These can be overcome with a good knowledge of trigonometry.

2.11.1 Extension of Base

The base line of a triangle in any class of triangulation is usually much shorter in comparison to the sides of a triangle. The base line is normally selected shorter because of the following reasons:

1. It is usually not possible to get a flat site (level or uniformly sloping) for a long base line.
2. It is often not possible to get a favourable site for a longer base.
3. It is difficult and expensive to measure a long base line.

To avoid the ill-conditioned triangles, resulting from a short base, the base line is expanded in a series of stages.

The group of triangles meant for the expansion of the base form what is known as a 'base net'. Figure 2.25 represents a base net, which can be used to extend the base AB . The base AB is extended to GH in a series of steps, viz., $AB \rightarrow CD \rightarrow EF \rightarrow GH$.

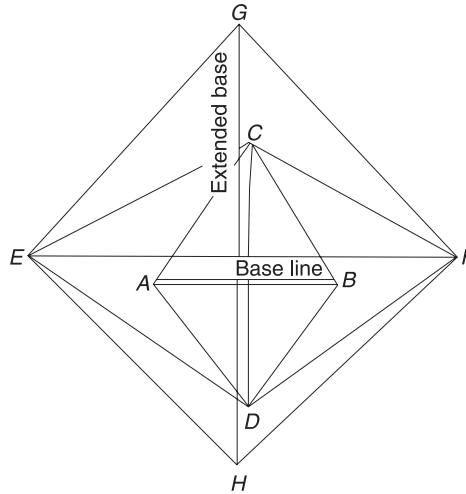


Fig. 2.25 Extension of base line

(a) Extension $AB \rightarrow CD$ The triangles involved are ABC , ABD , ACD , and BCD . Four values of CD are obtained, the average of which is taken.

(b) Extension $CD \rightarrow EF$ The triangles involved are CDE , CDF , CEF , and DEF . Four values of EF are obtained, the average of which is taken.

(c) Extension $EF \rightarrow GH$ The triangles involved are EFG , EFH , GHE and GHF . Four values of GH are obtained, the average of which is taken.

Thus, the base line AB is expanded to GH . Base extension can also be done by any other suitable method.

2.11.2 Broken Base

Sometimes it becomes essential to break a base because of the intervening obstructions. In Fig. 2.26 the base AB had to be broken into two sections AC and CB , because an obstruction C intervenes. Let C be an accessible point.

The length $AC' = b$ and $C'B = a$ are measured in the usual way and the length AB is calculated by the following formula obtained from the cosine rule:

$$\begin{aligned}
 AB &= a + b - \frac{ab\alpha^2}{a+b} \times \frac{1}{2} \times (\sin 1')^2 \\
 &= a + b - \frac{ab\alpha^2}{a+b} \times 4.2308 \times 10^{-8}
 \end{aligned}$$

where α is the deflection angle measured in minutes. It is desirable that this angle should not exceed 5° . But if, when necessary, it exceeds 5° , the sine rule should be used to compute the base.

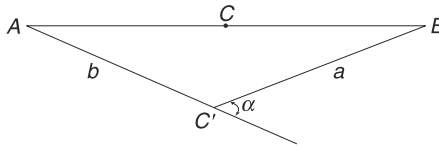


Fig. 2.26 Broken base

2.11.3 Interpolating Base

Sometimes, due to some reason or the other, it becomes impossible to measure a part of a base. In Fig. 2.27, it was not possible to measure CD (x), a portion of the base AB . A suitable point E is so selected such that the angles θ , α and β corresponding to the portions of the base can be measured. The problem is solved by simple trigonometry, and the following expression is used:

$$x = \frac{1}{2}(a + b) + \sqrt{\left(\frac{ab \sin(\theta + \alpha) \sin(\alpha + \beta)}{\sin \theta \sin \beta}\right) + \left(\frac{a - b}{2}\right)^2}$$

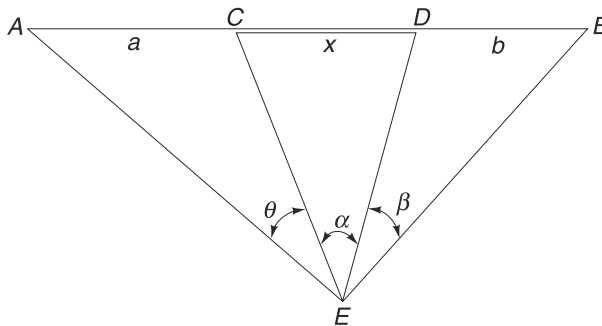


Fig. 2.27 Base interpolation

2.12 MEASUREMENT OF ANGLES

The angular measurements are made with a transit theodolite or with any other precise theodolite depending upon the desired degree of precision. For the highest precision, direction theodolites are used. A direction theodolite is similar to a transit theodolite and has a least count of $1''$. It does not have a lower plate tangent screw; verniers are not provided for reading the scales, and the readings are made with the help of two opposite micrometer microscopes. The transit theodolite and the methods of making precise angular measurements with it have been explained in Surveying Vol. 1 and need no further elaboration. Usually, the method of repetition is used for measuring angles.

Here it will be necessary only to consider the precautions to be exercised while making the observations.

Care should be taken in centring the instrument over the station mark. Let A be the true position of the instrument and suppose that it is required to observe angle BAC (Fig. 2.28). Due to false centring, the angle observed was $\angle BA'C$. The resulting error in angle ABC will be $\angle A'BA$, which is AA'/AB radians. If the error in centring is 10 cm and the length AB is 1 km, the resulting error is $20.62''$. The error in centring can be reduced by selecting the stations far away, thus making the sides of the triangle longer.

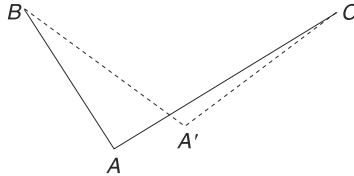


Fig. 2.28 Effect of false centring

Note It is preferable to observe horizontal angles in the mornings and evenings, whereas the vertical angles should be observed at noon to avoid errors due to refraction.

2.13 SATELLITE STATION AND REDUCTION TO CENTRE

Satellite stations are also called *eccentric* or *false stations*. To have intervisibility of stations and to achieve well-conditioned triangles, sometimes high objects like church spires, towers, etc., are selected as triangulation stations. It becomes impossible to set up the instrument over such stations for lack of sufficient space. In such cases some nearby stations, called *satellite stations*, are fixed and all the observations are made from it.

The angles measured at satellite stations are corrected and reduced to the values they would have been if the true station was occupied. The operation is known as *reduction to centre*. For example, in Fig. 2.29 let A , B and C be the triangulation

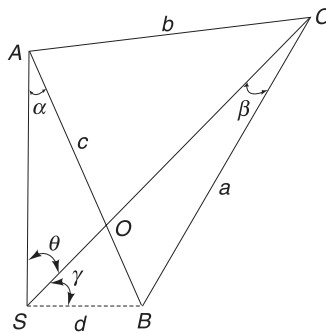


Fig. 2.29 Reduction of centre

stations and let B be the triangulation station on which it was impossible to set up the instrument. S is the satellite station at a distance d from the triangulation station B . a , b and c are the distances between the triangulation stations. AC (b) is the known distance (by computation), or is a base line and α and β are the corrections to be applied to the observed angle θ at the satellite station.

In triangle ABC

$$AB = c = CA \times \frac{\sin(\angle BCA)}{\sin(\angle ABC)} = b \frac{\sin(\angle BCA)}{\sin(\angle ABC)}$$

and
$$BC = a = CA \times \frac{\sin(\angle BAC)}{\sin(\angle ABC)} = b \frac{\sin(\angle BAC)}{\sin(\angle ABC)}$$

Knowing the sides a and c , apply sine rule to the triangles ASB and CSB to get

$$\sin \alpha = d \frac{\sin(\theta + \gamma)}{c}$$

and,
$$\sin \beta = d \frac{\sin \gamma}{a}$$

Since, α and β are small angles, we may write

$$\alpha = \frac{\sin \alpha}{\sin 1''} = d \frac{\sin(\theta + \gamma)}{c \sin 1''} = d \frac{\sin(\theta + \gamma)}{c} \times 206\,265 \text{ (in seconds)}$$

and
$$\beta = \frac{\sin \beta}{\sin 1''} = d \frac{\sin \gamma}{a \sin 1''} = d \frac{\sin \gamma}{a} \times 206\,265 \text{ (in seconds)}$$

Having determined the values of α and β , the true angle ABC can be found by

$$\angle AOC = \angle BAS + \angle CSA = \alpha + \theta$$

also,
$$\angle AOC = \angle ABC + \angle BCS = \angle ABC + \beta$$

or
$$\angle ABC + \beta = \alpha + \theta$$

or
$$\angle ABC = \theta + \alpha - \beta$$

The equation for angle ABC can be modified, depending upon the position of the satellite station. It should not be a practise to choose many satellite stations, since the accuracy of the work is affected and also the computation part is increased. These should, therefore, be selected only when necessary.

Example 2.10 Directions are observed from a satellite station S , 62.18 m from station C . Following were the results:

$$\angle A = 0^\circ 0' 00'', \angle B = 71^\circ 54' 32'', \angle C = 296^\circ 12' 02''$$

The approximate lengths of AC and BC were 8240.60 m and 10 863.60 m. Calculate the angle ACB .

Solution From triangle BSC in Fig. 2.30,

$$\angle CSB = 360^\circ - (296^\circ 12' 02'' - 71^\circ 54' 32'') = 135^\circ 42' 30''$$

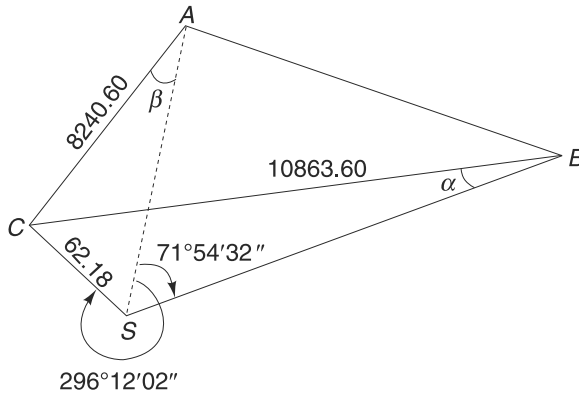


Fig. 2.30

$$\frac{62.18}{\sin \alpha} = \frac{10863.60}{\sin 135^{\circ}42'30''}$$

$$\text{or} \quad \sin \alpha = \frac{62.18 \times \sin 135^{\circ}42'30''}{10863.60}$$

$$\text{or} \quad \alpha = 0.00399 \times 206\,265 = 822.99''$$

From triangle CSA,

$$\angle CSA = 360^{\circ} - 296^{\circ}12'02'' = 63^{\circ}47'58''$$

$$\frac{62.18}{\sin \beta} = \frac{8240.60}{\sin 63^{\circ}47'85''}$$

$$\text{or} \quad \sin \beta = \frac{62.18}{8240.6} \sin 63^{\circ}47'58''$$

$$\text{or} \quad \beta = 0.00677 \times 206\,265 = 1396.47''$$

$$\therefore \angle ACB = 71^{\circ}54'32'' - 1396.47'' + 822.99'' = 71^{\circ}44'58.5''$$

Example 2.11 It was not possible to observe directions of some of the main triangulation stations from the triangulation station A. A satellite station S was selected 6 m away from the triangulation station A and to its south. The directions of stations B, C and D were observed from the satellite station S and are given below. Also the distances AB, AC and AD were reported to be 3100 m, 4000 m and 3000 m respectively. Determine the directions of AB, AC and AD.

Station	Whole circle bearing (WCB)
A	0°00''
B	130°20'00''
C	230°25'00''
D	300°00'00''

Solution The correction to any direction is given by

$$\beta = d \frac{\sin \theta}{D} \times 206\,265$$

where

θ = the angle reduced to the direction of line SA joining satellite station and triangulation station A

d = distance SA

D = the distance of the triangulation station from the observed station

Note The sign of correction is positive if the main station is in I and II quadrant and -ive if the same is in III or IV quadrant.

Correction to the direction of station B

$$\beta_B = \frac{6 \times \sin 130^\circ 20' 00''}{3100} \times 206\,265 = 5' 4.3''$$

Direction of AB = $130^\circ 20' + 5' 4.3'' = 130^\circ 25' 4.3''$

Correction to the direction of station C

$$\beta_C = \frac{6 \times \sin 230^\circ 25' 00''}{4000} \times 206\,265 = 3' 58.4''$$

Direction of AC = $230^\circ 25' 00'' + 3' 58.4'' = 230^\circ 28' 58.4''$

Correction to the direction of station D

$$\beta_D = \frac{6 \times \sin 300^\circ 00' 00''}{3000} \times 206\,265 = -5' 57.2''$$

Direction of AD = $300^\circ 00' 00'' - 5' 57.2'' = 299^\circ 54' 08''$

Example 2.12 A satellite station was established near a main triangulation station C and the following observations were recorded:

Satellite station	Sighting	Mean horizontal angle	Computed distance
	A	$0^\circ 00' 00''$	CB = 25 820 units
S	B	$61^\circ 48' 20''$	AC = 22 440 units
	C	$308^\circ 24' 40''$	

Determine angle ACB, if the eccentric distance was 273.70 units.

Solution

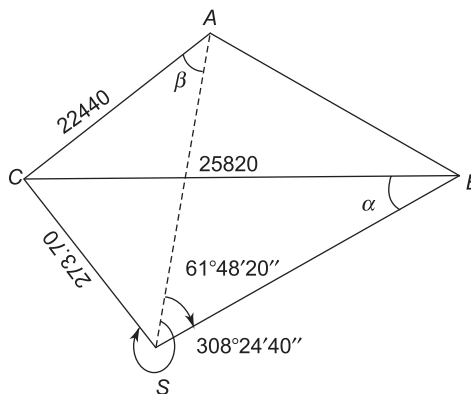


Fig. 2.31

From triangle BSC:

$$\begin{aligned}\angle CSB &= 360^\circ - 308^\circ 24' 40'' + 61^\circ 48' 20'' \\ &= 113^\circ 23' 40''\end{aligned}$$

Using sine formula:

$$\frac{273.70}{\sin \alpha} = \frac{25820}{\sin 113^\circ 23' 40''}$$

$$\therefore \alpha = 0^\circ 33' 26.76''$$

From triangle CSA:

$$\begin{aligned}\angle CSA &= 360^\circ - 308^\circ 24' 40'' \\ &= 51^\circ 35' 20''\end{aligned}$$

Using sine formula:

$$\frac{273.70}{\sin \beta} = \frac{22440}{\sin 51^\circ 35' 20''}$$

$$\Rightarrow \beta = 0^\circ 32' 51.35''$$

$$\begin{aligned}\therefore \angle ACB &= 61^\circ 48' 20'' + 0^\circ 33' 26.76'' - 0^\circ 32' 51.35'' \\ &= 61^\circ 48' 55.36''\end{aligned}$$

2.14 FIELD CHECKS IN TRIANGULATION

The only linear measurement made in triangulation for computation purposes is of the base line and is done with adequate precision. Sufficient check lines should be provided to check the computed sides. Regarding the angular measurements, before leaving a station the following checks should be ensured:

1. The angles are measured with face left and face right. The two observations should have a difference of 180° .
2. If the reiteration method of measuring angles is used, the sum of all the measured values should be 360° .
3. If the repetition method of measuring angles is used, each angle should be measured on different parts of the scale and the values should be compared.
4. The sum of the angles of a triangle should be 180° .
5. For quadrilaterals (Fig. 2.32(a)) and polygons (Fig. 2.32(b)), the following side check may be applied:

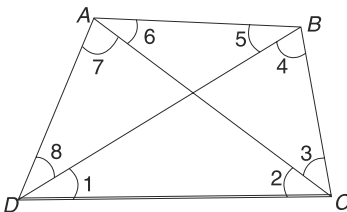


Fig. 2.32(a)

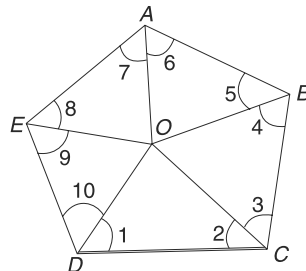


Fig. 2.32(b)

Let CD be the base line, then

$$\frac{\sin(1) \times \sin(3) \times \sin(5) \times \dots}{\sin(2) \times \sin(4) \times \sin(6) \times \dots} = 1$$

2.15 TRILATERATION

In this technique of triangulation, no angular measurement is made. The three sides of the triangles are measured precisely using the electromagnetic distance measuring equipment. This technique is useful when angular measurement is difficult or impossible due to any reason.

The geometric figures used are similar to those employed in triangulation. It was originally considered that trilateration will supersede triangulation as a horizontal control method. However, it is not so because of the following reasons:

1. Trilateration appears to be less time consuming and economical since there is no horizontal measurement required. However, much depends on the lengths of the lines involved and the accuracy desired.
2. Since all the EDMs measure slope distances, the relative levels of the control points and heights of the instrument, or the vertical angles will be required.
3. EMD instruments are calibrated for the velocity of electromagnetic waves under standard metrological conditions, which may not exist along the measured path. With the present sophistication, the temperature and pressure can be measured at each end of the line only.
4. Precise measurement of vertical angles will be required.
5. Triangulation has superior inertial checks, e.g., the sum of three internal included angles of a triangle should be 180° , whereas no such check is available in trilateration.

From the above considerations, trilateration possibly may be less accurate and economical than triangulation. In order to obtain sufficient redundancy for checks on the accuracy, all trilaterations must include the measurement of some of the angles. It follows from this that the modern practise is to combine trilateration and triangulation, thereby producing very strong networks. The coordinates of the networks stations may be obtained by the following approach.

The three angles of a triangle can be determined by using the half-angle formula if the sides a , b , c of the triangle are known. Thus,

$$\tan (A/2) = \frac{s(s-a)}{\sqrt{(s-b)(s-c)}}$$

where
$$s = \frac{1}{2}(a + b + c)$$

These angles are used to calculate bearings around the triangulation network. A closed traverse can thus be produced and adjusted to give the final coordinates.

The direct coordinates of the control points can be obtained without the use of angles. If the coordinates of two points A and B are (E_A, N_A) and (E_B, N_B) , then that of point C will be

$$E_C = \frac{1}{2} (E_A + E_B) + \frac{a^2 - b^2}{2c^2} (E_A - E_B)^2 + \frac{\Delta}{c^2} (N_A - N_B)$$

$$\text{and } N_C = \frac{1}{2} (N_A + N_B) + \frac{a^2 - b^2}{2c^2} (N_A - N_B)^2 + \frac{\Delta}{c^2} (E_A - E_B)$$

where A, B, C are the points in clockwise order, and

$$\Delta = \sqrt{s(s-a)(s-b)(s-c)}$$

2.16 FLARE TRIANGULATION

This is a method of triangulation over wide gaps—2 to 3 times the length of normal visual observation, i.e., 100–300 km—by synchronised theodolite observations to parachute flares. The flares are dropped near predetermined positions by aeroplanes.

Let A, B and C be three fixed stations and let D, E and F be the other three stations whose positions are to be fixed. The parachute flares are dropped in turn by an aeroplane at suitable points X, Y, Z (Fig. 2.33) lying on a great circle approximately half-way between the fixed stations and the stations to be fixed. The pilot of the aeroplane drops a flare at X and simultaneous observations to this flare are made with theodolites at A, B, C, D, E and F . Similarly, observations are made on flares dropped at Y and Z . The complete operation is repeated by dropping flares as closely as possible to the previous positions with change of face of the theodolite. This method of triangulation is highly expensive, requires most careful organisation, involves extensive equipments and a large number of trained personnels.

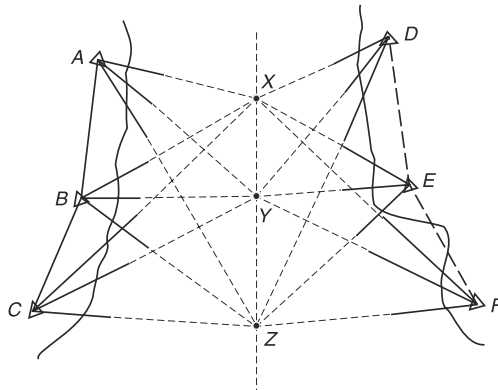


Fig. 2.33 Flare triangulation

2.17 FLOATING TRIANGULATION

Sometimes, the triangulation stations are fixed on floats in hydrographic surveying. Such triangulation is called floating triangulation or *triangulation afloat*. It may be required for any of the following purposes:

1. Coastal survey

2. For transferring the base line from one country to another separated by a channel or a strait.
3. For connecting an island with the main land.

Summary

Triangulation is concerned with involving the curvature of the earth's surface. This method of surveying is used when extensive areas and long distances are involved. The object is to establish horizontal control to accurately determine the relative positions of widely spaced points on the surface of earth and their absolute positions as well.

To accomplish the task the basic geometrical figure used is triangle and hence the method of surveying is termed triangulation. The stations form the vertices of a triangle and a series of mutually connected triangles form the triangulation system. Other geometrical figures such as quadrilaterals, pentagons and hexagons with central station are used appropriately depending upon the extent of the terrain to be surveyed; these figures are further subdivided into triangles for making necessary surveying measurements. All the triangles so made must be well conditioned; no angle should be less than 30° or more than 120° . The best shape of the triangle is equilateral and that of quadrilateral is square, though it is seldom possible to satisfy this condition in the field. This necessitates involvement of criteria called strength of figure, which provides guidance for selection of the course of triangles to be used for making calculations for the lengths of the mutually connected triangles. The triangulation systems are classified based on degree of accuracy required and magnitude of work involved, and are accordingly classified as primary, secondary, and tertiary triangulation; the former being the most precise.

The field procedure in triangulation survey consists in measuring the side of one of the triangles with all possible cautions, accuracy, and precision and is called base line. The two angles of the triangle involved in the computation of the sides of the triangle are also measured with all possible accuracy and precision. The remaining lengths of the sides of the triangle are calculated trigonometrically and these serve for calculation of the sides of the mutually connected triangles of the triangulation system. The so calculated lengths of the last line and of some selected intermediate lines are used to check the accuracy of the field work. Such lines are called check or proof lines. Sometimes the field conditions do not permit measurement of the base line at a stretch, the treatment and techniques to circumvent such problems are presented.

Selection of ground points as stations in triangulation has great bearing over accuracy achieved in the final output. This involves their intervisibility to facilitate measurement of horizontal and vertical angles from each station to the others. To overcome the intervisibility problem concept of satellite station is introduced. Then, methods of reduction of the real angles from the observed angles are discussed. The types of towers and signals erected at stations and the associated errors in sighting are discussed and described. The methods of measuring accurately the distances and angles are outlined. For the measurement of distances the flexible or rigid apparatus, or EDM technique, or radar ranging may be used. The procedures for making measurements with them are described.

With all the precision and accuracy exercised, still the survey carried out may contain errors. The field checks in the triangulation system are presented. The chapter ends with an introduction to trilateration and flare triangulation techniques.

Exercises

- 2.1 How would you determine the intervisibility of triangulation stations? Give in a tabular form the classification of triangulation based upon accuracy.
- 2.2 What is meant by triangulation? In what way does it differ from traversing?
- 2.3 Compare the various triangulation figures mentioning their suitability. How does triangulation differ from trilateration?
- 2.4 What are the points to be borne in mind for selecting a (a) base line and (b) triangulation station.
- 2.5 Describe a method of extending a base line and explain its necessity.
- 2.6 List out the different corrections that may be necessary for the measured length of a base line. Discuss briefly the conditions for consideration of the combined effect of the corrections.
- 2.7 What precautions will you like to exercise in measuring a base line?
- 2.8 What is meant by a satellite station? Why is it required? How are the observed angles to and from a satellite station reduced to their true value?
- 2.9 Explain briefly the various types of signals giving their merits and demerits.
- 2.10 Explain eccentricity of a signal. How would you correct the observations made upon an eccentric signal?
- 2.11 Define the following:
 - (a) Strength of figure
 - (b) Heliotropes
 - (c) Spherical excess
 - (d) Distance angles
- 2.12 Write short notes on the following:
 - (a) Luminous signals
 - (b) Station marks
 - (c) Base line measurement
 - (d) Station preparation in triangulation
- 2.13 Write short notes on the following:
 - (a) E.D.M. techniques of measuring distances
 - (b) Trilateration
 - (c) Distomat
 - (d) Jaderin's method of measuring base line
- 2.14 Discuss briefly the field checks in triangulation.
- 2.15 Write notes on the following:
 - (a) Signals used in triangulation
 - (b) Satellite station
 - (c) Corrections to be applied in base line measurements
 - (d) Intervisibility of stations
- 2.16 Discuss briefly the principle of radar ranging. Explain how the observed distances are reduced. Also discuss the correction to be applied to the variations of velocity of radio waves.
- 2.17 Write short notes on the following:
 - (a) Mekenometer
 - (b) Distomat
 - (c) Tellurometer
- 2.18 Compute the strength of the figure $ABCD$ (Fig. 2.34) for each of the routes by which the length BD can be computed from the known side AC .

(Ans. The best route is by using triangles ACD and BCD)

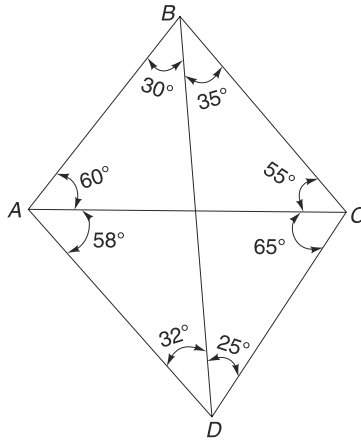


Fig. 2.34

- 2.19** The elevations of points *A*, *B* and *C* are 100, 110 and 125 m. The distance *AB* is 6 km and that of *BC* is 8 km. Will *B* create an obstruction to the line of sight from *A* to *C*? What will be the height of the signal required?
[Ans. Yes; by 2.52 m, 12.88 m']
- 2.20** In a triangulation survey, a theodolite was placed 10 cm out of centre from the station. What is the error in the measurement of a horizontal angle if the signal sighted is at 3 km.
[Ans. 6.88"]
- 2.21** Two triangulation stations *A* and *B* 100 km apart have elevations of 140 m and 406 m, respectively. A point *C*, 60 km from *A* has an elevation of 150 m. Check the intervisibility of *A* and *B* and, if required, determine the height of the signal at *B* so that the line of sight clears by 3.0 m.
[Ans. 25.1 m]
- 2.22** Two triangulation stations *A* and *B* are 100 km apart and have elevation of 40 m and 425 m respectively above mean sea level. A peak *C*, 60 m from *A* has an elevation of 150 m. Check the intervisibility of *A* and *B* and if required, determine the height of signal at *B* so that the line of sight clears the ground by 3.00 m. Take the radius of earth as 6375 km and coefficient of refraction as 0.07.
[Ans. 72.61 m]
- 2.23** Calculate the temperature correction for a 30 m long tape standardised at 28°C. It was used to measure a distance of 221.65 m. The mean temperature during measurement was 14°C. The coefficient of thermal expansion is $116 \times 10^{-7}/^{\circ}\text{C}$.
[Ans. 0.036 m (-)]
- 2.24** A base line was measured in four catenary lengths of 30.126, 29.9809, 30.06 and 24.56 m. The difference of levels were respectively 0.45, 0.55, 0.4, 0.5 m. The temperature during observation was 12°C and the straining mass was 15 kg. The tape was standardised as 30 m at 22°C with a straining mass of 5 kg. The coefficient of expansion was $0.000011/^{\circ}\text{C}$, the mass of tape = 1 kg, cross-sectional area = 3 mm^2 , $E = 210 \times 10^3 \text{ N/mm}^2$, and $g = 9.806 \text{ m/s}^2$. Calculate the length of the base.
[Ans. 114.502 m]
- 2.25** Sketch neatly one type of arrangement used to measure a base accurately using tapes and wires suspended in catenary form. Explain the procedure followed and the corrections which have to be applied.

- 2.26** Show that in base line measurement with tapes and wires in a flat catenary shape with the supports at different levels, the total correction will be $-(X + C)$, where X is the parabolic approximation for sag between level supports, and C is the level or slope correction taken permissibly to the first approximation.
- 2.27** Derive expressions for the relevant corrections for tension, sag, temperature and difference in level, applicable to base line measurement by tape in catenary form. (Properties of the parabola may be used if desired.)
- 2.28** A steel tape was exactly 30 m in length on a plane surface under a pull of 89 N at 20° C. A base line 780 m long was measured with it in catenary of 26 spans of equal lengths. The first 15 spans were measured at a temperature of 15°C and the remaining spans at 27°C. The pull applied during measurements was 178 N. Find the true base length reduced to m.s.l. for the following data:
- | | |
|---|--------------------------------------|
| Cross-sectional area of the tape | 3.40 mm ² |
| Mass of the tape | 7.8 N |
| E for steel | 2.07×10^5 N/mm ² |
| Thermal coefficient of expansion of steel | $0.11 \times 10^{-4}/^\circ\text{C}$ |
| Mean height of base above m.s.l. | 61 m |
| Radius of earth | 6367 km |
- The first and the last 5 spans were level, whereas the remaining 20 spans were on a uniform slope of 1 in 20.

[Ans. 779.281 m]

- 2.29** Show that for a flexible, uniform, inextensible tape of total weight W , hanging freely between two supports at different levels so that the low point lies outside the support, the tape distance between the low point and the support is

$$\frac{hT}{W} \pm \frac{s}{2}$$

where s is the total length of the tape, h the difference of support levels, and T the average tension at the supports.

- 2.30** Directions were observed from a satellite station 200 units from C with the following results:
- $\angle A = 30^\circ 48' 00''$ $\angle B = 75^\circ 42' 32.25''$ $\angle C = 300^\circ 00' 00''$
- The approximate lengths of AC and BC are 54 000 and 71 000 units, respectively. Compute the angle subtended at station C .

[Ans. $44^\circ 48' 1.94''$]

- 2.31** A satellite station S was occupied, during triangulation, at 62.18 m from the station A . The following observations were recorded at the satellite station:

Bearing	Length
$A = 296^\circ 12' 2''$	$AB = 10864$ m
$B = 71^\circ 54' 32''$	$AC = 8041$ m
$C = 0^\circ 00' 00''$	

Calculate the angle CAB [Ans. $71^\circ 44' 25''$]

- 2.32** The following observations were made on a satellite station S to determine angle BAC . Calculate the angle BAC .

Line	Length (m)	Line	Bearing
SA	9.435	SA	$0^\circ 00' 00''$
AB	2925	SB	$78^\circ 46' 00''$
AC	3426	SC	$100^\circ 12' 00''$

[Ans. $21^\circ 24' 26''$]

- 2.33** From a satellite station S near the main triangulation station A , the following observations were made on the three stations A, B, C .

$$\angle A = 309^\circ 17' 30'' \quad \angle B = 0^\circ 00' 00'' \quad \angle C = 63^\circ 32' 20''$$

A station D was fixed towards the station A at about 21 m from satellite stations S so that S, D and A are in the same vertical plane. The vertical angles observed at D and S to the station A were $21^\circ 53'$ and $14^\circ 24'$, respectively. The staff readings on a reference point, with the line of sight horizontal, from S and D were 0.996 m and 0.852 m. The lengths AB and AC were about 6415.8 m and 7129.2 m, respectively. Determine the angle BAC . [Ans. $63^\circ 33' 48.73''$]

- 2.34** In a triangle ABC , the station C was a church spire and hence a satellite station S was selected 12.10 m from C inside the triangle ABC . The angles CSA and ASB were observed to be $135^\circ 40' 30''$ and $71^\circ 29' 30''$ respectively. The distances AC and CB were measured as 2511 m and 1894 m, respectively. Compute the angle ABC . [Ans. $71^\circ 07' 54''$]

Objective-type Questions

- 2.1** The permissible average and maximum triangular closure in seconds in a second order triangulation is
 (a) $1''$ and $3''$ (b) $2''$ and $3''$ (c) $3''$ and $4''$ (d) $3''$ and $5''$
- 2.2** The best shape of a triangle in triangulation is
 (a) equilateral
 (b) isosceles with base angle $56^\circ 14'$
 (c) isosceles with base angle $65^\circ 14'$
 (d) isosceles with base angle 30°
- 2.3** The limiting strength of figure for first order and third order triangulations are
 (a) 25 and 30 (b) 25 and 50 (c) 30 and 25 (d) 50 and 25
- 2.4** How high should a helicopter pilot rise at a point A just to see the horizon at point B , if the distance AB is 40 km?
 (a) 101.75 m (b) 110.50 m (c) 107.75 m (d) 105.50 m
- 2.5** If two triangulation signals of 6.75 m height each, are to be just visible over the ground mutually, what is the maximum distance between their locations on the ground surface?
 (a) 10 km (b) 20 km (c) 30 km (d) 40 km
- 2.6** A sailor, standing on the deck of a ship, just sees the light beam from a lighthouse on this shore. If the height of the sailor's eye and of the light beam at the lighthouse, above the sea level, are 9 m and 25 m respectively, what is the distance between the sailor and the lighthouse?
 (a) 29.8 km (b) 31.1 km (c) 31.9 km (d) 33.2 km
- 2.7** If two ends of the tape are not at the same level and are inclined at an angle θ with horizontal, the sag correction will be (when pull is on higher end)
 (a) $C_s \cos^2 \theta$ (b) $C_s \cos^2 \theta \left(1 + \frac{Wl}{P} \sin \theta \right)$
 (c) $C_s \cos^2 \theta \left(1 - \frac{W}{P} \sin \theta \right)$ (d) $C_s \cos^2 \theta \left(1 - \frac{W}{Pl} \sin \theta \right)$
- 2.8** The distance angle in triangulation should not be small, since
 (a) a small angle cannot be observed accurately
 (b) extension of base net will not be possible

- (c) the error in computed side by using it will be large
- (d) the theodolites used are not that precise
- 2.9** Station marks are used in triangulation to provide
 - (a) a surface mark
 - (b) a good target point
 - (c) complete description to enable the station to be recovered even after years
 - (d) All of the above
- 2.10** In triangulation survey, the intervisibility of stations can be checked by
 - (a) method of least squares
 - (b) heliotropes
 - (c) Captain G.T. McCaw's solution
 - (d) three-point problem
- 2.11** The number of stations for a triangulation framework can be reduced, yet maintaining the required accuracy, by using a chain of
 - (a) triangles
 - (b) pentagons
 - (c) parallelograms
 - (d) polygons
- 2.12** In triangle ABC , the side opposite the angle C is known. It is required to compute the side opposite the angle B . The distance angle(s) will be
 - (a) $\angle A$ and $\angle B$
 - (b) $\angle B$ and $\angle C$
 - (c) $\angle C$ and $\angle A$
 - (d) $\angle A$ only
- 2.13** For a distance of 45 km in triangulation, the type of signal used is (are)
 - (a) heliotrope
 - (b) beacon
 - (c) heliograph
 - (d) all the above
- 2.14** The towers used in triangulation are known as
 - (a) heliotropes
 - (b) Bilby
 - (c) Captain McCaw
 - (d) Hunter
- 2.15** Most important factor in selecting a survey station is
 - (a) distance between the two stations
 - (b) intervisibility
 - (c) intervisibility and well-conditional triangle
 - (d) accessibility and distance
- 2.16** For the Survey of India, the instrument used for measuring bases was
 - (a) Colby apparatus
 - (b) Jaderin's apparatus
 - (c) Hunter's short bases
 - (d) Tellurometer
- 2.17** Pick out the correct statement.
 - (a) It is preferred to observe horizontal angles at noon to avoid refraction.
 - (b) For the angles with unequal weights, the correction is inversely proportional to the square of the weights.
 - (c) Figure adjustment in triangulation is best achieved by the method of least squares.
 - (d) Triangulation is a method of providing horizontal as well as vertical control.
- 2.18** Pick out the incorrect statement.
 - (a) Triangulation covers a large area in less time.
 - (b) Triangulation is most suitable for hilly areas.
 - (c) Triangulation is suited more than traversing for surveying long narrow belts.
 - (d) The only linear measurement made in triangulation is that of a base line.
- 2.19** Pick out the correct statement.
 - (a) A three-legged signal is preferred over a four-legged one.
 - (b) Heliotropes are luminous signals.

- (c) A beacon lamp signal is a type of electric lamp.
 (d) All the above are correct.
- 2.20** Pick out the correct statement.
- (a) Triangulation is most suited for flat areas.
 (b) In triangulation, the length of only one side is measured.
 (c) Small angles are essential in triangulation.
 (d) Limiting strength of figure for first and third order triangulations are 50 and 25, respectively.
- 2.21** Pick out the correct statement.
- (i) A base line may be used in triangulation to determine the scale of triangulation.
 (ii) Colby apparatus is a flexible type of apparatus for measuring bases.
 (iii) Bases can be measured electronically by a tellurometer.
 (a) only (i) is correct (b) only (ii) is correct
 (c) (i) and (ii) are correct (d) (i) and (iii) are correct
- 2.22** Pick out the correct statement.
- (a) Wheeler's method uses flexible method of base measurement
 (b) Jaderin's method uses radar technology
 (c) Tellurometer is most commonly used for third order triangulation
 (d) None of the above
- 2.23** Pick out the incorrect statement.
- (a) A base is extended to avoid ill-conditioned triangles.
 (b) In triangulation, satellite stations are the false stations.
 (c) Reduction to centre means reducing the observations to the centre of the earth.
 (d) The error in centring of instrument over a triangulation station can be reduced by selecting the stations far away.
- 2.24** Match List-I (Measurements) with List-II (Instruments) and select the correct answer using the codes given below the lists:
- | List-I (Measurement) | List-II (Instrument) |
|--|-----------------------------|
| (A) Displacement measurement from photograph | 1. Box sextant |
| (B) Electronic distance measurement | 2. Subtense bar |
| (C) Base line measurement | 3. Tellurometer |
| (D) Horizontal angle measurement | 4. Parallax bar |
- Codes:**
- | A | B | C | D |
|----------|----------|----------|----------|
| (a) 4 | 3 | 2 | 1 |
| (b) 1 | 3 | 2 | 4 |
| (c) 4 | 2 | 3 | 1 |
| (d) 1 | 2 | 3 | 4 |
- 2.25** Select the correct statement(s).
- (i) Triangulation is used for filling in the details.
 (ii) Triangulation is referred to as a system of multiplying ground control points.
 (iii) For a highway, a system of quadrilaterals is most suitable.
 (a) (i) is correct (b) (ii) is correct
 (c) (i) and (iii) are correct (d) (ii) and (iii) are correct
- 2.26** Select the correct statement(s).
- (i) In triangulation, angles less than 30° are not permitted.
 (ii) A quadrilateral figure is most suited for highway triangulation

- (iii) A triangle figure is most suited for railway triangulation.
 (a) (ii) is correct (b) (iii) is correct
 (c) (i) and (ii) are correct (d) (i) and (iii) are correct

2.27 Select the correct statement(s).

- (i) The value of spherical excess is 1" for an area of about 196 km².
 (ii) In triangulation, for the best results, the angle of a triangle which is not used in side computation should be as small as possible.
 (a) only (i) is correct
 (b) only (ii) is correct
 (c) both (i) and (ii) are correct
 (d) (ii) follows (i)

2.28 Match List-I (Phenomenon) with List-II (Method of survey) and select the correct answer using the code given below the list:

List-I (Phenomenon)

- (A) Crab and drift
 (B) Stadia intercept
 (C) Culmination and elongation
 (D) Baseline measurement

List-II (Method of survey)

1. Triangulation
 2. Astronomical survey
 3. Aerial
 4. Tacheometric survey

Codes:

	A	B	C	D
(a)	3	1	2	4
(b)	2	4	3	1
(c)	3	4	2	1
(d)	2	1	3	4

2.29 Match List-I (Triangulation Station) with List-II (Purpose) and select the correct answer using the code given below the lists;

List-I (Triangulation station)

- (A) Main stations
 (B) Subsidiary stations

 (C) Satellite stations

 (D) Pivot stations

List-II (Purpose)

1. Control points of triangulation network
 2. Points not for observation but for continuation of triangulation network
 3. Points to provide additional rays to intersected points
 4. Points close to main stations to avoid intervening obstructions

Codes:

	A	B	C	D
(a)	1	3	4	2
(b)	1	3	2	4
(c)	3	1	4	2
(d)	3	1	2	4

2.30 Match List-I with List-II and select the correct answer using the codes given below the list:

List-I

- (A) Satellite station
 (B) Gales traverse-method
 (C) Invar steel tape

 (D) Intervisibility of adjacent station

List-II

1. Measurement of Baseline
 2. Characteristics of triangulation station
 3. Auxiliary theodolite station near inaccessible main triangulation station.
 4. Computation of coordinates of traverse/triangulation stations.

Codes:

	A	B	C	D
(a)	3	1	4	2
(b)	2	1	4	3
(c)	3	4	1	2
(d)	2	4	1	3

Answers to Objective-type Questions

2.1 (d)	2.2 (b)	2.3 (b)	2.4 (c)	2.5 (b)	2.6 (d)
2.7 (b)	2.8 (c)	2.9 (d)	2.10 (c)	2.11 (c)	2.12 (b)
2.13 (d)	2.14 (b)	2.15 (c)	2.16 (a)	2.17 (c)	2.18 (c)
2.19 (d)	2.20 (b)	2.21 (d)	2.22 (d)	2.23 (c)	2.24 (a)
2.25 (b)	2.26 (d)	2.27 (c)	2.28 (c)	2.29 (a)	2.30 (c)

3

THEORY OF ERRORS AND TRIANGULATION ADJUSTMENT

Introduction

In surveying measurements, there are all possibilities for errors to creep in, irrespective of the precautions and precision exercised. This necessitates a proper control, assessment and distribution of errors. Error, mistake and discrepancy are few terms frequently encountered in surveying measurements. *True error* may be defined as the difference in measured and true value of a quantity. But, since the true value of a measured quantity is not known, true error of the measurement is never known. However, when a number of measurements are to fulfill a known condition, true error can be obtained. For example, the theoretical sum of exterior included angles in a traverse of N sides is $(2N + 4) \times 90^\circ$. Here, the error in the sum of observed angles is known, but the error in individual angles is not known.

The observed difference between two like measurements each of which may have an error, of a quantity, is termed as *discrepancy*. A discrepancy should not be regarded as an error. A small discrepancy between two measurements does not reflect that error is small. Whereas, a large discrepancy indicates a mistake in the observations.

Mistakes are the errors arising from carelessness, inexperience, poor judgment and confusion of the observer.

The basis of a sound survey procedure depends upon the accuracy of field observations, and for an intelligent analysis of field data it is necessary to apply theory of errors. Some important aspects of the study of errors are nature of error, its assessment, i.e., how it propagates, and its distribution. The scope of the chapter is a combination of reduction and adjustment of field measurements to provide most probable and consistent values of the field observations (data).

Nature of Error

Errors in a measurement may be positive if the measurement is too large, or negative if the same is too small as compared to its true value. Errors are classified as systematic errors and accidental errors.

Systematic errors These are the errors which occur from well-understood causes and can be reduced by adopting suitable methods. For example, the error due to sag of a tape supported at its ends can be calculated and subtracted from each measurement. However, the tape can be supported throughout its length at short intervals and the sag error may be reduced to a negligible quantity. It always has the same magnitude and sign so long as the conditions remain the same, and such an error is called constant *systematic error*. Whereas, if the conditions change, the magnitude of the error changes and is known as *variable systematic error*. A systematic error follows a definite mathematical or physical law and, therefore, a correction can always be determined and applied. These errors are also known as *cumulative errors*.

Accidental errors These are the errors due to a combination of causes and are beyond the control of the surveyor. These can be positive or negative. Erroneous calibration of a chain is an example of an accidental error.

There is in reality no fixed boundary between the accidental and systematic errors. Every accidental error has some cause, and if the causes were perfectly understood and the amount and sign could be determined, it would cause an accidental error, but would be classed as systematic. On the other hand, a constant or systematic error may be brought into the accidental class wholly or partially by varying the conditions, instruments, etc., such that the sign of the errors is frequently reserved.

3.1 DEFINITIONS

3.1.1 Observed Quantities

An observed quantity may be either an independent quantity or a conditioned quantity. An *independent quantity* is one, whose value is independent of the influence of the value of any other quantity. The examples include values of R.L. of a B.M. and length of base line. On the other hand, the value of a *conditioned quantity* depends on the value of one or more quantities. The angles of a polygon of N sides are conditioned quantities, since the theoretical sum of the angles should be $(2N \pm 4) \times 90^\circ$.

3.1.2 Value of a Quantity

The value of a quantity may be the true value, the observed value, or the most probable value. A *true value* is defined as the value absolutely free from all errors, whereas the *observed value* is obtained as a result of an observation, after applying corrections for all the known errors. When many observations are made for the same quantity, the *most probable value* is the value which is more likely to be the true value. The difference between the most probable value and the observed value of a quantity will be due to random errors, and is known as *residual error*.

3.1.3 Weight

The weight of an observation is measure of its relative trustworthiness. For a series of observations made under varying conditions (personal, instrumental or atmospheric), a uniform degree of reliability in the results would not be

expected. Such observations are called *unequal observations*. Obviously, to derive the ultimate value of the quantity measured, it is not enough in this case to simply take the arithmetic mean, but the quality of each measurement will have to be taken into account.

The quality of each measurement result is evaluated by means of a number known as *weight*. The more reliable the result, the higher its weight. The weight thus expresses the amount of confidence enjoyed by this result as compared to the other results. The weights are established either by estimation or in accordance with the measurement conditions. Since corresponding to the definite conditions of measurement there is a definite mean square error, it is most natural to establish the weight of the measurements with reference to the mean square error. The weights of the measurement results are assumed to be inversely proportional to the squares of their respective mean square error.

The following rules may be observed while assigning weights to the observations made:

1. The weight of an angle varies directly as the number of observations made on the angle.
2. For an angle repeated a great number of times, the weight is inversely proportional to the square of the probable error.
3. For a line of levels, the weight varies inversely as the length of the route.

Let w be the weight of the measurement, W the weight of the arithmetic mean of n such measurements, and m the mean square error. Then

$$\frac{W}{w} = \frac{m^2}{(m/\sqrt{n})^2} = n$$

Hence, the weight of an arithmetic mean is n times superior to that of the measurement results.

Numerical values assigned as the weights of a series of observations are simply ratios indicating the relative precision of the observations.

3.2 LAWS OF WEIGHT

When the reliability of different sets of measurements varies one to the other, then different considerations will have to be given to those sets; the method used is called *weighting* or *assigning weights*. Followings laws may be used to assess the weight of surveying measurements.

1. The weight of the arithmetic mean of observations of individual weight is equal to the number of observations.

From the equation $\frac{W}{w} = n$, if $w = 1$, then $W = n$

2. The weight of the weighted arithmetic mean of observations of unit weight is equal to the sum of the individual weights.

Let angle A be measured three times by the repetition method

$\angle A$	Weight
$48^\circ 20' 8''$	3
$48^\circ 20' 10''$	2
$48^\circ 20' 12''$	3

Sum of the weights = $3 + 2 + 3 = 8$

Weighted arithmetic mean

$$= 48^\circ 20' + \frac{1}{8} \times (8'' \times 3 + 10'' \times 2 + 12'' \times 3)$$

$$= 48^\circ 20' 10''$$

Weight of the weighted arithmetic mean $48^\circ 20' 10''$ is 8.

3. If two or more quantities are added algebraically, the weight of the result is equal to the reciprocal of the sum of the reciprocals of the individual weights.

For example,

Angle	Weight
$\alpha = 46^\circ 30' 30''$	3
$\beta = 24^\circ 42' 20''$	2

$$\text{Weight of } \alpha + \beta (71^\circ 12' 50'') = \frac{1}{\frac{1}{3} + \frac{1}{2}} = \frac{6}{5}$$

$$\text{Weight of } \alpha - \beta (21^\circ 28' 10'') = \frac{1}{\frac{1}{3} + \frac{1}{2}} = \frac{6}{5}$$

4. If a quantity is multiplied by a factor, the weight of the product is equal to the weight of that quantity divided by the square of the factor.

For example, if angle $\alpha = 32^\circ 12' 12''$ and weight = 2, then

$$\text{weight of } 3\alpha (96^\circ 36' 36'') = \frac{2}{3^2} = \frac{2}{9}$$

5. If a quantity is divided by a factor, the weight of the result equals the weight of that quantity multiplied by the square of that factor.

For example, if angle $\alpha = 41^\circ 31' 15''$ and weight = 2, then

$$\text{weight of } \frac{\alpha}{3} (13^\circ 50' 45'') = 2 \times 3^2 = 18$$

6. If an equation is multiplied by its own weight, the weight of the resulting equation is equal to the reciprocal of the weight of the equation.

For example, let angle $(A + B) = 189^\circ 40' 20''$ and weight of angle $(A + B) = 4/9$, then

$$\text{weight of } \frac{4}{9} (A + B), \text{ i.e., of } 84^\circ 17' 55.55'' = \frac{9}{4}$$

7. If an equation is added to or subtracted from a constant or if all the signs of the equation are changed, the weight remains unchanged.

For example, if $\angle A + \angle B = 78^\circ 45' 22''$ and its weight = 3 then, weight of $180^\circ - (\angle A + \angle B)$, i.e., of $101^\circ 14' 38'' = 3$

Example 3.1 An angle A was observed by the method of repetition by three different observers and the values were as below.

Observer No.	$\angle A$	Number of measurements
1	$40^\circ 20'$	2
2	$40^\circ 30'$	3
3	$40^\circ 45'$	4

Find the most probable value of the angle.

Solution

The most probable value of the angle is equal to its weighted arithmetic mean.

$$40^\circ 20' \times 2 = 80^\circ 40'$$

$$40^\circ 30' \times 3 = 121^\circ 30'$$

$$40^\circ 45' \times 4 = 163^\circ 00'$$

$$\text{Sum} = 365^\circ 10'$$

$$\Sigma \text{weight} = 2 + 3 + 4 = 9$$

$$\text{weighted arithmetic mean} = \frac{365^\circ 10'}{9} = 40^\circ 34' 26.67''$$

The most probable value of angle A is $40^\circ 34' 26.67''$.

3.3 LAWS OF ACCIDENTAL ERRORS

A close examination of a set of observations reveals that the accidental errors follow the law of probability and it is because of this accidental errors are also known as *probable errors* or *standard errors*. This law defines the occurrence of errors and when expressed in the form of an equation, it can be used to compute the probable value or the probable precision of a quantity. In a series of observations, the probable error is a value such that the number of errors numerically greater than it is the same as that those less than it. It is always written after the observed quantity with the plus and minus sign, e.g., $25^\circ 42' 30'' \pm 3.16''$. Probable error serves two important purposes: (i) as a measure of the precision of any series of observations, and (ii) as a means of assigning weights to two or more quantities, and thus to find the weighted mean or the most probable value of each.

It should be noted that the probable error is never applied as a correction to the mean value of a series. This is evidently impossible because the probable error carries the double sign, plus or minus. The mean value still remains the most probable value of the quantity being measured, and the only way in which its precision can be increased is (i) by increasing the number of observations in the field, or (ii) by more refined and careful field methods.

For understanding the laws of accidental errors, it is important to discuss the concept underlying the theory of errors. In surveying to establish the value of a

particular measurement, it is usual to take number of measurements. The result will be grouped around the true value. When all systematic errors and mistakes have been removed from the measurements made, the residual between the true value and the measurements made will be due to *random errors*; random errors will have an equal chance of being plus or minus. These errors when plotted take the shape of the curve shown in Fig. 3.1, which is called *probability curve*. This curve forms the basis for the mathematical derivation of the theory of errors. Geometrically the area under the probability curve is one, and the probability that a residual (error) falls within certain limits is equal to the area under the curve between those limits. For the laws of probability to be applicable, it is necessary that a large number of observations have been made, which generally in surveying is not possible. But, if the results when plotted take the form of a probability curve, the errors associated with the measurements made can be treated using random error distribution techniques.

The probability curve of the form shown in Fig. 3.1 has equation of the form,

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-v^2/2\sigma^2}$$

where y is the frequency of residual error, v is the size of the residual, and σ is the standard deviation; a measure of precision.

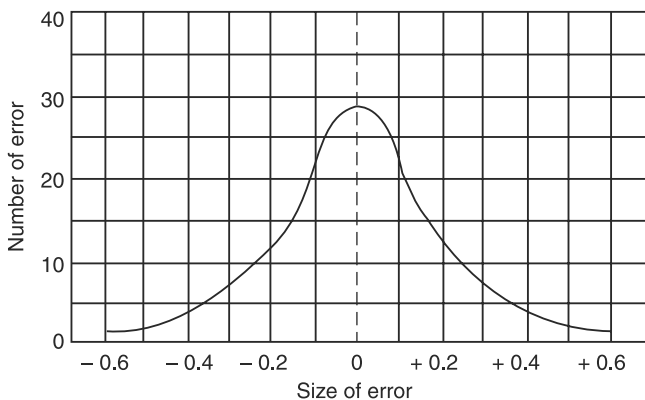


Fig. 3.1 Probability curve

Notes

1. In surveying true value of a measurement is never known. However, for the purpose of calculating errors, the arithmetic mean is taken to be the true value, or most probable value.

2. Since in surveying it is not possible to have very large (infinite) numbers of repetitive measurements, the arithmetic mean itself contains an error.

Mean is generally given by

$$\bar{x} = \frac{\sum x}{n}$$

$\sum x$ is the sum of the individual measurement values, and n is the number of individual measurements.

In statistics, precision is measured by the standard deviation, also called standard error, or root mean square error given by

$$\sigma = \pm \sqrt{\frac{\sum v^2}{n}}$$

where σ is the standard deviation of a very large sample, v is the true residual, and n is the large sample size.

In surveying, and otherwise also, the standard error of a set of repetitive measurements is given by equation

$$E_s = \sqrt{\frac{\sum v^2}{n-1}}$$

where v is the error ($x - \bar{x}$) and n is the number of repetitions. Since the use of $\bar{x}$ instead of the true value always results in an underestimation of the standard deviation, in the denominator ($n - 1$) is used instead of n , with an increase in number of repetitions the difference between n and ($n - 1$) becomes less significant.

From the foregoing discussion, it is apparent that arithmetic mean of repetitive measurements contains some uncertainty. The uncertainty can be expressed as the standard or probable error of the mean and is given by

$$E_m = \frac{E_s}{\sqrt{n}}$$

Figure 3.2 shows a probability curve with the confidence level the surveyor wishes to enjoy. In statistics it is known as 68-95-99.7 rule. It states that for a normal distribution, nearly all values lie within the three standard deviations of the mean. About 68.27%, or 95.45%, or 99.73% of the values are supposed to lie within the three standard deviations of the mean within the limits of the area under the curve, respectively. The area under the probability curve is directly proportional to the probability of expected result. Therefore, probability of measurements deviating from the mean for about 68.27%, 95.45%, and 99.73% will be $\bar{x} \pm \sigma$, $\bar{x} \pm 2\sigma$, and $\bar{x} \pm 3\sigma$, respectively. In theory of errors, however, the probable errors are associated with 50% confidence level and the limit is $\pm 0.6745E_s$.

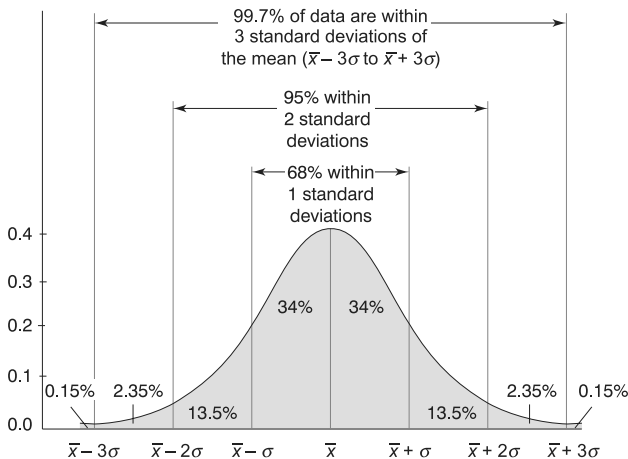


Fig. 3.2 68-95-99.7 rule in statistics

3.3.1 Direct Observations of Equal Weight

Probable error of a single observation,

$$E_s = 0.6745 \sqrt{\frac{\sum v^2}{n-1}}$$

also

$$E_s = 0.845 \bar{v} \quad (\text{approx.})$$

Probable error of an average (arithmetic mean),

$$E_m = \frac{E_s}{\sqrt{n}}$$

Probable error of a sum of measurements,

$$E_{sm} = \sqrt{E_1^2 + E_2^2 + E_3^2 + \dots + E_n^2}$$

where

n = number of observations in the series

v = difference between any single observation and the mean of the series

$E_1, E_2, \dots, E_n$ = probable errors of several observations which sum up in a measurement

3.3.2 Direct Observations of Unequal Weight

Probable error of a single observation of unit weight,

$$E_{su} = 0.6745 \sqrt{\frac{\sum (wv^2)}{n-1}}$$

Probable error of an observation of weight w ,

$$E_{suw} = \frac{E_{su}}{\sqrt{w}}$$

Probable error of weighted arithmetic mean,

$$E_{smw} = \frac{E_s}{\sqrt{\sum w}}$$

where w = weight of an observation

3.3.3 Indirect Observations on Independent Quantities

Probable error of an observation of unit weight,

$$E_{si} = 0.6745 \sqrt{\frac{\sum (wv^2)}{n-q}}$$

Probable error of an observation of weight w ,

$$E_{siw} = \frac{E_{si}}{\sqrt{w}}$$

where q = number of unknown quantities.

3.3.4 Indirect Observations Involving Conditional Equations

Probable error of an observation of unit weight,

$$E_{\text{sic}} = 0.6745 \sqrt{\frac{\Sigma(wv^2)}{n+p-q}}$$

The probable error of an observation of weight w ,

$$E_{\text{sicw}} = \frac{E_{\text{sic}}}{\sqrt{w}}$$

where p = number of conditional equations.

3.3.5 Computed Quantities

In surveying, measurements such as distances and angles are often used to derive other quantities using some mathematical relationships. Such measurements will be random variables and will have errors. Thus, the quantities derived from them will also have errors.

The law of propagation of standard errors for a quantity x is as below.

Let the computed quantity x be a function of $a, b, c, \dots$ measurements.

$$x = f(a, b, c, \dots)$$

$$e_x^2 = \left(\frac{\partial x}{\partial a}\right)^2 e_a^2 + \left(\frac{\partial x}{\partial b}\right)^2 e_b^2 + \dots$$

where

$e_{x_1}^2, e_{x_2}^2, \dots$ are the probable errors in the measurements $a, b, c, \dots$

Let x be a computed quantity with e_x as its probable error, and let a be an observed quantity with e_a as its probable error and k be a constant.

Case 1 The computed quantity is the sum of the difference of an observed quantity and a constant.

$$x = \pm a \pm k \quad \text{and} \quad e_x = e_a$$

Case 2 The computed quantity is the product of an observed quantity and a constant.

$$x = ka \quad \text{and} \quad e_x = ke_a$$

Case 3 The computed quantity is the algebraic sum of two or more independent observed quantities.

Let $a, b, c, \dots$ be the indirectly observed quantities and $e_a, e_b, e_c, \dots$ be their respective probable errors.

$$x = \pm a \pm b \pm c \pm \dots$$

$$e_x^2 = \left(\frac{\partial x}{\partial a}\right)^2 e_a^2 + \left(\frac{\partial x}{\partial b}\right)^2 e_b^2 + \left(\frac{\partial x}{\partial c}\right)^2 e_c^2 + \dots$$

$$e_x^2 = 1 \cdot e_a^2 + 1 \cdot e_b^2 + 1 \cdot e_c^2 + \dots$$

$$e_x = \sqrt{e_a^2 + e_b^2 + e_c^2 + \dots}$$

Case 4 The computed quantity is a function of a single observed quantity.

$$x = \phi(a) \quad \text{and} \quad e_x = \left(\frac{\partial x}{\partial a} \right) e_a$$

The computed quantity is a function of more than one observed quantity.

$$x = a \cdot b$$

$$e_x^2 = \left(\frac{\partial x}{\partial a} \right)^2 e_a^2 + \left(\frac{\partial x}{\partial b} \right)^2 e_b^2$$

Example 3.2 Show that for any reading taken to the nearest whole number, its magnitude is $\pm 1/4$.

Solution The probable error of a measured quantity has a value such that the true error is equally likely either to be greater than or to be less than the probable error.

When a reading is taken to the nearest whole number, the implication is that the true reading lies within the limits $\pm 1/2$ unit. Therefore, by definition, the probable error will be $\pm 1/4$ unit.

Example 3.3 Assess the errors that may be expected in ordinary angular measurement with vernier theodolites having a least count of $20''$.

Solution When a theodolite in perfect adjustment is accurately levelled, the main sources of accidental error will be

- (a) Reading the vernier or microscopes, and
- (b) Bisecting the signal or the point sighted.

(a) If the least count of the vernier be x , the maximum possible error of a reading is $(1/2)x$; but there is also the likelihood of the value being anything between $0''$ and $(1/2)x$ and $(1/2)x$ to x . Consequently, the probable error (p.e.) will be $\pm (1/4)x$ for a single vernier which is $\pm 5''$ for a least count of $20''$. An allowance of $1''$ may be made for imperfect division.

Hence,

- (i) Probable error of single vernier reading = $\pm \alpha = \pm 6''$ ($x = 20''$).
- (ii) Probable error when both verniers are read = $\pm \alpha / \sqrt{2} = \pm 6'' / \sqrt{2}$ ($x = 20''$).
- (iii) Probable error when both verniers are read with both the faces of the theodolites = $\pm \alpha / 2 = \pm 6'' / 2 = 3''$.

But in any angular measurement, the reading of an angle can be had for two sights or pointings.

Hence, probable error per angle = (i) $\pm \sqrt{2} \alpha$

$$(ii) \sqrt{2} \times \left(\pm \frac{\alpha}{\sqrt{2}} \right) = \pm \alpha$$

$$(iii) \sqrt{2} \times \left(\pm \frac{\alpha}{2} \right) = \pm \alpha / \sqrt{2}$$

(b) The error of sighting a signal occurs once for each sight in (i) and (ii) and twice for each sight in (iii) that is the error in (i) and (ii) will be $\sqrt{2} \beta$ and in (iii) $\sqrt{2} \beta / \sqrt{2} = \beta$.

The error β will vary with the nature of the signal, distance, atmospheric conditions, and the power and quality of the telescope. It may vary from 2" to 10" in ordinary work, and may be taken to be 6" on the average with a 15 cm vernier theodolite. In addition, there is an error of centring the theodolite over the station. Though it is negligible in triangulation, it can be serious when traversing with short lines.

The probable error from both sources will be the square root of the sum of the squares of the errors from the individual sources:

$$(i) \sqrt{(\sqrt{2}\alpha)^2 + (\sqrt{2}\beta)^2} = \sqrt{2\alpha^2 + 2\beta^2} = \sqrt{2 \times 6^2 + 2 \times 6^2} = 12''$$

$$(ii) \sqrt{(\alpha)^2 + (\sqrt{2}\beta)^2} = \sqrt{\alpha^2 + 2\beta^2} = \sqrt{6^2 + 2 \times 6^2} = 10.4''$$

$$(iii) \sqrt{(\alpha/\sqrt{2})^2 + (\beta)^2} = \sqrt{\alpha^2/2 + \beta^2} = \sqrt{6^2/2 + 6^2} = 7.3''$$

Example 3.4 The following observations were recorded for an angle under identical conditions:

162°20'00"	162°21'20"	162°21'40"
162°20'40"	162°19'40"	162°21'20"

Calculate (a) the probable error of a single observation, (b) the probable error of the mean, and (c) the most probable value of the angles.

Solution

Value	Mean	v	v^2
162°20'00"		-46.67"	36'18.08"
162°21'20"		+33.34"	18'31.56"
162°21'40"	162°20'46.66"	+53.33"	47'24.08"
162°20'40"		-6.67"	44.49"
162°19'40"		-1'6.67"	74'4.88"
162°21'20"		+33.34"	18'31.56"
		$\Sigma v = 0$	$\Sigma v^2 = 3^\circ 15' 34.57''$

$$E_s = \pm 0.6745 \times \sqrt{3^\circ 15' 34.57'' / (6-1)} = 32.676''$$

$$\text{Probable error of the mean } E_m = E_s / \sqrt{n} = 32.676' / \sqrt{6} = 13.34'$$

Most probable value of the angle = 162°20'46.66" $\pm$ 13.34' = 162°34' 7.06" or 162°7'26.26".

Example 3.5 The most probable value of an angle A was found to be 44°32'40" $\pm$ 2".

Calculate the most probable value of its supplement.

Solution We know, $A + A' = 180^\circ$

$$\text{Hence, } A' = 180^\circ - 44^\circ 32' 40'' = 135^\circ 27' 20''$$

Since the probable error of $A' =$ the probable error of A ,
the most probable value of $A' = 135^\circ 27' 20'' \pm 2''$

Example 3.6 The most probable values of two angles A and B are as given below:

$$\angle A = 47^\circ 32' 30'' \pm 2''$$

$$\angle B = 24^\circ 16' 40'' \pm 3''$$

Find the probable error of the angle $(A + B)$.

Solution Let sum of the two angles be C .

$$\angle C = \angle A + \angle B = 47^\circ 32' 30'' + 24^\circ 16' 40'' = 71^\circ 49' 10''$$

$$e_C = \sqrt{e_A^2 + e_B^2} = \sqrt{2^2 + 3^2} = \pm 3.605''$$

Hence,

$$\angle C = 71^\circ 49' 10'' \pm 3.605''$$

Example 3.7 Calculate the most probable value and the most probable error of the area of a circle of radius 12.25 ± 0.03 m.

Solution Area, $A = \pi r^2 = \pi \times (12.25)^2 = 471.4353 \text{ m}^2$

$$\frac{\partial A}{\partial r} = 2\pi r$$

$$e_r = \pm 0.03$$

$$e_A = e_r \times \frac{\partial A}{\partial r} = \pm 0.03 \times 2 \times \pi \times 12.25 = \pm 2.309$$

Hence, the most probable value of the area = $471.4353 \pm 2.309 \text{ m}^2$

Example 3.8 The radius of a circle is 21 m with a probable error ± 0.21 . What will be the probable error of its circumference?

Solution Circumference, $C = 2\pi r$
 $= 2 \times \pi \times 21 = 132 \text{ m}$
 $e_C = 2\pi e_r$
 $= 2 \times \pi \times 0.21 = \pm 1.32$

The probable error of the circumference is ± 1.32

Alternatively

$$c = 2\pi r$$

$$\frac{\partial c}{\partial r} = 2\pi$$

$$e_c = e_r \times \frac{\partial c}{\partial r}$$

$$= 0.21 \times 2\pi = \pm 1.32$$

Example 3.9 Following observations were recorded for horizontal angles of a triangle:

$$\angle A = 20^\circ 10' \pm 0.2$$

$$\angle B = 100^\circ 40' \pm 0.1$$

$$\angle C = 59^\circ 10' \pm 0.2$$

Determine the probable error of its summation.

Solution Sum of angles, $S = \angle A + \angle B + \angle C$

Let e_A , e_B , and e_C be the probable errors of angles A , B , and C .

$$\begin{aligned}\text{Then, } e_S &= \sqrt{e_A^2 + e_B^2 + e_C^2} \\ &= \sqrt{0.2^2 + 0.1^2 + 0.2^2} = \pm 0.3\end{aligned}$$

The probable error of summation is ± 0.3

Example 3.10 Calculate the most probable value and the probable error of the area of a rectangle whose sides are as follows.

Side $a = 100 \pm 0.02$ m side $b = 150 \pm 0.01$ m

Solution Area, $A = a \times b = 100 \times 150 = 15\,000 \text{ m}^2$

$$\frac{\partial A}{\partial a} = b = 150$$

$$\frac{\partial A}{\partial b} = a = 100$$

$$e_a = \pm 0.02, e_b = \pm 0.01$$

$$\begin{aligned}e_A &= \sqrt{\left(e_a \frac{\partial A}{\partial a}\right)^2 + \left(e_b \frac{\partial A}{\partial b}\right)^2} \\ &= \sqrt{(0.02 \times 150)^2 + (0.01 \times 100)^2} = 3.162\end{aligned}$$

The most probable value of the area $15\,000 \pm 3.162 \text{ m}^2$

Example 3.11 A base line was measured under certain conditions and was found to be 250 ± 0.04 m. What will be the probable error of another base line, 1000 m long, measured under the same conditions?

Solution $\frac{\text{probable error of first line}}{\text{probable error of second line}} = \frac{\sqrt{\text{length of first line}}}{\sqrt{\text{length of second line}}}$

$$\text{Probable error of second line} = 0.04 \times \sqrt{\frac{1000}{250}} = \pm 0.08 \text{ m}$$

Example 3.12 The internal angles of a n -sided polygon were all measured using the same equipment and methods such that each angle had the same probable error. What is the probable error for the sum of these angles?

Solution Let $A_1, A_2, \dots, A_n$ be the interior angles of a n -sided polygon and $e_{A1}, e_{A2}, \dots, e_{An}$ be their probable errors.

$$e_{A1} = e_{A2} = \dots = e_{An} = e$$

$$\text{now, } e_A^2 = e_{A1}^2 + e_{A2}^2 + \dots + e_{An}^2$$

$$e_A^2 = ne^2$$

$$\text{or } e_A = \pm \sqrt{n} \cdot e$$

Note The propagated error for the sum of a set of measurements with the same probable error is proportional to the square root of the number of measurements.

Example 3.13 Determine the volume of a box whose sides are:

$$3 \pm 0.02 \text{ m}, 4 \pm 0.01 \text{ m}, 3 \pm 0.02 \text{ m}$$

Solution Volume, $v = l \times b \times h = 3 \times 4 \times 3 = 36 \text{ m}^3$

$$\frac{\partial v}{\partial l} = b \times h = 4 \times 3 = 12$$

$$\frac{\partial v}{\partial b} = l \times h = 3 \times 3 = 9$$

$$\frac{\partial v}{\partial h} = l \times b = 3 \times 4 = 12$$

Now,

$$\begin{aligned} e_v &= \sqrt{\left(e_l \frac{\partial v}{\partial l}\right)^2 + \left(e_b \frac{\partial v}{\partial b}\right)^2 + \left(e_h \frac{\partial v}{\partial h}\right)^2} \\ &= \sqrt{(0.02 \times 12)^2 + (0.01 \times 9)^2 + (0.02 \times 12)^2} \\ &= \pm 0.351 \approx \pm 0.35 \end{aligned}$$

The probable error of volume is ± 0.35 .

Example 3.14 The height of a transit station has been computed from the measurement of a slope distance, $L = 279.1 \pm 0.06 \text{ m}$ and of a vertical angle, $\alpha = 3^\circ 20' \pm 30''$. What is the probable error in the calculated elevation of station A?

Solution Height of transit station, $H = L \sin \alpha$

$$\frac{\partial H}{\partial L} = \sin \alpha = \sin 3^\circ 20'$$

$$\frac{\partial H}{\partial \alpha} = L \cos \alpha = 279.1 \times 0.9983 = 278.6255$$

$$\begin{aligned} e_H &= \sqrt{\left(e_L \frac{\partial H}{\partial L}\right)^2 + \left(e_\alpha \frac{\partial H}{\partial \alpha}\right)^2} \\ &= \sqrt{(0.6 \times \sin 3^\circ 20')^2 + (0.000148 \times 278.6255)^2} \\ &= \pm 0.0540 \text{ m} \end{aligned}$$

3.4 PRINCIPLE OF LEAST SQUARES

The principle of distributing errors by the method of least squares is of great help to find the most probable value of a quantity which has been measured for several times, perhaps by different methods and different observers and in calculating the trust worthiness of such a value. In the method of least squares, the discrepancies or errors of the discrepant observations are assumed to be of accidental nature only. According to this principle, the most probable value of a quantity is the one for which the sum of the squares of the errors is a minimum.

The equation of the probability curve is

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{v^2}{2\sigma^2}} \frac{he^{-h^2v^2}}{\sqrt{\pi}} \quad \left(\text{where } h^2 = \frac{1}{2\sigma^2} \right)$$

where h is the index of precision and y the probability of occurrence of an error.

Let, $1/\sqrt{\pi} = C$, then

$$y = Che^{-h^2v^2}$$

Differentiating the above equation with respect to h and putting it equal to zero for maximum y gives

$$v^2 = \frac{1}{2h^2}$$

Considering errors

$$v_1^2 + v_2^2 + \dots + v_n^2 = \frac{1}{2h_1^2} + \frac{1}{2h_2^2} + \dots + \frac{1}{2h_n^2}$$

Then,
$$\Sigma v^2 = \Sigma \frac{1}{2h^2}$$

As h increases, $1/2h^2$ decreases, thus maximum accuracy is achieved when

$$v^2 = 1/(2h^2) \text{ is minimum}$$

The main objects of the method of least squares are (i) to determine the best values which can possibly be obtained from a given set of measurements, (ii) to determine the degree of dependence which can be placed upon these values, (iii) to enable us to trace to their sources the various errors affecting the measurements, and consequently, (iv) to increase the accuracy of the result by a proper modification of the methods and the instruments used.

It should be kept in mind that the method of least squares deals with accidental errors only and can tell us nothing about the constant or systematic errors which may affect the results of observations. A computation of the most probable value improves the result with respect to the accidental errors, but leaves the more serious form of error untouched. Multiplying observations and adjusting them for the purpose of removing small accidental errors, and at the same time failing to remove a large constant error may defeat the very purpose of the exercise.

Though the method of least squares may not show directly the presence of constant errors, yet a study of the precision of results, and a knowledge of the law governing the behaviour of accidental errors, may enable the observer to at least decide whether it is probably present. Consequently, the observer can modify his methods of observing so as to reduce the effect of such errors. Variations in the result, which are greater than the error of observation shown by precision measures, is likely to mean that systematic error is present. This tracing of errors to their sources, and the consequent modification of instruments and methods, may constitute the most important application of the method of least squares.

Notes

1. Smaller the sum of squared errors $\Sigma(x - \bar{x})^2$ the more precise will be the measurements. For assigning weights to such measurements, it is logical to weight each error $(x - \bar{x})$ such that sum of their squares $(x - \bar{x})^2$ is a minimum, which is the principle of least squares. $\Sigma w(x - \bar{x})^2$ is a minimum.
2. For least square adjustment to be valid, a reliable estimate of the E_s of the various measuring techniques must be available to identify the individual weights properly.

Example 3.15 The following observations were made in the rating tests of a current meter:

Number of revolution (N)	Velocity of flow (m/s)
0.3	0.50
1.0	1.50
1.5	2.30
2.1	2.78

The rating formula is $V = aN + b$

Find the constants a and b using the theory of least squares.

Solution The observation equations are

$$0.50 = 0.3 a + b$$

$$1.50 = 1.0 a + b$$

$$2.30 = 1.5 a + b$$

$$2.78 = 2.1 a + b$$

If a and b are the most probable values, then the error of observations are

$$0.50 - 0.3 a - b$$

$$1.50 - 1.0 a - b$$

$$2.30 - 1.5 a - b$$

$$2.78 - 2.1 a - b$$

By the theory of least squares

$$(0.5 - 0.3 a - b)^2 + (1.5 - 1.0 a - b)^2 + (2.3 - 1.5 a - b)^2 + (2.78 - 2.1 a - b)^2 = \text{minimum} \quad (\text{i})$$

For normal equation in a , differentiating Eq. (i) w.r.t a , we get

$$2(-0.3)(0.5 - 0.3 a - b) + 2(-1.0)(1.5 - 1.0 a - b) + 2(-1.5)(2.3 - 1.5 a - b) + 2(-2.1)(2.78 - 2.1 a - b) = 0$$

$$\text{or } 10.938 - 7.75 a - 4.9 b = 0 \quad (\text{ii})$$

Now, differentiating Eq. (i) w.r.t b to get normal equation in b , we get

$$2(-1)(0.5 - 0.3 a - b) + 2(-1)(1.5 - 1.0 a - b) + 2(-1)(2.3 - 1.5 a - b) + 2(-1)(2.78 - 2.1 a - b) = 0$$

$$\text{or } 7.08 - 4.9 a - 4b = 0 \quad (\text{iii})$$

Equations (ii) and (iii) are simultaneous. These may be solved for a and b .

$$a = 1.296, \quad b = 0.182$$

Hence, $V = 1.296 N + 0.182$

Example 3.16 The telescope of a theodolite is fitted with stadia wires. It is required to find the most probable values of the constants C and K of tacheometer. The staff was kept vertical at three points in the field and with line of sight horizontal the staff intercepts observed were as follows:

Distance of staff from tacheometer D (m)	Staff intercept S (m)
150	1.495
200	2.000
250	2.505

Solution The distance equation is

$$D = KS + C$$

The observation equations are

$$150 = 1.495 K + C$$

$$200 = 2.000 K + C$$

$$250 = 2.505 K + C$$

If K and C are the most probable values, then the error of observations are:

$$150 - 1.495 K - C$$

$$200 - 2.000 K - C$$

$$250 - 2.505 K - C$$

By the theory of least squares

$$(150 - 1.495 K - C)^2 + (200 - 2.000 K - C)^2 + (250 - 2.505 K - C)^2 = \text{minimum} \quad (\text{i})$$

For normal equation in K , differentiating Eq. (i) w.r.t. K , we get

$$2(-1.495)(150 - 1.495 K - C) + 2(-2.0)(200 - 2.00 K - C) + 2(-2.505)(250 - 2.505 K - C) = 0$$

$$\text{or} \quad 1250.50 - 12.51 K - 6C = 0$$

$$\text{or} \quad 208.41667 - 2.085 K - C = 0 \quad (\text{ii})$$

For normal equation in C , differentiating Eq. (i) w.r.t. C , we get

$$2(-1.0)(150 - 1.495 K - C) + 2(-1.0)(200 - 2.0 K - C) + 2(-1.0)(250 - 2.505 K - C) = 0$$

$$600 - 6K - 3C = 0$$

$$200 - 2K - C = 0 \quad (\text{iii})$$

Equations (ii) and (iii) are simultaneous. These may be solved for K and C .

$$K = 99.0196$$

$$C = 1.9608$$

$\therefore$ the distance equation is

$$D = 99.0196 S + 1.9608$$

Ans.

Example 3.17 The following levels along the centre line of a road were obtained at chainages of 100 m intervals under difficult conditions:

Chainage (m)	200	300	400	500	600
Level (m)	102.25	105.29	108.20	111.28	114.22

Assuming the grade to be a straight line, compute the most probable values of the level at chainage 100 m, and of the gradient.

Solution Level at any point will be

$$h = aC + b \quad (i)$$

where C = chainage of the point

a, b = constants

The observation equations are given by

$$102.25 = 200 a + b$$

$$105.29 = 300 a + b$$

$$108.20 = 400 a + b$$

$$111.28 = 500 a + b$$

$$114.22 = 600 a + b$$

If a and b are the most probable values, then the error of observations are

$$102.25 - 200 a - b$$

$$105.29 - 300 a - b$$

$$108.20 - 400 a - b$$

$$111.28 - 500 a - b$$

$$114.22 - 600 a - b$$

By the theory of least squares

$$(102.25 - 200 a - b)^2 + (105.29 - 300 a - b)^2 + (108.20 - 400 a - b)^2 + (111.28 - 500 a - b)^2 + (114.22 - 600 a - b)^2 = \text{minimum} \quad (ii)$$

For normal equation in a , differentiating Eq. (ii) w.r.t a , we get

$$2(-200)(102.25 - 200 a - b) + 2(-300)(105.29 - 300 a - b) + 2(-400)(108.20 - 400 a - b) + 2(-500)(111.28 - 500 a - b) + 2(-600)(114.22 - 600 a - b) = 0$$

$$\text{or } 900\,000 a + 2000 b - 219\,489 = 0 \quad (iii)$$

Now differentiating Eq. (ii) w.r.t b to get normal equation in b ,

$$2(-1)(102.25 - 200 a - b) + 2(-1)(105.29 - 300 a - b) + 2(-1)(108.20 - 400 a - b) + 2(-1)(111.28 - 500 a - b) + 2(-1)(114.22 - 600 a - b) = 0$$

$$\text{or } 2000 a + 5 b - 541.24 = 0 \quad (iv)$$

Equations (iii) and (iv) are simultaneous. These may be solved for a and b .

$$a = 0.02993 \text{ m} \quad \text{and} \quad b = 96.276 \text{ m}$$

Equation (i) can be rewritten as

$$h = 0.02993 C + 96.276$$

Now, the most probable level at chainage 100 m is

$$= 0.02993 \times 100 + 96.276 = 99.27 \text{ m}$$

$$\text{The most probable grade} = \frac{2.993}{100} = 1 \text{ in } 33.41$$

3.5 DISTRIBUTION OF ERROR TO THE FIELD MEASUREMENT

It is customary to check the observations made in the field for the closing error, if any. The closing error is then distributed to the various observed quantities as per the following rules.

1. The correction to be applied to an observation is inversely proportional to the weight of the observation.
2. The correction to be applied to an observation is directly proportional to the square of the probable error.
3. The correction to be applied to the line of levels is proportional to the length of the route.

Example 3.18 On a station O , three horizontal angles, closing the horizon, were measured as follows:

$$\angle A = 34^\circ 10' 20'' \pm 3'', \angle B = 176^\circ 40' 32'' \pm 4'', \angle C = 149^\circ 09' 04'' \pm 5''$$

Calculate the corrected angles.

Solution Sum of angles $= 34^\circ 10' 20'' + 176^\circ 40' 32'' + 149^\circ 09' 04''$
 $= 359^\circ 59' 56''$

$$\text{Discrepancy} = 360^\circ - 359^\circ 59' 56'' = 04''$$

This error of $04''$ is to be distributed in proportion to the square of the probable error in each angle. Let c_1 , c_2 and c_3 be the corrections to be applied to the angles A , B and C , respectively.

$$c_1:c_2:c_3 = (3)^2:(4)^2:(5)^2 = 9:16:25$$

$$\text{Hence, } c_2 = (16/9) c_1 \text{ and } c_3 = (25/9) c_1$$

$$\text{Also, } c_1 + c_2 + c_3 = 4''$$

$$\text{or } c_1 + (16/9) c_1 + (25/9) c_1 = 4''$$

$$\text{or } (50/9) c_1 = 4''$$

$$\text{or } c_1 = (4 \times 9)/50 = 0.72''$$

$$\therefore c_2 = (16/9) \times 0.72 = 1.28''$$

$$\text{and } c_3 = (25/9) \times 0.72 = 2.00''$$

Check,

$$c_1 + c_2 + c_3 = 0.72'' + 1.28'' + 2.00'' = 4''$$

Hence, the corrected angles are:

$$\angle A = 34^\circ 10' 20'' + 0.72'' = 34^\circ 10' 20.72''$$

$$\angle B = 176^\circ 40' 32'' + 1.28'' = 176^\circ 40' 33.28''$$

$$\angle C = 149^\circ 09' 04'' + 2.00'' = 149^\circ 09' 06''$$

Example 3.19 Neglecting spherical excess, adjust the angles of a triangle of which the observed values are:

$$\angle A \quad 48^\circ 18' 22'' \quad \text{Weight} = 3$$

$$\angle B \quad 76^\circ 32' 47.2'' \quad \text{Weight} = 1$$

$$\angle C \quad 55^\circ 08' 53.8'' \quad \text{Weight} = 3$$

Solution The sum of the angles of a triangle should be 180° .

$$\begin{aligned} \angle A + \angle B + \angle C \\ = 48^\circ 18' 22'' + 76^\circ 32' 47.2'' + 55^\circ 08' 53.8'' = 180^\circ 0' 3'' \end{aligned}$$

There is an error of $3''$ which should be distributed to the various angles inversely as their weights. Let C_A , C_B and C_C be the respective corrections to the angles A , B and C .

$$\text{Then,} \quad C_A : C_B : C_C = \frac{1}{3} : \frac{1}{1} : \frac{1}{3} \quad (i)$$

$$\text{and} \quad C_A + C_B + C_C = -3'' \quad (ii)$$

From Eq. (i)

$$\begin{aligned} C_A &= C_A \\ C_B &= 3C_A \\ C_C &= C_A \end{aligned}$$

Substituting these values in Eq. (ii)

$$C_A + 3C_A + C_A = -3''$$

Hence,

$$\begin{aligned} C_A &= -0.6'' \\ C_B &= -3 \times 0.6 = -1.8'' \\ C_C &= -0.6'' \end{aligned}$$

Corrected values of the angles are

$$\begin{aligned} \angle A &= 48^\circ 18' 22'' - 0.6'' = 48^\circ 18' 21.4'' \\ \angle B &= 76^\circ 32' 47.2'' - 1.8'' = 76^\circ 32' 45.4'' \\ \angle C &= 55^\circ 08' 53.8'' - 0.6'' = 55^\circ 08' 53.2'' \end{aligned}$$

Check:

$$\text{Total} = 180^\circ 00' 00''$$

3.6 NORMAL EQUATIONS

An 'observation equation' is the symbolic equality of a quantity and its observed value, e.g., $\angle A = 145^\circ 32' 20''$, or $\angle B + \angle C = 67^\circ 35' 17''$. The former value is observed by direct single measurement and the latter one is observed indirectly when simultaneous observations or conditions are involved.

Consider a measurement of angles A , B and C closing the horizon at a station O . The rigorous geometrical condition to be fulfilled is $\angle A + \angle B + \angle C = 180^\circ = -k$ (say).

Now, if the angles are measured with the same precision, the error e is $A + B + C + k$ and the most probable value of each of the angles will be obtained by applying $(1/3)e$ to each of the observed values. However, if one angle is measured directly and the others indirectly by subtraction from the observed values of two or three angles together (reiteration method of measuring horizontal angles), or in a way that the conditions of measurements are varied, the error equation takes the form

$$e = aA + bB + cC + k$$

and if the measurements are repeated giving separate observations A_1, B_1, C_1 and A_2, B_2, C_2 , etc., the errors will be

$$e_1 = aA_1 + bB_1 + cC_1 + k$$

$$e_2 = aA_2 + bB_2 + cC_2 + k$$

From the theory of errors, $\Sigma(e)^2 = \Sigma(aA + bB + cC + k)^2$ should be minimum.

To minimise this expression, differentiate it with respect to A, B and C and equating each result to zero. We get

$$\Sigma a(aA + bB + cC + k) = 0$$

$$\Sigma b(aA + bB + cC + k) = 0$$

$$\Sigma c(aA + bB + cC + k) = 0$$

This shows that the fundamental equation when multiplied by coefficients of $\angle A, \angle B$ and $\angle C$ will provide the equations, the simultaneous solution of which will lead to the most probable values of $\angle A, \angle B$ and $\angle C$. Such equations are known as *normal equations*.

If the three angles in the above example are measured with different precision and are accorded weights w_A, w_B and w_C , the error equations will be

$$\Sigma w_A a(w_A aA + w_B bB + w_C cC + k) = 0$$

$$\Sigma w_B b(w_A aA + w_B bB + w_C cC + k) = 0$$

$$\Sigma w_C c(w_A aA + w_B bB + w_C cC + k) = 0$$

For the solution of the normal equations, the direct method or the method of differences may be used.

3.6.1 Direct Method

In the direct method, each side of the observation equation is multiplied by the weight of the observation. The weighted observations containing that quantity are then multiplied by their respective coefficients to yield normal equations. These equations are then solved simultaneously to determine the individual and the most probable values of the quantities.

3.6.2 Method of Differences

In the beginning, the reported single observations are assumed to be the most probable values, and for the remaining quantities, the most probable value is obtained by subtracting it from two or three variable equations. Let the discrepancy be $k_1, k_2, \dots$, etc. The observation equations are then formulated in terms of $k_1, k_2, \dots$, etc., to express the discrepancies between the observed results and those given by the assumed values, always subtracting the latter from the former. The equations thus obtained are multiplied by the respective weights, and the normal equations are formed by multiplying each side by the coefficients of $k_1, k_2, \dots$, etc.,. These simultaneous equations are then solved for $k_1, k_2, \dots$, etc., and are added algebraically to the assumed values of the quantities.

The arithmetic work is reduced in this method and hence it is particularly advantageous when the number of conditions are more.

Example 3.20 Angles were measured on a station and the observations were recorded as follows:

Angle	Value	Weight
A	$45^{\circ}30'10''$	2
B	$40^{\circ}20'20''$	3
$A + B$	$85^{\circ}50'10''$	1

Find the most probable values of the angles A and B .

Solution

(a) Direct method

$$\angle A = 45^{\circ}30'10'' \quad \text{Weight} = 2$$

$$\angle B = 40^{\circ}20'20'' \quad \text{Weight} = 3$$

$$\angle A + \angle B = 85^{\circ}50'10'' \quad \text{Weight} = 1$$

Multiply both the sides of the weighted observations by algebraic coefficients of A , thus forming the normal equations in A .

$$2A = 91^{\circ}00'20''$$

$$A + B = 85^{\circ}50'10''$$

$$\text{Hence, } 3A + B = 176^{\circ}50'30'' \quad (i)$$

Likewise, the normal equation in B is formed

$$3B = 121^{\circ}01'00''$$

$$A + B = 85^{\circ}50'10''$$

$$\text{Hence, } A + 4B = 206^{\circ}51'10'' \quad (ii)$$

Solve the above two simultaneous equations for the values of A and B .

Multiply Eq. (i) by 4 and subtract Eq. (ii) from it to get

$$12A + 4B = 707^{\circ}22'00''$$

$$A + 4B = 206^{\circ}51'10''$$

$$\text{Hence, } 11A = 500^{\circ}30'50''$$

$$\text{or } \angle A = 45^{\circ}30'4.55''$$

Likewise for angle B , multiply Eq. (ii) by 3 and subtract Eq. (i) from it to get

$$3A + 12B = 620^{\circ}33'30''$$

$$3A + B = 176^{\circ}50'30''$$

$$\text{Hence, } 11B = 443^{\circ}43'00''$$

$$\text{or } \angle B = 40^{\circ}20'16.36''$$

The most probable value of $\angle A$ is $45^{\circ}30'4.55''$ and that of $\angle B$ is $40^{\circ}20'16.36''$.

(b) Method of differences

Let k_1 and k_2 be the most probable corrections to angles A and B . Then,

$$\angle A = 45^{\circ}30'10'' + k_1 \quad (i)$$

$$\angle B = 40^{\circ}20'20'' + k_2 \quad (ii)$$

Add Eqs. (i) and (ii)

$$\angle A + \angle B = 85^\circ 50' 30'' + k_1 + k_2 \quad (\text{iii})$$

Subtract these equations from the observed equations

$$k_1 = 0 \quad \text{Weight} = 2$$

$$k_2 = 0 \quad \text{Weight} = 3$$

$$k_1 + k_2 = -20'' \quad \text{Weight} = 1$$

Normal equation for k_1 is

$$2k_1 = 0$$

$$k_1 + k_2 = -20''$$

$$\text{Hence, } 3k_1 + k_2 = -20'' \quad (\text{iv})$$

Normal equation for k_2 is

$$3k_2 = 0$$

$$k_1 + k_2 = -20''$$

$$\text{Hence, } k_1 + 4k_2 = -20'' \quad (\text{v})$$

From the two simultaneous Eqs. (iv) and (v)

$$k_1 = -5.45'' \quad \text{and} \quad k_2 = -3.63''$$

Hence, the most probable value of angle A

$$= 45^\circ 30' 10'' - 5.45'' = 45^\circ 30' 4.55''$$

and the most probable value of angle B

$$= 40^\circ 20' 20'' - 3.63'' = 40^\circ 20' 16.37''$$

3.7 METHOD OF CORRELATES

Correlates are the unknown multiples or independent constants used for finding most probable values of unknowns.

We have seen that in solving the normal equations, the condition equation was used to eliminate one of the unknowns thus giving one more observation equation. However, in the method of correlates, the equation of condition imposed by the theory of least squares is used in addition to all the other available condition equations.

This method is very advantageous as the computations are reduced to a large extent. This method is illustrated by the following example.

Example 3.21 Determine the most probable values of the angles of a triangle PQR using the method of correlates.

$$\angle P = 58^\circ 24' 36'' \quad \text{Weight} = 3$$

$$\angle Q = 52^\circ 12' 43'' \quad \text{Weight} = 3$$

$$\angle R = 69^\circ 22' 45'' \quad \text{Weight} = 2$$

Solution

$$\begin{aligned} \text{Sum of angles } (\angle P + \angle Q + \angle R) &= 58^\circ 24' 36'' + 52^\circ 12' 43'' + 69^\circ 22' 45'' \\ &= 180^\circ 0' 4'' \end{aligned}$$

$$\text{Sum of angles of a triangle} = 180^\circ 0' 0''$$

$$\text{Error} = 180^\circ 0' 4'' - 180^\circ 0' 0'' = 0^\circ 0' 4''$$

$$\text{Hence, correction to be applied} = -0^\circ 0' 4''$$

Let C_1, C_2, C_3 be the respective corrections in the angles $\angle P, \angle Q$ and $\angle R$.

$$\Rightarrow C_1 + C_2 + C_3 = -4'' \quad (i)$$

From principle of least square $\sum wC^2$ is a minimum.

$$\Rightarrow 3C_1^2 + 3C_2^2 + 2C_3^2 \quad (ii)$$

Differentiating (i) and (ii), we get

$$\delta C_1 + \delta C_2 + \delta C_3 = 0 \quad (iii)$$

$$\text{and} \quad 6C_1\delta C_1 + 6C_2\delta C_2 + 4C_3\delta C_3 = 0 \quad (iv)$$

$$\text{or} \quad 3C_1\delta C_1 + 3C_2\delta C_2 + 2C_3\delta C_3 = 0 \quad (v)$$

Multiplying Eq. (iii) by $-\lambda$ and adding to Eq. (iv)

$$\delta C_1(3C_1 - \lambda) + \delta C_2(3C_2 - \lambda) + \delta C_3(2C_3 - \lambda) = 0$$

Since, C_1, C_2, C_3 must vanish independently,

$$3C_1 - \lambda = 0$$

$$\Rightarrow C_1 = \lambda/3$$

$$3C_2 - \lambda = 0$$

$$\Rightarrow C_2 = \lambda/3$$

$$2C_3 - \lambda = 0$$

$$\Rightarrow C_3 = \lambda/2$$

Substituting the above values in Eq. (i),

$$\lambda/3 + \lambda/3 + \lambda/2 = -4''$$

$$\Rightarrow \lambda = -3.43''$$

Thus, $C_1 = -1.143'', C_2 = -1.143'', C_3 = -1.715''$

Most probable values of the angles P, Q and R are:

$$\angle P = 58^\circ 24' 36'' + C_1 = 58^\circ 24' 34.86''$$

$$\angle Q = 52^\circ 12' 43'' + C_2 = 52^\circ 12' 41.86''$$

$$\angle R = 69^\circ 22' 45'' + C_3 = 69^\circ 22' 43.29''$$

Example 3.22 A surveyor carried out levelling operations of a closed circuit $ABCD$ starting from A and found that

B was 6.71 m above A

C was 5.59 m above B

D was 3.48 m above C

D was 13.72 m above A

The accuracy of all the four levelling operations is to be assumed as equal. Determine the probable heights of B, C and D above A by the method of correlates.

Solution Let c_1, c_2, c_3 and c_4 be the required corrections to the four height differences.

$$\text{Total error} = 6.71 + 5.59 + 3.48 - 13.72 = 2.06 \text{ m}$$

$$\therefore \text{Total correction} = -2.06 \text{ m}$$

$$c_1 + c_2 + c_3 + c_4 = -2.06 \quad (i)$$

From the principle of least squares

$$c_1^2 + c_2^2 + c_3^2 + c_4^2 = \text{a minimum} \quad (\text{ii})$$

Differentiating Eqs. (i) and (ii), we get

$$\delta c_1 + \delta c_2 + \delta c_3 + \delta c_4 = 0 \quad (\text{iii})$$

$$\text{and } 2c_1 \delta c_1 + 2c_2 \delta c_2 + 2c_3 \delta c_3 + 2c_4 \delta c_4 = 0$$

$$\text{or } c_1 \delta c_1 + c_2 \delta c_2 + c_3 \delta c_3 + c_4 \delta c_4 = 0 \quad (\text{iv})$$

Multiplying Eq. (iii) by $-\lambda$ and adding to Eq. (iv), we get

$$\delta c_1(c_1 - \lambda) + \delta c_2(c_2 - \lambda) + \delta c_3(c_3 - \lambda) + \delta c_4(c_4 - \lambda) = 0$$

Since the coefficients of c_1, c_2, c_3 and c_4 must vanish independently

$$c_1 - \lambda = 0 \quad \text{or} \quad c_1 = \lambda$$

$$c_2 - \lambda = 0 \quad \text{or} \quad c_2 = \lambda$$

$$c_3 - \lambda = 0 \quad \text{or} \quad c_3 = \lambda$$

$$c_4 - \lambda = 0 \quad \text{or} \quad c_4 = \lambda$$

Substituting these values in Eq. (i), we get

$$\lambda + \lambda + \lambda + \lambda = -2.06$$

$$\text{or } \lambda = -0.515$$

$$\text{Hence, } c_1 = c_2 = c_3 = c_4 = -0.515 \text{ m}$$

Hence, the corrected height differences are:

Height difference of

$$B \text{ from } A = 6.71 - 0.515 = 6.195 \text{ m}$$

$$C \text{ from } B = 5.59 - 0.515 = 5.075 \text{ m}$$

$$D \text{ from } C = 3.48 - 0.515 = 2.965 \text{ m}$$

$$D \text{ from } A = 13.72 + 0.515 = 14.235 \text{ m}$$

$\therefore$ Probable height of

$$B \text{ above } A = 6.195 \text{ m}$$

$$C \text{ above } A = 6.195 + 5.075 = 11.270 \text{ m}$$

$$D \text{ above } A = 11.270 + 2.965 = 14.235 \text{ m}$$

3.8 TRIANGULATION ADJUSTMENT

Even after exercising care and taking all precautions, errors creep into the observations made. In order to compute the coordinates of the triangulation stations, the observations made are adjusted by distributing the observational error. The most accurate method is that of least squares but is very complicated since all the angles are simultaneously involved. However, using an approximate method, the adjustment can be achieved by adjusting angles, stations and figures separately. After adjusting the triangulation figure, the sine rule is applied for computing sides. Then the positions of the points are determined by calculating the geodetic coordinates.

3.9 ANGLE ADJUSTMENT

In any system of triangulation, there are a number of geometrical conditions which the angles should fulfil, but because all measured angles are affected by the errors

of measurement, they never will meet all the conditions perfectly. For this reason it is necessary to adjust the angles so that the best, or the most probable, values may be found.

Usually, many observations are made for a single angle; for example, face left and face right, vernier *A* and vernier *B*, and reading an angle on different parts of the scale. The correction to be applied is directly proportional to the weight and also to the square of the probable error. The angles can be measured with equal or unequal weights. In the former case, the most probable value is the arithmetic mean of the observations, whereas in the latter case, it is the weighted arithmetic mean of the observed angles.

Example 3.23 An angle was measured by different persons and the following values were observed. Find the most probable value of the angle.

<i>Angle</i>	<i>Number of measurements</i>
54°30'20"	2
54°29'40"	2
54°30'10"	2

Solution Since the angles are measured with equal weights, the most probable value of the angle will be the arithmetic mean of the observations. Hence, the most probable value

$$= \frac{54^\circ 30' 20'' + 54^\circ 29' 40'' + 54^\circ 30' 10''}{3} = 54^\circ 30' 3.33''$$

Example 3.24 An angle was measured by different observers with the following values. Find the most probable value of the angle.

<i>Angle</i>	<i>Number of measurements</i>
54°30'20"	2
54°29'40"	3
54°30'10"	4

Solution The most probable value will be the weighted arithmetic mean. Hence, the most probable value

$$= \frac{(54^\circ 30' 20'' \times 2) + (54^\circ 29' 40'' \times 3) + (54^\circ 30' 10'' \times 4)}{2 + 3 + 4} = 54^\circ 30' 2.2''$$

Example 3.25 Find the most probable value of the angle *A* from the following observation equations:

$$\begin{aligned} A &= 42^\circ 14' 20'' & \text{Weight} &= 2 \\ 2A &= 85^\circ 05' 40'' & \text{Weight} &= 3 \end{aligned}$$

Solution

$$\begin{aligned} A &= 42^\circ 14' 20'' & \text{(i)} \\ 2A &= 85^\circ 05' 40'' & \text{(ii)} \end{aligned}$$

In Eq. (i) the coefficient of A is 1 and the weight of the observation is 2. Therefore, multiply it with 2 ($= 1 \times 2$) to get

$$2A = 84^\circ 28' 40'' \quad (\text{iii})$$

In Eq. (ii) the coefficient of A is 2 and the weight of the observation is 3. Therefore, multiply it with 6 ($= 2 \times 3$) to get

$$12A = 510^\circ 34' 00'' \quad (\text{iv})$$

Add Eqs. (iii) and (iv) to get

$$2A + 12A = 84^\circ 28' 40'' + 510^\circ 34' 00''$$

$$\text{or} \quad 14A = 595^\circ 02' 40''$$

$$\text{or} \quad A = 42^\circ 30' 11.4''$$

The most probable value of angle A is $42^\circ 30' 11.4''$.

3.10 STATION ADJUSTMENT

The station adjustment consists of determining the most probable values of the angles measured at a station so as to satisfy the geometric consistency. The various conditions can be (i) closing the horizon, (ii) measuring the angles with equal or unequal weights, and (iii) measuring different angles at a station individually or in combination. In the first case, the error if any is distributed equally to all the three angles. In the second case it is distributed inversely as the respective weights. Whereas in the last case, normal equations are formed and are solved simultaneously.

Example 3.26 Following angles were measured at a station. Find their most probable values:

Angle	Value	Weight
AOB	$126^\circ 30' 20''$	2
BOC	$74^\circ 20' 10''$	2
COA	$159^\circ 12' 20''$	2

Solution Refer to Fig. 3.3. The sum of the observed angles

$$= 126^\circ 30' 20'' + 74^\circ 20' 10'' + 159^\circ 12' 20'' = 360^\circ 2' 50''$$

$$\text{Discrepancy} = 360^\circ 2' 50'' - 360^\circ = 2' 50''$$

$$\therefore \text{Correction} = -2' 50''$$

$$\text{Correction to each angle} = \frac{-2' 50''}{3} = -56.66''$$

Most probable values of angles are

$$\begin{aligned} \angle AOB &= 126^\circ 30' 20'' - 56.66'' \\ &= 126^\circ 29' 23.34'' \end{aligned}$$

$$\begin{aligned} \angle BOC &= 74^\circ 20' 10'' - 56.66'' \\ &= 74^\circ 19' 13.34'' \end{aligned}$$

$$\begin{aligned} \angle COA &= 159^\circ 12' 20'' - 56.66'' \\ &= 159^\circ 11' 23.34'' \end{aligned}$$

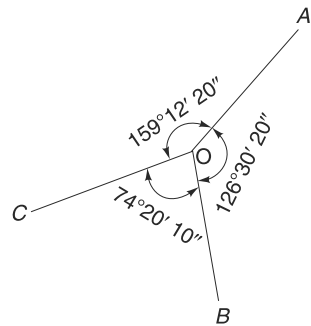


Fig. 3.3

Example 3.27 Following angles were measured at a station. Find their most probable values.

Angle	Value	Weight
$AOB (\alpha)$	$125^\circ 30' 20''$	2
$BOA (\beta)$	$234^\circ 28' 40''$	4

Solution Refer to Fig. 3.4. Sum of the observed angles
 $= 125^\circ 30' 20'' + 234^\circ 28' 40'' = 359^\circ 59'$

Discrepancy $= 360^\circ - 359^\circ 59' = 1'$

$\therefore$ correction required $= 1'$

The error is distributed in an inverse proportion to the weights.

Let c_1 and c_2 be the most probable corrections in angles α and β .

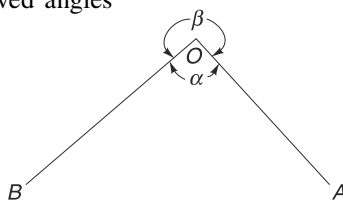


Fig. 3.4

Then, $c_1 : c_2 = \frac{1}{2} : \frac{1}{4}$

or $c_1 : c_2 = 1 : \frac{1}{2}$

also $c_1 + c_2 = 1'$ (i)

From Eq. (i)

$$c_2 = \frac{1}{2} c_1 \quad \text{(ii)}$$

Substituting this value in Eq. (ii), we get

$$c_1 + \frac{1}{2} c_1 = 1' \quad \text{(iii)}$$

or

$$c_2 = \frac{40}{2} = 20''$$

From Eq. (iii),

Therefore, the respective corrections in α and β are $40''$ and $20''$, hence

$$\angle AOB (\alpha) = 125^\circ 30' 20'' + 40'' = 125^\circ 31' 00''$$

$$\angle BOA (\beta) = 234^\circ 28' 40'' + 20'' = 234^\circ 29' 00''$$

3.11 FIGURE ADJUSTMENT

In any system of triangulation, determination of the most probable values of the angles so as to fulfill the geometrical conditions is called *figure adjustment*. There can be a number of geometrical conditions which the angles should fulfill, but since all the measured angles are affected by errors, they never will meet all the conditions perfectly. Therefore, it is necessary to adjust the angles so as to obtain the best possible and most probable value. The best solution can be obtained by the method of least squares, also known as the *rigid method*, which is a little complex and therefore, the adjustments are usually done by an approximate method. The geometrical figures encountered in triangulation are a triangle, a quadrilateral or a polygon with a central station.

3.12 ADJUSTMENT OF A TRIANGLE

A triangle is the basic figure of any triangulation system. All the three angles of a triangle are adjusted. Some of the rules for applying corrections to the observed angles are as follows. Let

A, B, C = angles of the triangle

n = number of observations for an angle

w = weight of the angle

d = discrepancy (error of closure)

c = correction to observed angle

Rule 1 For angles of equal weight, the discrepancy d is distributed equally among all the three angles:

$$c_A = c_B = c_C = \frac{1}{3} d$$

Rule 2 For angles of unequal weight, the discrepancy is distributed in inverse proportions to the weights:

$$c_A : c_B : c_C = \frac{1}{w_A} : \frac{1}{w_B} : \frac{1}{w_C}$$

$$c_A = \frac{\frac{1}{w_A}}{\frac{1}{w_A} + \frac{1}{w_B} + \frac{1}{w_C}} d, \quad c_B = \frac{\frac{1}{w_B}}{\frac{1}{w_A} + \frac{1}{w_B} + \frac{1}{w_C}} d, \quad c_C = \frac{\frac{1}{w_C}}{\frac{1}{w_A} + \frac{1}{w_B} + \frac{1}{w_C}} d$$

Rule 3 If instead of weights, number of observations is given, then the discrepancy is distributed in inverse proportions to the number of observations:

$$c_A = \frac{\frac{1}{n_A}}{\frac{1}{n_A} + \frac{1}{n_B} + \frac{1}{n_C}} d, \quad c_B = \frac{\frac{1}{n_B}}{\frac{1}{n_A} + \frac{1}{n_B} + \frac{1}{n_C}} d, \quad c_C = \frac{\frac{1}{n_C}}{\frac{1}{n_A} + \frac{1}{n_B} + \frac{1}{n_C}} d$$

Rule 4 The corrections are inversely proportional to the square of the number of observations:

$$c_A = \frac{\left(\frac{1}{n_A}\right)^2}{\left(\frac{1}{n_A}\right)^2 + \left(\frac{1}{n_B}\right)^2 + \left(\frac{1}{n_C}\right)^2} d, \quad c_B = \frac{\left(\frac{1}{n_B}\right)^2}{\left(\frac{1}{n_A}\right)^2 + \left(\frac{1}{n_B}\right)^2 + \left(\frac{1}{n_C}\right)^2} d,$$

$$c_C = \frac{\left(\frac{1}{n_C}\right)^2}{\left(\frac{1}{n_A}\right)^2 + \left(\frac{1}{n_B}\right)^2 + \left(\frac{1}{n_C}\right)^2} d$$

Rule 5 The corrections are proportional to the square of the probable errors:

$$c_A = \frac{E_A^2}{E_A^2 + E_B^2 + E_C^2} d, \quad c_B = \frac{E_B^2}{E_A^2 + E_B^2 + E_C^2} d, \quad c_C = \frac{E_C^2}{E_A^2 + E_B^2 + E_C^2} d$$

Rule 6 When the weights of the observations are not given directly, then if v is the difference between the mean observed value and the observed value of an angle, the weight of the angle is given by

$$w = \frac{\frac{1}{2}n^2}{\sum v^2}$$

Thus,
$$\frac{1}{w_A} = \frac{\sum v_A^2}{\frac{1}{2}n_A^2}, \quad \frac{1}{w_B} = \frac{\sum v_B^2}{\frac{1}{2}n_B^2}, \quad \frac{1}{w_C} = \frac{\sum v_C^2}{\frac{1}{2}n_C^2}$$

$$c_A = \frac{\frac{1}{w_A}}{\frac{1}{w_A} + \frac{1}{w_B} + \frac{1}{w_C}} d, \quad c_B = \frac{\frac{1}{w_B}}{\frac{1}{w_A} + \frac{1}{w_B} + \frac{1}{w_C}} d, \quad c_C = \frac{\frac{1}{w_C}}{\frac{1}{w_A} + \frac{1}{w_B} + \frac{1}{w_C}} d$$

Usually it is found that rules 1, 2 and 6 are sufficient to adjust angles of the triangle.

Example 3.28 Adjust the following angles of a triangle:

$\angle A$	$\angle B$	$\angle C$
34°22'13"	69°32'48"	76°03'18"
12"	44"	22"
16"	45"	21"
17"	49"	17"
11"	46"	
9"		

Solution Mean value of angle $A = 34^\circ 22' 13''$

Mean value of angle $B = 69^\circ 32' 46.4''$

Mean value of angle $C = 76^\circ 03' 19.5''$

$$\begin{aligned} \text{Sum of the angles} &= 34^\circ 22' 13'' + 69^\circ 32' 46.4'' + 76^\circ 03' 19.5'' \\ &= 179^\circ 58' 18.9'' \end{aligned}$$

$$\text{Discrepancy} = 180^\circ - 179^\circ 58' 18.9'' = 1' 41.1''$$

$$\text{Weight of angle } A, w_A = \frac{\frac{1}{2}n_A^2}{\sum v_A^2}$$

$$n_A = 6$$

$$\Sigma v_A^2 = (M_A - L_A)^2 = [(0)^2 + (+1)^2 + (-3)^2 + (-4)^2 + (+2)^2 + (+4)^2] \\ = 46$$

$$\therefore w_A = \frac{\frac{1}{2} \times (6)^2}{46} = 0.391$$

Similarly, for angle B

$$w_B = \frac{\frac{1}{2} n_B^2}{\Sigma v_B^2} \\ n_B = 5$$

$$\Sigma v_B^2 = (M_B - L_B)^2 = [(-1.6)^2 + (2.4)^2 + (1.4)^2 + (-2.6)^2 + (0.4)^2] = 17.20$$

$$\therefore w_B = \frac{\frac{1}{2} \times (5)^2}{17.20} = 0.727$$

And for angle C

$$w_C = \frac{\frac{1}{2} n_C^2}{\Sigma v_C^2} \\ n_C = 4$$

$$\Sigma v_C^2 = (M_C - L_C)^2 = [(1.5)^2 + (-2.5)^2 + (1.5)^2 + (2.5)^2] = 17.0$$

$$\therefore w_C = \frac{\frac{1}{2} \times (4)^2}{17.0} = 0.47$$

Hence,

$$c_A = \frac{\frac{1}{0.391}}{\frac{1}{0.391} + \frac{1}{0.727} + \frac{1}{0.47}} \times 1'41.1'' = 42.66''$$

$$c_B = \frac{\frac{1}{0.727}}{\frac{1}{0.391} + \frac{1}{0.727} + \frac{1}{0.47}} \times 1'41.1'' = 22.94''$$

$$c_C = \frac{\frac{1}{0.47}}{\frac{1}{0.391} + \frac{1}{0.727} + \frac{1}{0.47}} \times 1'41.1'' = 35.49''$$

Check Sum of corrections = $42.66'' + 22.94'' + 35.49'' = 1'41.09''$
 (M = Mean value of the angle, L = observed value of the angle)

3.13 SPHERICAL EXCESS

In triangulation of a small area with sides of the triangles less than 3.5 km, the triangles are considered to be plane. But in a geodetic survey, the curvature of the earth is given due consideration and the sides of triangles are considered to be arcs. The sum of three angles will always be more than 180° by an amount (by which the sum of observed angles exceeds 180°) known as *spherical excess*. Its value is approximately $1''$ for every 196 km^2 area. However, the exact value may be calculated from the following formula.

$$\text{Spherical excess in degrees} = \frac{A}{\pi R^2} \times 180^\circ$$

$$\text{or Spherical excess in seconds} = \frac{A}{\pi R^2} \times 180 \times 60 \times 60$$

$$= \frac{A}{\pi R^2} \times 648\,000$$

$$= \frac{A}{R^2} 206\,265$$

$$\text{or} \quad = \frac{A}{R^2 \sin 1''}$$

where A = area of spherical triangle (km^2), taken as the same as that of the plane triangle for calculation of spherical excess,

R = radius of the earth (km).

Note In geodetic work, or where triangles are large, the spherical excess must be determined and the reduced angles corrected before adjusting.

3.14 ADJUSTMENT OF A GEODETIC QUADRILATERAL

In a geodetic quadrilateral, all the eight angles are measured independently along the two diagonals. In case the quadrilateral is of large size, the angles should be corrected for spherical excess to deduce the plane angles for further adjustment. While standing on the intersection of the diagonals looking towards the sides, angles 1, 3, 5, 7 (Fig. 3.5) are on the right and are called *right-hand angles*, whereas angles 2, 4, 6, 8 will be on the left and are called *left-hand angles*.

There are three methods of adjusting the quadrilaterals, viz. method of least squares, approximate method, and method of equal shifts. The method of least squares of adjusting the observed angles is presented briefly without going into the details of the derivation.

The conditions to be fulfilled by the adjusted values of angles are:

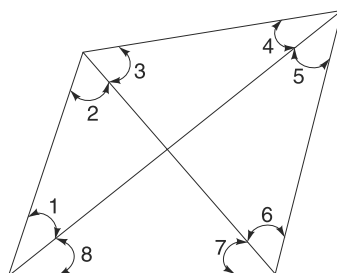


Fig. 3.5

1. Angle equations

$$\angle 1 + \angle 2 + \angle 3 + \angle 4 + \angle 5 + \angle 6 + \angle 7 + \angle 8 = 360^\circ \quad (3.1)$$

$$\angle 1 + \angle 2 = \angle 5 + \angle 6 \quad (3.2)$$

$$\angle 3 + \angle 4 = \angle 7 + \angle 8 \quad (3.3)$$

2. *Side equations:* If any one side of a quadrilateral is known, then any other side can be computed by different combinations of triangles. For a perfect quadrilateral, the computed side should be the same, irrespective of the combination of triangles used. The condition is expressed as a side equation which is

$$\Sigma \log \sin (\angle 1, \angle 3, \angle 5, \angle 7) = \Sigma \log \sin (\angle 2, \angle 4, \angle 6, \angle 8) \quad (3.4)$$

Note Since the sine of an angle is always less than unity, the logarithm of its sine is always negative. To avoid the trouble of printing proper sign to the logarithms of the trigonometric functions, the logarithms as tabulated in the standard log tables are not the true logarithms; but the true logarithms increased by 10. For example, $\sin 30^\circ = 1/2$. Hence, $\log \sin 30^\circ = \log(1/2) = -\log 2 = -0.30103 = 7.69897$. The logarithm tabulated is therefore $10 + \log \sin 30$, i.e., 9.69897.

From the principle of least squares for weighted quantities, the most probable value is that for which the sum of the weighted squares of the residuals is the least. Therefore, if

$c_1, c_2, c_3, \dots, c_n$ are the corrections to the angles 1, 2, 3, ... n , then

$$c_1^2 + c_2^2 + c_3^2 + c_4^2 + c_5^2 + c_6^2 + c_7^2 + c_8^2 = \text{minimum} \quad (3.5)$$

By solving the above five equations, the values of the corrections can be found and the angles can be corrected. The method is illustrated by the following example.

Example 3.29 Following are the values of the eight angles of a geodetic quadrilateral (Fig. 3.5) after being adjusted for spherical excess. The angles may be assumed to be of equal weights. Adjust the quadrilateral.

Angle	Value	Angle	Value
1	29°25'34"	5	30°29'25"
2	58°41'20"	6	57°37'33"
3	69°36'20"	7	50°35'44"
4	22°17'02"	8	41°17'26"

Solution Refer to Fig. 3.5.

Angle equations

$$\angle 1 + \angle 2 + \angle 3 + \angle 4 + \angle 5 + \angle 6 + \angle 7 + \angle 8 = 360^\circ \quad (i)$$

$$\angle 1 + \angle 2 = \angle 5 + \angle 6 \quad (ii)$$

$$\angle 3 + \angle 4 = \angle 7 + \angle 8 \quad (iii)$$

Sum of angles designated as 1, 2, 3, 4, 5, 6, 7, 8 is

$$\sum_{n=1}^8 n = 360^\circ 00' 24''$$

$$\text{Correction} = 360^\circ - 360^\circ 00' 24'' = -24''$$

$$\angle 1 + \angle 2 = 88^\circ 6' 54'' \quad \text{and} \quad \angle 5 + \angle 6 = 88^\circ 6' 58''$$

According to Eq. (ii), these must be equal. Therefore,

$$\text{Discrepancy} = (\angle 1 + \angle 2) - (\angle 5 + \angle 6)$$

$$\text{Correction to be applied} = + 4''$$

Similarly,

$$\angle 3 + \angle 4 = 91^\circ 53' 22'' \quad \text{and} \quad \angle 7 + \angle 8 = 91^\circ 53' 10''$$

$$\text{Discrepancy} = (\angle 3 + \angle 4) - (\angle 7 + \angle 8)$$

$$\text{Correction to be applied} = - 12''$$

Side equations

$$\Sigma \log \sin (1, 3, 5, 7) = \Sigma \log \sin (2, 4, 6, 8)$$

$$\begin{aligned} \text{L.H.S.} &= \log \sin(1) + \log \sin(3) + \log \sin(5) + \log \sin(7) \\ &= 9.6913474 + 9.9718859 + 9.7053437 + 9.8880021 \\ &= 39.256579 \end{aligned}$$

$$\begin{aligned} \text{R.H.S.} &= \log \sin(2) + \log \sin(4) + \log \sin(6) + \log \sin(8) \\ &= 9.9316399 + 9.5788637 + 9.9266354 + 9.8194635 \\ &= 39.256603 \end{aligned}$$

$$\text{Difference, R.H.S.} - \text{L.H.S.} = 39.256603 - 39.256579 = + 234''$$

If $d_1, d_2, \dots, d_8$ are the log sin differences for $1''$ in the value of angles 1, 2, 3, 4, 5, 6, 7, 8 and $c_1, c_2, \dots, c_8$ are corrections in angles 1, 2, ..., 8, then

$$c_1 + c_2 + c_3 + c_4 + c_5 + c_6 + c_7 + c_8 = - 234'' \quad (\text{iv})$$

$$c_1 + c_2 - c_5 - c_6 = + 4'' \quad (\text{v})$$

$$c_3 + c_4 - c_7 - c_8 = - 12'' \quad (\text{vi})$$

and from log tables we have

$$\begin{aligned} d_1 &= 37.3, d_2 = 12.8, d_3 = 7.8, d_4 = 51.4 \\ d_5 &= 35.7, d_6 = 13.3, d_7 = 17.3, d_8 = 23.9 \end{aligned}$$

Now, we have

$$c_1 d_1 - c_2 d_2 + c_3 d_3 - c_4 d_4 + c_5 d_5 - c_6 d_6 + c_7 d_7 - c_8 d_8 = + 24$$

$$\begin{aligned} \text{or} \quad 37.3 c_1 - 12.8 c_2 + 7.8 c_3 - 51.4 c_4 + 35.7 c_5 \\ - 13.3 c_6 + 17.3 c_7 - 23.9 c_8 = + 234 \end{aligned} \quad (\text{vii})$$

From the least squares condition

$$c_1^2 + c_2^2 + c_3^2 + c_4^2 + c_5^2 + c_6^2 + c_7^2 + c_8^2 = \text{minimum} \quad (\text{viii})$$

To form the equation in $c_1, c_2, \dots, c_8$, differentiate all equations and multiply Eqs. (iv) to (vii) by $-\lambda_1, -\lambda_2, -\lambda_3, -\lambda_4$, respectively, and add to Eq. (viii) and equate the coefficient of $\delta c_1, \delta c_2, \dots, \delta c_8$ to zero. We get,

$$c_1 = \lambda_1 + \lambda_2 + 37.3 \lambda_4$$

$$c_2 = \lambda_1 + \lambda_2 - 12.8 \lambda_4$$

$$c_3 = \lambda_1 + \lambda_3 + 7.8 \lambda_4$$

$$c_4 = \lambda_1 + \lambda_3 - 51.4 \lambda_4$$

$$c_5 = \lambda_1 - \lambda_2 + 35.7 \lambda_4$$

$$c_6 = \lambda_1 - \lambda_2 - 13.3 \lambda_4$$

$$c_7 = \lambda_1 - \lambda_3 + 17.3 \lambda_4$$

$$c_8 = \lambda_1 - \lambda_3 - 23.9 \lambda_4$$

We now have

$$8 \lambda_1 - 3.3 \lambda_4 = \sum_{i=1}^n c_i = - 24'' \quad [\text{from Eq. (iv)}]$$

$$4 \lambda_2 + 2.1 \lambda_4 = c_1 + c_2 - c_5 - c_6 = + 4'' \quad [\text{from Eq. (v)}]$$

$$4 \lambda_3 - 37 \lambda_4 = c_3 + c_4 - c_7 - c_8 = - 12'' \quad [\text{from Eq. (vi)}]$$

$$\begin{aligned} \Rightarrow \quad 37.3 (\lambda_1 + \lambda_2 + 37.3 \lambda_4) - 12.8 (\lambda_1 + \lambda_2 - 12.8 \lambda_4) \\ + 7.8 (\lambda_1 + \lambda_3 + 7.8 \lambda_4) - 51.4 (\lambda_1 + \lambda_3 - 51.4 \lambda_4) \\ + 35.7 (\lambda_1 - \lambda_2 + 35.7 \lambda_4) - 13.3 (\lambda_1 - \lambda_2 - 13.3 \lambda_4) \end{aligned}$$

$$+ 17.3 (\lambda_1 - \lambda_3 + 17.3 \lambda_4)$$

$$- 23.9(\lambda_1 - \lambda_3 - 23.9 \lambda_4) = + 234$$

$$\text{or} \quad - 3.3 \lambda_1 + 2.1 \lambda_2 - 37 \lambda_3 + 6579.81 \lambda_4 = + 234$$

$$\text{or} \quad - 3.3 \left(3 + \frac{3.3}{8} \lambda_4 \right) + 2.1 \left(1 - \frac{2.1}{4} \lambda_4 \right)$$

$$- 37 \left(3 + \frac{37}{4} \lambda_4 \right) + 6579.81 \lambda_4 = 234$$

$$\text{or} \quad 9.9 - 1.36 \lambda_4 + 2.1 - 1.1 \lambda_4 + 111 - 342.25 \lambda_4 + 6579.81 \lambda_4 = 234$$

$$\text{or} \quad 6235.096 \lambda_4 = 111$$

$$\text{or} \quad \lambda_4 = 0.0178,$$

$$\lambda_1 = + - 2.9926, \quad \lambda_2 = + 0.99065, \quad \lambda_3 = - 2.8353$$

Therefore, we have

$$c_1 = 1.338'', \quad c_2 = - 2.2298'', \quad c_3 = - 5.6891'', \quad c_4 = - 6.742'',$$

$$c_5 = - 3.3478'', \quad c_6 = - 4.22'', \quad c_7 = 0.1507'', \quad c_8 = - 4.22''.$$

Correct angles

$$1 = 29^\circ 25' 32.66''$$

$$2 = 58^\circ 41' 17.77''$$

$$3 = 69^\circ 36' 14.31''$$

$$4 = 22^\circ 17' 55.26''$$

$$5 = 30^\circ 29' 21.65''$$

$$6 = 57^\circ 37' 28.78''$$

$$7 = 50^\circ 35' 44.15''$$

$$8 = 41^\circ 17' 25.41''$$

Now, for a check

$$\begin{aligned} \Sigma \log \sin (1, 3, 5, 7) &= \log \sin(1) + \log \sin(3) + \log \sin(5) + \log \sin(7) \\ &= 9.6913424 + 9.9718814 + 9.7053317 + 9.8880023 \\ &= 39.26557 \end{aligned}$$

$$\begin{aligned} \Sigma \log \sin (2, 4, 6, 8) &= \log \sin(2) + \log \sin(4) + \log \sin(6) + \log \sin(8) \\ &= 9.9316371 + 9.5788291 + 9.926629 + 9.819462 \\ &= 39.256557 \end{aligned}$$

$$\text{Hence, } \Sigma \log \sin (1, 3, 5, 7) = \Sigma \log \sin (2, 4, 6, 8)$$

3.15 ADJUSTMENT OF QUADRILATERAL WITH CENTRAL STATION

A quadrilateral $ABCD$ with a central station O (Fig. 3.6) involves the independent measurement of all the twelve angles. The left-hand angles are denoted by 1, 3, 5, 7 and the right-hand angles by 2, 4, 6, 8, whereas 9, 10, 11, 12 are the central angles.

The condition equations are:

$$\angle 1 + \angle 8 + \angle 12 = 180^\circ$$

$$\angle 2 + \angle 3 + \angle 9 = 180^\circ$$

$$\angle 4 + \angle 5 + \angle 10 = 180^\circ$$

$$\begin{aligned}\angle 6 + \angle 7 + \angle 11 &= 180^\circ \\ \angle 9 + \angle 10 + \angle 11 + \angle 12 &= 360^\circ \\ \Sigma \log \sin (\angle 1, \angle 3, \angle 5, \angle 7) &= \Sigma \log \sin (\angle 2, \angle 4, \angle 6, \angle 8)\end{aligned}$$

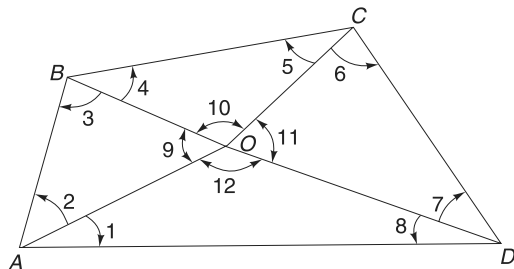


Fig. 3.6

If c_1, c_2, c_3 , etc., are the corrections in seconds to the angles 1, 2, 3, etc., and d_1, d_2, d_3 , etc., are the tabular differences for one second for $\log \sin 1, \log \sin 2, \log \sin 3$, etc., then

$$\begin{aligned}c_1 + c_8 + c_{12} &= \pm k_1 \\ c_2 + c_3 + c_9 &= \pm k_2 \\ c_4 + c_5 + c_{10} &= \pm k_3 \\ c_6 + c_7 + c_{11} &= \pm k_4 \\ c_9 + c_{10} + c_{11} + c_{12} &= \pm k_5\end{aligned}$$

and
$$\sum_{i=1}^{12} c_i^2 = \text{a minimum}$$

The process of finding the corrections is explained by the following example.

Example 3.30 The following are angles of a quadrilateral $ABCD$ with a central station O , measured independently. Determine the most probable values of the corrections.

Triangle	R.H. angle	L.H. angle	Central angle
AOB	$37^\circ 40' 12.5''$	$42^\circ 23' 34.2''$	$99^\circ 56' 03''$
BOC	$60^\circ 19' 22.8''$	$39^\circ 36' 46.6''$	$80^\circ 03' 55.6''$
COD	$31^\circ 52' 31.5''$	$48^\circ 11' 17.8''$	$99^\circ 56' 15.7''$
DOA	$54^\circ 30' 2.7''$	$45^\circ 26' 8.3''$	$80^\circ 03' 59''$

Solution Refer to Fig. 3.6.

$$\begin{aligned}\angle 1 &= 54^\circ 30' 2.7'' & \angle 7 &= 31^\circ 52' 31.5'' \\ \angle 2 &= 42^\circ 23' 34.2'' & \angle 8 &= 45^\circ 26' 8.3'' \\ \angle 3 &= 37^\circ 40' 12.5'' & \angle 9 &= 99^\circ 56' 03'' \\ \angle 4 &= 39^\circ 36' 46.6'' & \angle 10 &= 80^\circ 03' 55.6'' \\ \angle 5 &= 60^\circ 19' 22.8'' & \angle 11 &= 99^\circ 56' 15.7'' \\ \angle 6 &= 48^\circ 11' 17.8'' & \angle 12 &= 80^\circ 03' 59''\end{aligned}$$

The conditions are:

$$\angle 1 + \angle 8 + \angle 12 = 180^\circ$$

$$\angle 2 + \angle 3 + \angle 9 = 180^\circ$$

$$\angle 4 + \angle 5 + \angle 10 = 180^\circ$$

$$\angle 6 + \angle 7 + \angle 11 = 180^\circ$$

$$\angle 9 + \angle 10 + \angle 11 + \angle 12 = 360^\circ$$

$$\Sigma \log \sin (\angle 1, \angle 3, \angle 5, \angle 7) = \Sigma \log \sin (\angle 2, \angle 4, \angle 6, \angle 8)$$

$$c_1 = 180^\circ - (\angle 1 + \angle 8 + \angle 12) = -10''$$

$$c_2 = 180^\circ - (\angle 2 + \angle 3 + \angle 9) = +10.3''$$

$$c_3 = 180^\circ - (\angle 4 + \angle 5 + \angle 10) = -5''$$

$$c_4 = 180^\circ - (\angle 6 + \angle 7 + \angle 11) = -5''$$

$$c_5 = 360^\circ - (\angle 9 + \angle 10 + \angle 11 + \angle 12) = -13.3''$$

$$c_6 = \Sigma \log \sin (\angle 2, \angle 4, \angle 6, \angle 8) - \Sigma \log \sin (\angle 1, \angle 3, \angle 5, \angle 7)$$

<i>Right</i>				<i>Left</i>			
	<i>Angle</i>	<i>log sine</i>	<i>d_r</i>		<i>Angle</i>	<i>log sine</i>	<i>d_l</i>
		(<i>l</i>)				(<i>l</i>)	
1	54°30'2.7"	9.9106901	15.01	2	42°23'34.2"	9.8287952	23.06
3	37°40'12.5"	9.7861227	27.27	4	39°36'46.6"	9.804547	25.43
5	60°19'22.8"	9.9389350	11.99	6	48°11'17.8"	9.8723542	18.83
7	31°52'31.5"	9.7226948	33.85	8	45°26'8.3"	9.8527621	20.73
	$\Sigma l = 39.358443$		$\Sigma d_r =$		$\Sigma l = 39.358459$		$\Sigma d_l =$
			88.12				88.05

$$\Sigma d^2 = \Sigma (d_r^2 + d_l^2) = 4221.29$$

$$c_6 = \Sigma l \text{ (Right)} - \Sigma l \text{ (left)} = 39.358459 - 39.358443 = 160$$

If $c_1, c_2, c_3 \dots$ are the corrections of angles 1, 2, 3, ..., then

$$c_1 + c_8 + c_{12} = -10'' \quad \text{(i)}$$

$$c_2 + c_3 + c_9 = +10.3'' \quad \text{(ii)}$$

$$c_4 + c_5 + c_{10} = -5'' \quad \text{(iii)}$$

$$c_6 + c_7 + c_{11} = -5'' \quad \text{(iv)}$$

$$c_9 + c_{10} + c_{11} + c_{12} = -13.3'' \quad \text{(v)}$$

$$c_1 d_1 - c_2 d_2 + c_3 d_3 - c_4 d_4 + c_5 d_5 - c_6 d_6 + c_7 d_7 - c_8 d_8 = 160$$

$$\begin{aligned} \text{or } 15.01 c_1 - 23.06 c_2 + 27.27 c_3 - 25.43 c_4 + 11.99 c_5 \\ - 18.83 c_6 + 33.85 c_7 - 20.73 c_8 = 160 \end{aligned} \quad \text{(vi)}$$

Differentiate Eqs. (i) to (vi) and then multiply Eqs. (i) to (vi) by $-\lambda_1, -\lambda_2, -\lambda_3, -\lambda_4, -\lambda_5, -\lambda_6$ respectively.

c_1	+1	0	0	0	0	+15.01
c_2	0	+1	0	0	0	-23.06
c_3	0	+1	0	0	0	+27.27
c_4	0	0	+1	0	0	-25.43
c_5	0	0	+1	0	0	+11.99
c_6	0	0	0	+1	0	-18.83
c_7	0	0	0	+1	0	+33.85
c_8	+1	0	0	0	0	-20.73
c_9	0	+1	0	0	+1	0
c_{10}	0	0	+1	0	+1	0
c_{11}	0	0	0	+1	+1	0
c_{12}	1	0	0	0	+1	0

The following equations can be obtained from the above table:

$$c_1 = \lambda_1 + 15.01 \lambda_6 \quad c_7 = \lambda_4 + 33.85 \lambda_6$$

$$c_2 = \lambda_2 - 23.06 \lambda_6 \quad c_8 = \lambda_2 - 20.73 \lambda_6$$

$$c_3 = \lambda_2 + 27.27 \lambda_6 \quad c_9 = \lambda_2 + \lambda_5$$

$$c_4 = \lambda_3 - 25.43 \lambda_6 \quad c_{10} = \lambda_3 + \lambda_5$$

$$c_4 = \lambda_3 + 11.99 \lambda_6 \quad c_{11} = \lambda_4 + \lambda_5$$

$$c_6 = \lambda_4 - 18.83 \lambda_6 \quad c_{12} = \lambda_1 + \lambda_5$$

Substituting these values of $c_1, c_2, c_3, \dots, c_{12}$ in Eqs. (i) to (vi) leads to

$$3 \lambda_1 + \lambda_5 - 5.72 \lambda_6 = -10'' \quad (\text{vii})$$

$$3 \lambda_2 + \lambda_5 + 4.21 \lambda_6 = +10.3'' \quad (\text{viii})$$

$$3 \lambda_3 + \lambda_5 - 13.44 \lambda_6 = -5'' \quad (\text{ix})$$

$$3 \lambda_4 + \lambda_5 + 15.02 \lambda_6 = -5'' \quad (\text{x})$$

$$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + 4 \lambda_5 = -13.3'' \quad (\text{xi})$$

$$-5.72 \lambda_1 + 4.21 \lambda_2 - 13.44 \lambda_3 + 15.02 \lambda_4 + 4221.29 \lambda_6 = 160 \quad (\text{xii})$$

Solution of Eqs. (vii) to (xii)

Neglecting the term λ_6 , add Eqs. (vii) to (x) and subtract the resulting equation from $3 \times \text{Eq. (xi)}$ to get

$$\lambda_5 = \frac{3c_5 - (c_1 + c_2 + c_3 + c_4)}{8} = -3.775''$$

Substituting the value of λ_5 in Eq. (vii) to (x) and neglecting the term λ_6 , we get

$$\lambda_1 = -2.075'' \quad \lambda_3 = -0.408''$$

$$\lambda_2 = +4.692'' \quad \lambda_4 = -0.408''$$

Substituting these approximate values in Eq. (xii), we get

$$\lambda_6 = \frac{c_6 - \Sigma \lambda(d_1 - d_r)}{\Sigma d^2} = 0.0379$$

Second approximation using the value of λ_6 as 0.0379:

$$\begin{aligned}
c'_1 &= c_1 - \lambda_6(d_1 - d_2) \\
&= -10 - (0.0379)(-5.27) = -9.78 \\
c'_2 &= c_2 - \lambda_6(d_3 - d_4) \\
&= +10.3 - (0.0379)(4.21) = +10.14 \\
c'_3 &= c_3 - \lambda_6(d_5 - d_6) \\
&= -5 - (0.0379)(-13.44) = -4.49 \\
c'_4 &= c_4 - \lambda_6(d_7 - d_8) \\
&= -5 - (0.0379)(15.02) = -5.57 \\
\lambda_5 &= \frac{3c_5 - (c'_1 + c'_2 + c'_3 + c'_4)}{8} = -3.775''
\end{aligned}$$

Since the value of λ_5 turns out to be the same, hence second approximation is not necessary.

$$\begin{aligned}
\lambda_1 &= \frac{c'_1 - \lambda_5}{3} = -2.002'' \\
\lambda_2 &= \frac{c'_2 - \lambda_5}{3} = +4.638'' \\
\lambda_3 &= \frac{c'_3 - \lambda_5}{3} = -0.238'' \\
\lambda_4 &= \frac{c'_4 - \lambda_5}{3} = -0.598'' \\
\lambda_6 &= \frac{c_6 - \Sigma \lambda(d_1 - d_r)}{\Sigma d^2} = 0.0379
\end{aligned}$$

Hence, the most probable values of correction are:

$$\begin{aligned}
c_1 &= \lambda_1 + 25.01 \quad \lambda_6 = -2.002 + 25.01(0.0379) = -1.4331'' \\
c_2 &= \lambda_2 - 23.06 \quad \lambda_6 = 4.638 - 23.06(0.0379) = 3.764'' \\
c_3 &= \lambda_2 + 27.27 \quad \lambda_6 = 4.638 + 27.27(0.0379) = 5.6715'' \\
c_4 &= \lambda_3 - 25.43 \quad \lambda_6 = -0.238 - 25.43(0.0379) = -1.2018'' \\
c_5 &= \lambda_3 + 11.99 \quad \lambda_6 = -0.238 + 11.99(0.0379) = -0.2164'' \\
c_6 &= \lambda_4 - 18.83 \quad \lambda_6 = -0.598 - 18.83(0.0379) = -1.3116'' \\
c_7 &= \lambda_4 + 33.85 \quad \lambda_6 = -0.598 + 33.85(0.0379) = -0.6849'' \\
c_8 &= \lambda_1 - 20.73 \quad \lambda_6 = -2.002 - 20.73(0.0379) = -2.78766'' \\
c_9 &= \lambda_2 + \lambda_5 = 4.638 - 3.775 = 0.863'' \\
c_{10} &= \lambda_3 + \lambda_5 = -0.238 - 3.775 = -4.013'' \\
c_{11} &= \lambda_4 + \lambda_5 = -0.598 - 3.775 = -4.373'' \\
c_{12} &= \lambda_1 + \lambda_5 = -2.002 - 3.775 = -5.777''
\end{aligned}$$

3.16 METHOD OF EQUAL SHIFTS

The method of equal shifts indicates that any shift which is necessary to satisfy the local equation should be the same for each triangle of the polygon; and any shift necessary to satisfy the side equation should be the same for each triangle.

The method of equal shifts not only simplifies the computation, but also renders the method systematic, though unfortunately the method is devoid of the mathematical justification, so far as theory of errors is concerned. In this method, the discrepancy in the angular measurements is equally divided between the angles of a triangle.

The condition equations are the same as those described in Section 3.15. The method is explained with the help of the following steps.

1. The various triangles formed in a polygon with a central station, are numbered (say) 1, 2, 3 and so on.
2. Find the sum of the observed angles of various triangles and calculate the discrepancy of the angles over the theoretical sum (180°). Distribute one-third of the total discrepancy in each of the central angles of each triangle in such a way so as to make the sum of the angles of the triangle equal to 180° .
3. Find the sum of the observed central angles and calculate the discrepancy over 360° . Calculate the correction to be applied to the central angles.
4. Find the algebraic sum of the respective corrections obtained for the central angles in steps 2 and 3. Calculate the required shift (*first trial correction*) to each of the central angles so as to make the sum of the central angles 360° . The shift being equal to all these angles.
5. The first trial correction in the central angles so obtained is deducted from the total correction in the corresponding triangle and half of the remainder will be the trial corrections to the left- and right-hand angles of each triangle. While looking from the central station to the middle of the side of a triangle, angles of the triangle on the left-hand side are designated as left-hand angles and those on the right as right-hand angles.
6. Find the difference between the summation of the log sines of the left- and right-hand angles. The effect of the trial corrections obtained in Step 5 is checked to see whether these help to satisfy the side equation. A correction multiplied by the difference of $\log \sin 1''$ for the corresponding angle gives the corrected difference. Hence, obtain the corrected difference for each right- and left-hand angle and check how far the trial corrections to the left- and right-hand angles have helped to satisfy the side equation.
7. Find the algebraic sum of the log difference for $\sin 1''$ of left- and right-hand angles. This value reflects that a shift of $1''$ in each angle of a triangle from left to right and vice versa will affect their log sines by that amount. Find out proportionately the shift necessary to account for the balance of the difference in the summation of the log sines of the left- and right-hand angles to satisfy the side equation. Add the shift to all those angles, the sum of whose log sines is less and deduct from those for which it is greater.
8. Check whether the final corrections after applying the shift in Step 7, when multiplied by $\log \sin 1''$ corresponding to the angles to which the corrections are applied, make the sum of the log sines of the left-hand angles equal to log sines of the right-hand angles.

Example 3.31 Solve Example 3.30 by the method of equal shifts.

Solution

<i>Triangle</i>	<i>Sum of observed angles</i>	<i>Total corrections (a)</i>
<i>AOB</i>	179°59'49.7"	+10.3"
<i>BOC</i>	180°00'05"	–5"
<i>COD</i>	180°00'05"	–5"
<i>DOA</i>	180°00'10"	–10"

One-third of the total corrections will be the correction for the corresponding central angle, as tabulated below. The sum of the Ist trial corrections is – 3.23". Correction to the sum of central angles is – 13.3" (Example 3.30).

So, to make corrections to – 13.3", apply a correction of

$$- \frac{(13.3'' - 3.23'')}{4} = - 2.52''$$

to each Ist trial correction.

<i>Triangle</i>	<i>Ist trial correction (a/4)</i>	<i>IInd trial correction (b)</i>	<i>Corrected central angles</i>	<i>Remaining correction for angles (a – b)</i>
<i>AOB</i>	3.42"	0.92"	99°56'03.92"	9.38"
<i>BOC</i>	– 1.67"	– 4.19"	80°03'51.47"	– 0.81"
<i>COD</i>	– 1.67"	– 4.19"	99°56'11.5"	– 0.81"
<i>DOA</i>	– 3.33"	– 5.84"	80°03'53.1"	– 4.16"
	– 3.23"	–13.3"	360°00'00"	

This remaining correction is to be distributed equally to the left- and right-hand angles of each triangle.

Knowing the difference (d) for 1", find the corresponding differences for the corrections.

<i>Triangle</i>	<i>Trial correction for angles</i>		<i>Trial correction difference</i>	
	<i>Left</i>	<i>Right</i>	<i>Left</i>	<i>Right</i>
<i>AOB</i>	+ 4.69"	+ 4.69"	108.15	127.9
<i>BOC</i>	– 0.405"	– 0.405"	– 10.303	– 4.86
<i>COD</i>	– 0.405"	– 0.405"	– 7.63	–13.713
<i>DOA</i>	– 2.08"	– 2.08"	– 43.14	– 31.24
			47.077	78.087

$$\begin{aligned} \text{Total shift required} &= (39.358459 + 47.077) - (39.358443 + 78.087) \\ &= - 31.009 = 31 \text{ from right to left} \end{aligned}$$

Hence, the left-hand angles will have to be increased and the right-hand angles will have be decreased.

$$\Sigma d = 176.17 \quad (\text{from Example 3.30})$$

$$\therefore \frac{31}{176.17} = 0.1759 = 0.176''$$

Hence, increase the left-hand angles by $0.176''$ and decrease the right-hand angles by $0.176''$.

<i>Triangle</i>	<i>Final corrections</i>	
	<i>Left</i>	<i>Right</i>
<i>AOB</i>	4.866''	4.514''
<i>BOC</i>	– 0.229''	– 0.581''
<i>COD</i>	– 0.229''	– 0.581''
<i>DOA</i>	– 1.904''	– 2.256''

<i>Corrected angles</i>			
<i>Left</i>	<i>Right</i>	<i>Central</i>	<i>Sum</i>
42°23'17.3''	37°40'17''	99°56'3.92''	180°00'00''
39°36'46.37''	60°19'22.22''	80°3'51.47''	180°00'00''
48°11'17.57''	31°52'30.92''	99°56'11.5''	180°00'00''
45°26'6.39''	54°30'0.44''	80°3'53.1''	180°00'00''

3.17 METHOD OF LEAST SQUARES BY MATRICES

When the observations are of equal weight, Σr^2 is to be a minimum. For m such observations

$$r_1^2 + r_2^2 + r_3^2 + \dots + r_m^2$$

is to be a minimum.

In matrix form, this can be written as

$$(r_1, r_2, r_3, \dots, r_m) \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ \vdots \\ r_m \end{pmatrix}$$

If the column vector is denoted by r , the sum of the squares of the residuals is then $r^T r$. Now assume that n sets of independent unknowns $X_1, X_2 \dots X_n$ have been related to a set of observations $k, k_2 \dots k_m$ by the following observational equations, m being larger than n :

$$a_{11} X_1 + a_{12} X_2 + a_{13} X_3 + \dots + a_{1n} X_n = k_1$$

$$\begin{array}{ccccccc} a_{21} X_1 + a_{22} X_2 + a_{23} X_3 + \cdots + a_{2n} X_n = k_2 \\ \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ a_{m1} X_1 + a_{m2} X_2 + a_{m3} X_3 + \cdots + a_{mn} X_n = k_m \end{array}$$

These equations can be written in the form $\mathbf{AX} = \mathbf{K}$, where $\mathbf{A}$ is the matrix of the coefficients and is of the order $m \times n$, and the residual $\mathbf{r}$ can be expressed as

$$\mathbf{r} = -\mathbf{AX} + \mathbf{K}$$

so that the sum of the squares is

$$(-\mathbf{AX} + \mathbf{K})^T (-\mathbf{AX} + \mathbf{K})$$

Now

$$(-\mathbf{AX} + \mathbf{K})^T (-\mathbf{AX} + \mathbf{K}) = \mathbf{X}^T \mathbf{A}^T \mathbf{AX} - \mathbf{K}^T \mathbf{AX} - \mathbf{X}^T \mathbf{A}^T \mathbf{K} + \mathbf{K}^T \mathbf{K}$$

For this to be a minimum, differentiate with respect to $\mathbf{X}$, and equate to zero to obtain

$$2\mathbf{A}^T \mathbf{AX} - \mathbf{A}^T \mathbf{K} - \mathbf{A}^T \mathbf{K} = 0$$

$$\mathbf{A}^T \mathbf{AX} = \mathbf{A}^T \mathbf{K}$$

It will be noted that the observation equations have been premultiplied on each side by $\mathbf{A}^T$, and that

$$\mathbf{X} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{K}$$

When the observations are not of equal weight, the weight matrix ($\mathbf{W}$) has to be introduced since

$$w_1 r_1^2 + w_2 r_2^2 + w_3 r_3^2 + \cdots + w_m r_m^2$$

is to be a minimum. This expression can be written as

$$(r_1, r_2, r_3, \dots, r_m) \begin{pmatrix} w_1 & 0 & 0 & \dots & 0 \\ 0 & w_2 & 0 & \dots & 0 \\ 0 & 0 & w_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & w_m \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ \vdots \\ r_m \end{pmatrix}$$

It will be noted that $\mathbf{W}$ is symmetric. The above expression may be written as

$$\mathbf{R}^T \mathbf{WR}$$

or

$$(-\mathbf{AX} + \mathbf{K})^T \mathbf{W}(-\mathbf{AX} + \mathbf{K})$$

For this to be a minimum check that

$$\mathbf{A}^T \mathbf{WAX} = \mathbf{A}^T \mathbf{WK}$$

Both sides of the observation equation have been premultiplied by $\mathbf{A}^T \mathbf{W}$.

As an alternative

$$(w_1^{1/2} r_1)^2 + (w_2^{1/2} r_2)^2 + \cdots + (w_m^{1/2} r_m)^2$$

can be minimised.

If we denote the matrix of the square roots of the weights as $\mathbf{V}$, then

$$\mathbf{V} = \begin{pmatrix} \sqrt{w_1} & 0 & 0 \dots 0 \\ 0 & \sqrt{w_2} & 0 \dots 0 \\ \cdot & \cdot & \cdot \cdot \\ \cdot & \cdot & \cdot \cdot \\ \cdot & \cdot & \cdot \cdot \\ 0 & 0 & 0 \dots \sqrt{w_m} \end{pmatrix}$$

We now minimise $(\mathbf{V}\mathbf{R})^T (\mathbf{V}\mathbf{R})$ or $(-\mathbf{V}\mathbf{A}\mathbf{X} + \mathbf{V}\mathbf{K})^T (-\mathbf{V}\mathbf{A}\mathbf{X} + \mathbf{V}\mathbf{K})$ which produces the previous result of

$$\mathbf{A}^T \mathbf{W} \mathbf{A} \mathbf{X} = \mathbf{A}^T \mathbf{W} \mathbf{K}$$

and

$$\mathbf{X} = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} \mathbf{K}$$

Example 3.32 In making a round of levels $ABCD$ (Fig. 3.7) from A the following results were obtained:

A 6.92 m above B (mean of two measurements)

B 7.26 m below C

C 2.91 m above D

D 2.60 m below A (mean of two measurements)

C 0.37 m above A

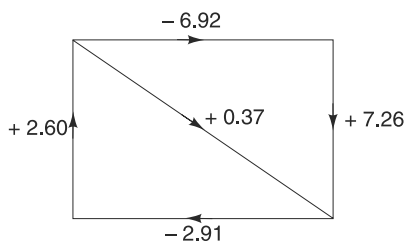


Fig. 3.7

Compute the most probable heights of B , C and D above A using the method of least squares

- direct
- matrices.

(a) Using least squares direct If the most probable heights of B , C , D above A are b , c , d , then the errors are $(-6.92 - b)$, $(7.26 - c - b)$, $(-2.91 - d - c)$, $(2.60 - (-d))$, $(0.37 - c)$. In drawing up a table of coefficients the lines AB and DA must be given a weight of 2.

Weight	C_b	C_c	C_d	N
2	-1	0	0	-6.92
1	+1	-1	0	+7.26
1	0	+1	-1	-2.91
2	0	0	+1	+2.60
1	0	-1	0	+0.37

Then by the principles of least squares,

$$2 \times (-6.92 - b)^2 + (7.26 - c + b)^2 + \dots = \text{minimum}$$

Differentiating partially with respect to b , and dividing by 2,

$$2 \times (-1) \times (-6.92 - b) + (7.26 - c + b) + \dots = 0$$

or more conveniently from the table,

$$b \times (+2 + 1 + 0 + 0 + 0) + c \times (-1 + 0 + 0 + 0 + 0) + d \times (0) + (2 \times 6.92 + 1 \times 7.26 + 0 + 0 + 0) = 0$$

i.e., $+3b - c + 21.10 = 0$

Similarly, differentiating partially with respect to c ,

$$b \times (0 - 1 + 0 + 0 + 0) + c \times (0 + 1 + 1 + 0 + 1) + d \times (0 + 0 - 1 + 0 + 0) + (-7.26 - 2.91 + 0 - 0.37) = 0$$

i.e., $-b + 3c - d - 10.54 = 0$

Similarly, differentiating partially with respect to d ,

$$b \times (0) + c \times (-1 + 0 + 0) + d \times (1 + 2) + (2.91 + 2 \times 2.60) = 0$$

i.e., $-c + 3d + 8.1 = 0$

Thus we have the normal equations,

$$\begin{aligned} +3b - c + 21.10 &= 0 \\ -b + 3c - d - 10.54 &= 0 \\ -c + 3d + 8.11 &= 0 \end{aligned}$$

Hence,

$$\begin{aligned} b &= -6.919 \text{ m, } \approx -6.92 \text{ m} \\ c &= +0.343 \text{ m, } \approx +0.34 \text{ m} \\ d &= -2.589 \text{ m, } \approx -2.59 \text{ m} \end{aligned}$$

(b) Using matrices The error given in method (a) using least squares direct imply the following observation equations for the unknown b , c and d :

$$\begin{aligned} b &= -6.92 \\ (c - b) &= +7.26 \\ (d - c) &= -2.91 \\ -d &= +2.60 \\ c &= +0.37 \end{aligned}$$

thus

$$\mathbf{A} = \begin{pmatrix} +1 & 0 & 0 \\ -1 & +1 & 0 \\ 0 & -1 & +1 \\ 0 & 0 & -1 \\ 0 & +1 & 0 \end{pmatrix}$$

$$\mathbf{W} = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{A}^T \mathbf{W} = \begin{pmatrix} +1 & -1 & 0 & 0 & 0 \\ 0 & +1 & -1 & 0 & +1 \\ 0 & 0 & +1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} +2 & -1 & 0 & 0 & 0 \\ 0 & +1 & -1 & 0 & +1 \\ 0 & 0 & +1 & -2 & 0 \end{pmatrix}$$

$$\mathbf{A}^T \mathbf{W} \mathbf{A} = \begin{pmatrix} +2 & -1 & 0 & 0 & 0 \\ 0 & +1 & -1 & 0 & +1 \\ 0 & 0 & +1 & -2 & 0 \end{pmatrix} \begin{pmatrix} +1 & 0 & 0 \\ -1 & +1 & 0 \\ 0 & -1 & +1 \\ 0 & 0 & -1 \\ 0 & +1 & 0 \end{pmatrix} = \begin{pmatrix} +3 & -1 & 0 \\ -1 & +3 & -1 \\ 0 & -1 & +3 \end{pmatrix}$$

$$(\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} = \frac{1}{\begin{vmatrix} +3 & 1 & 0 \\ -1 & +3 & -1 \\ 0 & -1 & +3 \end{vmatrix}} \text{adj} \begin{bmatrix} +3 & 1 & 0 \\ -1 & +3 & -1 \\ 0 & -1 & +3 \end{bmatrix} = \begin{bmatrix} 8/21 & 1/7 & 1/21 \\ 1/7 & 3/7 & 1/7 \\ 1/21 & 1/7 & 8/21 \end{bmatrix}$$

$$\mathbf{A}^T \mathbf{W} \mathbf{K} = \begin{pmatrix} +2 & -1 & 0 & 0 & 0 \\ 0 & +1 & -1 & 0 & +1 \\ 0 & 0 & +1 & -2 & 0 \end{pmatrix} \begin{pmatrix} -6.92 \\ +7.26 \\ -2.91 \\ +2.60 \\ +0.37 \end{pmatrix} = \begin{pmatrix} -21.10 \\ +10.54 \\ -8.11 \end{pmatrix}$$

$$\mathbf{X} = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} \mathbf{K}$$

$$\therefore \begin{pmatrix} b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 8/21 & 1/7 & 1/21 \\ 1/7 & 3/7 & 1/7 \\ 1/21 & 1/7 & 8/21 \end{pmatrix} \begin{pmatrix} -21.10 \\ +10.54 \\ -8.11 \end{pmatrix}$$

gives

$$b = -6.92 \text{ m}$$

$$c = +0.34 \text{ m}$$

$$d = -2.59 \text{ m}$$

Example 3.33 The following angular observations were made at a certain station:

$$A = 32^\circ 02' 41'' \quad \text{weight 1}$$

$$B = 24^\circ 21' 15'' \quad \text{weight 1}$$

$$A + B = 56^\circ 23' 54'' \quad \text{weight 2}$$

$$C = 47^\circ 12' 31'' \quad \text{weight 1}$$

$$B + C = 71^\circ 33' 49'' \quad \text{weight 2}$$

Find the most probable values of A , B and C , using matrices.

Solution In this case

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad W = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

$$K = \begin{pmatrix} 32^\circ 02' 41'' \\ 24^\circ 21' 15'' \\ 56^\circ 23' 54'' \\ 47^\circ 12' 31'' \\ 71^\circ 33' 49'' \end{pmatrix}$$

$$A^T W = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}$$

$$A^T W A = \begin{pmatrix} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 0 \\ 2 & 5 & 2 \\ 0 & 2 & 3 \end{pmatrix}$$

$$A^T W K = \begin{pmatrix} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 32^\circ 02' 41'' \\ 24^\circ 21' 15'' \\ 56^\circ 23' 54'' \\ 47^\circ 12' 31'' \\ 71^\circ 33' 49'' \end{pmatrix}$$

$$= \begin{pmatrix} 144^\circ 50' 29'' \\ 280^\circ 16' 41'' \\ 190^\circ 20' 09'' \end{pmatrix}$$

$$X = [A^T W A]^{-1} [A^T W K]$$

$$(A^T W A)^{-1} = \frac{1}{\begin{vmatrix} 3 & 2 & 0 \\ 2 & 5 & 2 \\ 0 & 2 & 3 \end{vmatrix}} \quad \text{adj} \begin{pmatrix} 3 & 2 & 0 \\ 2 & 5 & 2 \\ 0 & 2 & 3 \end{pmatrix}$$

$$= 1/21 \begin{pmatrix} 11 & -6 & 4 \\ -6 & 9 & -6 \\ 4 & -6 & 11 \end{pmatrix}$$

$$[A^T W A]^{-1} [A^T W K] = 1/21 \begin{pmatrix} 11 & -6 & 4 \\ -6 & 9 & -6 \\ 4 & -6 & 11 \end{pmatrix} \begin{pmatrix} 144^\circ 50' 29'' \\ 280^\circ 16' 41'' \\ 190^\circ 20' 09'' \end{pmatrix} = \begin{pmatrix} 32^\circ 2' 39.48'' \\ 24^\circ 21' 15.29'' \\ 47^\circ 12' 32.81'' \end{pmatrix}$$

Hence

$$\begin{pmatrix} A \\ B \\ C \end{pmatrix} = \begin{pmatrix} 32^\circ 2' 39.48'' \\ 24^\circ 21' 15.29'' \\ 47^\circ 12' 32.81'' \end{pmatrix}$$

Summary

Despite of exercising all the precision and precautions, errors are bound to creep in the surveying measurements. This chapter deals with assessment and distribution of such errors; errors can not be eliminated, but can be minimised. In surveying it is usual to take multiple observations for individual measurements to reduce error. The concept of assigning weight to each measurement and laws governing them are presented. These laws are explained with the help of examples.

Probability theory governing the accidental errors is described in a little detail as it forms base for sound understanding of theory of errors. Principle of least squares, its application and method of distributing errors of the field measurements are presented. The concept of formation of normal equations is discussed. This technique is important when some of the observations are made directly and the others are derived using the geometrical conditions imposed by the geometry of the figure.

To compute the coordinates of the triangulation stations, the observations made are adjusted by distributing the observational errors. The method of least squares for minimizing the errors is explained. Since the method is complex in application and cumbersome to use, an approximate method consisting in adjustment of angles, stations, and figures is presented and elaborated with the help of solved examples. The chapter ends with description of method of least squares by matrices and method of equal shifts.

Exercises

- 3.1 Explain what is meant by the probable error of a measurement and state what you consider to be the suitable values for the probable error of the following operations when conducted with the utmost refinement:
 - (a) The measurement of a base line 8 km long.
 - (b) Running a line of precise level 80 km long.

List the individual errors which contribute to the total probable error in the above operations.

- 3.2** Define mean square error, residual error and weight. From the definitions deduce the mean square error E_m of an observation as deduced from a series of observations on a single quantity as

$$E_m = \pm 0.6745 \sqrt{\frac{\sum v^2}{n(n-1)}}$$

where v = residual error, and n = number of observations.

- 3.3** A rectangular plot of land has sides 60.45 ± 0.02 m and 50.84 ± 0.03 m. What value should be assigned to its area?

[Ans. 3073.28 ± 2.08 m]

- 3.4** Enumerate the principle of least squares. Show how this principle is used for determining two unknowns in linear equations. Find expressions for the unknowns in terms of the coefficients and measured quantities.

- 3.5** Define (a) spherical excess, (b) accidental and systematic errors and (c) condition equation.

- 3.6** Detail the process of adjustment in accordance with the theory of least squares in the case of a triangle subdivided into triangles by means of a central station, the observations at the four stations being of equal weight.

- 3.7** In order to investigate precision of changing, a line 500 m in length was measured with an invar precision tape under the same conditions, the discrepancy from 500 m being as follows:

+ 0.34, - 0.45, + 0.32, + 0.61, - 0.11, 0, + 0.21, - 0.56, - 0.18

Determine the most probable length of the line and the probable error of a single measurement.

[Ans. $L = 500 \pm 0.2603$, $E_s = 0.2603$]

- 3.8** Explain clearly what is meant by side equation. How would you adjust a geodetic quadrilateral?

- 3.9** Find the most probable values of angles A , B and C from the following data using the method of correlates:

Angle	Value	Weight
A	$32^\circ 02' 41''$	1
B	$24^\circ 21' 15''$	1
$A + B$	$56^\circ 23' 54''$	2
C	$47^\circ 12' 31''$	1
$B + C$	$71^\circ 33' 49''$	2

[Ans. $32^\circ 02' 39.48''$, $24^\circ 21' 15.28''$, $47^\circ 12' 32.8''$]

- 3.10** Find the most probable values of the angles α , β and γ from the following observations at one station:

$\alpha = 38^\circ 12' 26.5''$ Weight = 1

$\beta = 32^\circ 45' 13.2''$ Weight = 2

$\alpha + \beta = 70^\circ 57' 38.6''$ Weight = 2

$\alpha + \beta + \gamma = 126^\circ 28' 0.6''$ Weight = 3

$\beta + \gamma = 88^\circ 15' 37.8''$ Weight = 1

[Ans. $\alpha = 38^\circ 12' 25.09''$, $\beta = 32^\circ 45' 13.35''$, $\gamma = 55^\circ 30' 22.73''$]

- 3.11 Observations were made for two angles α and β as follows:

$$\alpha = 20^\circ 10' 10'' \quad \text{Weight} = 6$$

$$\beta = 30^\circ 20' 30'' \quad \text{Weight} = 4$$

$$\alpha + \beta = 50^\circ 30' 50'' \quad \text{Weight} = 2$$

Determine the adjusted values of α and β .

$$[\text{Ans. } \alpha = 20^\circ 10' 11.82'', \beta = 30^\circ 20' 32.73'']$$

- 3.12 The following mean values of the three angles of a triangle were observed

$$\angle A = 54^\circ 12' 25'' \quad \text{Weight} = 8$$

$$\angle B = 48^\circ 46' 16.25'' \quad \text{Weight} = 4$$

$$\angle C = 77^\circ 02' 10.83'' \quad \text{Weight} = 6$$

Determine the corrected values of the angles by the method of least squares.

$$[\text{Ans. } \angle A = 54^\circ 12' 12.98'', \angle B = 48^\circ 45' 52.21'', \angle C = 77^\circ 01' 54.81'']$$

- 3.13 While measuring horizontal angles by a theodolite at station O , the following observations were recorded:

$$\angle AOB = \alpha = 28^\circ 01' 28.2''$$

$$\angle BOC = \beta = 37^\circ 59' 12.6''$$

$$\angle AOC = \gamma = 66^\circ 00' 44.4''$$

Determine the most probable values of α , β and γ .

Note | Angles are subjected to the condition $\gamma = \alpha + \beta$.

$$[\text{Ans. } \alpha = 28^\circ 11' 29.4'', \beta = 37^\circ 59' 13'', \gamma = 66^\circ 00' 43.2'']$$

- 3.14 In Fig. 3.8, α and β were measured as $4^\circ 45' 25''$ and $3^\circ 04' 42''$, respectively, when AB was 2 m long. Find the probable error in d given that probable errors of ± 1 mm in the length of AB and of $\pm 2''$ in the measurement of each angle were obtained.

$$[\text{Ans. } 0.034 \text{ m}]$$

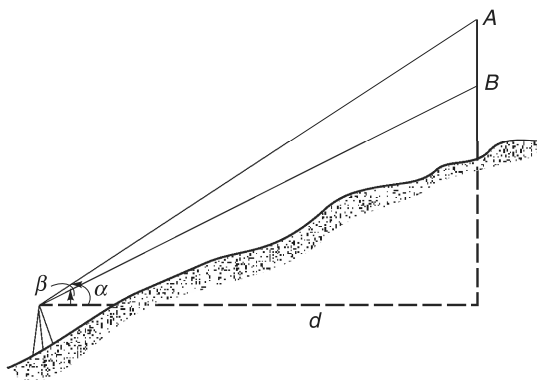


Fig. 3.8

- 3.15 Find the most probable values of angles of a quadrilateral $ABCD$ with O as the central station. The data for the quadrilateral $ABCD$ is given below:

Triangle	Central angles	Left-hand angles	Right-hand angles
AOB	$110^\circ 45' 20''$	$26^\circ 12' 20''$	$43^\circ 02' 35''$
BOC	$79^\circ 30' 20''$	$54^\circ 00' 40''$	$46^\circ 28' 35''$
COD	$83^\circ 25' 10''$	$51^\circ 50' 10''$	$44^\circ 45' 10''$
DOA	$86^\circ 19' 10''$	$53^\circ 21' 05''$	$40^\circ 19' 30''$

Ans.

Central angles	L.H. angles	R.H. angles
$- 5''$	$+ 4.95''$	$- 14.95''$
$+ 10''$	$+ 19.95''$	$+ 0.05''$
$- 10''$	$- 0.05''$	$- 19.95''$
$+ 5''$	$+ 14.95''$	$- 4.95''$

3.16 Solve Question 3.15 by the method of equal shifts.

3.17 A round of levels *ABCDEA* gave the following average rises, all of equal accuracy, for successive lines:

$+ 3.27, + 10.21, - 11.34, + 2.93, - 5.21$ (in metres)

Determine the most probable levels of *B, C, D, E*, above *A* using

- (a) Method of least squares
(b) Method of correlates

[Ans. (a) 3.30 m, 13.54 m, 2.22 m, 5.18 m;
(b) 3.298 m, 13.536 m, 2.224 m, 5.182 m]

Objective-type Questions

- 3.1** An angle measured with a theodolite is α with weight 2. The weight of $\alpha/4$ will be
(a) $2/4$ (b) 2×4 (c) 2×4^2 (d) $4/2$
- 3.2** The weight of an angle α is 2. The weights of 2α will be
(a) $2/2$ (b) $2^2/2$ (c) $2/\sqrt{2}$ (d) $2/2^2$
- 3.3** The most probable values of two angles are $40^\circ 20' 30'' \pm 2''$ and $84^\circ 44' 20'' \pm 3''$. The probable error of the sum of these two angles will be
(a) ± 5 (b) ± 1 (c) $\pm \sqrt{5}$ (d) $\pm \sqrt{2^2 + 3^2}$
- 3.4** The radius of a circle is 10 m with a probable error of ± 0.2 . The probable error of its circumference is
(a) $\pm 0.2 \times 2\pi$ (b) $\pm 0.2^2 \times 2\pi$ (c) ± 0.2 (d) $\pm 2\pi/0.2$
- 3.5** The two sides of a rectangle are 50 ± 0.02 m and 60 ± 0.03 m. The probable error of the area of the rectangle is
(a) ± 0.05 (b) $\sqrt{(50 \times 0.02)^2 + (60 \times 0.03)^2}$
(c) $\sqrt{0.02^2 + 0.03^2}$ (d) $\sqrt{(0.02 \times 60)^2 + (0.03 \times 50)^2}$
- 3.6** The probable error of a line 250 m long was ± 0.02 m. The probable error of another line 500 m long under the same conditions will be
(a) $\pm .02^2 \times \frac{500}{250}$ (b) $\pm 0.02 \times \frac{250}{500}$
(c) $\pm .02 \times \frac{500}{250}$ (d) $\pm 0.02^2 \times \frac{250}{500}$
- 3.7** The most probable value of angle *A* from the following observation equations is
 $A = 30^\circ 00' 20''$ Weight = 2
 $2A = 60^\circ$ Weight = 1
 (a) $30^\circ 00' 6.66''$ (b) $30^\circ 00' 10''$ (c) $30^\circ 00' 2.22''$ (d) $30^\circ 00' 30''$

- 3.8** An angle measured by two observers has the following values:

Angle	No. of observations
$30^{\circ}00'20''$	2
$30^{\circ}00'10''$	4

The most probable angle will be

- (a) $30^{\circ}00'10''$ (b) $30^{\circ}00'16.66''$
 (c) $30^{\circ}00'13.33''$ (d) $30^{\circ}00'18''$
- 3.9** The most probable value of angle A from the following observation equation is
- $$A = 30^{\circ}00'20'' \quad \text{weight} = 2$$
- $$2A = 60^{\circ}00'00'' \quad \text{weight} = 1$$
- (a) $30^{\circ}00'6.66''$ (b) $30^{\circ}00'10''$ (c) $30^{\circ}00'2.22''$ (d) $30^{\circ}00'30''$
- 3.10** The most probable value of an angle was $39^{\circ}29'30'' \pm 3''$. The value of its supplement will be
- (a) $140^{\circ}30'30'' \pm 3''$ (b) $320^{\circ}30'30'' \pm 3''$
 (c) $140^{\circ}30'30'' \pm 9''$ (d) $140^{\circ}30'30'' \pm 1$
- 3.11** Residual error of a measurement is the difference of
- (a) true and observed values (b) most probable and observed values
 (c) true and most probable values (d) standard and probable values
- 3.12** If a quantity of a given weight is multiplied by a factor, the weight of the result is obtained by dividing its given weight by
- (a) the factor (b) the square of the factor
 (c) the square root of the factor (d) the cube root of the factor
- 3.13** Errors arising from carelessness of the observer are known as
- (a) compensating errors (b) systematic errors
 (c) mistakes (d) discrepancy
- 3.14** In the case of line of levels, the weight varies
- (a) directly as the length of the route
 (b) inversely as the length of the route
 (c) directly as square root of the length of the route
 (d) inversely as square root of the length of the route
- 3.15** The method of least squares deals with discrepant observations having
- (a) negative cumulative errors (b) positive cumulative errors
 (c) compensating errors (d) systematic errors
- 3.16** If an equation is subtracted from a constant K , the weight of the resulting equation will be
- (a) weight of the equation divided by K
 (b) weight of the equation multiplied by K
 (c) weight of the equation multiplied by K^2
 (d) unchanged
- 3.17** Which of the following method is known as a rigid method of figure adjustment?
- (a) Method of differences (b) Method of least squares
 (c) Method of correlates (d) Direct method
- where A is the area of a spherical triangle and R is the radius of earth.
- 3.18** In the triangulation adjustment for closing the horizon condition, the error is distributed
- (a) equally to all the angles
 (b) inversely as the respective weights

- (c) by forming and solving normal equations
 (d) All the above are true
- 3.19** Log sine correction is made in triangulation to satisfy
 (a) apex condition (b) opposite angles
 (c) side condition (d) peripheral sum
- 3.20** The rules to be observed while assigning weights to the observations made are:
 (i) The weight of an angle varies directly as the number of observations made on an angle.
 (ii) For an angle repeated for a great number of times, the weight is inversely proportional to the square of the probable error.
 (iii) For a line of levels, the weight varies inversely as the length of the route.
 Of the above rules, the correct ones are:
 (a) only (i) (b) (i) and (ii) (c) (i), (ii) and (iii) (d) (ii) and (iii)

Answers to Objective-type Questions

3.1 (c)	3.2 (d)	3.3 (d)	3.4 (a)	3.5 (d)	3.6 (c)
3.7 (a)	3.8 (c)	3.9 (a)	3.10 (a)	3.11 (b)	3.12 (b)
3.13 (c)	3.14 (b)	3.15 (c)	3.16 (d)	3.17 (b)	3.18 (a)
3.19 (c)	3.20 (c)				

4

HYDROGRAPHY

Introduction

It is the branch of surveying which deals with water bodies, e.g., bay, harbour, lake, river, etc. The usual fundamental principles of surveying and levelling are adopted for acquiring data for determination of water areas, volumes, rate of flow and to determine the shape of the area underlying the water surface.

The purpose for making the hydrographic survey is to (i) determine the quantities of subsqueous excavations, (ii) measure areas subjected to scouring or silting in harbours and docks, (ii) locate rocks and other objects such as buoys, lighthouses, etc., to aid safe navigation, (iv) prepare navigation charts exhibiting the depths available for navigation, (v) control floods, and to plan water supply and storage from rivers, and (vi) to develop water resources for power, irrigation and recreation.

Hydrographic surveying consists both on-shore and off-shore surveying. In on-shore survey, the field work consists of both horizontal as well as vertical control. The horizontal control is established by traversing or triangulation. In case of narrow rivers, control is achieved by traversing on the shore, whereas a chain of triangles or network of triangles is most suited for large water bodies (more than 7300 m wide). Where the shoreline is obstructed by woods, a system of triangulation is the only choice. If the shore is inaccessible, a base is measured on the water, usually by fixing its ends by intersection from the shore, and then the shoreline is surveyed by the usual methods. For vertical control, the tide gauges are kept in operation continuously since the water level at the gauge must also be known when the soundings are recorded. It is also necessary to connect all the tide gauges set on the various points by vertical control prior to making the soundings. As such, a number of benchmarks are established near the shoreline at close spacing to serve as vertical control. These benchmarks are used for setting and checking the levels of gauges to which the soundings are referred. All the irregularities in the shoreline, prominent features of topography, lighthouses, buoys, etc., are established so as to use them for references in sounding work.

Off-shore surveying may be categorized as measurement of vertical depth (sounding), horizontal control of the sounding position (locating sounding) reduction of sounding to a datum and finally plotting the sounding. The measurement of

depth of water at various points is termed *sounding*. In other words, the depth of sounding is referred to the water level at the time it is made. Thereafter, the soundings are reduced to the datum water level, to account for the tidal waters which undergo continual change of elevation, with the help of gauges. As already stated in the above paragraph, a number of benchmarks (B.M.) are established at frequent intervals along the shorelines, and gauges are set on them. These are so set that the surface level of water may be read on the scale at any time. By observations, the zero of the scale is fixed relative to the low water level of the spring tides or in some cases, where a survey is to be joined with the topographic one, to any arbitrary datum and can be related to the nearest B.M. by accurate levelling.

The main object of hydrographic survey is showing the topographic configuration of the underwater features. Hence, the underwater topographic map produced is different than the topographic maps of land survey because of the reasons to follow. Moreover, they are less accurate than the land survey topographic maps.

1. The depth of the sea bed below the survey vessel, known as sounding, is subject to variations due to wave and tidal action, though it is corrected after making sounding and is known as *reduced sounding*.
2. The position of a sounding vessel is fixed in the horizontal plane known as *locating sounding*, which can not be achieved accurately due to poor weather conditions such as winds, waves, and fog.
3. The land surveyor can see the features while the hydrographic surveyor can not.

The methods used for plotting the soundings are similar to those used for plotting details in topographic surveying. The soundings are usually represented in metres and sometimes in fathoms (measure of depth used in nautical charts: 1 fathom = 6 ft Appendix I). Where the mean low water level is datum the soundings below the datum are lettered in black whereas those above the datum, in any other colour. The figures are written with their decimal points exactly coinciding with the sounding points. From the plotted soundings, submarine contours are drawn similar to those on a topographic map. The colour used is blue. Marsh and beach areas are shown green and the land areas in yellow colour. The high water line on the finished map is drawn as the heaviest with the low water line next to it. Lighthouses and buoys are shown by the conventional signs.

4.1 TIDES

These are periodical variations in the water surface of oceans due to the attraction of celestial bodies. The principal tide-producing agents are the sun and the moon of which the moon is a more powerful tide producer.

Tides produced by unbalanced attracting forces between the moon and earth are known as *lunar tides* and those by the unbalanced attracting forces between the sun and earth are known as *solar tides*. Since the sun and the moon act simultaneously, the lunar and solar tides are superimposed, and the ocean assumes an equilibrium figure in obedience to the combined tide-producing force.

Tidal theory is based on the equilibrium theory of Newton. Newton's law of universal gravitation states that every body in the universe attracts every other body

with a force proportional to the product of their masses and inversely proportional to the square of the distance between them. The force acts in a straight line between the bodies.

4.1.1 Spring and Neap Tides

At new moon, the sun and moon have the same celestial longitude and cross a meridian of the earth at the same instant. The three bodies are in one plane, with the sun and moon on the same side of the earth (Fig. 4.1), the crests of the lunar and solar tides coincide. The high water level of the resulting tide is above the average, whereas the low water level is below the average. The tide is known as *spring tide of new moon*. In practise, the tide occurs for some small interval (1-2 days) after the theoretical time; this is called the *age of the tide*. Equinoctical spring tides are exceptionally high and occur during the equinoxes when the sun and moon are vertically over the equator.

Thereafter, the moon falls behind the sun. In about 7.5 days, when the moon is in quadrature (Fig. 4.2), the crest of a lunar tide coincides with the trough of a solar tide. High water level is below the average, whereas low water level is above the average. Such a tide is known as *neap tide of the first quarter*.

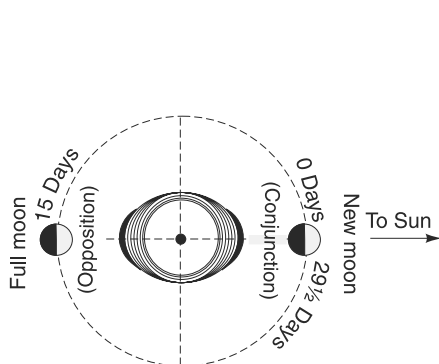


Fig. 4.1 Spring tides

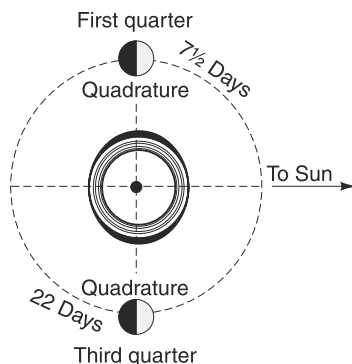


Fig. 4.2 Neap tides

Then, in about 15 days, when the moon's elongation is 180° , the three bodies are in the same plane with the sun and moon on the opposite side of the earth. The resulting tide is known as *spring tide of full moon*. In about 22 days, the moon reaches quadrature with 270° elongation (Fig. 4.2), the resulting tide is known as *neap tide of the third quarter*.

In about 29.5 days from the previous new moon, the moon again returns to the meridian and the same cycle is repeated.

4.2 MEAN SEA LEVEL AS A DATUM PLANE

For a survey of small extent, an arbitrary horizontal plane, at a certain distance below a benchmark, may be used as a datum plane. But for large surveys, or where the survey done is to be joined with another one, it is essential to reduce the observations to the mean sea level (m.s.l.) at that place. The mean sea level is defined as the mean level of the sea obtained by taking the mean of all the

heights of the tides measured at regular intervals of one hour over a stated period covering the entire number of complete tides. It is found that mean sea level shows appreciable variations from day-to-day, from month-to-month and even from year-to-year. Consequently, the period of time to be covered by the observations from which it is to be derived as a datum for levelling will depend on the purpose for which levels are required, e.g., for geodetic surveys, a period covering 12 lunar months can be accepted as a minimum. However, a good estimate of mean sea level can be had by extending the observations for about 19 years, by which time the moon's nodes will have completed one entire revolution. The day-to-day variations in mean sea level values are irregular owing to the irregular variations in the force and direction of the wind and in the atmospheric pressure. The monthly variations are more or less periodic, so that, at any given place, mean sea level tends to be low on certain months of the year and high on the others. Mean sea level for the year may also show fairly considerable variations.

The advantage of adopting mean sea level as a datum for levels is that it is a natural level surface which can always be re-established if benchmarks on the land should be disturbed or moved in any way. Also some connection between the levels on the land and the level of the sea is often required for different engineering purposes such as drainage, reclamation, hydro-electric, harbour works, gradual subsidences or elevations of the land relative to the sea, etc. For all such works, mean sea level is the best datum to use.

In India, the vertical datum for heights has been chosen as the mean sea level at a group of nine tidal observatories at Indian ports. It was assumed that the mean sea level at these ports belongs to the same level surface. All these ports serve as a reference for the level net.

4.3 TIDE GAUGES

These are used to determine the exact water surface level, i.e., the vertical movements of tides during the time soundings are made. A typical gauging station is shown in Fig. 4.3.

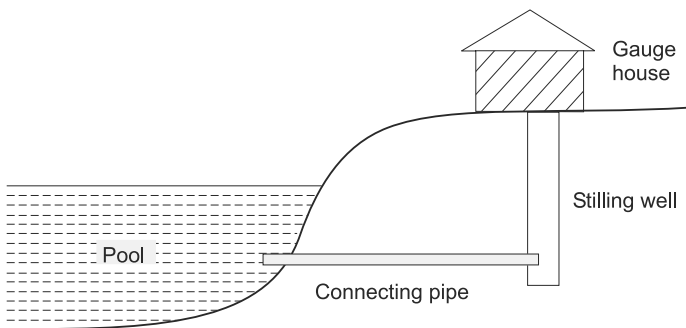


Fig. 4.3 *A typical gauging station*

The gauges are read at regular intervals, varying from 10 to 30 min, depending upon the circumstances. As the variation of water level is not linear, a graph should

be plotted of tide gauge readings against time. The exact reading at the time of making the sounding can thus be interpolated. Whenever irregular gauge readings have been made, the soundings may be reduced by the following formula:

$$h = h_0 + \frac{y}{2} \cos\left(180^\circ \times \frac{t}{T}\right)$$

where h is the height of water level above datum at the time of sounding, h_0 the height of mean sea level above datum, y the range of the tide, t the time interval between time of high water and time of sounding, and T the time interval between high and low water.

The most simple tide gauge is a graduated pole, erected with its zero mark below the lowest expected water level and of sufficient length to cover the tidal range.

The gauges are broadly classified as non self-registering and self-registering. The latter being better and preferred because these are automatic, require no observer and produce a continuous record in the form of a graph known as *marigram*. The examples of the non self-registering gauges are staff gauge, float or box gauge and weight gauge.

4.4 SOUNDINGS

Soundings in hydrography are similar to reduced levels in topography. These provide necessary data to represent the submarine (underwater) surface conventionally on a plan or chart. They may be defined as the depths measured below the water surface of a water body.

The mean low water of the ordinary spring tides (L.W.O.S.T.) is the datum to which the soundings are referred to. But, L.W.O.S.T. is base for only short time and varies considerably in case of extensive coastal surveys and in tidal waters. Thus, soundings are commonly referred to the nearest ordnance benchmark. These are represented in metres.

Soundings are accomplished by measuring, from a boat or launch, the depth of water at various points. The sounding points should be selected keeping in mind that all the important irregularities in the submarine surface are recorded. The soundings are thus made along a series of straight lines at right angles to the shoreline.

The spacing between the sounding lines and between the sounding points depends upon the nature of the submarine surface as well as on the object of the survey. Usually, the former is kept as 30 m and the latter as 7.5–15 m.

4.4.1 Range

A range or a range line is the line along which soundings are made. These are usually fixed perpendicular to the shoreline and are parallel to each other.

4.4.2 Shore Signal

Each range line is marked with two signals, at some distance apart, along it on each shore. The signals are usually wooden tripods with a white and coloured flag on the top.

4.5 SOUNDING EQUIPMENT

Essential equipments are a boat or a launch, apparatus for making sounding, and instruments for locating soundings.

4.5.1 Sounding Boat

To carry out the sounding operations in small water bodies generally a flat bottom, low draft boat known as canoe is used. It should be stable and should have sufficient space to accommodate at least three persons. It should also have sufficient space for position fixing and plotting. The sounding boat should be stable and manoeuvrable at low speeds (up to 6 km/h). The minimum cruising speed should be at least 15 km/h. These boats are generally provided with openings called well through which soundings are made. Boats of large size equipped with motor are used in sea, or rough waters. Plate-1 shows such a large size motorised boat. It also shows a recent 500 tonnes capacity ship MAKAR for hydrographic surveying, and other activities such as coast guarding, etc., recently acquired by Indian Navy.

4.5.2 Apparatus for Making Sounding

These are sounding rods, lead lines, sounding machine, etc.

1. *Sounding rods*: Usually 25–75 mm in diameter and 3–5 m long timber rods are used for shallow waters, say up to 5 m. To facilitate plunging, the rod is weighted with lead at the base. If the underwater surface is soft, a circular metal disk can be attached to the base of the rod. The rod should be light but stiff enough to prevent bending. One edge of the rod is graduated and the other is streamlined to facilitate handling in water. A sounding boat with sounding rod is shown in Fig. 4.4.

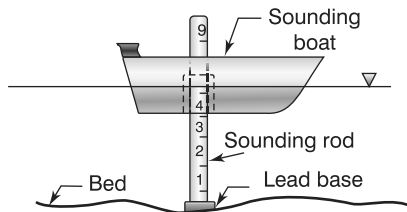


Fig. 4.4 Sounding boat with sounding rod

2. *Lead line*: It consists of a graduated line or chain to which a lead is attached (Fig. 4.5). The most commonly used lead line is made of cotton or hemp. Under ordinary tension, when wet, the line should not change its length. Every one feet of the lead is marked with a cloth. Every fifth marker is made of different colour and every tenth marker carries notches. This system facilitates in taking readings. Such a lead should be immersed half an hour before the sounding is made.

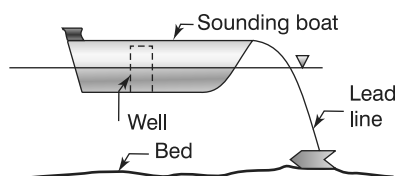


Fig. 4.5 Sounding boat with lead line

3. *Echo sounding machine*: Where the depth of water is too much, an echo sounding machine known as 'fathometer' is used. It measures the depth below the boat on which it is installed. A fathometer works on the echo property of sound. It consists of a transmitter and a receiving oscillator, recorder unit, and a power unit. The sound waves emitted at the surface of the water are recorded back after these return from the underwater surface (Fig. 4.6). Since the velocity of sound waves in water is known, the distance travelled by sound waves can be calculated. If V is the velocity of sound waves in water, t the time interval between the transmitting and receiving of waves, and D the depth of water, then

$$D = \frac{1}{2} Vt$$

This method is very accurate and can be used in strong winds. It is more sensitive than other methods, and produces a continuous record in the form of curves when the boat is in motion.

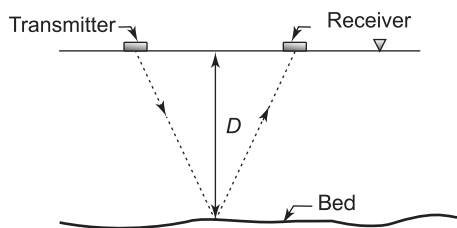


Fig. 4.6 Principle of echo sounding

Echo sounding has now established itself as an invaluable aid to the hydrographical surveyor and is in extensive use by many harbour, river conservation and improvement authorities all over the world. It has the following advantages over ordinary sounding methods:

- (a) It is more accurate than the lead line, as a truly vertical sounding is obtained, and swiftly moving water or the motion of the vessel, does not cause it to be deviated appreciably from the vertical.
- (b) It can be used in strong currents or streams, where accurate soundings with the lead line are almost impossible.
- (c) It is, as a rule, more sensitive than the lead line method.
- (d) It can be used on days, or in weather, when the ordinary lead line method would be impossible.
- (e) It provides a continuous record of the bottom as the ship moves forward and this record can be examined at leisure later.
- (f) It often enables information concerning the sea or river bed, which could not be obtained by ordinary simple means, to be obtained simply and cheaply.
- (g) It is much more rapid in use than the ordinary methods.

4. *Water glass*: During sounding, an inspection of the submarine surface or of underwater construction is sometimes required, and may be made by means of water glass or water telescope. The instrument consists of a water tight box of wood or sheet iron open at the top and having a sheet of plain glass fitted at the

bottom. Two handles are provided for steadying the instrument. The observer on the boat, lowers the bottom a few centimeters below the surface of water, and looks in at the upper end. The clearness of the view presented depends upon the depth and transparency of the water and the brightness of the day.

4.5.3 Instruments for Locating Sounding

1. *Sounding sextant:* When the location is made from the shore, a transit theodolite is used, whereas locations from the boat are made by a sextant. The sextant used is somewhat different from the astronomical one and is designed to suit the requirements of sighting from a moving small boat.

It is of specially strong construction, and is fitted with a large index glass to minimize the difficulty of sighting from a small boat, due to its motion. The horizon glass has no transparent part, and is entirely supported by a metal tray. The object viewed directly is therefore sighted over the top of the horizon glass, and the reflected image is separated from the direct one by the small thickness of the frame. The arc is graduated bold and is read by vernier to single minutes. The angle which can be measured directly with it is limited to 120° . The ring carries a disc with pin-hole sight, but a telescope of wide field is also provided for long sights. The constructional details of a sounding sextant are shown in Fig. 4.7.

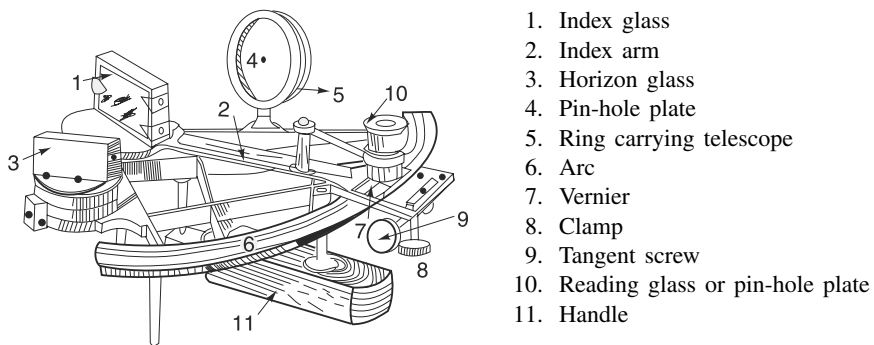


Fig. 4.7 Sounding sextant

4.6 LOCATING SOUNDING

The soundings can be located with respect to the fixed shore station by observation made entirely from either the shore or the boat, or from both. These are discussed in the foregoing. Besides the methods mentioned herein, there are several other methods, e.g., electronic methods, which may also be used for determining the locations and depths of soundings.

4.6.1 Location by Cross-Rope

For narrow rivers and lakes, this method is very advantageous. A cross-rope or a wire with markers spaced at suitable intervals is stretched along the width of the water body. The soundings are then taken on these markers.

4.6.2 Location by Range and Time Intervals

It is assumed that the distance of soundings are proportional to the time intervals. Soundings are made at different intervals, as and when required, with the boat at constant speed. Since the boat runs at constant speed and since the time elapsed at the instant of making a sounding is known, the distance of the sounding points along the range can be found.

This method can be used to its full advantage in still waters, where a constant speed of the boat can be maintained. To ensure a regular rate of speed of the boat the first sounding point should be at a sufficient distance from the shore.

4.6.3 Location by Range and Angle from the Shore

In this method, soundings are located by sailing the boat along the range line and fixing the position of soundings by a theodolite on some traverse station at the shore (Fig. 4.8). The traverse station should be so selected that the line of sight from it intersects the range line as nearly at right angles as possible.

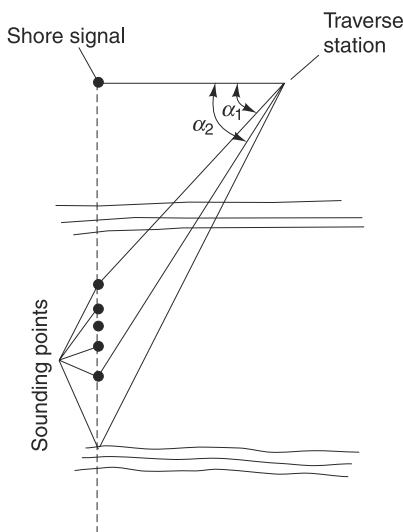


Fig. 4.8 Locating sounding by angle from shore

The transit man sights the boat, signals at definite interval, and reads the horizontal angle from a known reference line. Both the sounding man and the transit man should record the time of observations so as to identify the readings. Usually, every fifth sounding is located by angular observation and the intermediate points are fixed by time interval.

4.6.4 Location by Range and Angle from Boat

The method is similar to the previous method except that the angle is measured with a sextant from the boat (Fig. 4.9). This is superior to the previous method since it has a better control over the operations. Thus mistakes in booking the readings are also reduced.

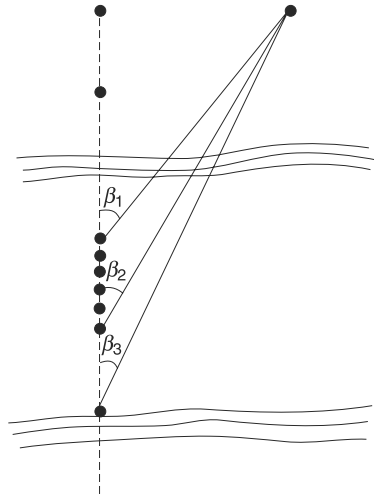


Fig. 4.9 Locating sounding by angle from boat

4.6.5 Location by Two Angles from Shore

Where it is difficult to keep the boat along the range in strong current or if it is impracticable to establish ranges, the soundings are located by the measurement of two angles from the shore, as shown in Fig. 4.10. The transit stations should be so selected that a good intersection of the sounding points may be achieved. Both the transit man as well as the sounding man should also record the time of reading so as to identify the observation. The coordinates of the sounding point O may be located from the relations:

$$x = \frac{l \tan \beta}{\tan \alpha + \tan \beta} \quad \text{and} \quad y = \frac{l \tan \alpha \tan \beta}{\tan \alpha + \tan \beta}$$

where α and β are the angles measured by the transit men, and l is the distance between the transit stations.

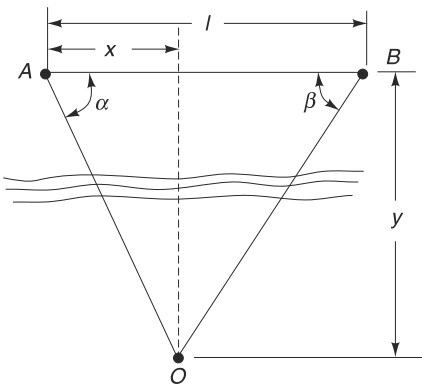


Fig. 4.10 Locating sounding by two angles from shore

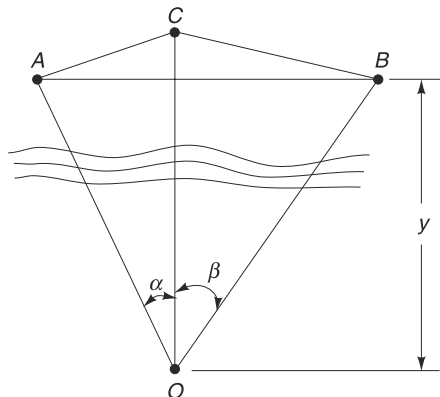


Fig. 4.11 Locating sounding by two angles from boat

4.6.6 Location by Two Angles from Boat

Two angles are read simultaneously with a sextant from the boat, on three points located on the shore (Fig. 4.11). This is an example of the three-point problem. The three shore points should be clearly visible and well defined. Problems similar to the three-point problem in plane tabling are associated here also. If possible, the boat position should be plotted as the work proceeds. This helps to identify the weak locations at once.

Note *In a fast flowing stream or current, the boat from which the soundings are being taken is pulled in the direction of the current and then while a reading is being observed, allowed to float freely with the stream. The boat is thus carried downstream at approximately the same velocity as the sounding line, so that any great deflection of this from the vertical is obviated, and a more correct reading is obtained than would otherwise be the case.*

4.6.7 Location by One Angle from Shore and the Other from the Boat

This method is a combination of the previous two methods. Out of the two shore points, one point is the transit station and on the other point a shore signal is placed to facilitate observation of the horizontal angle by a sextant from the boat (Fig. 4.12). If the distance l between the shore points is known, the coordinates of the boat at O can be found.

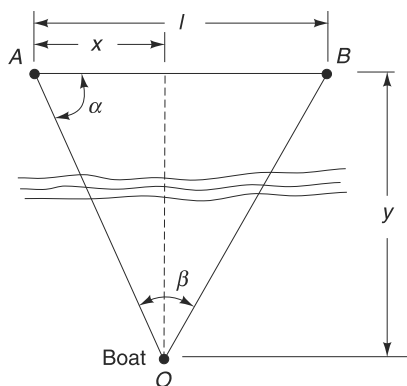


Fig. 4.12 Locating sounding by one angle from shore and other from boat

4.6.8 Location by Intersecting Ranges

Sometimes it is necessary to repeat the sounding at fixed points over a period of time. This method is best suited for situations such as those in harbours, reservoirs, etc., where it is desired to determine the character of changes in the bed due to scour, silting, or to measure the dredged material. Permanent ranges are established preferably at nearly right angles (Fig. 4.13). Soundings are then made at the intersections of the ranges.

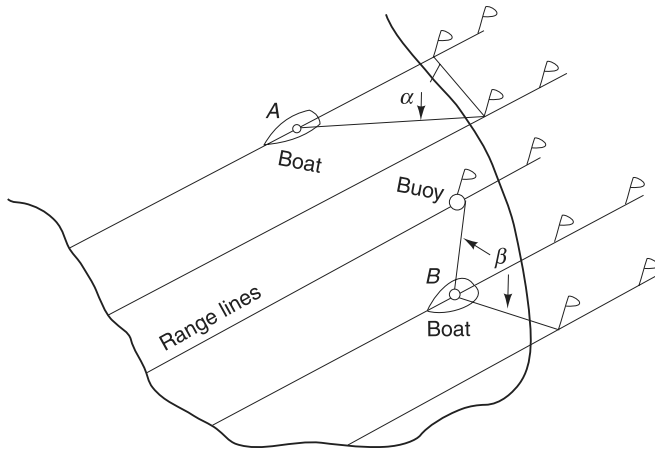


Fig. 4.13 Locating sounding by range lines

4.6.9 Location by Stadia Method

In this method, the location of the sounding along the range is fixed by stadia observations from the shore and the sounding is measured with a sounding rod. As far as possible, the line of sight is kept horizontal and the measurement of vertical angles is avoided to save calculations. If s is the staff intercept at the time the sounding is made and D is the distance of the boat from the stadia, then

$$D = Ks + C$$

where K and C are the tacheometer constants.

4.6.10 Location by Radio Acoustical Method

This method is used in locating ships offshore. It is based on the principle that if a ship detonates a charge of explosives near it, the sound of the explosion reaches a shore station via radio waves. With known velocity of sound waves in water, the distance of the ship from the offshore station can be computed and the ship can be located with respect to the station. In a similar manner, radar may also be used.

4.6.11 Location by Electromagnetic Method

Hydrographic surveys are increasingly conducted these days using aircrafts/satellites and sophisticated electronic sensor systems as against traditional boats/vessels with a sounding line or echo sounding. Electromagnetic position-fixing systems determine the sounding vessel position by the intersection of a minimum of two range distances measured to shore stations.

1. *Short range:* The equipment is a portable, microwave equipment, comprising two on-shore 'remote' units which form a base line and give continuous distances to a receiver system on-board ship. The ship's position is thus at the apex of a triangle whose three sides are known.

Two well-known systems are the Decca Trisponder 202 A (range = 80 km, accuracy = ± 3 m) and the Tellurometer Hydrodist MRB 201 (range = 50 km, accuracy = ± 1.5 m). The systems operate at speeds up to 30 knots and can be linked to dynamic position fixing systems capable of automatic operation of computers, plotters, data recorders, etc.

2. *Medium range:* Figure 4.14 shows a 'master' station A (also called reference station), combined with two 'slaves' B and C, which, in effect, generate a hyperbolic lattice of electromagnetic wave pattern over the area. With the aid of on-board phase meters, the ship's position within the lattice can be defined by hyperbolic coordinates. Correlation of the on-shore units into an appropriate survey system enables the hyperbolic coordinates to be converted to geographical or rectangular coordinates.

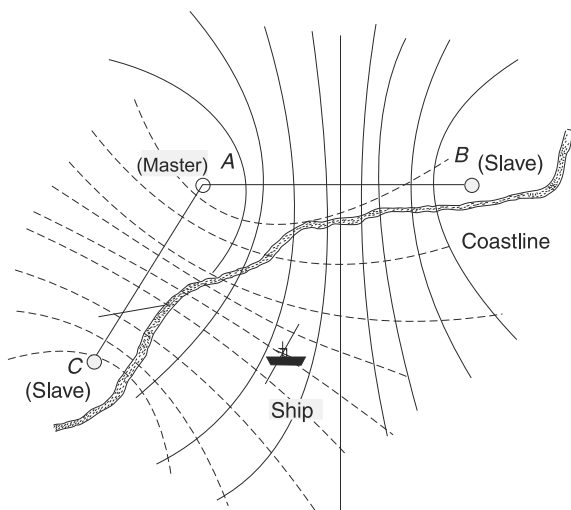


Fig. 4.14 Locating sounding by electro-magnetic method

A well-known example of such a system is Decca Hi-Fix 6 (range = 300 km, accuracy = 0.01 lanes). The lane width along the base lines AB and AC is 75 m giving an accuracy of ± 0.75 m, but this falls off rapidly as the lines diverge from each others. This fault can be eliminated by using two on-shore 'slaves' and a 'master' on board ship, resulting in lattice of intersecting circles with a constant lane width. The disadvantage is that only one ship at a time may operate in this latter arrangement.

3. *Satellite doppler:* Position fixing on a world-wide basis, by the use of artificial satellites, became generally available in 1967. GPS is a continuous, all weather, and worldwide satellite-based electronic positioning system and is discussed later in Chapter 9. The technique to process signals from two GPS receivers operating simultaneously to determine 3-D line vector between the two receivers is known as differential GPS. This powerful technique can produce real-time positions of moving vessels.

When a sound source and a receiver are in motion relative to each other, the frequency of a transmitted wave at the receiver will differ from the frequency at the source according to a definite mathematical form. This is called the *Doppler effect* and is the principle used in satellite fixation. The orbiting satellites are tracked from accurately located ground stations and their positions are continuously updated. Thus, the satellite becomes a navigational beacon which regularly transmits information defining its position. A vessel possessing the appropriate equipment receives the transmitted data, measures the Doppler shift and converts the information into geographical coordinates defining the ship's position. The application of the DGPS technique for hydrographic surveying is shown in Fig 4.15.

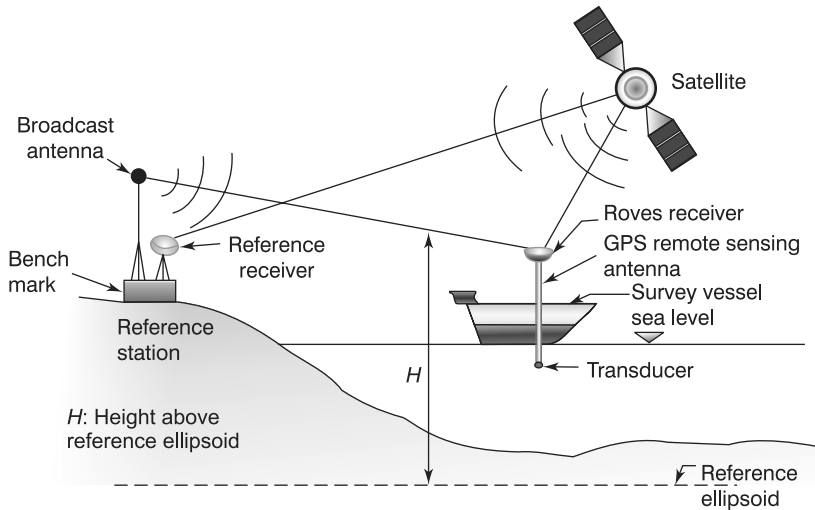


Fig. 4.15 Differential positioning of survey vessel

4.7 REDUCTION OF SOUNDINGS

The soundings are measured with respect to the water level at the time of making them. Thereafter, these are reduced with reference to the adopted datum, usually the mean level of the low water of spring tides, with the help of gauge readings. A correction equal to the difference of level between the actual water level (read by gauges) and the datum is applied to the observed soundings to obtain the reduced soundings.

4.8 PLOTTING THE SOUNDINGS

The plotting methods for soundings depend upon the method used for locating the soundings. Where the soundings are located by ranges, their plotting needs no explanation. If two angles are measured from the shore, two protector method can be used with one each over the plotted position of the transit station. The angles are plotted in pairs and their intersection fixes the sounding location. However, where the soundings are located by two angles from the boat by observations to three shore points, the solution can be had by mechanical, graphical or analytical methods.

The mechanical or tracing paper method is similar to the solution of the three-point problem in plane tabling. The analytical method is similar to the problems dealt with in triangulation with satellite stations. Therefore, these methods need no elaboration. As far as the graphical method is concerned, there can be many graphical solutions but the one described below is commonly employed.

Let a, b, c be the plotted positions of the three shore signals A, B, C and let α and β be the angles subtended at the boat O (Fig. 4.16). To locate the boat point o , join a and c and draw ad making an angle β with ac . Draw cd making an angle α with ca . Let both of these lines intersect in d . Draw a circle passing through the three points a, d and c and join d and b . Prolong db to meet the circle. The point o thus obtained represents the boat position. From the properties of the circle,

$$\angle aod = \angle acd = \alpha$$

$$\text{and} \quad \angle cod = \angle cad = \beta$$

In practise, however, the most popular method is by plotting *circle charts*, the principle of which is shown in Fig. 4.17. The intersection of two circles passing through shore stations A, B and C gives the position of boat O . It can be clearly seen that the radius of circle BC is $1/2 (BC) \operatorname{cosec} \beta$. Thus, for various values of α and β , a chart can be produced of a large number of such circles having their respective centres on the perpendicular bisectors of the chords AB and BC . Interpolation between the circles for the measured angles α_n, β_n will fix the boat position.

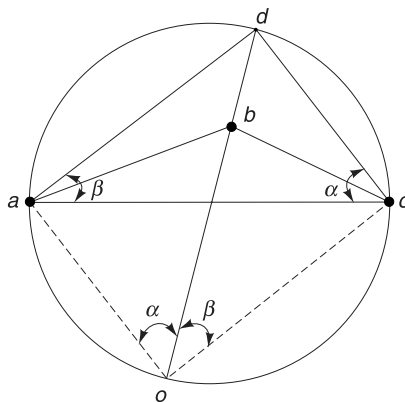


Fig. 4.16 Plotting sounding graphically

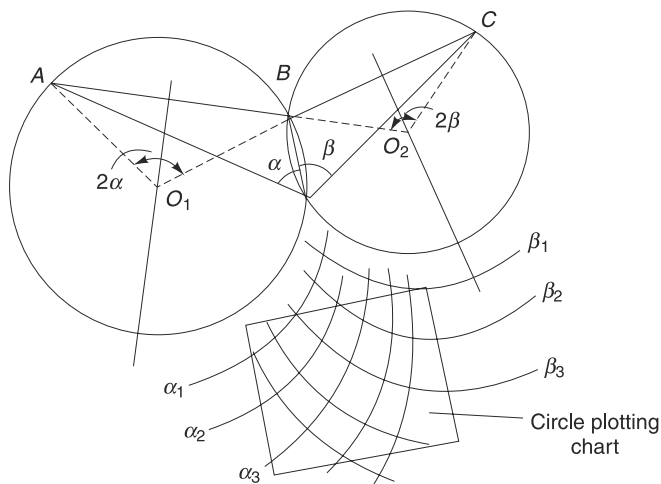


Fig. 4.17 Plotting sounding using circle charts

If by chance O should fall on the circumference of a circle passing through A , B and C , no solution is possible. This is referred to as the 'danger circle' and should be carefully avoided.

4.9 STREAM MEASUREMENTS

A surveyor is often called upon to measure the stream discharge for the design of water supply, irrigation and power schemes. Since the stream discharge varies from day to day, the investigations should extend over a considerable period. The unit of discharge is m^3/s or l/s .

The discharge may be measured either by the area of cross section of the stream at the place and the mean velocity of flow, or by a weir or dam across the stream and observing the head, or by the chemical method.

4.9.1 Area Velocity Method

The site at which the measurement is to be made is known as *discharge station*. It should be so selected that the streamlines at it are as regular as possible. On either side of the station, the stream should be straight with its bed uniform in shape and character.

The cross section is measured with soundings, and for measuring the average velocity of the current, any of the following methods may be used:

1. By floats
2. By current meter

1. *Floats*: These are usually wooden rods about 1 m long which drift with the current and are located from time to time. These should be easily identified from a distance. They are classified into surface, subsurface and rod types of floats.

(a) *Surface floats* The surface floats should be light so as to move freely with the current and should expose but little surface above the water which would be subjected to the action of the wind or surface waves. Specially designed floats are available, but lacking these, a corked bottle nearly submerged, and carrying a small mast and flag if desired, makes a good float. Blocks of wood or other objects may also be used. These are best suited for measuring velocities in high floods but their use is restricted in winds.

The surface floats should be sufficiently distinct to enable an observer on the banks to clearly and easily locate the course they follow, but otherwise they should be as small as possible in order to minimise the effect which the wind may have upon their progress. An upstream wind considerably retards the velocity of surface floats, while a downstream wind increases it, so that, unless the weather is favourable, very misleading results may be obtained. The mean velocity on any vertical section may be obtained by Bazin formula.

(b) *Subsurface floats* These are also known as *submerged floats* or *double floats*, and yield a result much more independent of the wind. It consists of a cylinder, slightly heavier than water, attached by a chord to a surface float. The length of the chord is adjusted so that the submerged float will move with that layer of water whose velocity is desired to be measured. They are most suited in sea or deep

rivers and are never used on small streams. A typical subsurface float is shown in Fig. 4.18.

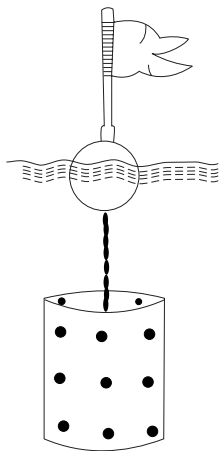


Fig. 4.18 Subsurface float

The mean velocity may be calculated by Cunningham equation:

$$u = \frac{1}{4} (V + 3V_1)$$

where V_1 = the velocity at 0.6 times the depth at which the float lies

If the velocity is determined at half the depth, the mean velocity may be taken as

$$u = 0.98 V_{0.5}$$

while for 0.6 of the depth, the mean velocity is approximately equal to the observed velocity.

(c) Rod floats These are made of wooden rods or hollow tubes weighted at the bottom to submerge the float nearly to the bed of the stream and to maintain it in a vertical position (Fig. 4.19). Rod floats produce better results than either of the above types, provided that the stream is of fairly regular cross section and the bed is not choked with weeds. To ascertain the velocity of a current by means of a float, it is necessary to determine the exact length of the path followed by the float during an observed interval of time.

i. Measurement of velocity To measure velocity by a float, the time taken by a float to travel a known distance down the stream is observed. A number of floats are observed across the stream and the mean velocity is calculated.

ii. Measurement of discharge The cross-sectional area is obtained along the ranges with the help of soundings. The mean velocity of the float is obtained between the ranges. A curve is plotted to show the variation of mean velocity across the stream. The discharge through any

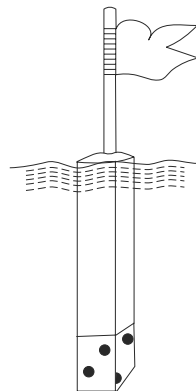


Fig. 4.19 Rod float

section is thus the area at the section multiplied by the mean velocity and the sum of all such quantities will be the total discharge.

2. *Current meter*: A current meter may be used to measure the stream velocity and gives better results than floats. It essentially consists of either a spindle mounted on a fork and carrying a wheel, with helical vanes, which is rotated by the flowing water or a small wheel mounted upon a vertical spindle and fitted with a number of cup-shaped or conical vanes as shown in Fig. 4.20. In the latter case, the action of the current causes the wheel to rotate in the opposite direction to that in which the concavity of the faces of cups.

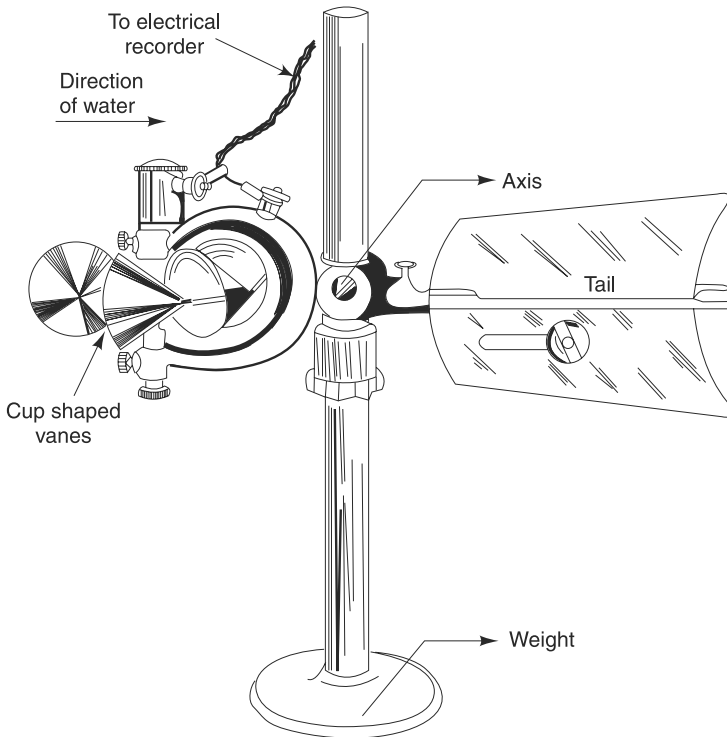


Fig. 4.20 Current meter

The number of revolutions made by the wheel are recorded by gearing or by means of an electric arrangement. To obtain the velocity by a current meter, the boat is towed or sculled over a known distance in still water at as constant a speed as possible, the number of revolutions recorded by the meter and the time taken being also noted. The boat, being in rear of the meter, does not disturb the water until the meter has passed through it, and consequently does not affect the reading.

From these data are calculated the relative velocities of the meter and the water, and also the number of revolutions per second, or the time taken to complete, say 50 revolutions, and from a series of such experiments, a law may be discovered connecting these quantities.

The equation, table, or curve so ascertained will be applicable to observations made in a flowing stream from a stationary boat or from a bridge, as it is the velocity relative to the meter which is measured in each case.

To measure the discharge, on account of the observed velocities, by a current meter, the mean velocity is found for each of a series of vertical sections spaced uniformly across the stream. Taking soundings and observing velocities is done simultaneously. The area of each section is known from the soundings and from the spacing between the sections, and thus the discharge of each section can be computed. The sum of the quantities for all the sections gives the total discharge.

Note *To calibrate the meter, it is suspended from a projection jutting out in front of the bow of a boat, and immersed at least 60 cm to eliminate any surface effects.*

4.9.2 Weir Method

For smaller streams, very accurate results may be obtained by constructing a weir and passing all the water over it. A weir is an artificial obstruction created in the passage of flow of the stream. From the cascade formed over the weir, the discharge measurements are done. This is very costly method and is used only when very accurate results are required.

Weirs have a sharp crest, virtually a line over which flow takes place. The notch (opening) through which water flows may be rectangular, triangular, trapezoidal or stepped (Fig. 4.21), of which the rectangular one is the most commonly used. The length of the notch should be at least three times the head of water over it. The rectangular notch is most suitable for shallow streams, whereas the triangular one is suitable for small discharges. The stepped notch is most suited for floods. Weirs with end constraints and without end constraints are shown in Fig. 4.22.

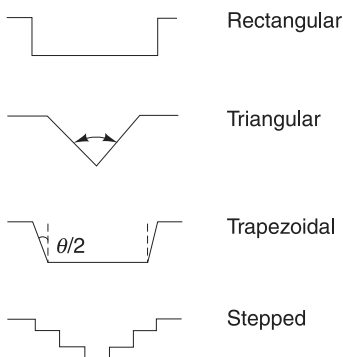


Fig. 4.21 *Types of notches*

The dimensions of the weir being known, the only additional data required for the calculation of discharge is the head which is measured by a hook gauge. Head is defined as the difference of level between the crest and the surface of sensibly still water (Fig. 4.23). This surface normally lies upstream at more than $3 H$ distance from the weir.

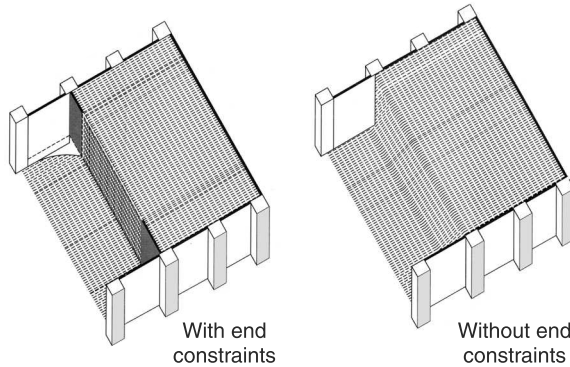


Fig. 4.22 *Types of weirs*

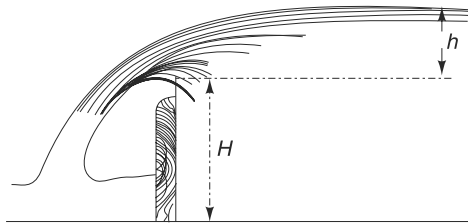


Fig. 4.23 *Diagrammatic representation of head*

$$\text{Rectangular } Q = \frac{2}{3} C_{d1} \sqrt{2g} b H^{3/2}$$

$$\text{Triangular } Q = \frac{8}{15} C_{d2} \sqrt{2g} \tan \frac{\theta}{2} H^{5/2}$$

$$\text{Trapezoidal } Q = \frac{2}{3} C_{d1} \sqrt{2g} b H^{3/2} + \frac{8}{15} C_{d2} \sqrt{2g} \tan \frac{\theta}{2} H^{5/2}$$

Stepped Q = Sum of discharge of separate rectangular notches

where C_{d1} , C_{d2} = coefficient of discharge for rectangular and triangular weirs

b = effective width of crest

H = head over the crest, and

θ = apex angle

4.9.3 Chemical Method

A fairly concentrated solution of a chemical is introduced at a known rate in the stream and the analysis of water is compared before and after its immersion. This method is particularly suitable for turbulent streams with rocky beds.

Example 4.1 A , B and C are three visible stations in a hydrographic survey. The computed sides of the triangle ABC are $AB = 1200$ m, $BC = 1442$ m and $CA = 1960$ m. A station O is established outside the triangle and its position is to be determined by resection on A , B and C , the angle AOB and BOC being

respectively $45^\circ 30'$ and $52^\circ 15'$. Determine distances of OA and OC , if O and B are on the opposite sides of line AC .

Solution Refer to Fig. 4.24

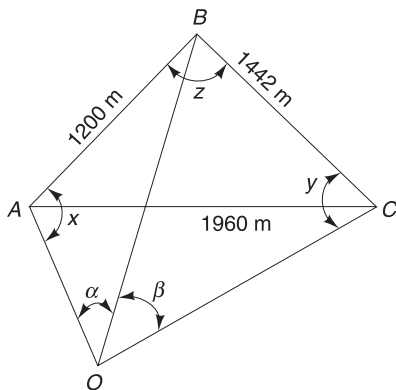


Fig. 4.24

$$a = BC = 1442 \text{ m}$$

$$b = CA = 1960 \text{ m}$$

$$c = AB = 1200 \text{ m}$$

Then,

$$\angle ABC = Z \text{ is given by}$$

$$b^2 = c^2 + a^2 - 2ac \cos Z$$

or

$$\begin{aligned} \cos Z &= \frac{c^2 + a^2 - b^2}{2ac} \\ &= \frac{(1200)^2 + (1442)^2 - (1960)^2}{2(1200)(1442)} = -0.0931102 \end{aligned}$$

or

$$Z = 95^\circ 20' 33''$$

Now

$$\theta = x + y$$

$$\theta = 360^\circ - (\alpha + \beta + Z)$$

$$= 360^\circ - (45^\circ 30' + 52^\circ 15' + 95^\circ 20' 33'')$$

$$\theta = 166^\circ 54'$$

$$\cot x = \cot \theta + \frac{c \sin \beta}{a \sin \alpha \sin \theta}$$

$$= \cot 166^\circ 54' + \frac{1200 \times \sin 52^\circ 15'}{1442 \times \sin 45^\circ 30' \times \sin 166^\circ 54'}$$

$$= -4.2972 + \frac{948.82749}{233.11249} = -0.226944$$

or

$$x = 102^\circ 47'$$

$$y = 166^\circ 54' - 102^\circ 47' = 64^\circ 07'$$

$$\angle ABO = 180^\circ - x - \alpha$$

$$= 180^\circ - 102^\circ 47' - 45^\circ 30' = 31^\circ 43'$$

$$\begin{aligned}\angle CBO &= 180^\circ - y - \beta \\ &= 180^\circ - 64^\circ 07' - 52^\circ 15' = 63^\circ 38'\end{aligned}$$

Hence,

$$\begin{aligned}AO &= \frac{c}{\sin \alpha} \sin (\angle ABO) \\ &= \frac{1200}{\sin 45^\circ 30'} \times \sin 31^\circ 43' = 884.49 \text{ m} \\ CO &= \frac{a}{\sin \beta} \sin (\angle CBO) \\ &= \frac{1442}{\sin 52^\circ 15'} \times \sin 63^\circ 38' = 1634 \text{ m}\end{aligned}$$

Example 4.2 The following observations were made on three stations A , B and C from a boat at O with the help of a sextant. Station B and O being on the same side of AC .

$$\begin{aligned}\angle AOB &= 30^\circ 25', & \angle BOC &= 45^\circ 25', & \angle ABC &= 130^\circ 10' \\ AB &= 4000 \text{ m}, & BC &= 4995 \text{ m}\end{aligned}$$

Calculate the distances of the boat from the three stations.

Solution Refer to Fig. 4.25.

$$\alpha = 30^\circ 25', \quad \beta = 45^\circ 25', \quad c = 4000 \text{ m}, \quad a = 4995 \text{ m}, \quad Z = 130^\circ 10'$$

$$\begin{aligned}\text{Now} \quad \theta &= x + y \\ &= 360^\circ - [\alpha + \beta + (360^\circ - Z)] \\ &= 360^\circ - [30^\circ 25' + 45^\circ 25' + (360^\circ - 130^\circ 10')] = 54^\circ 20'\end{aligned}$$

$$\begin{aligned}\cot x &= \cot \theta \left(1 + \frac{c \sin \beta \sec \theta}{a \sin \alpha} \right) \\ &= \cot 54^\circ 20' \times \left(1 + \frac{4000 \times \sin 45^\circ 25' \times \sec 54^\circ 20'}{4995 \times \sin 30^\circ 25'} \right) \\ &= 2.10433\end{aligned}$$

or

$$\begin{aligned}x &= 25^\circ 25' 3.1'' \\ y &= 54^\circ 20' - 25^\circ 25' 3.1'' = 28^\circ 54' 56.9''\end{aligned}$$

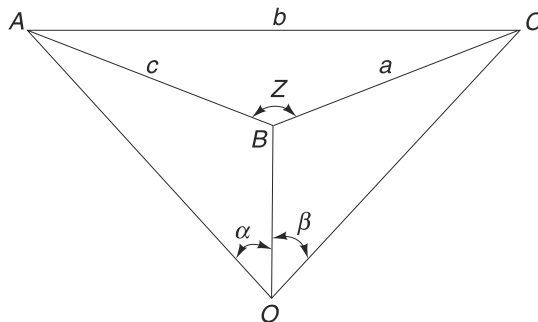


Fig. 4.25

From triangle AOB , applying sine rule

$$\begin{aligned} AO &= \frac{c}{\sin \alpha} \sin [180^\circ - (x + \alpha)] \\ &= \frac{4000}{\sin 30^\circ 25'} \times \sin [180^\circ - (25^\circ 25' 3.1'' + 30^\circ 25')] = 6537.159 \text{ m} \end{aligned}$$

and

$$\begin{aligned} BO &= \frac{c}{\sin \alpha} \sin x \\ &= \frac{4000}{\sin 30^\circ 25'} \times \sin 25^\circ 25' 3.1'' = 3391.068 \text{ m} \end{aligned}$$

From triangle CBO , applying sine rule

$$\begin{aligned} CO &= \frac{a}{\sin \beta} \sin [180^\circ - (y + \beta)] \\ &= \frac{4000}{\sin 30^\circ 25'} \times \sin [180^\circ - (28^\circ 54' 56.9'' + 45^\circ 25')] = 6752.60 \text{ m} \end{aligned}$$

Example 4.3 Calculate the distances and bearings of OA , OB and OC from the following data:

Line	Bearing	Length (m)	Angle	Remarks
AB	$76^\circ 54' 58''$	1741.5	$\angle AOB = 85^\circ 40' 15''$	Boat at O
BC	$187^\circ 23' 48''$	2728.5	$\angle BOC = 146^\circ 32' 20''$	
CA	$329^\circ 09' 04''$	2703.9	$\angle COA = 127^\circ 47' 35''$	

Solution

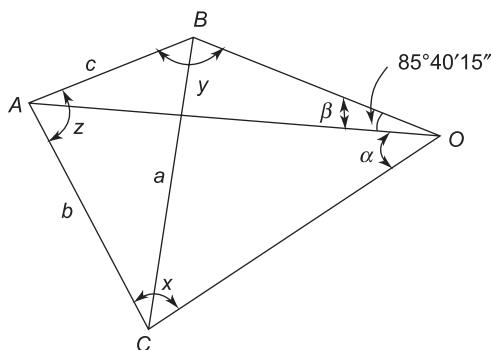


Fig. 4.26

$$a = 2728.5 \text{ m}, \alpha = 127^\circ 47' 35''$$

$$b = 2703.9 \text{ m}, \beta = 85^\circ 40' 15''$$

$$c = 1741.5 \text{ m}$$

Refer to Fig. 4.26.

$$\begin{aligned}
 \text{Bearing of } AC &= \text{Bearing of } CA - 180^\circ \\
 &= 329^\circ 9' 4'' - 180^\circ \\
 &= 149^\circ 9' 4'' \\
 \angle Z &= \text{Bearing of } AC - \text{Bearing of } AB \\
 &= 149^\circ 9' 4'' - 76^\circ 54' 58'' \\
 &= 72^\circ 14' 6'' \\
 \theta &= 360^\circ - (\alpha + \beta + Z) = x + y
 \end{aligned}$$

Sum of angles at point O

$$\begin{aligned}
 &= 85^\circ 40' 15'' + 146^\circ 32' 20'' + 127^\circ 47' 35'' \\
 &= 360^\circ 0' 10''
 \end{aligned}$$

$$\text{Correction of each angle} = -\frac{10''}{3} = 0^\circ 0' 3.33''$$

$$\begin{aligned}
 \text{Corrected value of } \alpha &= 127^\circ 40' 35'' - 0^\circ 0' 3.33'' \\
 &= 127^\circ 47' 31.67''
 \end{aligned}$$

$$\begin{aligned}
 \text{Corrected value of } \beta &= 85^\circ 40' 15'' - 0^\circ 0' 3.33'' \\
 &= 85^\circ 40' 11.67'' \\
 \theta &= 360^\circ - (127^\circ 40' 31.67'' + 85^\circ 40' 11.67'' + 72^\circ 14' 6'') \\
 &= 74^\circ 18' 10.66''
 \end{aligned}$$

$$\begin{aligned}
 \cot x &= \cot \theta \left(1 + \frac{b \sin \beta \sec \theta}{c \sin \alpha} \right) \\
 &= \cot 74^\circ 18' 10.66'' \\
 &\quad \times \left[1 + \frac{2703.9 \times \sin 85^\circ 40' 11.67'' \times \sec 74^\circ 18' 10.66''}{1741.5 \times \sin 127^\circ 47' 31.67''} \right]
 \end{aligned}$$

$$\text{or } \cot x = 2.316$$

$$\Rightarrow x = 23^\circ 21' 13.21''$$

From $\triangle AOC$; applying sine rule,

$$\begin{aligned}
 OA &= \frac{AC}{\sin \alpha} \sin x = \frac{2703.9}{\sin 127^\circ 47' 31.67''} \times \sin 23^\circ 21' 13.21'' \\
 &= 1356.35 \text{ m}
 \end{aligned}$$

$$\begin{aligned}
 OC &= \frac{AC}{\sin \alpha} \sin (180^\circ - \alpha - x) \\
 &= \frac{2703.9}{\sin 127^\circ 47' 31.67''} \times \sin (180^\circ - 127^\circ 47' 31.67'' - 23^\circ 21' 13.21'') \\
 &= 1651.21 \text{ m}
 \end{aligned}$$

From $\triangle AOB$; applying sine rule,

$$\begin{aligned}
 OB &= \frac{AB}{\sin \beta} \sin (180^\circ - \beta - y) \\
 &= \frac{1741.5}{\sin 85^\circ 40' 11.67''} \times \sin [180^\circ - 85^\circ 40' 11.67'' \\
 &\quad - (74^\circ 18' 10.66'' - 23^\circ 21' 13.21'')] \\
 &= 1199.56 \text{ m}
 \end{aligned}$$

$$\begin{aligned}
 \text{Bearing of } CO &= \text{Bearing of } CA + \angle x \\
 &= 329^\circ 9' 4'' + 23^\circ 21' 13.21'' \\
 &= 352^\circ 30' 17.21''
 \end{aligned}$$

$$\text{Bearing of } OC = 352^\circ 30' 17.21'' - 180^\circ = 172^\circ 30' 17.21''$$

$$\begin{aligned}
 \text{Bearing of } AO &= \text{Bearing of } AB + \angle BAO \\
 &= 76^\circ 54' 58'' + [180^\circ - 85^\circ 40' 11.67'' \\
 &\quad - (74^\circ 18' 10.66'' - 23^\circ 21' 13.21'')] \\
 &= 120^\circ 17' 48.88''
 \end{aligned}$$

$$\begin{aligned}
 \text{Bearing of } OA &= 180^\circ + 120^\circ 17' 48.88'' \\
 &= 300^\circ 17' 48.88''
 \end{aligned}$$

$$\begin{aligned}
 \text{Bearing of } BO &= \text{Bearing of } BA - \angle y \\
 &= (180^\circ + 76^\circ 54' 58'') - (74^\circ 18' 10.66'' - 23^\circ 21' 13.21'') \\
 &= 205^\circ 58' 0.55''
 \end{aligned}$$

$$\begin{aligned}
 \therefore \text{bearing of } OB &= 205^\circ 58' 0.55'' - 180^\circ \\
 &= 25^\circ 58' 0.55''
 \end{aligned}$$

Example 4.4 To determine the coordinates of the position of a ship O an observer in the ship measured the angles of the stations A , B and C located on the sea shore: $\angle AOB = 60^\circ 20' 00''$ and $\angle BOC = 40^\circ 30' 00''$. The ship was south of the stations at the time of observation. The coordinates of the stations A , B and C were as follows:

Station	Latitude	Departure
A	0	0
B	500 N	1000 E
C	300 N	1500 E

Solution Refer to Fig. 4.27.

$$\text{Latitude of } AB = 500 - 0 = 500$$

$$\text{Departure of } AB = 1000 - 0 = 1000$$

$$\text{Latitude of } BC = 300 - 500 = -200$$

$$\text{Departure of } BC = 1500 - 1000 = 500$$

$$\text{Latitude of } CA = 0 - 300 = -300$$

$$\text{Departure of } CA = 0 - 1500 = -1500$$

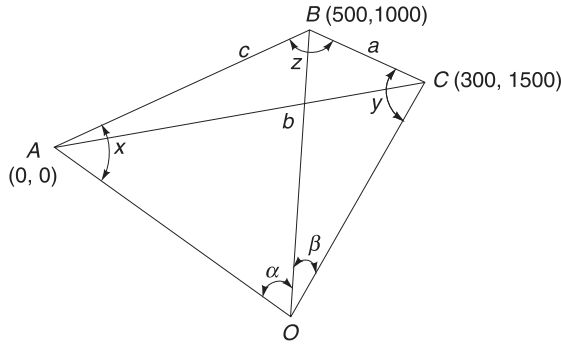


Fig. 4.27

$$\text{Length of } AB = \sqrt{500^2 + 1000^2} = 1118.0 \text{ m}$$

$$\text{Bearing of } AB = \tan^{-1}\left(\frac{1000}{500}\right) = 63^\circ 26' 5.8''$$

$$\text{Length of } BC = \sqrt{(-200)^2 + 500^2} = 538.5 \text{ m}$$

$$\text{Bearing of } BC = \tan^{-1}\left(\frac{500}{-200}\right) = 68^\circ 11' 54''$$

Since the latitude of BC is negative and the departure is positive, the line lies in the second quadrant.

$$\text{Bearing of } BC = 180^\circ - 68^\circ 11' 54'' = 111^\circ 48' 6''$$

$$\text{Length of } CA = \sqrt{(-300)^2 + (-1500)^2} = 1529.7 \text{ m}$$

$$\text{Bearing of } CA = \tan^{-1}\left(\frac{1500}{-300}\right) = 78^\circ 41' 24''$$

Since both the latitude and departure of CA are negative, the line lies in the third quadrant.

$$\text{Hence bearing of } CA = 180^\circ + 78^\circ 41' 24'' = 258^\circ 41' 24''$$

$$c = 1118 \text{ m}, b = 1529.7 \text{ m}, a = 538.5 \text{ m}, \alpha = 60^\circ 20', \beta = 40^\circ 30'$$

$\angle Z$ can be computed from the bearings of the lines BA and BC .

$$\text{Bearing of } BA = 180^\circ + \text{Bearing of } AB$$

$$= 180^\circ + 63^\circ 26' 5.8'' = 243^\circ 26' 5.8''$$

$$\angle Z = \text{Bearing of } BA - \text{Bearing of } BC$$

$$= 243^\circ 26' 5.8'' - 111^\circ 48' 6'' = 131^\circ 37' 59.8''$$

$$\theta = x + y = 360^\circ - (\alpha + \beta + Z)$$

$$= 360^\circ - (60^\circ 20' + 40^\circ 30' + 131^\circ 37' 59.8'') = 127^\circ 32' 0.2''$$

Now
$$\cot x = \cot \theta \left(1 + \frac{\sin \beta \sec \theta}{a \sin \alpha} \right)$$

$$= \cot 127^\circ 32' 0.2'' \times \left(1 + \frac{1118 \times \sin 40^\circ 30' \times \sec 127^\circ 32' 0.2''}{538.5 \times \sin 60^\circ 20'} \right)$$

$$= 1.188556$$

or $x = 40^\circ 4' 32.9''$

Hence, $y = 127^\circ 32' 0.2'' - 40^\circ 4' 32.9'' = 87^\circ 27' 27.3''$

From triangle AOB , by sine rule

$$AO = \frac{c}{\sin \alpha} \times \sin [180^\circ - (x + \alpha)]$$

$$= \frac{1118}{\sin 60^\circ 20'} \times \sin [180^\circ - (40^\circ 4' 32.9'' + 60^\circ 20')]$$

$$= 1265.48 \text{ m}$$

Bearing of $AO = \text{Bearing of } AB + \angle x$

$$= 63^\circ 26' 5.8'' + 40^\circ 4' 32.9'' = 103^\circ 30' 38.7''$$

Departure of $O = 1265.48 \times \sin 103^\circ 30' 38.7'' = 1230.46 \text{ m}$

Latitude of $O = 1265.48 \times \cos 103^\circ 30' 38.7'' = -295.65 \text{ m}$

Therefore, the coordinates of O are:

Latitude 295.65S, Departure 1230.46E.

Summary

The object of hydrographic survey is to depict topographic configuration of the under water features and their description. In hydrography measurements are also made for tides, currents and waves, which interact to create the overall pattern of the water's behaviour. The position of sounding boat/vessel in the horizontal plane when a particular sounding is being measured is a fundamental requirement. There are a number of position-fixing techniques and have been described. The depth of the bed of the water body is measured at intervals and is known as sounding. Weighted sounding lines are used up to 30 m; echo sounding machines have no limits. With the latest advancements in computer technology the trend is shifting to acquire data with the aid of aircrafts and electronic sensors to GPS. Although the individual data acquired by the latter technique may not be as accurate as by the traditional survey, but the high density data acquired to result in digital terrain model (under water) is more accurate.

The process of sounding consists in measuring, reducing, and plotting of soundings. To accomplish the task both on-shore and off-shore survey is required. For measurement of sounding a boat, sounding rod or lead line or echo sounding machine is required. A small boat (canoe), which can accommodate at least three persons and instruments, is required for small water bodies, whereas large vessels such as ships are used in sea. Sounding are located by sextant, or electromagnetic methods, or radio acoustic method, or GPS. The techniques that are in use are introduced. The measured soundings are reduced to account for the difference in the actual water level during the observations made and the datum. Various methods of plotting soundings are elaborated.

The chapter ends with the description of methods used for the measurement of velocity and discharge of rivers and streams.

Exercises

- 4.1 Define sounding, and state the equipment and personnel used for locating and making soundings. What is meant by reduction of sounding?
- 4.2 Describe briefly the methods of locating soundings.
- 4.3 Describe the various methods of plotting the soundings.
- 4.4 Explain briefly how would you determine the levels of the points on a river bed and fix the positions of the soundings (a) by use of sextant in a boat, and (b) by a theodolite on shore.
- 4.5 Describe briefly how the soundings are located by (a) two angles from the shore and (b) intersecting ranges.
- 4.6 To locate the soundings, observations were made from a boat at O with a sextant on three shore stations A , B and C . The observed angles AOB and BOC were $28^{\circ}42'40''$ and $30^{\circ}28'20''$, respectively. The distances AB and BC were 918 m and 1074 m, respectively. Points B and O were to the opposite side of AC . The angle ABC was $60^{\circ}50'40''$. Compute the distance of the boat from the three shore stations.
- [Ans. 1133.6 m, 1733.3 m, 876.75 m]
- 4.7 From the boat O offshore readings were taken to three shore signals A , B and C with the help of sextant. The angles AOB and BOC were $32^{\circ}30'$ and $62^{\circ}30'$, respectively. The distances AB and BC were measured to be 360 m and 615 m, respectively, and the angle ABC on the landward side was $233^{\circ}30'$. Determine the distance of the boat from the signal B .

[Ans. 636.63 m]

- 4.8 The coordinates of the three shore stations on a hydrographic survey are as follows:

Station	Northing	Easting
A	200	10
B	219	855
C	252	677

If from a point D south of the shore stations, the angles ADB and BDC were observed to be $92^{\circ}56'$ and $11^{\circ}59'$, calculate the coordinates of D .

[Ans. 152.48 N, 253.44 E]

- 4.9 The coordinates of three shore stations are as follows:

Station	Coordinates	
	South	East
A	0	0
B	0	750
C	500	1250

The angles AOB and BOC as observed with a sextant from a boat are $52^{\circ}12'$ and $70^{\circ}36'$, respectively. Find the coordinates of O .

[Ans. 642.67S, 693.56E]

- 4.10** Calculate the distance PA , PB and PC and their bearings from the following data for the boat position P as shown in Fig. 4.28.

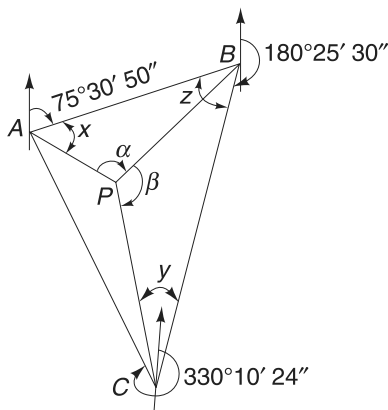


Fig. 4.28

<i>Line</i>	<i>bearing</i>	<i>length (m)</i>
<i>AB</i>	$75^{\circ}30'50''$	5750
<i>BC</i>	$180^{\circ}25'30''$	11016.83
<i>CA</i>	$330^{\circ}10'24''$	11028.79

[Ans. 4639.838 m, 3691.036 m, 7697.01 m]

Objective-type Questions

- 4.1** The observations to establish m.s.l., are recorded for about
 (a) 9 years (b) 19 years (c) 29 years (d) 39 years
- 4.2** The unit of sounding is
 (a) m/s (b) cm^2/s (c) Fathom (d) cycle/s
- 4.3** The major tide producing force is due to
 (a) sun (b) mars (c) venus (d) moon
- 4.4** The branch of surveying which deals with water bodies is known as
 (a) aqueous surveying (b) topographic surveying
 (c) hydrographic surveying (d) sea surveying
- 4.5** To estimate the dredged material from harbours, the best method for locating sounding is
 (a) two angles from shore (b) two angles from boat
 (c) fixed intersecting ranges (d) one angle from shore
- 4.6** The most common weir to measure stream discharge is
 (a) rectangular notch (b) triangular notch
 (c) cipolletti notch (d) stepped notch
- 4.7** In floods the most suited weir is
 (a) rectangular notch (b) triangular notch
 (c) cipolletti notch (d) stepped notch
- 4.8** The purpose of making a hydrographic survey is
 (a) to determine the quantities of subaqueous excavations
 (b) to measure areas subjected to scouring and silting in harbours

- (c) to measure soundings and preparing navigation charts
- (d) All the above

4.9 In hydrography

- (a) only horizontal control is required
- (b) only vertical control is required
- (c) both horizontal and vertical control are required
- (d) no such control is required.

4.10 Sounding in hydrography is defined as the

- (a) distance travelled by sound waves in a water body
- (b) depth of the water body at the point of measurement
- (c) difference of gauge reading and water level at the time of observation
- (d) All the above

Answers to Objective-type Questions

4.1 (b)	4.2 (c)	4.3 (d)	4.4 (c)	4.5 (c)	4.6 (a)
4.7 (d)	4.8 (d)	4.9 (c)	4.10 (b)		

5

AERIAL SURVEY

Introduction

The terms aerial surveying, aerial photography and aerial photogrammetry are used synonymously to designate the taking of photographs from the air, depicting a portion of the earth's surface. Advancements in instrumentation and techniques in photogrammetry in the past 70 years have enabled aerial survey, the most accurate and efficient method for compiling maps and generating topographic details and information. The expression *aerial survey* is used to cover both photogrammetry and photo interpretation. Photogrammetry may be defined as the science and art of producing a mosaic or map by compiling the photographs. The word photogrammetry is derived from the Greek words 'photos', 'gramma', and 'metron' meaning 'light', 'that which is drawn or written' and 'to measure', respectively. Therefore, its literal meaning is measuring from a photograph and thus it may be defined as the science of obtaining reliable measurements by means of photographs in order to primarily determine geometric characteristics such as size, form and position of the object. When the measurements are of such a nature so as to give descriptive information about the photographed object, the term photo interpretation is used. It is the task of drawing inferences from the qualities of photographed objects.

A photograph taken from an altitude will be a perspective projection. The concept of using the perspective projection to depict the terrain was first utilised by a French engineer Colonel Amie Laussedat of Turich in 1861 even before the advent of cameras. Laussedat used *Lucida*, an apparatus resembling the usual cameras from outside. It was not designed to take pictures; instead an arrangement was made to deflect the light rays, from the object, by 90° on to a transparent medium where the image formed was sketched. A number of French, German and Russian engineers worked for many years to develop the methods and equipments to substitute sketching by photographs. In 1909 and thereafter, Pulfrich, a German engineer, conducted experiments to use stereoscopic pairs of photographs for mapping purposes. He laid the foundation for most of the modern methods involving the use of stereoscopic work with both the terrestrial and aerial photographs.

In terrestrial or ground photogrammetry, the instrument used is a photo-theodolite, a combination of a photographic camera fitted on a tripod with its axis horizontal and a theodolite. Terrestrial photographs are taken with the photo-theodolite, the position and orientation of which are often measured directly at the time of exposure. On the other hand, a camera with its axis vertical, mounted on an aircraft, is used for aerial photogrammetry. Mapping by aerial photogrammetry has a wider scope as compared to terrestrial photogrammetry since the latter is limited to the plotting of special features, e.g., vertical cliffs, mountainous terrain, etc.

Terrestrial photogrammetry was first introduced in 1886 by Dr. E. Deville who was the Surveyor General of Canada. It is almost similar to plane tabling. Various ground features surveyed are shown in views photographed from ground stations. The features are thus located by the intersection of sight lines from two or more such stations. The plotting work is done in the field. The accuracy attained is not as great as with a plane table but the speed is increased.

Aerial photogrammetry was first utilised in the first world war. Since then, lot of improvements have been made in the development of aerial cameras, and in reducing the photographic information to maps, which has led to the replacement of terrestrial photogrammetry by the aerial one. Also, the speed of work and the degree of precision achieved are much higher in aerial photogrammetry. The aerial surveys are invaluable in terrains difficult and impossible for ground surveys.

In aerial photogrammetry, vertical photographs (a photograph with the camera axis along the direction of gravity) of the ground are taken with a camera fixed in an aircraft flying along predetermined parallel flight lines. It is already mentioned, that a photograph taken from an aircraft is not the true representation of the earth's surface, since it is a perspective view rather than an orthographic projection. For precise work, it is therefore required to correct any error that might otherwise arise due to distortion, tilt and variations in flying heights. For this purpose, it is essential to establish some ground control. A few points on the ground are located, which can be easily identified on the aerial photographs. These points are fixed on the ground to control the scale of the plot. Some of the other associated problems are that of the direct measurement of base length between two camera positions, different elevations of ground points and angle of exposure of photographs. The aircraft should fly horizontally at the predetermined altitude so as to maintain a uniform scale of the photographs. Since the scale of the photographed object and its position varies according to the distance of the corresponding actual object relative to the camera, these differences are also measured to determine the location and elevation of the details photographed.

For an area to be surveyed by aerial photography, a series of parallels, called *flight strips* are planned. The photographs are normally exposed in such a way that the area covered by each successive photograph along a flight strip overlaps part of the coverage of previous photo. The overlaps of the photographs is observed for the two consecutive passes as well. Sufficient number of photograph along with longitudinal and side overlaps (60% and 30%, respectively) are thus taken covering the entire area to be surveyed. The overlap is necessary to examine the photographs stereoscopically. Each pair of photographs (i.e., every two consecutive

photographs) taken with sufficient overlap are known as *stereo pair*. A number of stereo pairs are used to compile a controlled or photogrammetric map. The former provides a photographic picture of the terrain while the latter represents the country by means of symbols. When a stereo pair is viewed through a stereoscope, the picture seen is the same as viewed by an aircraft camera.

The most important aspect of aerial photogrammetry, perhaps, is the interpretation of the photograph. The factors which aid in identification of objects on the photographs have been dealt with separately. A successful and correct interpretation requires a thorough study, experience and an understanding of geology, engineering and other related sciences.

Photographic surveying is a valuable method of obtaining relatively small-scale maps of large areas and large-scale maps of relatively small areas. It is very useful for topographical surveys, preliminary route surveys, i.e., highways, railways pipelines, etc., forest and agricultural surveys, and the like. Its use for various purposes needs no emphasis. Exploration, geology, city zoning and planning, parks, traffic studies and especially the route selection are some of the fields in which it is most suitable and is gaining popularity. With the development of geographic information system (GIS) the importance of application of photogrammetry has increased tremendously. Several topographic maps are prepared for use in providing spatial data for geographical information systems. Since the latest development of photogrammetric products, orthophotos, and digital elevation models are in digital format, these are compatible for direct entry into GIS database.

The main advantages of aerial photogrammetry are the speed with which an area is covered, the ease with which topography of inaccessible areas can be detailed, no possibility of omitting any field data, the tremendous amount of details shown at an economy.

5.1 TERRESTRIAL STEREOPHOTOGRAPHY

Terrestrial stereophotography is performed with a photo-theodolite which is a combination of a theodolite and a precise camera. A high-quality objective is mounted on one side of the camera, and on the opposite side, a plate carrier to which a sensitive plate or film measuring 13 cm × 18 cm fits closely when the photographs are taken.

The plate carrier is provided with *fiducial marks* in the shape of teeth or small round holes. These marks and the focal length of the objective, which is about 20 cm, are printed on the plate or film, as well as on the photograph. By joining the opposite fiducial marks, we obtain the axes of the coordinates with respect to which the position of any point on the photograph can be determined. The fiducial marks are made so that the point of intersection of the coordinate axes on the photograph should coincide with the foot of a perpendicular dropped from the centre of the objective onto the plane of the plate carrier. The foot of this perpendicular on the photograph gives its 'principal point'. Hence, the origin of the coordinates on the photograph should coincide with the principal point of the photograph. The line passing through the optical centre of the objective and the principal point represents the optical axis of the camera.

Figure 5.1 shows two stations (S_1 and S_2) of the photo-theodolite on the ends of a base line of length B . The position of the camera optical axes with respect to the base line may vary, but the frequent practise is to use the horizontal axes perpendicular to the base line (the case of 'normal axes') or equally deflected from the base line to both sides. If the centre of the left objective is assumed to be the origin of coordinates for the given pair of stereoscopic photographs, from the focal length f (S_1, k_1) of the camera objective, the length of base line B , coordinates x_1 and z_1 of the image m_1 on the left photograph of a given point M on the ground, it is possible to determine three space coordinates X, Y, Z of this point in the system of coordinates called *photogrammetric system*.

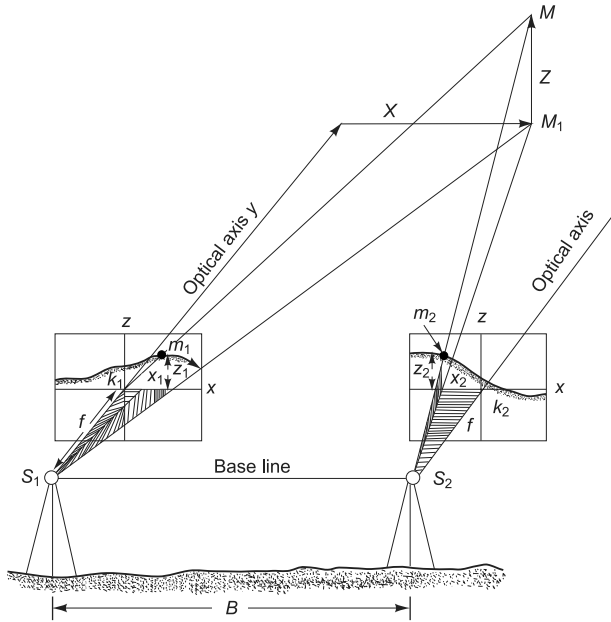


Fig. 5.1 Terrestrial photogrammetry

The direction of the optical axis is accepted in this system as the Y -axis, and the horizontal line perpendicular to the Y -axis (in a particular case, if the base line is horizontal, the X -axis coincides with the base line) is accepted as the X -axis, the vertical line passing through the centre of the objective is accepted as the Z -axis. In order to calculate these three space coordinates of different points of the ground with respect to the centre of the photo-theodolite left objective, it is necessary to measure on both negatives the difference of the abscissae and coordinate z_1 on the left negative for each point. Then, in case of normal axes, as shown in Fig. 5.2,

$$\frac{Y}{f} = \frac{B}{p}$$

where $p = x_1 - x_2$

(5.1)

is called *linear horizontal parallax* of the space point M . From Fig. 5.2 we can also find the space coordinates.



$$\frac{X}{x_1} = \frac{Z}{z_1} = \frac{Y}{f} = \frac{B}{p}$$

$$\left. \begin{aligned} X &= x_1 \frac{B}{p} \\ Y &= f \frac{B}{p} \\ Z &= z_1 \frac{B}{p} \end{aligned} \right\}$$

(5.2)

In order to compile a general plan it is necessary to interconnect all the base lines, and the coordinates and azimuth or the grid bearing of the optical axis in the general geodetic system of coordinates must be determined for every left station. However, the calculation of space coordinates of individual points is a slow and tedious process. Therefore, in stereophotogrammetry the procedure of office processing is usually mechanised by special instruments which are called 'stereoautographs' and 'stereoplanigraphs'.

These instruments produce a model of the ground comparable to the one produced by a stereocomparator. Placing the instrument mark in contact with representative points of the model, the position of the corresponding point of the ground is obtained automatically to a selected scale on the plan attached to the instrument table. If the mark is moved over the model on a given height, the pencil will draw a corresponding contour line on the plan. For stereophotogrammetric control it is necessary to have a sufficient number of control points. By comparing the position of such points, obtained by geodetic measurements, with the position of the same points, calculated from the measurements on photographs, it is possible to judge the quality of aerial photography.

Terrestrial stereoscopic surveys may be performed on any scale from 1:20 to 1:50 000. The larger the scale, the shorter must be the distance Y of the camera from the objects being taken. The accuracy of the photograph is increased with the length of the base line, but at the same time, the usable size of the photographs is reduced correspondingly. Table 5.1 presents approximately Y and B values for certain scales.

Table 5.1

<i>Scale</i>	$Y_{max} - Y_{min}$ (m)	<i>B</i> (m)
1:1000	300 – 60	15
1:5000	1500 – 300	75
1:10000	3000 – 600	150

During terrestrial stereoscopic surveys the field work consumes about 25% of the time and the remainder is used for office processing; during plane table surveys, nearly all work is performed in the field and only a small part of it is done in the office. In engineering, terrestrial stereoscopic photography is employed in inaccessible areas with a sharply defined relief where field location survey is difficult. Terrestrial stereoscopic survey is not inferior in accuracy to that of plane table surveying and enables a stereo pair to be processed at any time and for a variety of practical purposes. Stereo pairs enable location directly from a spatial model.

Terrestrial stereoscopic photography can be employed in studying deformation of structures under the effect of loads and temperature, particularly so when the structures are inaccessible for direct measurement. The structure photographs are taken several times from the same points under different loads and temperatures. Then, using a stereo comparator, the coordinates of the same points of the structure are measured and their changes are used to determine the value of the deformation and its tendency.

5.2 AERIAL PHOTOGRAPH

These are the photographs obtained as a result of photography of ground from air with a camera mounted on an aircraft. The image of the ground photographed is formed on the focal plane of the camera's objective, where a sensitive film is placed. Various objects in the photograph are distinguished due to the distribution

of light and shade and with the aid of some other factors, the details of which will be discussed in Sec. 5.10.

Any representation of objects on a plane constructed according to a definite mathematical law is formed by straight rays of light emanating from ground points and passing through the camera objective with the optical centre of the objective as the *centre of projection*. Therefore, an aerial photograph may be regarded as a *gnomonic* or *perspective* projection. From the mathematical point of view, however, the plan of the ground or a map is an orthographic projection, which is produced on a plane by projecting the ground points through straight lines perpendicular to the plane.

In a particular case, both of these projections, the perspective and the orthographic, produce similar images. This is when the ground is flat and the film is horizontal, i.e., the optical axis of the camera is vertical. Such a photograph is known as a *vertical photograph* and it may be considered to represent the plane of the ground.

In spite of all the sophistication, it is impossible to maintain the optical axis of the camera truly vertical in practice resulting in tilted photographs. In case the vertical axis of the camera is unintentionally inclined to the vertical by not more than 3° , the photograph is called *tilted*. However, sometimes the vertical axis of the camera is intentionally inclined to the vertical and then the resultant photograph is known as *oblique*. A high-oblique photograph is one that includes the horizon, whereas a low-oblique photograph does not include it.

In general, vertical air photography is resorted to. However, in practice, most of the photographs contain tilt, despite all the precautions taken to keep the camera axis vertical. Also, it is almost impossible to have a ground exactly flat. Therefore, because of the pronounced ground relief and associated tilt, an aerial vertical photograph does not represent an image similar to the orthographic projection of the photographed area and consequently cannot be considered as a true plan of the ground.

5.2.1 Map versus Aerial Photograph

The natural and cultural features shown on a photograph are similar to those on a planimetric map, but with the following differences:

1. Map is an orthogonal projection, whereas an aerial photograph is a central projection, i.e., perspective projection.
2. Map has a single constant scale, whereas it varies from point to point depending upon their elevations on an aerial photograph.
3. The amount of details on a map are selective, but in an aerial photograph a wealth of detail is there.
4. Due to symbolic representation the clarity of details is more on maps than on a photo.

5.3 DEFINITIONS

1. *Exposure station*: The position of the aircraft at the point of each camera exposure is known as the *exposure station* or *air station*. More precisely, an

exposure station is the point in space occupied by the camera lens at the instant of exposure.

2. *Flying height*: It is the elevation of the exposure station above the datum.

3. *Altitude*: It is the vertical distance of the aircraft above the earth's surface.

4. *Tilt*: It is the rotation of the aerial camera about the line of flight. In Fig. 5.3, angle kOn (t) is a vertical angle obtained by the intersection of the optical axis with the plumb line at the exposure station. It lies in the principal plane.

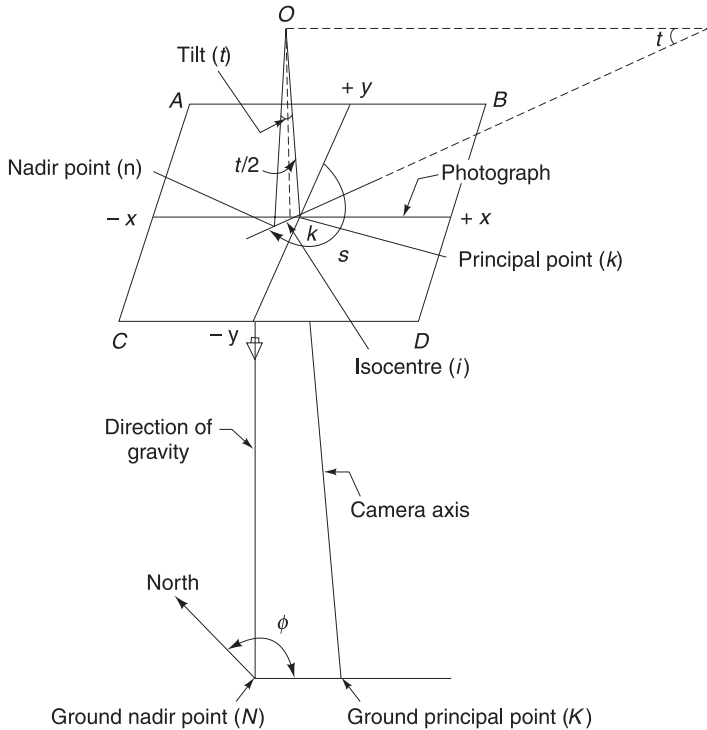


Fig. 5.3 Tilted photograph

5. *Tip*: It is the rotation of the aerial camera about a horizontal axis normal to the line of flight. This is also known as *swing*. In Fig. 5.3, the angle s measured in the plane of the photograph from the positive Y -axis clockwise to the nadir point is the swing.

6. *Principal point*: It is the point of intersection of the optical axis of the aerial camera with the plane of the aerial photograph and is shown as k in Fig. 5.3. It is considered to coincide with the intersection of x and y axes of the photograph. The point K is known as *ground principal point* where the line Ok when produced meets the ground.

7. *Isocentre*: It is the point on an aerial photograph in which the bisector of the angle of tilt meets the photograph. In Fig. 5.3, i is isocentre. It lies on the principal line at a distance of $f \tan(t/2)$ from the principal point.

8. *Nadir point*: It is the point on an aerial photograph where a plumb line dropped from the front nodal point pierces the photograph. In Fig. 5.3, n is the nadir point. It is also known as *photo plumb point*. The point N immediately below it on the ground is known as *ground nadir point*.

Note For a vertical photograph, the isocentre and the photo nadir point coincide with the principal point.

9. *Homologous points*: In perspective projections, rays originating from one plane pass through a point before projecting on another plane. The pairs of points representing ground points and their photo points are called homologous points.

5.4 SCALE OF VERTICAL PHOTOGRAPH

Photographs taken from an aircraft with the optical axis of the camera vertical are the vertical photographs. It is seldom that the terrain to be photographed is flat and the photographs taken are vertical. Therefore, because of the ground relief, tilt and tip the aerial photographs do not have uniform scale and vary from point to point. However, a datum may be selected and if the altitude of the camera is known, the scale of the photography may be computed for the datum chosen. Any change in altitude of the airplane alters the scale of the photograph. Most of the scale (in unit equivalents) are between 1 cm = 80 m and 1 cm = 800 m. The scale 1 cm = 80 m can also be expressed in unit fraction (1 cm/80 m), dimensionless representative fraction (1/8000), or dimensionless ratio 1:8000.

Figure 5.4(a) shows the side view of a vertical photograph taken over flat terrain. The scale of the photograph,

$$S = \frac{\text{photo distance}}{\text{ground distance}} = \frac{ab}{AB} = \frac{f}{H'} = \frac{f}{H - h} \quad (5.3)$$

Figure 5.4(b) shows the side view of vertical photograph on terrain with varying elevations. The distance of the object from the camera will vary and the photo scale will likewise vary; scale will increase with increasing terrain elevation and vice versa. As shown in Fig. 5.4(b), the line AB lies at an elevation h_1 above the datum. In similar triangles Oab and OAB , the scale of the photographic image of the line ab to the ground line AB is

$$\text{Scale, } S = \frac{ab}{AB} = \frac{f}{H - h_1}$$

The line CD lies at an elevation of h_2 above the datum. Therefore, the scale of the photographic image of the line cd to the ground line CD is

$$\text{Scale, } S = \frac{cd}{CD} = \frac{f}{H - h_2}$$

In general, the scale S of a photograph for any ground elevation is

$$S = \frac{f}{H - h} \quad (5.4)$$

where H is the altitude of the aircraft above mean sea level, h the elevation of the ground above mean sea level, and f the focal length of the camera.

Note

In Eq. (5.4) as in Eq. (5.3), scale of a vertical photograph is simply the ratio of image distance to object distance. For vertical photographs taken over variable terrain, there are an infinite number of different scales. This is one of the principal difference between a photograph and a map.

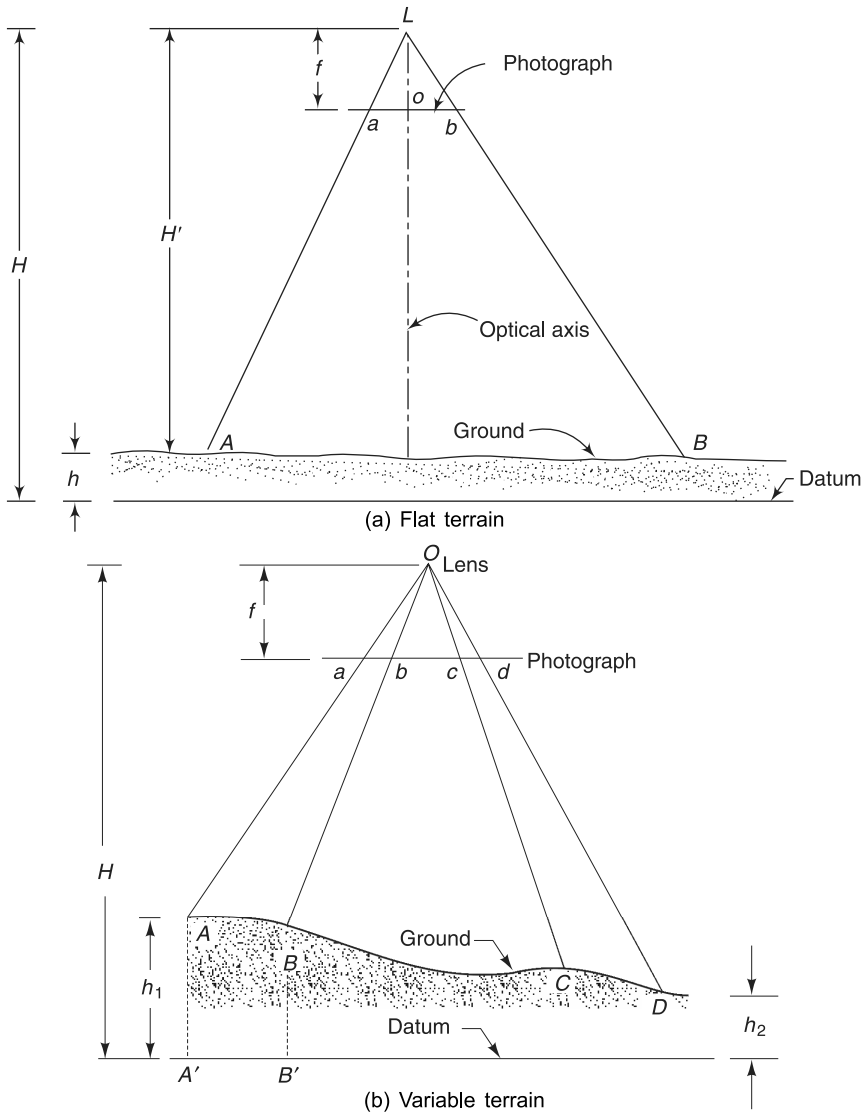


Fig. 5.4 Scale of a vertical photograph

It is often desirable to use an average photographic scale of a vertical photograph over a variable terrain, to define the overall mean scale of the vertical photograph. Such a scale can be determined from Eq. (5.4) using the average aircraft altitude for H , and the average elevation of the area for h . The average elevation of the area is determined from the contour map of the area and selecting the apparent

average elevation. However, the ground distances obtained by using the average scale will provide only an approximation.

The *datum scale* is the scale effective over the entire photograph and is determined as if all the ground points were projected vertically on the datum.

The datum scale, $S_d = \frac{f}{H}$

Length of a Line from a Vertical Photograph Although it is difficult to have a uniform scale for a vertical photograph, the lengths of the lines can be determined from the photographic coordinates (x, y) and ground coordinates (X, Y). Thus,

$$\frac{x}{X} = \frac{f}{H - h} \quad (5.5)$$

and
$$\frac{y}{Y} = \frac{f}{H - h} \quad (5.6)$$

The ground coordinates are referred to the plumb point as the origin, the coordinate axes being parallel to the geometrical axes of the photograph, the X-axis being in the direction of the line of the flight.

Let it be required to determine the length of a line AB on the ground from the coordinates of the end points a and b photographed using the principal point as the origin. From Eqs. (5.5) and (5.6)

$$\begin{aligned} X_A &= \frac{H - h_A}{f} x_A, & Y_A &= \frac{H - h_A}{f} y_A \\ X_B &= \frac{H - h_B}{f} x_B, & Y_B &= \frac{H - h_B}{f} y_B \end{aligned}$$

The true horizontal distance

$$AB = \sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2} \quad (5.7)$$

5.5 SCALE OF TILTED PHOTOGRAPH

In spite of all the sophistication, it is impossible to maintain the optical axis of the camera truly vertical in practice resulting in tilted photographs. For an aerial photograph of a flat ground, the downward half of the photograph will have a larger scale than the upward half. The problem becomes more complicated where the ground relief is also there and to determine the scale from point to point, the position of the points must be known with reference to the principal lines.

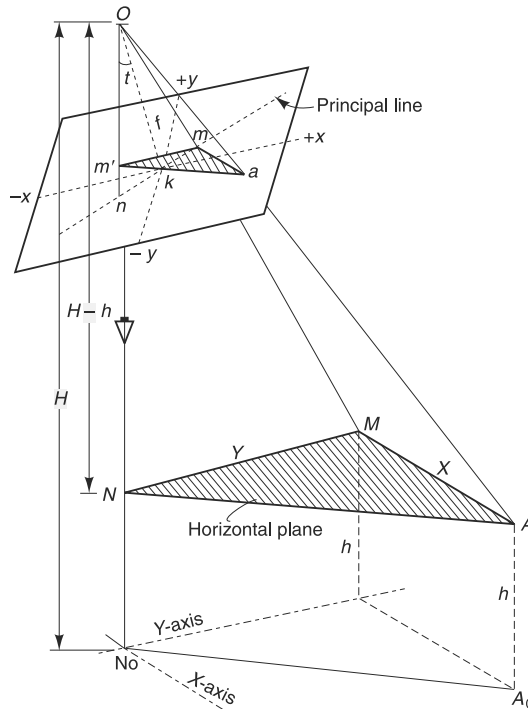
Figure 5.5 portrays a tilted photograph in which a is the image of the ground point A , the ground point being at a height h above datum. The points n , and k are the nadir and principal points, respectively, and nk is the principal line. Draw a perpendicular am to the principal line. From m draw mm' perpendicular to the plumb line. Since am is horizontal, line mm' will also be horizontal. Therefore, triangle amm' lies in a horizontal plane.

From similar triangles $Om'a$ and ONA .

$$\frac{m'a}{NA} = \frac{Om'}{ON} = \frac{\text{Map distance}}{\text{Ground distance}} = \text{Scale of the point on a tilted photograph}$$

and

$$S = \frac{f \sec t - mn \sin t}{H - h}$$



Let θ be the angle between the y-axis and the principal line (Fig. 5.6), (counterclockwise θ being negative), and n and k are the nadir and principal points. Let the photographic coordinates of the point a be x and y .

$$y' = mn = km + kn = ed - eb + f \tan t = x \sin \theta - y \cos \theta + f \tan t$$

Hence,
$$S = \frac{f \sec t - y' \sin t}{H - h} \quad (5.8)$$

5.5.1 Ground Coordinates

The ground length of a line can be computed from the ground coordinates (X , Y) of ground point obtained from the coordinates (x' , y') of its image on the tilted photograph.

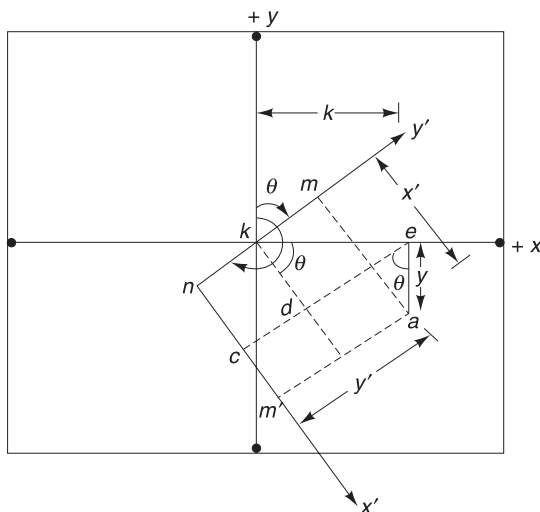


Fig. 5.6 Coordinate axes through plumb point

From the Figs. 5.5 and 5.6,

$$AM = X = \frac{H - h}{f \sec t - y' \sin t} x'$$

$$NM = Y = \frac{H - h}{f \sec t - y' \sin t} y' \cos t$$

Hence, length of the ground line = $\sqrt{(X_a^2 - X_b^2) + (Y_a^2 - Y_b^2)}$

Example 5.1 A line AB measures 11.00 cm on a photograph taken with a camera having a focal length of 21.5 cm. The same line measures 3 cm on a map drawn to scale of 1/45 000. Calculate the flying height of the aircraft, if the average altitude is 350 m.

Solution

$$\frac{\text{Photo scale}}{\text{Map scale}} = \frac{\text{Photo distance of line } AB}{\text{Map distance of line } AB}$$

$$\text{i.e.,} \quad \frac{S}{1/45000} = \frac{11.00}{3.00}$$

$$\Rightarrow S = \frac{11.00}{3.00} \times \frac{1}{45000} = \frac{1}{12272.73} \approx \frac{1}{12273}$$

$$\text{Photoscale} = \frac{f}{H - h_1} = \frac{0.215}{H - 350}$$

$$\text{Hence,} \quad \frac{0.215}{H - 350} = \frac{1}{12273}$$

$$\Rightarrow H = 2988.695 \text{ m}$$

Example 5.2 A line 2350 m long lying at an elevation of 500 m measures 10.50 cm on a vertical photograph. The focal length of the camera used is 20 cm. Determine the scale of the photograph for an area having an elevation of 1200 m.

Solution

$$\text{Scale at 500 m} = \frac{10.50}{2350} = \frac{20}{H - 500}$$

$$\text{i.e.,} \quad H - 500 = \frac{20 \times 2350}{10.50}$$

$$\text{or} \quad H = 4976.19 \text{ m}$$

$$\text{Scale at 1200 m} = \frac{20}{(4976.19 - 1200) \times 100} = \frac{1}{18880.95} \approx \frac{1}{18881}$$

$$\text{Scale is 1 cm} = 188.81 \text{ m or } 1:18881$$

Example 5.3 Two points *A* and *B* having elevations of 650 m and 250 m, respectively, above datum, appear on a vertical photograph obtained with a camera of focal length of 250 mm and flying altitude of 2700 m above datum. Their correlated photographic coordinates are as follows:

Point	Photographic coordinates	
	<i>x</i> (cm)	<i>y</i> (cm)
<i>a</i>	+3.65	+2.54
<i>b</i>	-2.25	+5.59

Determine the length of the ground line *AB*.

Solution

The ground coordinates are given by

$$X_a = \frac{H - h_a}{f} x_a = \frac{2700 - 650}{250} \times (+3.65 \times 10) = +299.3 \text{ m}$$

$$Y_a = \frac{H - h_a}{f} y_a = \frac{2700 - 650}{250} \times (+2.54 \times 10) = +208.28 \text{ m}$$

$$X_b = \frac{H - h_b}{f} x_b = \frac{2700 - 250}{250} \times (-2.25 \times 10) = -220.5 \text{ m}$$

$$Y_b = \frac{H - h_b}{f} y_b = \frac{2700 - 250}{250} \times (+5.59 \times 10) = +547.82 \text{ m}$$

$$\text{Now,} \quad (X_a - X_b)^2 = (299.3 + 220.5)^2 = 27.0192 \times 10^4$$

$$\text{and} \quad (Y_a - Y_b)^2 = (208.28 - 547.82)^2 = 11.5287 \times 10^4$$

$$\begin{aligned} \text{Hence,} \quad AB &= \sqrt{(X_a - X_b)^2 + (Y_a - Y_b)^2} \\ &= \sqrt{(27.0192 + 11.5287) \times 10^4} = 620.87 \text{ m} \end{aligned}$$

5.6 DISPLACEMENTS AND ERRORS IN AERIAL PHOTOGRAMMETRY

On an aerial photograph, the true perspective positions of points differ from their orthographic positions because of the inability of an aircraft to maintain the flying height at a constant altitude, changes in scale of the photograph caused by ground relief, and displacement caused by tilt of the camera axis. Consequently, such a photograph is not a map. If all such sources of errors are neglected, the aerial photograph may be regarded as a true orthographic projection. In practise, it is impossible to neglect such conditions. However, to account for distortions, some corrections are formulated which when applied provide the true measurements.

The principal causes of distortion are the displacement due to ground relief, and the displacement due to the tip and tilt of the camera.

5.6.1 Displacement Due to Ground Relief

Figure 5.7 shows the images of the top and bottom of a photographed ground object AA_2 (say, for example, a tower). It will be seen that the top and bottom are photographed at different places a and a_2 , whereas a map would show them at the same position. This displacement of the image due to the height of the object is known as *height distortion*. Since all the photographs are perspectives, all the objects except those in the datum plane and the plumb point, are displaced.

The distance on a photograph, from the image of any ground point to its fictitious image projected to a datum plane is the image displacement caused by topographic relief and is known as *relief displacement*. Relief displacement occurs when the point being photographed is not at an elevation of the mean datum. The practical aspects of relief displacement relate primarily to the proper horizontal location of all points, and secondarily to the assemblage of massaiacs—a series of overlapping aerial photographs that form one continuous picture.

On vertical photographs, the relief displacements are always radial from the principal point. The concept of relief displacement is illustrated in Fig. 5.7, which represents a vertical photograph taken from a height H above a datum. The image of the terrain point A , which has an elevation h , above datum, is located at a on the photograph and the corresponding imaginary point A_1 located vertically beneath A in the datum plane finds its image position on the photograph at a_1 . Here aa_1 is the relief displacement of the point A due to its elevation h . Similarly, the relief displacement of point A_2 of the object AA_2 is the line aa_2 on the photograph.

Let

r_0 = radial distance of a from the principal point k .

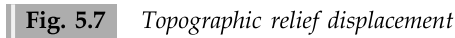
r_1 = radial distance of a_1 from the principal point k .

R = ground distance AK_0 .

From similar triangles a_1kO and A_1K_1O ,

$$\frac{f}{H} = \frac{r_1}{R}$$

$$\text{or} \quad r_1 = \frac{Rf}{H} = a_1 \quad (5.9)$$


$$\frac{f}{H - h_1} = \frac{r_0}{R}$$

$$d = ka - ka_1 = r_0 - r_1$$

$$d = ka_2 - ka = r_2 - r_0 \quad (5.12)$$

From Eq. (5.10), $r_0 = \frac{Rf}{H - h_1}$

and $\frac{f}{H - (h_1 + h_2)} = \frac{r_2}{R}$

or $r_2 = \frac{Rf}{H - (h_1 + h_2)}$ (5.13)

Now, $d = r_2 - r_0$

$$= \frac{Rf}{H - (h_1 + h_2)} - \frac{Rf}{H - h_1}$$

$$= Rf \left[\frac{(H - h_1) - \{H - (h_1 + h_2)\}}{\{H - (h_1 + h_2)\}(H - h_1)} \right]$$

$$= \frac{Rfh_2}{\{H - (h_1 + h_2)\}(H - h_1)}$$

or $d = \frac{r_2 h_2}{H - h_1}$ (5.14)

From Eqs. (5.11) and (5.14) it is observed that the relief displacement decreases with the increase in flying height, whereas it increases as the distance from the principal point increases. Also, it is positive and radially outwards for points above the datum and negative and radially inwards for the points below the datum. For the points vertically below the exposure station, the relief displacement is zero. If the ground is flat ($h = 0$), the relief displacement will be zero and thus the vertical photograph can be used as a map.

From the above observation, it can be concluded that the central portion of an aerial photograph, within the boundaries of which the relief displacements do not exceed the permissible limits, can be established and used for the measurements. The area of the photograph having the least permissible distortion is known as the *working area*.

Notes

1. Relief displacement often causes the linear structures such as straight roads, canals, fences, etc., on rolling grounds to appear crooked on a vertical photograph, especially near the edges of the photograph. Therefore, the interpreter prefers to use the working area only.
2. Image displacements on tilted photographs caused by photographic relief occur much the same as they do on vertical photos, except relief displacements on tilted photographs occur along radial lines from the nadir point. Equations (5.11) and (5.14) may be used satisfactorily for tilted photographs also as the error introduced will be insignificant.

Example 5.4 A tower AB 80 m high appears in a vertical photograph. The flying height of the aircraft above m.s.l. is 3000 m. The distance of the image of the top of the tower from the principal point is 7.25 cm. Compute the displacement of the top of the tower with respect to the image of its bottom if the elevation of the bottom of the tower is 1150.0 m.

Solution $H = 3000$ m, $h_1 = 1150$ m, $h_2 = 80$ m, $r_2 = 7.25$ cm.

The displacement d of the image of the top with respect to the image of the bottom is given by

$$d = \frac{r_2 h_2}{H - h_1}$$

$$d = \frac{7.25 \times 80}{(3000 - 1150)} = 0.31 \text{ cm}$$

Example 5.5 The distance from the principal point to an image on a photograph is 7.50 cm and the elevation of the object above the datum (sea level) is 350 m. What is the relief displacement of the point if the datum scale is 1/8000 and the focal length of the camera is 250 mm.

Solution $h_1 = 350$ m, $r_1 = 7.5$ cm, $S_d = 1/8000$.

The datum scale is given by

$$S_d = \frac{1}{8000} = \frac{250/1000}{H}$$

or
$$H = \frac{250}{1000} \times 8000 = 2000 \text{ m above m.s.l.}$$

Again, the relief displacement is given by

$$d = \frac{r_1 h_1}{H - h_1} = \frac{7.50 \times 350}{2000 - 350} = 1.590 \text{ cm}$$

Example 5.6 The images x and y of the base and top, respectively of a factory chimney 150 m high are observed in a truly vertical aerial photograph of scale 1 = 10000. Determine the position of x given that y is 70.0 mm from the principal point of the photograph. Take the focal length of the camera to be 125 mm and assume the chimney to be at datum level.

Solution Refer to Fig. 5.8. Let h_1 be the height of the point X above datum and let H be the height of the camera above datum.

$$h_2 = 150 \text{ m}, \quad S = 1:10\,000, \quad f = 125 \text{ mm}, \quad r_2 = 70 \text{ mm}$$

$$\text{Scale} = \frac{f}{H - h_1}$$

$$\frac{1}{10\,000} = \frac{125}{H - h_1}$$

or
$$H - h_1 = 1250 \text{ m}$$

From Eq. (5.13),

$$R = \frac{r_2 \{(H - h_1) - h_2\}}{f} = \frac{(70/1000) \times (1250 - 150)}{125/1000} = 616 \text{ m}$$

From Eq. (5.10),

$$r_0 = \frac{Rf}{H - h_1} = \frac{616 \times 125}{1250} = 61.6 \text{ mm}$$

$$\begin{aligned} \text{Now} \quad 0.08 - 0.07 &= \frac{0.08 \times h}{446.5} \\ \Rightarrow \quad h &= 55.81 \text{ m} \end{aligned}$$

5.6.2 Displacement Due to Tilt

We have seen that the images of the points photographed on vertical photographs, the plane of photographs being horizontal, are displaced due to ground relief. For a vertical photograph, the camera axis is truly vertical. For a tilted camera position, the plane of photographs taken no longer remain horizontal, consequently additional displacement causing photographic distortion. In ordinary surveying practise, the formulae derived for tilted photographs are seldom used. Instead, the displacement due to tilt is rectified by using special cameras and the process is known as *rectification*.

Tilt is the rotation of the camera axis about the line of flight and tip is the rotation of the camera axis normal to the flight line. However, when considering an aerial photograph, tilt is generally understood to mean the combined effect of both of these effects and it causes the image of the object on the film to be displaced. Displacements due to tilt are usually less than those produced due to ground relief. These displacements are errors which are zero at the centre of the photograph and increase towards the elevated and depressed edges; the displacements are radial about the isocentre.

If the amount of tilt is known, the amount of displacement may be computed. But it is very difficult to determine the angle of tilt analytically, graphically or by a rectifying instrument. To determine the tilt, the position of the images of at least three ground points, whose position have been determined by ground surveying, must appear on the photograph.

Let us consider two ground points A and B photographed as a and b on a tilted photograph, and a' and b' on a vertical photograph (Fig. 5.9). If the vertical photograph is rotated about the axis of tilt by the angle of tilt, the images a' and b' would fall on a'' and b'' , respectively. The corresponding tilt displacements will be $a'a''$ and $b'b''$.

The lines ab and $a'b'$ are the principal lines of the tilted and vertical photographs.

The tilt displacement of a with respect to a'

$$d_1 = ia' - ia$$

$$\text{But } ia' = n'a' - n'i = f \tan (t + \alpha) - f \tan (t/2)$$

$$\text{and } ia = ka + ki = f \tan \alpha + f \tan (t/2)$$

$$\text{Hence, } d_1 = f \tan (t + \alpha) - f \tan (t/2) - f \tan \alpha - f \tan (t/2)$$

$$\text{or } d_1 = f [\tan (t + \alpha) - \tan \alpha - 2 \tan (t/2)] \quad (5.15)$$

Similarly,

$$d_2 = f [\tan \beta - \tan (\beta - t) - 2 \tan (t/2)] \quad (5.16)$$

where d_2 is tilt displacement of b with respect to b' .

The angles α and β can be found by

$$\tan \alpha = \frac{ka}{f}, \quad \tan \beta = \frac{kb}{f}$$

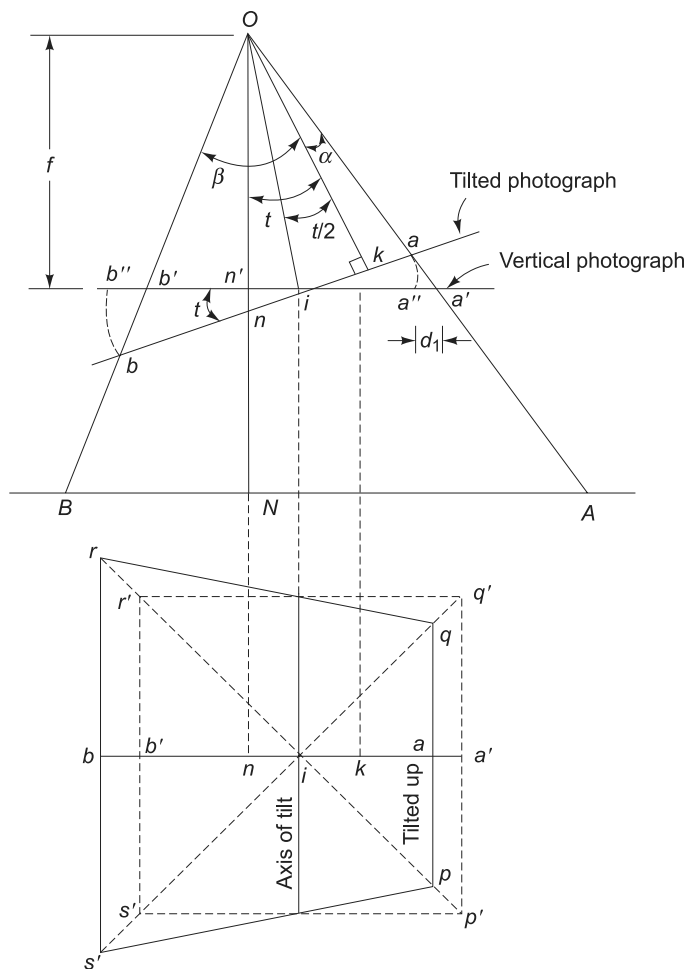


Fig. 5.9 Tilt displacement

Equations (5.15) and (5.16) are used for determining exact tilt displacements. For approximate values, the following expression may be used:

$$d = \frac{(ia)^2 \sin t}{f - ia \sin t} \quad (5.17)$$

All the above three expressions give tilt displacement for a point on the principal line. For a point not lying on the principal line, the displacement is always more and can be found from the following expression:

$$qq' = aa' \sec I$$

where qq' is the tilt displacement of the image q of a ground point Q , and I is the angle at the isocentre from the principal line to the point q . Hence, the ratio of the tilt displacement of a point not on the principal line to that of a point on the principal line is equal to the secant of the angle at the isocentre from the principal line to the point.

The corresponding tilt displacements will be

$$d_3 = f \sec I [\tan (t + \alpha) - \tan \alpha - 2 \tan (t/2)] \quad (5.18)$$

$$\text{and} \quad d_4 = f \sec I [\tan \beta - \tan (\beta - t) - 2 \tan (t/2)] \quad (5.19)$$

where d_3 and d_4 are the displacements on the upper and lower halves of the photograph, respectively.

Since, $ia'' = ia$ and $ib'' = ib$; therefore, the triangles aia'' and bib'' are both isosceles.

Also, $\angle aia'' = \angle bib'' = t$
in both the triangles, the remaining angles will each be

$$= \left(\frac{180^\circ - t}{2} \right) = \left(90^\circ - \frac{t}{2} \right)$$

But $\angle kiO$ is also equal to $[90^\circ - (t/2)]$

$$\therefore \angle ia'' = \angle kiO$$

But these are alternate angles; therefore, it follows that Oi is parallel to aa'' and similarly, Oi is parallel to bb'' .

Therefore, triangle Oia' is similar to the triangle $aa''a'$ and triangle Oib'' is similar to the triangle $bb''b'$.

$$\frac{a'a''}{aa''} = \frac{ia'}{Oi}$$

$$\text{and} \quad \frac{b'b''}{bb''} = \frac{ib'}{Oi}$$

$$\text{but} \quad Oi = f \sec (t/2)$$

$$\text{From triangle } ia'a'', \angle aia'' = t, \angle ia''a' = [90^\circ - (t/2)]$$

Hence, applying sine rule

$$\frac{ia}{\sin[90^\circ - (t/2)]} = \frac{aa''}{\sin t}$$

$$\begin{aligned} \text{or} \quad aa'' &= \frac{ia \sin t}{\sin [90^\circ - (t/2)]} \\ &= \frac{ia \times 2 \sin(t/2) \cos(t/2)}{\cos (t/2)} \end{aligned}$$

$$= 2ia \sin (t/2) \quad (5.20)$$

$$\text{Similarly,} \quad bb'' = 2ib \sin (t/2) \quad (5.21)$$

$$\frac{a'a''}{2ia \sin(t/2)} = \frac{ia'' + a'a''}{f \sec(t/2)} = \frac{ia + a'a''}{f \sec(t/2)}$$

$$\begin{aligned} a'a'' &= \frac{ia \times 2 \sin(t/2)}{f \sec(t/2)} (ia + a'a'') \\ &= \frac{ia \sin t}{f} (ia + a'a'') \end{aligned}$$

$$= \frac{ia^2 \sin t}{f - ia \sin t} \quad (5.22)$$

Similarly,
$$b'b'' = \frac{ib^2 \sin t}{f - ib \sin t} \quad (5.23)$$

Note *The combined effect of tilt and relief displacement is that these tend to cancel in the upper part of the photograph while they are cumulative in the lower part of the photograph.*

Example 5.9 An aerial photograph 22.5 cm × 22.5 cm is taken by a camera of focal length 150 mm and tilted at 1° to the vertical. Find the relative displacement of the two points *a* and *b* each 74.2 mm distant from the principal point and lying on the principal line.

Solution

$$\begin{aligned} ia &= 74.2 + 150 \tan (1/2)^\circ = 75.509 \text{ mm} \\ ib &= 74.2 - 150 \tan (1/2)^\circ = 72.890 \text{ mm} \\ a'a'' &= \frac{(75.509)^2 \times \sin 1^\circ}{150 - 75.509 \times \sin 1^\circ} = 0.668 \text{ mm} \\ b'b'' &= \frac{(72.890)^2 \times \sin 1^\circ}{150 - 72.890 \times \sin 1^\circ} = 0.623 \text{ mm} \end{aligned}$$

5.7 PROCEDURE OF AERIAL SURVEY

The general procedure of an aerial survey consists of establishing ground control; flight planning and photography; photo interpretation and stereoscopy; and the construction of map and cartography. These aspects are dealt with in the following sections.

5.8 GROUND CONTROL

To obtain results with sufficient accuracy, a certain amount of ground control is essential. It consists in establishing a framework of points, of known relative positions, around which the details in the photograph are plotted and through which the photographic data is correlated with the terrain surveyed. These points are known as *control stations* or *control points*. The extent of the ground control is determined by the objects and scale of the map, flight control, and the cartographical process by which the maps will be produced. At least three to four control points must appear in each photograph and these should be readily identifiable on the photographs. The methods for establishing control for aerial surveys are similar to those used for any other survey.

5.9 FLIGHT PLANNING

The information required to plan a flight mission consists of the area to be surveyed, focal length of the camera, scale of the photograph, longitudinal and

side overlap, and the approximate ground speed of the aircraft in still air. This information is used to compute the altitude of the aircraft above the datum, area covered by each photograph, time interval between exposures, number of strips, and number of photographs. All these factors have great bearing on the cost of aerial survey.

For photogrammetric mapping and to examine the terrain for air photo interpretation purpose, it is essential that each ground point appear in two adjacent photographs along a flight line so that all points can be viewed stereoscopically. Sufficient photographs with proper overlap must be taken to cover the entire terrain to be surveyed. The overlap of the photographs in the direction of flight line is known as *longitudinal* or *forward overlap* or fore-and-aft overlap and is kept at about 60%. The vertical profile through a flight line is shown in Fig. 5.10. The overlap between the adjacent flight lines is known as *side overlap* and is kept at about 30%. The desired overlap is obtained either with a *view finder* or an *intervalometer*, both of which are used to time the exposure, but the latter is more automatic and records the exposures at a constant time interval.

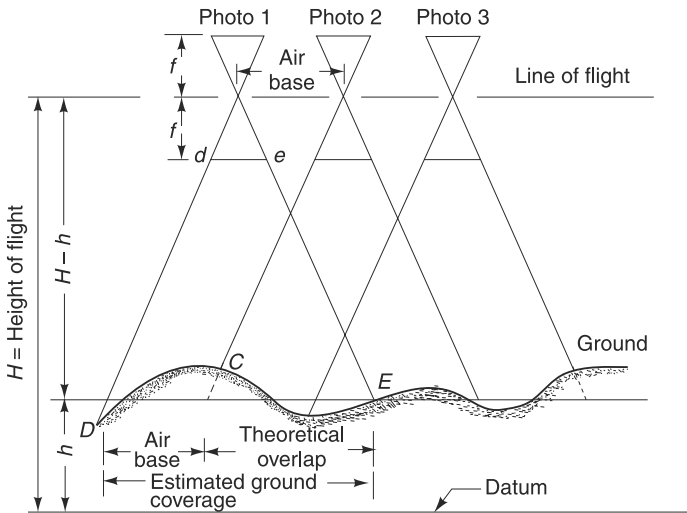


Fig. 5.10 Vertical profile through a flight line

The area to be surveyed is covered by aircraft flying regularly in strips. To photograph one strip, the aircraft is lined up so as to pass over the centre point of one end of the strip, flying in the correct direction and at the required height. Thereafter, it must continue to fly in a straight line, maintaining the same height and avoiding tilt to the greatest possible extent. The photographs are taken on a time basis ensuring that each photograph overlaps its predecessor not less than 50 per cent. However, at least 60% forward overlap is considered as standard as the aircraft is subjected to altitude variation, tip, and tilt as flight proceeds. Any adjacent pair of photographs in this strip may be considered to form a stereoscopic pair.

On the completion of the first strip, the aircraft turns and flies back on a parallel course, overlapping the first strip. The overlap aimed at is usually 15 to 20 per

cent. It will be appreciated that an increase of lateral or side overlap increases the amount of flying time. This is to ensure that no gaps for unphotographed ground exist.

5.9.1 Reasons for Overlap

To ensure a complete coverage of the area to be surveyed, there should be some overlap of the area covered by each aerial photograph. Some of the reasons for providing overlap are as follows:

1. To orient prints so as to form a continuous flight strip. For tying the different prints together, it is essential that the principal point of each print appears on the edges of as many prints as possible.
2. The central portion of the print is always less distorted than the outer edges. Since, due to the overlap, the same area is photographed twice or thrice, the highly distorted outer portion of the photograph can therefore be discarded.
3. For a stereoscopic vision only the overlapped portion is useful.
4. If a proper side overlap is provided, there will be no possibility of gaps left because of the deviation of the aircraft from the flight line. Moreover, the side overlap extends control between flight lines for photogrammetric methods.

Wind Effects It is assured that the aircraft fly in a straight course along the predetermined flight lines. This assumes no wind and navigational flight being strictly along the proposed straight and parallel strips. This implies the nose of aircraft to be in the line of the proposed flight direction and the body of the camera also aligned in the direction of the nose. However, if there is wind, the aircraft will move bodily and fly in a course inclined at an angle to the proposed course, the angle being called *drift angle*. The failure of the aircraft to stay along the flight line is called *drift* and the photographs will be as shown in Fig. 5.11. This reduces the effective forward and lateral overlap and creates possibilities of gaps in aerial photography. When the camera is not square with the direction of flight at the time of exposure, an angle is formed between the flight line and the edges of photographs. This angle is known as *crab angle*. The effect is known as *crab* and is shown in Fig. 5.12. Carabbing also causes loss of longitudinal and side

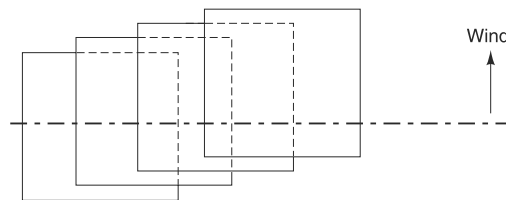


Fig. 5.11 Drift

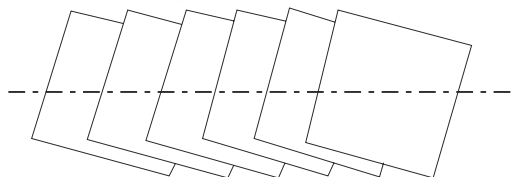


Fig. 5.12 Crab

overlaps. It should be remembered that crabbing is introduced while avoiding drift. If the aircraft tilts laterally, there is always a possibility of a gap being left unphotographed between adjacent strips, whereas for a fore- and -aft inclination of the aircraft along the flight lines, gaps may be left between the photographs along the strip.

5.9.2 Altitude of Aircraft

The flying altitude is related to the scale of the photograph, the accuracy of the process, and the contour interval to be used. It can either be computed by Eq. (5.4) or by the formula

$$H = (\text{Contour interval}) \times (C\text{-factor}).$$

where H is the flying height, and C -factor varies between 500 and 1500 depending upon the conditions surrounding the map compilation operation.

5.9.3 Area Covered by One Photograph

If the length and width of the exposed portion and the scale of the photograph are known, the area of ground covered by one photograph is given by

$$\text{Area covered} = (\text{Length} \times \text{Scale}) \times (\text{Width} \times \text{Scale})$$

5.9.4 Number of Photographs to Cover a Given Area

If there is no overlap of the photographs then the number of photographs required to cover an area will be the total area divided by the area of photograph. But the necessity of longitudinal and side overlaps has already been discussed and should be accounted for. Let the total area to be covered be A , longitudinal and side overlap be P_l and P_w , respectively, L and W be the ground length and width covered by each of the photographs, and N be the number of photographs required. Then,

$$L = (1 - P_l) Sl$$

and

$$W = (1 - P_w) Sw$$

where l = length of the photograph in the direction of flight, and

w = width of the photograph normal to the direction of flight.

Hence, the ground area covered by each photograph:

$$a = LW = (1 - P_l) Sl (1 - P_w) Sw$$

⇒

$$a = lwS^2 (1 - P_l) (1 - P_w) \quad (5.24)$$

and

$$N = \frac{A}{a}$$

where

$$A = \text{total area.}$$

If the dimensions of the area $A(L_1 \times B_1)$ are given, the number of photographs required is computed by calculating the number of photographs in each strip and the number of strips. Let N_1 be the number of photographs in each strip and N_2 be the number of strips required. Thus, net length covered by each photograph,

$$L = (1 - P_l) Sl$$

so that
$$N_1 = \frac{L_1}{(1 - P_l)Sl} + 1 \quad (5.25)$$

Similarly, net width covered by each photograph,

$$W = (1 - P_w) Sw$$

so that
$$N_2 = \frac{B_1}{(1 - P_w)Sw} + 1 \quad (5.26)$$

Hence, the number of photographs required = $N_1 \times N_2$

Note It is pertinent to note that the approximate number of photographs varies as the square of the photographic scale.

5.9.5 Interval between Exposures

If the ground speed of the aircraft is V km/h and the ground distance covered between exposures along the flight line is L km, the time interval between exposures, in seconds, is given by

$$T = \frac{3600 L}{V} \quad (5.27)$$

5.9.6 Air Base

The distance covered by an aircraft between any two successive exposures is known as the air base and is shown in Fig. 5.13.

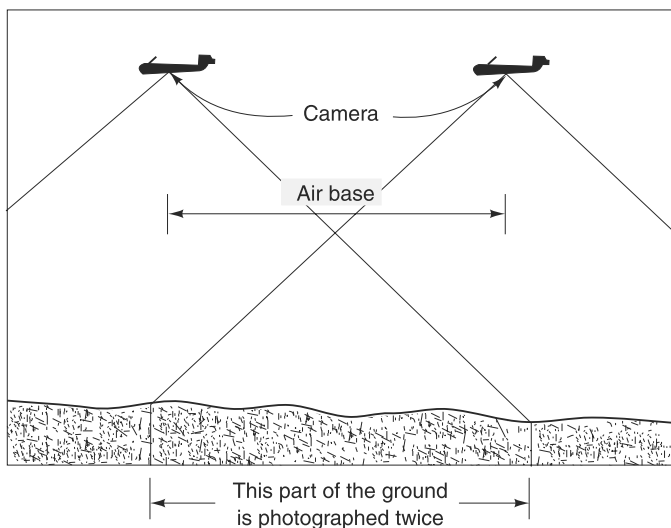


Fig. 5.13 Area covered twice by consecutive photographs

Example 5.10 The scale of an aerial photograph 25 cm $\times$ 25 cm is 1 km = 10 m. Determine the number of photographs required to cover an area of 20 km $\times$ 15 km, if the longitudinal overlap is 60% and the side overlap is 30%.

Solution

$$\begin{aligned}\text{Number of photographs per strip} &= \frac{\text{length of the area}}{(1 - P_l)Sl} + 1 \\ &= \frac{20 \times 1000}{(1 - 0.6) \times 100 \times 25} + 1 = 21\end{aligned}$$

$$\begin{aligned}\text{Number of strips} &= \frac{\text{width of the area}}{(1 - P_w)Sw} + 1 \\ &= \frac{15 \times 1000}{(1 - 0.3) \times 100 \times 25} + 1 = 11\end{aligned}$$

Hence, the number of photographs required = $21 \times 11 = 231$.

Example 5.11 Compare the scales of photography for the area recorded and the strip widths given by cameras *A* and *B* at the same flying heights.

	Camera A	Camera B
Format	180 mm × 180 mm	230 mm × 230 mm
Focal length	210 mm	150 mm

How many photographs would be taken by the camera *A* in covering a strip 16 km long at a flying height of 1350 m? The longitudinal overlap is 60%.

Solution Scale of photograph = $\frac{H}{f}$

For same flying height, the scale of photograph is inversely proportional to the focal length

$$\begin{aligned}\frac{\text{scale of photograph by camera A}}{\text{scale of photograph by camera B}} &= \frac{f_B}{f_A} \\ &= \frac{150}{210} = \frac{5}{7}, \text{ i.e., } 5:7\end{aligned}$$

Now,

$$L = 16 \times 10^3 \text{ m}, P_l = 60\% = 0.6, l = 18 \text{ cm}, H = 1350 \text{ m}, f = 21.0 \text{ cm}$$

$$\therefore S = \frac{1350}{21} = 64.286$$

$$\text{or } S = 1 \text{ cm} : 64.286 \text{ m}$$

Hence, number of photographs by camera *A* to cover 16 km length is

$$N_1 = \frac{L}{(1 - P_l)Sl} + 1 = \frac{16 \times 10^3}{(1 - 0.6) \times 64.286 \times 18} + 1 = 35.567 \approx 36$$

Example 5.12 A chimney 120 m high appears at the principal point of a vertical photograph. On the next vertical photograph, taken shortly after, the base of the chimney is on the *X*-axis and 82.5 mm to the left of the principal point.

The photographs are 200 mm $\times$ 200 mm. The flying height of the aircraft was 780 m and the focal length of the camera was 125 mm. Determine

- the distance of the top of the chimney from the Y -axis on the second photograph and
- the percentage overlap between the two photographs.

Solution In Fig. 5.14, P_1 and P_2 indicate the two consecutive photographs.
 $r_0 = 82.5$ mm, $f = 125$ mm, $H - h_1 = 780$ m, $h_2 = 120$ m

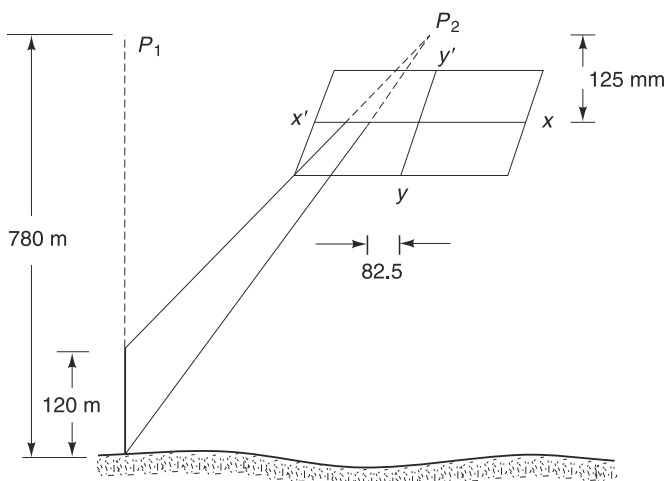


Fig. 5.14

- Radial distance of the base of chimney, $r_0 = \frac{Rf}{H - h_1}$

$$\begin{aligned} \text{The ground distance, } R &= r_0 \frac{H - h_1}{f} \\ &= \frac{(82.5/1000) \times 780}{(125/1000)} = 514.8 \text{ m} \end{aligned}$$

Radial distance of top of chimney,

$$\begin{aligned} r_2 &= \frac{Rf}{(H - h_2)} = \frac{514.8 \times (125/1000)}{780 - 120} \\ &= 0.0975 \text{ m} = 97.5 \text{ mm} \end{aligned}$$

- Size of photograph = 200 mm $\times$ 200 mm

Ground distance covered, $R = 514.8$ m

$$\text{Scale, } S = \frac{H - h_1}{f} = \frac{780}{(125/10)} = 62.4, \text{ or } 1 \text{ cm} = 62.4 \text{ m}$$

$$l = 200 \text{ mm} = 20 \text{ cm}$$

$$R = (1 - P_1) Sl$$

$$\text{or } 514.8 = (1 - P_1) \times 62.4 \times 20$$

or $P_1 = 0.5875$
Hence, percentage overlap = 58.75%

Example 5.13 It is desired to compute a flight mission for an area 8 km wide and 16 km long. The airplane has a speed of 192 km/h. A camera with a focal length of 21 cm is to be used. The approximate scale is 1:10 000, the average elevation of the ground is 366 m, and the photographs are to be 23 cm $\times$ 23 cm. The forward lap is 60% and the side lap 25%.

Solution

(a) The flight altitude H is found by formula,

$$H - h_a = \frac{f}{S}$$

$$\text{or } H - 366 = \frac{21/100}{1/10\,000}$$

$$\text{or } H - 366 = 2100$$

$$\text{or } H = 2466 \text{ m}$$

(b) The longitudinal and lateral coverage on the ground for each photograph is

$$= \frac{23}{100} \times 10\,000 = 2300 \text{ m}$$

The effective longitudinal coverage is $2300 \times 0.4 = 920 \text{ m} = 0.920 \text{ km}$

The effective lateral coverage is $2300 \times 0.75 = 1725 \text{ m} = 1.725 \text{ km}$

(c) The approximate time between exposures is therefore,

$$T = \frac{3600L}{V} = \frac{3600 \times 0.920}{192} = 17.25 \text{ s}$$

(d) The number of strips

$$= \frac{8 \times 1000}{1725} + 1 = 5.63 \approx 6$$

(e) The number of photographs per strip, theoretically, are

$$\frac{16 \times 1000}{920} + 1 = 18.39 \approx 19$$

As a factor of safety, add two photographs at each end, making a total of $19 + 2 + 2 = 23$ photographs.

Example 5.14 In a pair of overlapping vertical photographs, the mean distance between the principal points is 95.50 mm. At the time of photography the aircraft was 3000 m above the terrain, which was at datum level, and the camera had a focal length of 250 mm. In the common overlap, a tower is observed and the difference in parallax measurements of its top and base was found to be 3.50 mm. Estimate the height of the tower, the base of which is also at datum level.

Solution S = scale of the photograph for datum elevation

$$= \frac{f}{H} = \frac{25}{3000} = \frac{1}{120}, \quad \text{or } 1 \text{ cm} = 120 \text{ m}$$

For the datum elevation, $\frac{B}{b} = \frac{H}{f}$

or
$$B = \frac{H}{f} b$$

$$= S \times b = 12\,000 \times \frac{95.50}{1000} = 1146 \text{ m}$$

The parallax

$$P_{\text{bottom}} = \frac{Bf}{H-h} = \frac{1146 \times 250}{(3000-0)} = 95.5 \text{ mm}$$

$$P_{\text{top}} = 95.5 + 3.5 = 99 \text{ mm}$$

$$P_{\text{top}} = \frac{1146 \times 250}{3000-h} = 99$$

or

$$h = 106.06 \text{ m}$$

Example 5.15 Two consecutive photographs were taken with a camera of focal length 200 mm mounted on an airplane flying at a height of 1500 m. The overlap was exactly 60% and the size of prints was 250 mm $\times$ 250 mm. The flying height was same in case of both the exposures and the flight was balanced so that there was no drift. The ground was flat and was 250 m above the mean sea level. Determine the scale of the photograph and the length of the air base.

Solution Focal length of the camera, $f = 200 \text{ mm} = 20 \text{ cm}$.

Height at which the photographs were taken, $H = 1500 \text{ m}$.

Height of the ground above the mean sea level, $h = 250 \text{ m}$.

$$\text{Scale of the photograph} = \frac{f}{H-h} = \frac{20 \text{ cm}}{(1500-250) \text{ m}}$$

Thus, the scale of the photographs is 1 cm = 62.5 m

The actual ground length covered by each photograph:

$$L_1 = (1 - P_1) Sl = (1 - 0.6) \times 62.5 \times 25 = 625 \text{ m}$$

Hence, air base = 625 m

Alternatively, refer to Fig. 5.15.

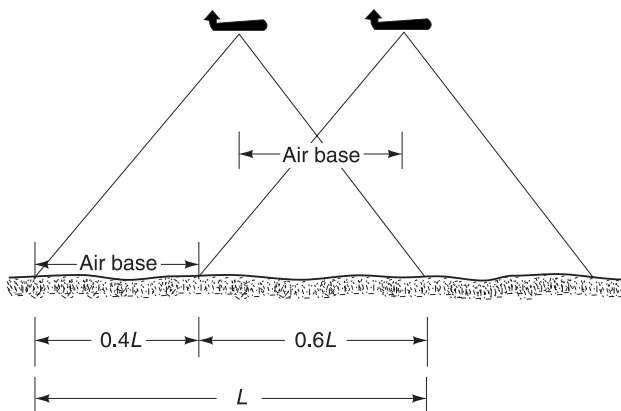


Fig. 5.15

$$L = Sl = 62.5 \times 25 = 1562.5 \text{ m}$$

$$\text{Air base} = 0.4 L = 0.4 \times 1562.5 = 625 \text{ m}$$

Example 5.16 An area of $150 \text{ km} \times 105 \text{ km}$ is to be surveyed using aerial photogrammetry. From the data given below determine number of photographs required to cover complete area; height of flight, spacing of flight lines, ground exposure distance and exposure interval:

Size of photograph = $23 \text{ cm} \times 23 \text{ cm}$

Average scale of photograph = 1:25000

Average elevation of terrain = 335 m

Longitudinal overlap = 65%

Side overlap = 28%

Ground speed of aircraft = 270 km/hr

Focal length of camera = 200 mm

Least count of intervalometer = 0.5 sec

Flying Height:

$$\frac{H}{f} = \frac{1}{S}$$

or

$$\frac{H}{0.2} = \frac{2500}{1}$$

$\Rightarrow$

$$H = 5000 \text{ m}$$

Theoretical ground spacing of flight lines:

$$W = (1 - P_w)Sw$$

$$= (1 - 0.28) \times 250 \times 23 = 4140 \text{ m}$$

Number of flight lines required

$$N_2 = \frac{L_2}{W} + 1 = \frac{105 \times 10^3}{4140} + 1 = 26.36 = 27$$

$$\text{Actual spacing of flight lines} = \frac{105 \times 10^3}{27} = 3888.89 \text{ m}$$

Ground distance between exposure:

$$L = (1 - P_l)Sl$$

$$= (1 - 0.65) \times 250 \times 23 = 2012.5 \text{ m}$$

$$\text{Exposure interval} = \frac{2012.5}{75}$$

$$= 26.83 \text{ sec (since least count of intervalometer} = 0.5S)$$

$$= 26.5 \text{ sec}$$

$$\left[270 \text{ km/hr} = \frac{270 \times 10^3}{3600} = 75 \text{ m/s} \right]$$

Adjusted ground distance between exposures:

$$L = 75 \times 26.5 = 1987.5 \text{ m}$$

Number of photographs per flight line,

$$N_1 = \frac{150 \times 10^3}{1987.5} + 1 = 76.47 = 77$$

$$\text{Total number of photographs required} = N_1 \times N_2 = 77 \times 27 = 2079$$

5.10 PHOTO-INTERPRETATION

Photo-interpretation means identifying and recognising objects in the aerial photograph and then judging their significance in the photograph. The main applications of this technique for civil engineering projects are: identification of land forms and consequently site considerations (soil types and depths, average topographic slopes, drainage conditions, etc.); areas of wet or unstable ground, density and type of vegetation cover; an overview of site and surrounding area which cannot be achieved through ground work and soil exploration planning for location of bore holes. In addition to civil engineering, it has been applied successfully in many specialised fields, viz. geography, geology, geotechnical engineering, forestry, natural resources management, agriculture, archaeology, environmental engineering, ecology, meteorology, military intelligence, oceanography, urban and regional planning, and in medical sciences.

Oblique photographs are normally easy to interpret as they depict the terrain in a more pictorial form. Local knowledge of the area, skill and ingenuity, practice and experience obtained by verification of ground, etc., all contribute towards correct interpretation. Colour photographs are easier to interpret as compared to black-and-white photographs because tonal variations are more pronounced here. Colour films have many colour tones, whereas black-and-white films have only grey tone.

The success of a project depends upon the ability of an engineer to interpret air photographs accurately. The details in the photograph can be interpreted directly by inspection or by association with familiar objects. A photo-interpreter should have a lot of training and experience in this field and should have keen powers of observation, coupled with imagination and a great deal of patience. However, the interpretation is greatly assisted when a stereo pair is viewed stereoscopically.

The interpretation of detail on air photographs is a matter which can only be learnt by experience. The best method of learning is to start with an area accurately mapped at a scale of not less than 1/25 000 and to compare photographs with detail as shown on the map. Thereafter, the same photographs should be studied in relation to the ground itself. Even when considerable experience has been gained in the interpretation of aerial photographs of one type of country it will be found that there will be many points which will require clearing up in interpreting photographs of an entirely different type of country. It is therefore desirable that a surveyor should have an opportunity of visiting the ground and comparing some of it with the aerial photographs. Whether this is possible or not, it is desirable to study the basic characteristics of photographic images, to make a systematic and correct photo-interpretation. The following characteristics of the photo images, should be studied:

1. *Shape:* Shape relates to the general form, configuration or outline of an object. Shape is one of the most important factors for recognising objects from their photographic images. For example, a railway is, usually, readily distinguished from a highway because its shape consists of long straight tangents and gentle curves as against the curved shape of a highway.

2. *Size*: Size of an object on the photograph is an important factor in the interpretation. Objects can be misinterpreted, if the sizes are not properly evaluated, e.g., a canal may be interpreted as a road side drain and a dog house may be confused with a barn.

3. *Pattern*: Pattern means the spatial arrangement of the objects photographed, e.g., buildings, roads, etc., have a particular pattern which can easily be recognised. An outdoor drive in a theatre has a particular layout and pattern of parking spaces which aid in its identification.

4. *Shadow*: The outline of a shadow gives the profile of an object, which aids in interpretation.

5. *Texture*: Texture is the frequency of the change in tone in the photographic image. Texture is produced by an aggregate of unit features, which individually may be too small to be discerned on photographs. It is a product of their individual shape, size, pattern, shadow and tone. As the photo scale is reduced, the texture of a given object becomes progressively finer and eventually disappears. On large-scale photographs, for example, large leaf tree species such as bass wood or oak is likely to be distinguished from small leaf species such as poplar, on the basis of their coarser texture.

6. *Site*: The location of an object in relation to its surroundings is very helpful in identification. A building in a forest might not be identified, whereas it can be easily identified in residential areas. A ferris wheel is easy to identify in an amusement park rather than in a field near a barn.

For a detailed interpretation it is necessary to have large-scale aerial photographs (1:5000 to 1:10 000). The quality of interpretation largely depends upon the conditions of aerial photography, the quality of photosensitive material, and the experience of the interpreter. As a precaution during examining the photographs, these should be so oriented that the direction of light is the same as that of the sun during exposure. If this is not observed, the depression and elevations may appear reversed. Therefore, as a rule, the orientation will be correct if shadows point towards an observer when an aerial camera moves forward between two exposures.

5.10.1 Air Photo-interpretation Keys

An air photo-interpretation key helps the interpreter to evaluate the information presented on aerial photographs in an organised and consistent manner. Ideally, a key consists of two basic parts:

1. A collection of annotated or captioned stereograms illustrative of the features or conditions to be identified,
2. A graphic or work description that sets forth in some systematic fashion the image recognition characteristics of those features or conditions. Two general types of air photo-interpretation keys exist:

- (a) A *selective key*, which contains numerous photographic examples with supporting text. The interpreter selects the example that most nearly resembles the feature or conditions found on the photograph under study.

- (b) An *elimination key*, which is arranged so that the interpretation proceeds step-by-step from the general to the specific and leads to the elimination of all features or conditions except the one being identified. Elimination keys often take the form of dichotomous keys where the interpreter makes a series of choices between two alternatives and progressively eliminates all but one possible answer.

In general, keys are more easily constructed and more reliably utilised for cultural feature identification (houses, bridges, roads, water towers) than for vegetation or landform identifications.

5.11 BASIC PHOTO-INTERPRETATION EQUIPMENT

Photo-interpretation is carried out with the help of magnifiers. The instrument used is known as *stereoscope*. It enables the observer to see the spatial model of the area photographed. It provides a means of measuring parallax, three-dimensional study of photographs, and drawing planimetric or topographic maps if sufficient ground control, lengths of lines, and elevations are available.

5.11.1 Stereoscopic Vision

The principle of stereoscopy is similar to that of a binocular vision enabling an observer to view an object or to be more precise, two different perspectives of an object, so as to obtain therefrom a mental impression of the three-dimensional image. Stereoscopic depth perception is of fundamental importance in photogrammetry, for it enables the formation of a three-dimensional stereomodel by viewing a pair of overlapping photographs. For stereoscopic vision of photographs, the third dimension, i.e., depth or height, is obtained by using the principle of parallax.

In natural vision, the two eyes see an object from slightly different angle and, therefore, the images at the retinas are different as shown in Fig. 5.16. The two images are, however, fused into one in imagination and an impression of relief is experienced. In Fig. 5.16, AB is the object, β_1 and β_2 are the parallax angles, α is the differential parallax, and b_e is the distance between the human eyes (generally 6.35 cm).

When the angle β_1 is less than $20''$ the observer no longer perceives the spatial or three dimensional image of the object. With these limitations the maximum distance EO for which spatial view can be obtained is 640 m. To overcome these limitations and to apply the principles to aerial photographs the instruments, e.g., stereoscope, etc., are designed so as to have an increased base distance b_e , and magnify the field of view.

In aerial photogrammetry, in order to have a spatial model of an area, it is necessary to have two photographs of the same area obtained from different points in space. It should, however, be remembered that only the overlapped portion of a photograph, providing two views, one for each eye, of the area under observation, lends itself to stereoscopic examination. This, therefore, necessitates that the photographs be taken with sufficient overlap. If two consecutive aerial photographs are placed before an observer in such a way that the left photograph is viewed only by the left eye and the right photograph by only the right eye, then the overlapped

portion will be perceived by the observer stereoscopically, i.e., a spatial model of the area will actually be in front of the observer. This is known as *stereoscopic fusion*. The stereoscopic effect provides a scale relief which is optically read and by means of the plotting devices of the projector, the stereoscopic model can be altered in size and measurements can be made from it.

Stereoscopic viewing makes it possible to measure elevations and plot contours and planimetric features from the photographs oriented in stereoscopic plotting instruments.

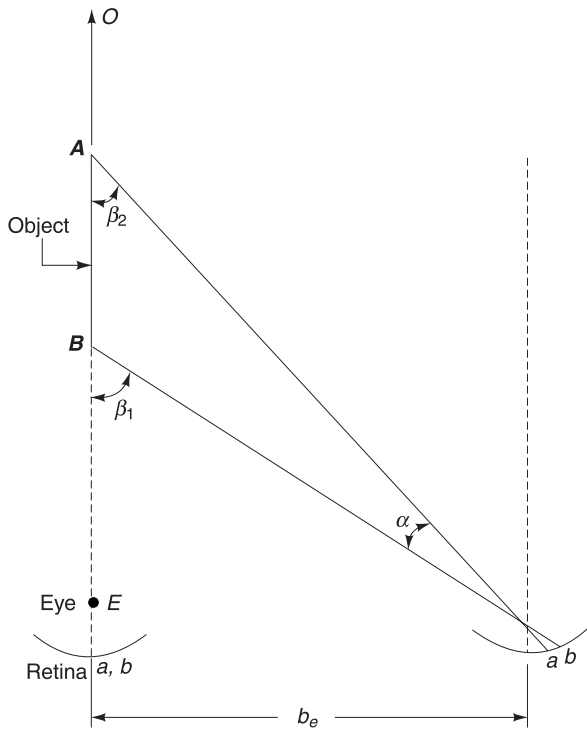


Fig. 5.16 Angle of parallax

5.11.2 Stereoscope

An aerial photograph when viewed by an unaided eye shows the surface of a terrain as if it were in a horizontal plane. In order to bring out the difference in elevation, a stereoscope is used so that the photograph stands out in relief. Thus, the primary function of a stereoscope is to accommodate the wide separation of the individual photograph of the stereopair, to the fixed length of the eye base. The simple stereoscope instrument enables the surveyor to observe a pair of nearly identical photographs, one with each eye, in such a manner that both the photographs appear to fuse into a single three-dimensional picture.

Following are the types of stereoscopes in general use:

1. Lens stereoscope, 2. Mirror stereoscope, 3. Scanning mirror stereoscope, 4. Zoom stereoscope.

In both lens and mirror types of stereoscopes, the photographs to be viewed under the stereoscope are first base lined. The line joining the principal points of the photographs represents the direction of the flight line. This line is then aligned parallel to the eye base, and on viewing through a stereoscope, a spatial view is observed. To obtain the difference in elevations from the aerial photographs under stereoscopic examination, the parallax equation is used.

1. *Lens stereoscope:* A lens stereoscope consists of one magnifying lens for each eye (Fig. 5.17). The two lenses are mounted on an assembly such that the separation between them is equal to the average eye base or the interpupillary distance of the human eyes. A provision is, however, made for minor adjustments of this separation according to the user. The distance between the nodal points of the lens and the photograph is decided by the focal length of the lens being used in the stereoscope. Lens stereoscopes have the disadvantage of causing strain to the eyes of the user but being compact, are small in size and can be conveniently handled and used in the field. A typical phenomenon associated with the lens stereoscope is the greatly exaggerated height of buildings, etc., that appear through the stereoscope. This is due to the fact that the lenses used have a magnifying effect and as such the heights seem larger than the actual heights.

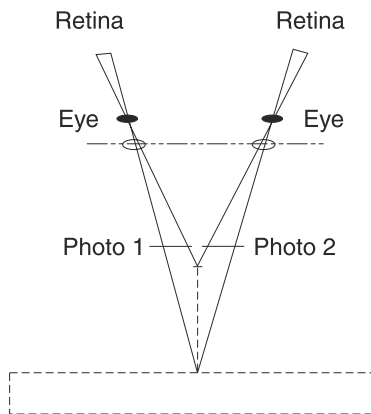


Fig. 5.17 Optical diagram of lens stereoscope

2. *Mirror stereoscope:* A mirror stereoscope (Fig. 5.18) consists of an arrangement of four mirrors, each of which is oriented at an angle of 45° with the plane of the photographs. There are two small eyepiece mirrors and two larger wing mirrors. The photographs to be viewed stereoscopically are placed at a certain distance from the wing mirrors and the light reaches the eyes exactly as it would come from the actual terrain. Wild ST4 mirror stereoscope is shown in Plate 2.

Using the Stereoscope The surveyor should be thorough with the operator's manual provided by the manufacturer. Stereoscopic viewing with a stereoscope requires that the eye base, the line of joining the centers of the stereoscope lenses, and the flight line all be parallel. First of all, the photos are inspected and laid out in the correct sequence and orientation. It is important to orient the photos so that the left eye sees the left photo and the right eye the right photo only. The principal points O_1 and O_2 (Fig. 5.19) are first located by joining the opposite fiducial marks on the photographs indicating the flight line. The corresponding principal points O_1 and O_2 , also known as *conjugate principal points*, are marked; this is achieved by estimation and careful observation of images surrounding the principal points.

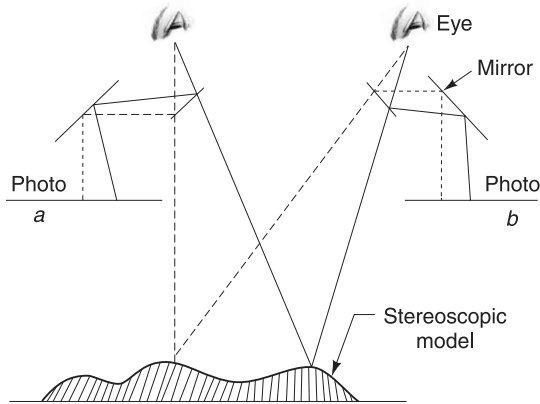


Fig. 5.18 Optical diagram of mirror stereoscope

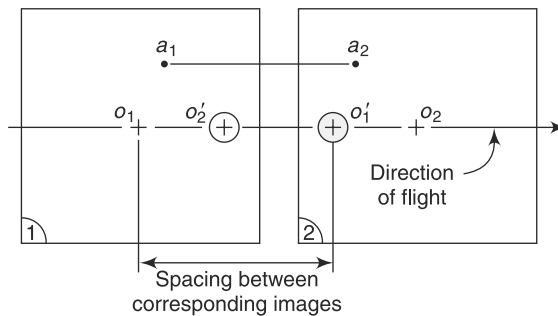


Fig. 5.19 Pair of photos properly oriented for stereoscopic viewing

One of the photos, say the left one is fastened down to the table and the right one is oriented so that principal points and the conjugate principal points defining the flight line be along a straight line. The right photo is then moved sideways until the spacing between corresponding images produces a stereoscopic view. Normally, the spacing between the corresponding images for a pocket stereoscope is 50 mm and that for a mirror stereoscope is 250 mm.

Notes

- i. It is not necessary to follow the procedure described above. The stereoscopic view may be had by trial till a clear stereoscopic view is obtained.
- ii. To keep the line joining the centers of lenses parallel to the flight line, the stereoscope may be rotated slightly if required since the photos are oriented.
- iii. The operator should look directly into the centers of lenses, thereby holding the eye-base parallel with flight line.

5.12 PARALLAX

An aerial photograph is a complete record of the ground covered and can be used for measurement of the coordinates of any ground point or object. However, for the third dimension (height/depth) two photographs, in which the same ground has been photographed from a different angle or viewpoint, are needed.

The apparent displacement of the position of a point with respect to a reference point or a system, caused by a shift in the point of observation is defined as

parallax. When an aerial camera moves forward between two exposures, with its focal plane as the reference plane, the positions of all the images travel across the focal plane from one exposure to another. The change in position of an image from one photograph to the next caused by aircraft's motion is termed as *parallax*, *stereoscopic parallax* or *x parallax*. Parallax exists for all images appearing on successive overlapping photographs. The images of points at higher elevation travel more than those of the points at lower elevation.

Figure 5.20 shows the two images of a point A : as a when the aircraft is at position O_1 ; and as a' when the aircraft is at O_2 . There is obviously a shift in the image because the image a is to the right of principal point whereas the image a' is to the left of principal point. Coordinates of the point image a are (x_a, y_a) and that of a' are (x'_a, y'_a) , while that of ground point A these are (X_A, Y_A) .

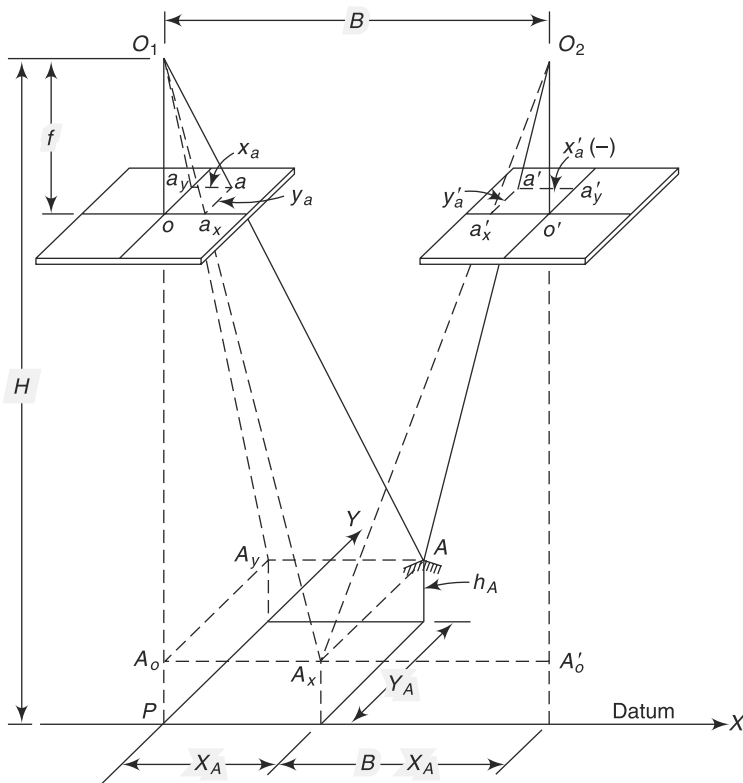


Fig. 5.20 Parallax relationship

It can be noticed that while going from position O_1 to O_2 , the image of A has shifted by an amount aa' . The *absolute parallax* of a point is the total movement of the image of that point in the focal plane between exposures. Thus aa' is the absolute photographic parallax. This absolute parallax is measurable on a pair of consecutive photographs and can be used to determine the elevations of ground points. The absolute parallax can be determined by first locating the position of the principal point of each of the two consecutive photographs containing images of the

point. The position of each principal point is then transferred to its corresponding position on the adjacent photographs. These principal points, both of which now appear on each photograph, define the flight line on each print. A line is then drawn on each photograph perpendicular to the flight line and passing through the principal point of that photograph. From the two sets of axes so constructed, the x -coordinate of the point is measured. The absolute parallax of the point is obtained from the relation

$$p_a = x_a - x'_a$$

where p_a is the parallax of the point A under consideration, x_a is measured photo-coordinate of image of the point on the left-hand photograph and x'_a is measured photo coordinate of image of the point on the right-hand photograph. These photo coordinates are not measured with respect to the fiducial axis system but are measured with respect to flight-line axis system. For a vertical photograph of a stereopair, the flight line is the line connecting the principal point and the corresponding (conjugate) principal point. In substituting the numerical values of x_a and x'_a their proper algebraic signs must be considered.

By considering similar triangles of Fig. 5.20, formulae for calculating h_A , X_A , and Y_A are derived as under.

From similar triangles O_1oa_y and $O_1A_oA_y$, we get

$$\frac{Y_A}{H - h_A} = \frac{y_a}{f}$$

or

$$Y_A = \frac{y_a}{f} (H - h_A) \quad (i)$$

Likewise, the similar triangles O_1oa_x and $O_1A_oA_x$, yield

$$\frac{X_A}{H - h_A} = \frac{x_a}{f}$$

or

$$X_A = \frac{x_a}{f} (H - h_A) \quad (ii)$$

Also, from similar triangles $O_2o'a'_x$ and $O_2A'_oA_x$,

$$\frac{B - X_A}{H - h_A} = \frac{-x_a}{f}$$

or

$$X_A = B + \frac{x_a}{f} (H - h_A) \quad (iii)$$

From Eqs. (ii) and (iii)

$$h_A = H - \frac{Bf}{x_a - x'_a} \quad (iv)$$

Substituting p_a for $x - x'_a$ into Eq. (iv), we get

$$h_A = H - \frac{Bf}{p_a} \quad (5.28)$$

Now substituting Eq. (5.28) into Eqs. (ii) and (i) and reducing, we get

$$X_A = B \frac{x_a}{p_a} \quad (5.29)$$

$$Y_A = B \frac{y_a}{p_a} \quad (5.30)$$

Equation (5.28), (5.29) and (5.30) are commonly called parallax equations.

5.12.1 Difference in Elevation by Stereoscopic Parallax

Parallax differences between one point and another are caused by different elevations of the two points. While parallax Eq. (5.30) serves to define the relationship of stereoscopic parallax to flying height, elevation, air base and camera focal length, parallax differences are more convenient for determining elevations.

Refer to Fig. 5.21.

The position of camera lens in the first and second photographs is indicated by O_1 and O_2 , respectively.

The negative is shown at a distance of f (the equivalent focal length) at the back of the camera lens and the positive or photograph is shown in front of the lens at the same distance.

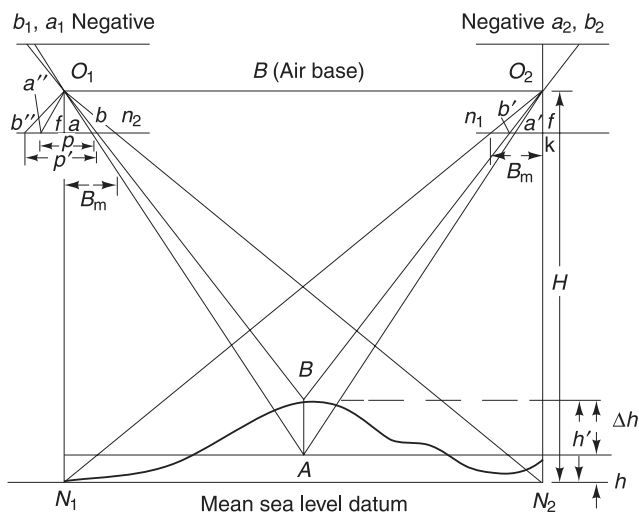


Fig. 5.21 Measurement of parallax

The points N_1 and N_2 are the nadir points determined by dropping a perpendicular to the sea level plane from O_1 and O_2 , respectively.

Rays are drawn from the ground points through O_1 and O_2 .

Let h = elevation of A above sea level

h' = elevation of B above sea level

Δh = the difference of elevation between A and B

H = the altitude of the camera lens at the time of exposure

n_2 = conjugate centre or the positive of N_2 on photograph 1

n_1 = conjugate centre or the positive of N_1 on photograph 2

a_1 = image of A on the negative 1

b_1 = image of B on the negative 1

a_2 = image of A on the negative 2

b_2 = image of B on the negative 2

a = position of A on the photograph 1

b = position of B on the photograph 1

a' = position of A on the photograph 2

b' = position of B on the photograph 2

p = absolute parallax of A

p' = absolute parallax of B

Δp = difference of parallax

B = air base (distance between the exposure stations)

B_m = air base projected on the photograph in millimeters

O_1a'' is parallel to O_2A and O_1b'' is parallel to O_2B .

From Fig. 5.21

$$p' = p + \Delta p$$

$$h' = h + \Delta h$$

Draw O_1a'' parallel to O_2a' and O_1b'' parallel to O_2b'' .

From the similar triangles O_1AO_2 and O_1aa''

$$\frac{p}{f} = \frac{B}{H - h}$$

or

$$p = \frac{Bf}{H - h} \quad (5.31)$$

From the similar triangles $O_2N_1N_2$ and O_2n_1k

$$\frac{B_m}{B} = \frac{f}{H}$$

or

$$Bf = B_m H \quad (5.32)$$

Combining these equations,

$$p = \frac{B_m H}{H - h} \quad (\text{for point } A)$$

and

$$p' = \frac{B_m H}{H - h'} \quad (\text{for point } B)$$

$\Rightarrow$

$$\begin{aligned} p + \Delta p &= \frac{B_m H}{H - h - \Delta h} \\ \Delta p &= \frac{B_m H}{H - h - \Delta h} - p \\ &= \frac{B_m H}{H - h - \Delta h} - \frac{B_m H}{H - h} \\ \Delta p &= \frac{B_m H \Delta h}{(H - h - \Delta h)(H - h)} \end{aligned} \quad (5.33)$$

Equation (5.33) may now be solved for Δh .

$$\Delta h = \frac{(H - h)^2 \Delta p}{B_m H + (H - h) \Delta p} \quad (5.34)$$

If point A is at sea level, i.e., at datum, $h = 0$, and Eq. (5.34) becomes

$$\Delta h = \frac{H \Delta p}{B_m + \Delta p} \quad (5.35)$$

In order to find the difference of elevation between two points, the parallaxes of the two points are to be measured very carefully. In case of mapping, this procedure becomes cumbersome since many such computations will be required. In such cases, *unit-change method* is used.

Equation (5.31) can be rewritten as

$$h = H - \frac{Bf}{p} \quad (5.36)$$

Differentiating the above equation

$$dh = \frac{Bf}{p^2} dp \quad (5.37)$$

Substituting the value of p from Eq. (5.31), we get

$$dh = \frac{(H-h)^2}{Bf} dp$$

$$\text{or} \quad dh = \frac{(H-h)^2}{B_m H} dp \quad \left(\frac{B}{b} = \frac{H}{f} \right) \quad (5.38)$$

Equation (5.38) expresses the rate of change of p for the infinitesimal change dh in the value of h .

Note In practise, ' B_m ' on one photograph is not equal to ' B_m ' on the consecutive photograph. The mean of the two values of B_m is called the mean principal base (b_m).

Mean Principal Base (b_m) The principal base or *photographic base* is the distance between the principal point of a photograph and the position of transferred principal point of its next photograph obtained as a result of stereoscopic fusion. If B_m and B'_m are the principal bases for the left and right photographs, respectively, the mean principle base is the average of the two bases.

5.13 MEASUREMENT OF PARALLAX

Floating Marks These are the marks which appear in the form of a dot or a pinpoint of light in the field of view of instruments used for stereoscopy. Two small identical dots, also called *half marks*, are etched on clear glass and placed over the photographs—one on the left photo and the other on the right photo. The left dot is seen with the left eye and right dot with the right eye. When viewed stereoscopically, both of them fuse, perceived as one measuring mark, called a *space mark* or *floating mark* and produce a third dimension, in the mind of the observer. Some of the types of floating marks in use are shown in Fig. 5.22. When it is viewed in the third dimension at a specific level in the stereomodel, it appears to float (depth) either above, below or on the surface of the landscape. If the half marks are

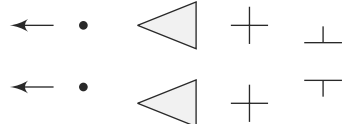


Fig. 5.22 Floating marks

moved closer together, the parallax of the half marks is increased and the fused mark will therefore appear to rise. Conversely, if the half marks are moved apart, parallax is decreased and the fused mark appears to fall. This apparent variation in the elevation of the mark as the spacing of half marks is varied is the basis for the term *floating mark*.

The spacing of the half marks, and hence the parallax of the half marks, may be varied so that the floating mark appears to rest exactly on the terrain. This produces the same effect as though an object of the shape of the half marks had existed on the terrain when the photos were originally taken. The floating mark may be moved about the stereomodel from point to point, and as the terrain varies in elevation, the spacing of the half marks may be varied to make the floating mark rest exactly on the terrain.

Floating marks are used in measuring differences in parallax. A micrometer screw is used to move one of the dots, usually the right one in the x -direction. When the dots fuse, the reading is noted. The instrument is moved to a different point on the photograph and the dots are fused again by turning the micrometer screw. The micrometer screw reading is a measure of the difference in parallax between the two points.

5.13.1 Parallax Bar

It is an instrument used to determine the parallax difference between any two points. The instrument, as shown in Fig. 5.23, consists of a pair of glass graticules held in frames attached to rigid bar such that the lower surface of each graticule is in contact with one of the pair of stereo photographs. A small opaque circular dot is marked on each graticule.

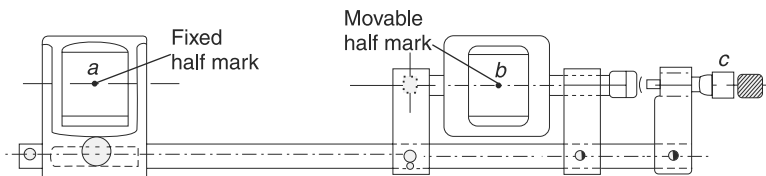


Fig. 5.23 Parallax bar

The left-hand frame a can be set at any fixed distance from the right-hand frame b according to the size of the stereo photographs. The right-hand frame can also be shifted through a range of 12.5 mm relative to the left-hand frame by means of a micrometer screw c . The scale can read up to a smallest value of 0.01 mm.

While using a parallax bar, the stereopair is adjusted for stereoscopic viewing under a stereoscope and the parallax bar is so placed on the photographs that the two graticule dots fuse together forming a floating mark, which on adjusting the micrometer appears to move vertically. This floating mark is adjusted to the level of the selected point and the reading is noted. This reading gives the parallax at that point. The 'floating mark' procedure is repeated for another selected point giving the parallax at that point. The difference in the reading at the two points gives the parallax difference.

The parallax bar finds application in determining flying heights. From the difference in parallax observed on aerial photographs, the difference in heights can be determined by the following equation:

$$\Delta h = \frac{(H - h)^2 \Delta p}{(H - h) \Delta p + b_m H} \quad (5.39)$$

where b_m is the mean principal base.

Example 5.17 A photographic survey is carried out to a scale of 1:20 000. A camera with a wide angle lens of $f = 170$ mm was used with $25 \text{ cm} \times 25 \text{ cm}$ plate size for a net 65% overlap along the line of flight. Find the error in height given by an error of 0.15 mm in measuring the parallax of the point.

Solution

$$\text{Scale} = \frac{f}{H}$$

$$\frac{1}{20000} = \frac{170/1000}{H}$$

or
$$H = \frac{170}{1000} \times 20000 = 3400 \text{ m}$$

The length of the air base is given by

$$B = (1 - P_1) IS = (1 - 0.65) \times \frac{25}{100} \times 20000 = 1750 \text{ m}$$

Also, we have
$$dh = \frac{(H - h)^2}{Bf} dp$$

Corresponding to the datum elevation, the error dh for $dp = 0.15$ mm is

$$dh = \frac{(3400 - 0)^2}{1750 \times 170} \times 0.15 = 5.83 \text{ m}$$

Example 5.18 A pair of photographs was taken with an aerial camera from an altitude of 4500 m above m.s.l. The mean principal base measured is equal to 95 mm. The difference in parallax between the two points is 2.50 mm. Find the difference in height between the two points if the elevation of the lower point is 450 m above datum. What will be the difference in elevation if the parallax difference is 20.0 mm?

Solution

(a) $dp = \Delta p = 2.50 \text{ mm}$

Since dp is extremely small, h will also be small. Hence, an approximate formula can be used to calculate h .

Thus,
$$dh = \frac{(H - h)^2}{b_m H} dp$$

$$= \frac{(4500 - 450)^2}{95 \times 4500} \times 2.50 = 95.92 \text{ m}$$

For more precise calculations, we have

$$\begin{aligned}\Delta h &= \frac{(H-h)^2 \Delta p}{(H-h)\Delta p + b_m H} \\ &= \frac{(4050)^2 \times 2.50}{4050 \times 2.50 + 95 \times 4500} = 93.70 \text{ m}\end{aligned}$$

$$(b) \quad \Delta p = 20.0 \text{ mm}$$

$$\begin{aligned}\text{Hence,} \quad \Delta h &= \frac{(H-h)^2 \Delta p}{(H-h)\Delta p + b_m H} = \frac{(4050)^2 \times 20.0}{4050 \times 20.0 + 95 \times 4500} \\ &= 645.13 \text{ m}\end{aligned}$$

5.14 WORKING UP DATA FROM THE PHOTOGRAPHS

After processing a film, the negatives are numbered in the order of exposure. The prints are made and are assembled to produce a *mosaic* or an *index map*. The photographs are so assembled and oriented that the ground details on the photograph match with their true relative positions. To achieve this, some *picture control points*, i.e., points which can be easily identified on the prints, are required. Some *ground control points*, which control the scale of the plot, are also plotted. In the preparation of a plainmetric map from aerial photographs, both the ground and picture control points are plotted on the transparent map sheet to the scale of the photographs. These plotted controls are then matched with the same points on the prints. Often the negative used for making a map have different scales. The prints of these negatives are enlarged or reduced to bring them to same scale. These prints are known as *rationed prints*. Much of the distortion may be eliminated by using only the central portion of the photographs and the rationed prints. All the oblique negatives are rectified, by a special camera known as *transforming printer*, so that the negatives have the same shape as that of the ground area covered. By a process of resection and adjustment, the true positions of all the principal points and control points are obtained. The plot of these points forms the framework to which the detail of the map is added.

Picture Control This consists of selecting carefully a few points on the photograph to be used for orienting the photographs. These points do not have any corresponding ground control points. These can be objects clearly defined and readily identifiable on each of at least two consecutive photographs. These objects may be trees, corners of buildings, well-defined ends of structures, e.g., bridges, culverts, etc. The picture control points are transferred from photographs to the mapping sheet by the radial line method. Therefore, the control points should be so selected that they can be used to the best advantage in the radial line method.

5.15 PLOTTING THE DETAILS

The methods used for plotting map details are graphical method (radial line method), mechanical method (slotted template method) and stereoscopic plotting method.

5.15.1 Radial Line Method (Arundel's Method)

This is also known as *photo triangulation*. A planimetric map can be prepared either by the radial line method or by using expensive instruments. It is a graphical method of plotting the map from rectified or vertical photographs. The following perspective properties are used:

1. The displacement due to ground relief and tilt is measured radially outwards from the principal point of the photograph. Further, these displacements are within the permissible limits of graphical measurements.
2. The images near the principal point are free from error of tilt regardless of ground relief.
3. The position of a point appearing on two overlapping aerial photographs may be located on the map where three rays from three known points intersect.

The method consists in locating the principal point of photographs on a map and transferring images from photographs to map.

To Locate Principal Point of Photographs on a Map Let a, b, c be the three photo control points as shown in Fig. 5.24. These appear in the two consecutive photographs: 1-1 and 1-2. A, B and C are their map positions already known by ground survey. It is required to determine the map position of the principal points k_1 and k_2 by three-point resection. Locate k_1 on photograph 1-1 and draw rays k_1a, k_1b, k_1c . Locate k_2 on photograph 1-2 and draw rays k_2a, k_2b, k_2c . Place a tracing paper on photo 1-1 and trace the rays drawn. The tracing paper is then transferred on to the map and is so oriented that the three rays pass through the plotted position A, B, C . The point of intersection of the three rays is pricked on the map which is the true map position of k_1 . Similarly, k_2 is also transferred on to the map as shown in Fig. 5.25.

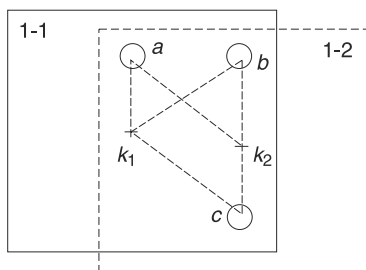


Fig. 5.24 Photo control points on a stereo pair

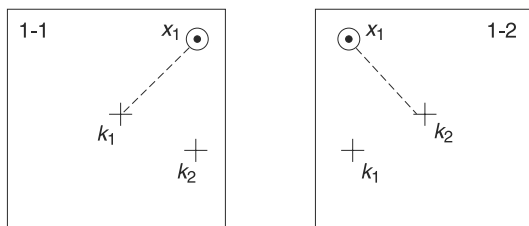


Fig. 5.25 Transferring images on map from stereopairs

To Transfer Images from a Photograph to Map Let there be an image x_1 on the two consecutive photographs and let it be required to transfer it to the map.

First, points k_1 and k_2 are transferred to the map as already discussed. On each of the photographs, the rays are drawn to image x_1 (Fig. 5.25) and are traced on tracing papers. The tracing papers are then placed over the map, and are oriented till the map positions of k_1 and k_2 coincide respectively with the traced positions. The intersection of the rays k_1x_1 and k_2x_1 gives the position of x_1 .

Note *If the lines do not intersect in a common point, it indicates that one or more of the pictures may be distorted by tilt. Such negatives are picked out and corrected for use.*

5.15.2 Slotted Template Method

This is a mechanical method of plotting the picture details. The method consists of using templates. Generally, these templates are made up of cardboard or transparent celluloid sheets (Fig. 5.26) of the size of the print. The principal and control points are established on templates by pricking through the prints with needles. Short slots are cut into the templates along the radial direction, passing through the various points, and a hole is punched through the point corresponding to the principal point of the print. The templates are then fastened together by small pins passing through the slots. The slots allow the templates to move in relation to the pins for proper orientation.

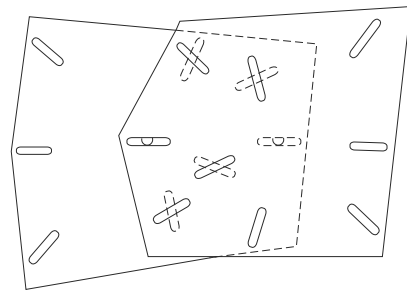


Fig. 5.26 Templates

5.15.3 Stereoscopic Plotting Method

The tracing of details is greatly simplified by the use of photographic apparatus such as a vertical reflecting projector, vertical sketch master and rectoplanigraph. A stereoplotter, also known as a multiplex, is an optical and mechanical apparatus used for preparing maps of high precision from aerial photographs. Its function is to create a precise three-dimensional scale optical model of the overlapped area, provide a means to view the spatial model, and provide a means for measuring the model and plotting its orthographic or map projection. It is extensively used by large mapping organisations.

5.16 RECTIFICATION

It is usual to perform aerial photogrammetry with the vertical photographs. When a greater coverage of the area is required, oblique photographs are taken. Since the camera axis is to be kept out of vertical for such photographs, the scale is distorted. To use the oblique negative in a map, it has to be rectified. Rectification is the process by which an oblique photograph is reconstituted so that the image is the equivalent of a truly vertical photograph. The rectification is achieved by a special

camera known as *transforming printer*. This has the same effect as stretching one edge of the negative until it has the same shape as the ground area covered by the oblique photograph. This produces a true scale on the rectified photograph, neglecting the effect of ground point elevation differences, or tilt of the camera from its assumed position at the time of exposure. These may also be rectified for tilt if desired. The rectified aerial photograph is equivalent to a vertical one and to the required scale. In the case of large relief features, the process of rectification is not that simple and the photograph has to be rectified several times. Figure 5.27 shows a vertical print surrounded by four rectified oblique prints.

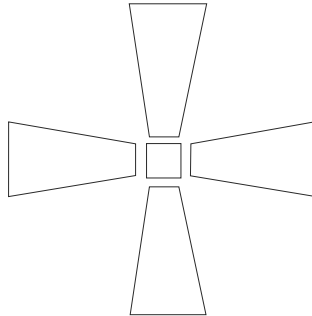


Fig. 5.27 Vertical print surrounded by four rectified oblique prints

5.17 FIDUCIAL MARKS

These are also known as *collimation marks* and are as shown in Fig. 5.28. The frame across which the film passes contains the fiducial marks which are printed on each photograph. These marks locate, on the photograph, the geometric axes whose intersection must be on the optical axis of the camera. This point of intersection is known as principal point k .

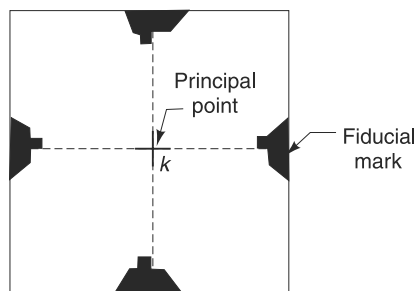


Fig. 5.28 Fiducial marks

5.18 PHOTOGRAMMETRIC PRODUCTS

Aerial photograph is the basic product of an aerial survey and has been discussed in details in the previous sections. Mosaic and topographic maps as well as the newer photographic products, the orthophotos and digital elevation models, which are replacing the traditional topographic maps in combination, are discussed as follows.

5.18.1 Mosaic

An assembly of individual aerial photographs is known as *mosaic*. It may be defined as placing the overlapping aerial photographs, all on approximately the same scale, adjusting one another along the edges by cutting the overlapped portion, and then pasting them together to form a single composite air photograph, a bird's eye view, of the area photographed.

In some cases, it may prove to be more useful than plans since a mosaic shows the actual ground conditions and many details which cannot be shown on plans. For example, for obtaining preliminary information for projects such as possible routes for railways and highways, pipelines, etc., river improvements, in estimating timber, in traffic studies, etc., the mosaic is often preferable to a map.

Following are some of the differences in a mosaic and a map:

1. The scale in a map is uniform throughout and ground objects appear in their azimuth or directions, while in a mosaic this is not so because of the distortions by tilt and relief.
2. Since the detail in mosaic appears in its natural form, a mosaic is more useful to a layman who is not conversant with map symbols and interpretations.
3. Spot height, coordinates, and other information like names are not in a mosaic, and hence it is not suitable for quantitative studies whereas in a map, measurements can be made.
4. A mosaic can be prepared in a much shorter time compared to a map.
5. A mosaic can be prepared only for flat terrain and of vertical photos. It is extremely difficult to make a mosaic of an oblique photo and also for vertical photo of highly mountainous terrain because of large scale difference.

Some of the advantages of a mosaic are low cost, rapid reproduction, completeness and better portrayal of details, and ease of understanding even by a non-technical man. The disadvantages of mosaics are that they do not show elevations, and horizontal scale measurement between any two points on mosaic are limited in accuracy, primarily due to relief displacement.

Mosaics are generally classified as *controlled* and *uncontrolled*. In a controlled mosaic, the rectified aerial photographs are placed on the sheet according to their control points and are matched by ground and picture control stations. In assembling the photographs, only their central portion is used and the overlapped portion is cut away. An uncontrolled mosaic is a representation of the ground, achieved by assembling unrectified photographs without any reference to the ground control points. These have different scales in different portions of the mosaic and are therefore inconvenient for measurements.

In addition to the controlled and uncontrolled mosaics, there are a few other types of mosaics, e.g., semi-controlled mosaic, strip mosaic and wet mosaic. If rectified prints are available and if the ground control points, as appearing on the original photographs, are matched to the plotted ground control points, so as to estimate the distances and directions, the resulting mosaic is called a *semi-controlled mosaic*. When a strip of terrain is photographed, as for highways, railways, canals, etc., the mosaic is called a *strip mosaic*. When the photographs

are placed in proper sequence with the overlap portion out, and if no glue is pasted, it is known as a *wet mosaic*.

After the photographs are arranged, these are pasted down with an adhesive or rubber cement. The latter reduces the shrinkage on drying but makes the re-adjustments of the prints difficult. The completed mosaic is then photographed and the prints are made, which are then used for study.

Mosaics are used when a pictorial view of the terrain and its culture is desired with a moderate degree of accuracy.

5.18.2 Photographic Maps

Mosaics are often wrongly called maps though they are not. Photo-maps are mosaic on which certain details have been accentuated by drawing and the result reproduced in the sheets of a map series as a combination of drawn detail and a background of photographic detail. The photographs are arranged under a transparent celluloid control sheet and then the details are transferred on it from the photographs which results in a map. The photographs are usually brought to some desired average scale by enlargement or reduction. Title information, place names, etc. may be superimposed on the photos in the same way as that for maps. The newer photogrammetric products, the orthophotos and digital elevation models are replacing topographic maps as are more friendly to be used for GIS. Depending upon the purpose and method of the plotting, the photographic maps are classified as below:

1. *Index map*: It is a photograph of the assembled prints arranged in the order of exposure. It is very useful since it makes possible a rapid selection of any particular negative from the file, as and when needed.
2. *Planimetric map*: This is also known as *line map*. It shows the natural and cultural features in the plan.
3. *Topographic map*: It shows contour lines in addition to the planimetric data. The contours are drawn by any of the ground survey methods. The best method to plot is to fix the planimetric map sheet to a plane table and do the plane tabling.
4. *Stereotopographic map*: When the contours are drawn by stereoscopic plotting instruments, the topographic maps are known as stereotopographic maps.

5.18.3 Orthophotos

Also called orthographic photographs, these do not contain the scale, tilt and relief distortions characterizing normal aerial photographs. An orthophoto is an aerial photograph that has been modified by differential rectification, which eliminates image displacements due to photographic tilt and terrain relief. Thus orthophotos are equivalent to planimetric maps having one scale but unlike planimetric maps which show features by means of lines and symbols. Orthophotos show the actual images of features like photographs. For this reason they are more easily interpreted than planimetric maps, and hence are preferred by many users. Hence, orthophotos give the surveyor the best of both—a product that can be readily interpreted like

a photograph but one where the true distances, angles, positions and areas, etc. can be measured directly without making corrections for image displacements. Because of these advantages, orthophotos make excellent base maps for compiling data to be input to GIS. Orthophotos when overprinted with contour information, are called *topographic orthophotomaps*.

5.18.4 Digital Elevation Model

A *Digital Elevation Model* (DEM) consists of an array of points in an area that have their x , y , as well as z coordinates determined. Until recently, elevations (z coordinate) have typically been represented by contours. However, the digital representation by elevations for a network of x , y , locations is called a digital elevation model; these are also called digital terrain model (DTM). Thus they provide a numerical representation of the topography in the area, and contours, cross sections, profiles, etc., can be computed from them. For more details, please refer to Appendix II.

5.19 COMPARISON OF PRODUCTS OF AERIAL PHOTOGRAPHY

Photographic prints, mosaics, air photo-maps and topographical maps are the products of aerial surveying.

Photographs can be examined stereoscopically and provide much more detailed information than the rest of the products. On the other hand, they are not so convenient as mosaics for indexing and ready reference. Photographs are of great value in studying details of particular points, whereas a mosaic is of more value for general consideration of an area. However, mosaics give little indication of the relief of the country and such indication as they give, is dependent on shadow effect and may be misleading, sometimes. Stereoscopic pairs of single photographs are excellent for studying the relief of the country. On the other hand, the extraction of a large amount of detail available from a stereoscopic pair is time consuming and cumbersome and it may be that particular points which are of interest may be picked up more quickly on a mosaic.

Photo-maps, on which some detail have been marked up, can be interpreted, as far as those details go, more readily than a mosaic. If contours or at least spot heights have been added, photo-maps are at a great advantage over mosaics in indicating relief. On the other hand, the drawing of conventional surface of a mosaic must necessarily obscure some of the details which could be obtained from an unmarked photograph. It must also be remembered that a photo-map will take much longer to produce than a mosaic and will be more expensive.

Topographical maps produced from air photographs, are intended to show all the important details and can be read accurately and instantaneously. Moreover, since the details display their correct plan position, the distances can be measured with better accuracy. This is not the case with a photo-map, which will retain some of these details.

Orthophotos and DEMs are both widely applied in all fields where maps are used, but because they are both in digital form, one of their most common applications

is their use in connection with geographic information systems. Orthophotomaps offer significant advantages over aerial photos and other photographic maps. They possess the advantages of both the aerial photos and maps—the pictorial qualities of air photos and direct measurements as from maps.

5.20 ERRORS

Some of the errors in aerial survey are as follows:

1. Failure to maintain the flying height of the aircraft.
2. Neglecting the relief displacements.
3. Inaccuracies in locating and transferring the principal points.
4. Failure to set the floating mark correctly on the point.
5. Inaccurate orientation of photographs in forming a mosaic.

5.21 MISTAKES

Some of the typical mistakes are:

1. Failing to provide proper control.
2. Improper orientation of photographs under the stereoscope.
3. Reading the scales incorrectly.

Summary

The chapter deals with the acquisition of aerial photographs, and the basic principles required to analyse and interpret them. Aerial survey consists in taking photographs of the ground from elevated position. The platform for aerial photography may be a fixed-wing aircraft, helicopter, balloon, kite, etc., but the most common use of photogrammetry is the preparation of topographic maps.

Flight planning is the most important operation in acquisition of aerial photographs and the coverage of area of interest. It consists in selecting an image scale, camera lens and format size, and image overlap. The approximate flying height of aircraft, the number of photographs required to cover the area under consideration are calculated in advance. These have been described and supported with solved examples and exercises to make the readers understand importance of flight planning. The basic geometrical concepts and the operations required to achieve the goals are described within the limitations of the scope of the book.

The aerial photographs may be vertical, or tilted, or oblique. When the camera axis is vertical the photograph is called a vertical photograph. Such a photograph is considered to represent plane of the ground photographed. However, due to the unintentional tilt of the camera axis, the resultant photograph will not represent the actual plane ground, and is called tilted photograph. Sometimes the camera axis is intentionally tilted resulting in photograph called oblique photograph. It is customary to resort to only vertical photographs as they are supposed to represent an image similar to the orthographic projection of the photographed area. In practise, however, it is impossible to maintain the camera axis vertical throughout the photography. Also, it is seldom that the terrain photographed will be flat all through. Therefore, the aerial photographs contain errors due to associated tilt and ground relief, which require appropriate corrections for any further measurements to be made from them.

The concepts of scale variation and relief displacement for photographs taken over area of varied relief are introduced, and the corrections involved are presented. The procedure for

finding ground distances and directions from measurements made on vertical photographs is outlined. The coordinates of image points on the photograph are measured. After obtaining the correct photo coordinates, the ground coordinates are obtained using the scale. Once the ground coordinates are determined and relief distortions eliminated, the lengths and bearings can be determined.

The height of an object can be determined either by measuring relief displacement or by the measurement of parallax. Both the methods have been discussed. The relief displacement causes features to appear to lean away radially from the centre of the photograph. The magnitude of the relief displacement depends on the flying height, the distance from the photo principal point to the feature, and the height of the object. By measuring relief displacement of the object and its radial position over the photograph, the height of the object can be determined.

One of the aims of acquisition of aerial photographs is to recognize and identify objects of interest. The success of a project largely depends upon the ability of the interpreter. For a systematic and correct interpretation, the basic characteristics of photographic images have been presented. Air photo-interpretation keys are also introduced. The parallax, parallax bar, stereoscopic vision, stereoscope (instrument for photo interpretation), and stereo pair (two consecutive overlapping photographs) are explained. Since with change in camera position the relative position of the features change (known as parallax), the characteristic when measured on stereo pair gives the object height and terrain elevations. For this ground control points—the points that can be accurately located on the photograph and whose ground coordinates are known and/or elevations are known—are used to calibrate photo measurements.

The chapter ends with introduction of photographic products the mosaic, orthophotos (the undistorted image), orthomaps (orthophotos showing cultural features or contours), and digital elevation model—spatially referred data files containing terrain elevation models used in GIS applications are introduced.

Exercises

5.1 Define the following:

(a) Scale	(b) Tilt and Tip
(c) Relief	(d) Principal point
(e) Isocentre	(f) Digital elevation model

5.2 Define the following:

(a) Air base	(b) Floating marks
(c) Fiducial marks	(d) Exposure station

5.3 Write short notes on the following:

(a) Aerial photograph	(b) Parallax
(c) Mosaic	(d) Aerial maps
(e) Rectification	(f) Orthophotos

5.4 Distinguish between the following:

- (a) Tip and tilt
- (b) Principal point and isocentre
- (c) Ground control and picture control points
- (d) Fiducial and floating marks

5.5 Derive an expression for the scale of a vertical photograph. Explain how the ground coordinates and the distances can be obtained from a vertical photograph.

5.6 Define relief. Derive an expression for displacement due to ground relief.

5.7 Define tilt and obtain an expression for tilt distortion.

- 5.8** Explain briefly the aspects of flight planning for an aerial survey and obtain an expression for the number of photographs required for a given area of length and width for such a survey.
- 5.9** Define and explain photo-interpretation. State and discuss briefly the factors that aid in photo-interpretation.
- 5.10** Explain the principle and working of a parallax bar.
- 5.11** (a) How do orthophotos differ from normal aerial photos?
(b) Outline some of the uses of photomaps.
- 5.12** Describe the various methods used for plotting map details in areal surveying.
- 5.13** (a) Define parallax and derive the parallax equation.
(b) What is the principle of stereoscopic vision?
- 5.14** A vertical photograph was taken from an aircraft flying at an altitude of 2000 m above m.s.l. Determine the scale of the photograph for a ground area at an elevation of (a) 80 m and (b) 200 m. The focal length of the camera is 200 mm.
[Ans. 1:9600, 1:9000]
- 5.15** The average scale of a photograph is equal to 1/6500. The minimum and maximum ground elevations are 170 m and 400 m, respectively. If the focal length of the camera is given to be 15 cm, calculate the flying height of the aircraft.
[Ans. 1260 m]
- 5.16** In an aerial photograph, the image of a line *AB*, 2330 m long at an elevation of 250 m, measures 6.66 cm. If the focal length of the camera lens is 25 cm, determine the scale of the photograph for another area lying at an elevation of 660 m.
[Ans. 1 cm = 333.63 m]
- 5.17** On an aerial photograph taken with a camera of 20 cm focal length, a line *AB* measures 12 cm. On a map drawn to a scale of 1/44 500, the same line measures 2.66 cm. If the average height of area above m.s.l. is 500 m, calculate the flying height of the aircraft.
[Ans. 2472.83 m]
- 5.18** The base position of a 850 m high mountain is at 9 cm from the flight line on a flight map. If the flying altitude is 5600 m above the datum, calculate the relief displacement of the image of the peak and its distance from the edge of the photograph, given that the size of the print is 22.5 cm × 22.5 cm.
[Ans. 1.61 cm, 0.64 cm]
- 5.19** From the following data, deduce the R.L. of points *A* and *B*.
- | | | | | |
|-------------------|----------|--------------------------------|----------|----------|
| Point | <i>C</i> | <i>D</i> | <i>B</i> | <i>A</i> |
| R.L. | 250 m | 330 m | ? | ? |
| Parallax readings | 9.66 mm | 10 mm | 8.36 mm | 12.01 mm |
| Average base | 106 mm | Focal length of camera 250 mm. | | |
- [Ans. 792.68 m, 19.33 m]
- 5.20** An aerial photograph is taken by a camera of focal length 200 mm and tilted at 1° to the vertical. Find the relative displacements of two image points each 85.6 mm distant from the principal point and lying on the principal line.
[Ans. 0.67 mm, 0.61 mm]
- 5.21** The image of a hill is 89 mm from the centre of a photograph. The elevation of the hill is 610 m and the flying height is 4270 m with respect to the datum. How much is the image displaced due to the elevation of the hill? [Ans. 12.71 mm]
- 5.22** Determine the number of photographs required to cover an area of 10 km × 10 km if the longitudinal overlap is 60% and the side overlap is 30%. The photo size is 20 cm × 20 cm and the scale is 1 cm = 100 m. [Ans. 110]

- 5.23** Two consecutive photographs were taken with a camera of focal length 37.5 cm, at a height of 7200 m. The overlap was exactly half and the prints were 22.5 cm $\times$ 22.5 cm. The height was same for both the exposures and the aircraft flew on even peel with no drift. The ground was flat at approximately 2500 m above m.s.l. Determine the scale of the photograph and length of the air base.
[Ans. 1 cm = 125 m, 1.41 km]
- 5.24** An aircraft was flying 2300 m above a cliff, the difference in parallax between whose top and bottom was 0.62 mm. Determine the height of the cliff, the base on the photograph measured was 50 mm.
[Ans. 28.52 m]
- 5.25** Explain why the heights of buildings, trees, etc., appear to be exaggerated when a pair of vertical air photographs is viewed stereoscopically.
- 5.26** What is meant by stereophotography? Two distant objects, at the left and right of photothedolite line of sight, make an angle of 47° on the instrument station. In the positive print, they are seen to be 70 mm left and 60 mm right of the photo centre. Calculate the focal length of the camera lens.
[Ans. $f = 149.35$ mm]

Objective-type Questions

- 5.1** A spatial view can be obtained with naked eyes for a distance up to
(a) 240 m (b) 640 m (c) 1140 m (d) 1540 m
- 5.2** Which of the following is an appropriate scale of an air photograph?
(a) 1 cm = 10 m (b) 1 cm = 100 m
(c) 1 cm = 500 m (d) 1 cm = 1000 m
- 5.3** The focal length of an aerial camera is 12.5 cm and the exposure station was 2500 m above the datum. The scale of the photograph is
(a) 1 cm = 200 m (b) 1 cm = 31 250 m
(c) 1:2000 (d) 1:3125
- 5.4** An aircraft is flying at a ground speed of 180 km/h. The focal length of the camera is 150 mm. The longitudinal coverage of the ground is 900 m. Which one is the exposure interval.
(a) 18 s (b) 50 s (c) 750 s (d) 270 s
- 5.5** If the distance of the image of a point A from the isocentre, on a tilted photograph, is 15.6 mm and the same distance is equal to 14.9 mm on a vertical photograph, tilt distortion is
(a) 2.9 mm (b) 0.007 mm (c) 0.7 mm (d) 0.29 mm
- 5.6** A vertical photograph was taken at an altitude of 1500 m above M.S.L. If the focal length of the camera is 20 m, the scale of photograph for a terrain lying at an elevation of 500 m is
(a) 1:50 (b) 1:100 (c) 1:1000 (d) 1:25
- 5.7** Two consecutive photographs were taken with an overlap of 55% and the print size was 22.5 cm $\times$ 22.5 cm for both the photographs. All the conditions were same and the scale of the photograph was 1 cm = 200 m. The length of the air base is
(a) 0.2025 km (b) 2.5 km (c) 2.25 km (d) 2.025 km
- 5.8** A vertical photograph of a chimney was taken from an elevation of 500 m above M.S.L. The elevation of the base of the chimney was 250 m. If the relief displacement of the top of chimney was 51.4 mm and the radial distance of the image of the top of the chimney was 110 mm, the height of the chimney is
(a) 233.64 m (b) 133.2 m (c) 116.82 m (d) 58.41 m

- 5.9** In an aerial photogrammetric survey, if the exposure interval is 20 seconds to cover a ground distance of 100 m between exposures, what would be the ground speed of the aircraft?
 (a) 90 km/hour (b) 120 km/hour (c) 150 km/hour (d) 180 km/hour
- 5.10** What is the actual ground area covered by a 20 cm × 20 cm size vertical aerial photograph, at an average scale of 1 cm = 200 m having 60% forward overlap and 0% side overlap?
 (a) 1.92 km² (b) 4.48 km² (c) 6.72 km² (d) 2.88 km²
- 5.11** A tower appears in two successive photographs taken at an altitude of 4000 m above datum. The focal length of the camera is 160 mm, the length of the air base is 300 m. The parallax for the top and bottom of the tower are 72 mm and 63 mm respectively. The height of the top of the tower above its bottom is nearly
 (a) 63 m (b) 72 m (c) 95 m (d) 135 m
- 5.12** A road section of length 1 km scales 8 cm on a vertical photograph. The focal length of the camera is 160 mm. If the terrain is fairly level, then the flying height will be
 (a) 20 m (b) 2000 m (c) 20 km (d) 200 km
- 5.13** The distances of images of the top and bottom of a tower, in a print taken from H above ground, are D and d from the principal point, respectively. The height is then
 (a) $H(D - d)/D$ (b) Hd/D
 (c) $HD/(D + d)$ (d) $HD/(D - d)$
- 5.14** If f is the focal length of camera and t is the tilt angle, distance of the photo nadir from the principal point will be
 (a) $t \sin \theta$ (b) $t \cos \theta$ (c) $t \tan \theta$ (d) $t \cot \theta$
- 5.15** If the net length covered by each photograph is 1.2 km and the length of strip is 20 km, the number of air photographs required are
 (a) 16 (b) 17 (c) 18 (d) 24
- 5.16** Aerial photographs are
 (a) perspective projections (b) orthographic projections
 (c) isometric projections (d) gnomonic projections
- 5.17** The perspective and orthographic projections produce similar images on a
 (a) vertical photograph (b) tilted photograph
 (c) high oblique photograph (d) low oblique photograph
- 5.18** The vertical axis of a camera is inclined by 1° to the vertical to cover large number of details while photographing from air. The resulting photograph will be a
 (a) vertical photograph (b) tilted photograph
 (c) oblique photograph (d) horizontal photograph
- 5.19** Vertical photographs do not represent the true plan of the ground because of
 (a) variation in speed of aircraft (b) ground relief
 (c) tilt displacement (d) image displacement
- 5.20** Which of the following coincides with the point of intersection of x and y axes of an aerial photograph?
 (a) Nadir point (b) Principal point
 (c) Isocentre (d) Air station
- 5.21** The scale of vertical photographs of a flat terrain, with a variation in the flying altitude, will be
 (a) uniform (b) non-uniform
 (c) uniform, if tilt is not there (d) non-uniform, if tilt is not there

- 5.22 On a vertical photograph the relief displacement is always radial from
(a) principal point (b) isocentre
(c) nadir point (d) homologous points
- 5.23 Ground control points are established in aerial photogrammetry to control
(a) height distortion (b) tilt distortion
(c) relief displacement (d) scale
- 5.24 When the effect of wind is not considered during flight planning, the result is a
(a) mosaic (b) drift (c) crab (d) tilt
- 5.25 When the aerial camera cannot be kept square with the direction of flight, the result is a
(a) mosaic (b) drift (c) crab (d) tilt
- 5.26 The spatial view of a photographed terrain can be had through
(a) photo-theodolite (b) parallax bar
(c) stereoscope (d) Both (b) and (c)
- 5.27 The elevations of the objects on air photographs can be measured because of the
(a) tilt (b) tip
(c) stereoscopic fusion (d) drift
- 5.28 In aerial photogrammetry, map details are plotted by
(a) graphical method (b) mechanical method
(c) stereoplotters (d) All of the above
- 5.29 Slotted template method of map plotting in aerial survey is
(a) graphical method (b) mechanical method
(c) based on three-point problem (d) known as Arundel's method
- 5.30 Rectification of aerial photographs is required for
(a) eliminating tilt (b) uniformity of scale
(c) Both (a) and (b) (d) vertical photographs
- 5.31 In air photogrammetry the desired overlap is obtained by a
(a) stereoscope (b) photo-theodolite
(c) parallax bar (d) intervalometer
- 5.32 For reconnaissance of preliminary location of highways, the aerial photographs used should preferably be
(a) vertical (b) tilted (c) oblique (d) Any of these
- 5.33 Aerial photograph in which the horizon is included is known as
(a) vertical (b) tilted
(c) low oblique (d) high oblique
- 5.34 Tilt displacements are radial about
(a) principal point (b) nadir point
(c) isocentre (d) homologous points
- 5.35 In a tilted photograph, the relief displacement is radial from the
(a) principal point (b) isocentre
(c) nadir point (d) homologous points
- 5.36 Parallax bar is used to measure
(a) parallax (b) height
(c) parallax difference (d) relief displacement
- 5.37 An example of air-photo interpretation key is
(a) photo key (b) ground key
(c) parallax key (d) elimination key
- 5.38 Observation of two photographs simultaneously is called
(a) orthography (b) spectomy
(c) spectrometry (d) stereoscopy

- 5.39** The relief displacement on an air photograph decreases with the
- (a) decrease in flying height
 - (b) increase in flying height
 - (c) decrease in the distance from the principal point
 - (d) Both (b) and (c)
- 5.40** Pick out the correct statement for relief displacement in aerial photographs.
- (a) It increases as the distance from the principal point decreases.
 - (b) It is positive and radially outwards for points above the datum.
 - (c) It is positive and radially outwards for points below the datum.
 - (d) It increases for the points vertically below the exposure station.
- 5.41** Overlap in aerial photograph is necessary
- (a) to tie the different points such that the principal point of each print appears on the adjacent prints
 - (b) because only central portion of the photographs which is less distorted is used
 - (c) because only the overlapped portion can be used for stereoscopic examination
 - (d) for all the above
- 5.42** The object of photo-interpretation is
- (a) identification of objects
 - (b) recognition of objects
 - (c) judging the significance of objects
 - (d) All of the above
- 5.43** Photo-interpretation depends upon
- (a) the quality of photosensitive material used
 - (b) the conditions of air photogrammetry
 - (c) scale of the photograph
 - (d) All of the above
- 5.44** Pick out the correct statement for picture control points in aerial photographs.
- (a) These are used for tying the photographs.
 - (b) These do not have corresponding ground control points.
 - (c) It is not essential for them to occur on stereopairs.
 - (d) None of the above is correct
- 5.45** Pick out the incorrect statement regarding mosaics in aerial photogrammetry.
- (a) Mosaics show the actual ground conditions and maximum details.
 - (b) Photographs forming a mosaic are reduced to same scale.
 - (c) Uncontrolled mosaics are obtained by rectified air photographs.
 - (d) A composite of processed photographs pasted together to form a single photograph of an area is known as a mosaic.
- 5.46** An aerial photograph is taken at a flying height of H_1 metre above the datum. The focal length of the camera was f cm and the average scale was $1 \text{ cm} = H_2$ metre such that $H_2 > H_1$. This implies that
- (a) the photographed terrain lies at an elevation $(H_2 - H_1)$ above the datum.
 - (b) the photographed terrain lies at an elevation $(H_2 - H_1)$ below the datum.
 - (c) the terrain is at the same elevation as the datum.
 - (d) it cannot be ascertained from the given data.
- 5.47** (i) The horizon point of an aerial photograph $ABCD$ lies on the photo itself.
(ii) The photograph $ABCD$ is a high oblique photograph.
- State whether
- (a) (i) follows from (ii)
 - (b) (ii) follows from (i)

- (c) (i) and (ii) have no relation
 - (d) can not be ascertained
- 5.48** In a tilted photograph
- (a) the scale is uniform
 - (b) the downward half has a smaller scale
 - (c) the downward half has larger scale
 - (d) either of the two halves can have a larger scale
- 5.49** (i) In an aerial photograph the principal and nadir points coincide,
(ii) The isocentre also coincides with the other two points.
Conclusion: It is a truly vertical photograph.
State whether
- (a) only (i) is enough to arrive at the conclusion
 - (b) only (ii) is enough to arrive at the conclusion
 - (c) the conclusion can be deduced only if both (i) and (ii) are given
 - (d) the conclusion cannot be deduced
- 5.50** An observer views two points *A* and *B* such that the angle of parallax of *B* is greater than that of *A*. Therefore,
- (a) object *A* is nearer to the observer
 - (b) object *B* is nearer to the observer
 - (c) both are at the same distance
 - (d) cannot be ascertained from the given data
- 5.51** The difference of parallax for a given difference in elevation is independent of
- (a) focal length of the camera
 - (b) percentage of overlap
 - (c) overall size of the photograph
 - (d) All of the above
- 5.52** The three-dimensional view or the depth perception is caused due to
- (a) viewing of object simultaneously by both the eyes.
 - (b) relative apparent size of the near and far objects.
 - (c) the effects of the light and shade.
 - (d) All of the above
- 5.53** Which of the following is true for a tilted photograph?
- (a) The scale is uniform for all the points on a line perpendicular to the principal line.
 - (b) The scale is uniform for all the points lying on the principal line.
 - (c) Both (a) and (b).
 - (d) Tilt is less than 1° .
- 5.54** Which of the following is true?
- (a) A mosaic is not the same as a map.
 - (b) A mosaic is a perspective view, whereas a map is an orthographic one.
 - (c) Both (a) and (b)
 - (d) None of the above
- 5.55** Homologous points are
- (a) opposite corners of the photograph
 - (b) nodal points of the camera lens
 - (c) corresponding points on the ground and the photograph
 - (d) plumb points of a stereopair
- 5.56** Setting off the 'proper principal distance' in the projectors of projection stereoplotters is a procedure for which one of the following?
- (a) Finding the focal length of the camera used
 - (b) Evaluating the digital terrain model

- (c) Carrying out the interior orientation
 (d) Filling the quotation for the purchase of a stereoplotter
- 5.57** The basic requirements of an aerial camera is
 (a) fast lens
 (b) high-speed shutter
 (c) high-speed emulsion for the film
 (d) All of the above
- 5.58** (i) Every point on an aerial photograph is displaced from its true orthographic position
 (ii) Aerial photographs shows perspective view.
 State whether
 (a) only (i) is true
 (b) only (ii) is true
 (c) both (i) and (ii) are true but (i) does not follow from (ii)
 (d) both (i) and (ii) are true and (i) follows from (ii)
- 5.59** Pick out the correct statement.
 (i) The tilt and relief displacement tend to cancel in the upper part of a tilted photograph.
 (ii) The tilt and relief displacement are cumulative in the lower part of the aerial photograph.
 (a) only (i) is true (b) only (ii) is true
 (c) both (i) and (ii) are true (d) none of these is true
- 5.60** Pick out the correct statement for aerial photogrammetry.
 (i) Relief displacement is zero for the points vertically below the exposure station.
 (ii) Relief displacement is zero if the ground photographed is flat.
 (a) only (i) is correct (b) only (ii) is correct
 (c) both (i) and (ii) are correct (d) none of (i) and (ii) is correct

Answers to Objective-type Questions

5.1 (b)	5.2 (b)	5.3 (a)	5.4 (a)	5.5 (c)	5.6 (a)
5.7 (d)	5.8 (c)	5.9 (d)	5.10 (b)	5.11 (c)	5.12 (b)
5.13 (a)	5.14 (c)	5.15 (c)	5.16 (a)	5.17 (a)	5.18 (c)
5.19 (b)	5.20 (b)	5.21 (b)	5.22 (a)	5.23 (d)	5.24 (b)
5.25 (c)	5.26 (c)	5.27 (c)	5.28 (d)	5.29 (b)	5.30 (c)
5.31 (d)	5.32 (c)	5.33 (b)	5.34 (c)	5.35 (c)	5.36 (c)
5.37 (d)	5.38 (d)	5.39 (d)	5.40 (b)	5.41 (d)	5.42 (d)
5.43 (d)	5.44 (b)	5.45 (c)	5.46 (b)	5.47 (b)	5.48 (c)
5.49 (c)	5.50 (b)	5.51 (d)	5.52 (d)	5.53 (a)	5.54 (c)
5.55 (c)	5.56 (b)	5.57 (d)	5.58 (d)	5.59 (c)	5.60 (c)

6

REMOTE SENSING

Introduction

The collection of information about the earth's surface from a considerable distance with aircrafts mounted with cameras—the aerial survey—has been described in the previous chapter. Aerial photogrammetry is the original form of remote sensing and still remains the most widely used method. Collection of such information from space by the development and deployment of satellites, are referred to as remote sensing, originated in 1960s at the Willow Run Laboratories of the University of Michigan (now renamed as Environment Research Institute), U.S.A. However, most people do not understand what remote sensing really means, even though it is used everyday in some or the other form.

Remote sensing is the science and art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object area, or phenomenon under investigation. It therefore, consists of collecting, processing, and interpreting information about a target without being in physical contact with it. Thus, remote sensing is the methodology employed to study from a distance the physical and/or chemical characteristics of objects.

The advantage of remote sensing is the bird's eye view or synoptic view it provides, so that data covering a large area of earth can be captured instantaneously and then can be processed to generate map-like products. The remotely collected data can be of many forms, including (i) variations in force distribution, e.g., gravity meter acquiring data on variations in distribution of the gravity force, (ii) acoustic wave distribution, e.g., the bat's guidance system and (iii) electromagnetic wave distribution, e.g., eyes. In this chapter, remote sensing is restricted to methods that employ electromagnetic radiation as the means of detecting and measuring the earth resources.

In a remote-sensing system, data is usually collected by highly sophisticated sensors, e.g., camera, multispectral scanner, radar, etc., mounted on a platform, e.g., aircraft, satellites, etc. The information received by the sensor is manipulated modified and sent back to the storage device, e.g., magnetic tapes, etc. The stored data is reformatted and processed on ground to produce photographs, computer-compatible magnetic tapes or other digital data-storage media. The photographs/digital data are interpreted visually/digitally to produce thematic maps and other

resource information. Thus remote sensing includes all the methods of obtaining pictures or other forms of electromagnetic records of the earth's surface from a distance, their processing and interpretation. Most of the remote-sensing methods make use of the reflected infrared bands, thermal infrared bands, or microwave portions of the electromagnetic spectrum.

The human eye is perhaps the most familiar example of a remote-sensing system. We can see an object only when it reflects light. During daytime, the source is the sun whose radiation is reflected by the object and which travels through the space between the object and the human eye. Here, the human eye is the sensor which detects the object by its image formed at the retina by the electromagnetic energy. Obviously, the human eye is a sensor (which is sensitive only to visible part of the electromagnetic spectrum) and the head is the platform on which the sensor is placed. The nervous system carries the information to the brain which acts as interpreter for identification of the object. In doing so, the brain recalls past experiences, acts like a computer/processing unit and completes the cycle of viewing and understanding.

The human-eye-brain system described above, however works only under visible light which is a very small part of the bandwidth of the electromagnetic spectrum. For obtaining information in the other parts of electromagnetic spectrum, sensors are required. In fact, sight, smell and hearing are all rudimentary forms of remote sensing.

The remote sensing of earth resources involves two basic processes as given below:

1. data-acquisition process, and
2. data analysis.

The data-acquisition process consists of (a) energy sources, (b) propagation of energy through the atmosphere, (c) energy interaction with earth surface features, (d) re-transmission of energy through the atmosphere, (e) airborne and/or space-borne sensors and (f) sensor's data in pictorial and/or digital form.

The data analysis process consists of (a) viewing and interpreting devices for pictorial data, and a computer to analyse digital sensor data, (b) completion of information in the form of hard-copy maps and tables, or as computer files and (c) presentation of the information to the users who apply it to their decision-making process.

Because of the extensive application of remote sensing for resource management a more appropriate definition of remote sensing than the one given earlier is as follows.

Remote sensing means sensing of the earth's surface from space by making use of properties of electromagnetic waves emitted, reflected, or defracted by the sensed objects, for the purpose of improving resource management, land use and the protection of environment.

6.1 REMOTE SENSING PROCESS

A remote sensing imaging system may be considered to be comprised of following elements, and are shown in Fig. 6.1.

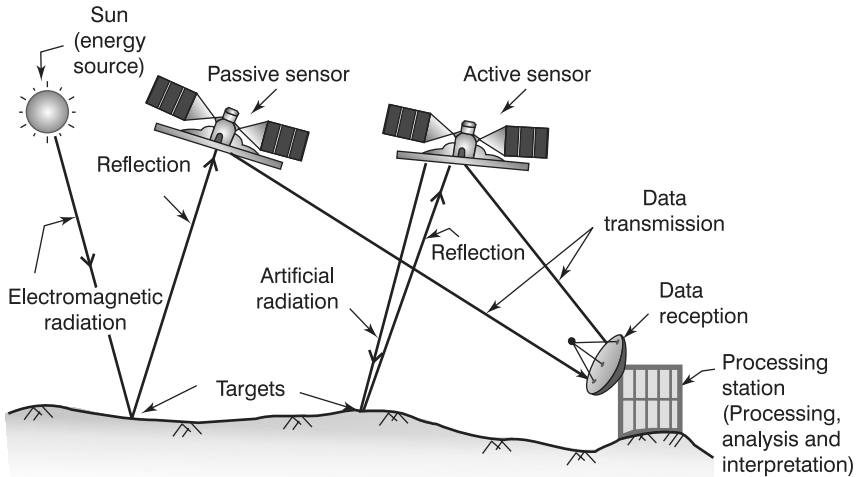


Fig. 6.1 Electromagnetic remote sensing process

1. **Energy source or illumination:** An energy source that provides illumination or electromagnetic energy to the object of interest.
2. **Radiation and the atmosphere:** As the energy propagates from the source to the object, it will come in contact with the atmosphere it passes, and will interact with the atmosphere. The interaction will again take place while travelling back from object to the sensor. The atmospheric effects are caused through the mechanism of scattering and absorption.
3. **Interaction with earth surface features:** The energy will interact with the object depending upon the properties of the object and the radiation. The effects of the interaction of the incident energy with earth surface features may cause the energy to be reflected, transmitted, scattered, absorbed, and emitted.
4. **Recording of energy by the sensor:** The scattered or emitted energy from the object is received, collected and recorded by the sensor. These are mounted on platforms, which are at a considerable height from the earth surface. The sensors may be passive or active.
5. **Processing:** The energy recorded by the sensor is transmitted in electronic form to a receiving and processing station where the data is processed into an image (photographic film and videotape or digital data on magnetic tape).
6. **Interpretation and analysis:** The processed image is interpreted visually and/or digitally or electronically to extract information about the object which was illuminated. The information about the object extracted from the image helps to understand it better, reveal some new information, or assist in solving some new problem.

To have a grasp of the general concepts and foundations of remote sensing, let us begin with the fundamentals of electromagnetic energy and its interaction with the atmosphere and the earth's surface features.

6.2 ELECTROMAGNETIC ENERGY

The sun is the most obvious source of electromagnetic radiation (EMR) remote sensing. In fact, all the terrestrial objects having temperatures above 0 K are also sources of electromagnetic radiations, but with different magnitudes and spectral compositions than that of the sun. The physical laws governing EMR are given in Appendix III.

The carrier of information in remote sensing is electromagnetic energy. It is a form of energy which moves with the velocity of light (3×10^8 m/s) in a harmonic pattern consisting of sinusoidal waves as shown in Fig. 6.2. It has two fields—the electrical and the magnetic fields—both being orthogonal to each other. Visible light, for example, is a particular range of electromagnetic radiation. Some of the other familiar forms are radio waves, ultraviolet rays, X-rays and heat. Remote sensing makes use of electromagnetic radiation which is not visible to the human eye; it can supply information during night also. This radiation can be detected only when it interacts with matter whereby a change in the electromagnetic energy takes place, which is detected by remote sensing. The data obtained is used for the determination of the characteristics of the objects.

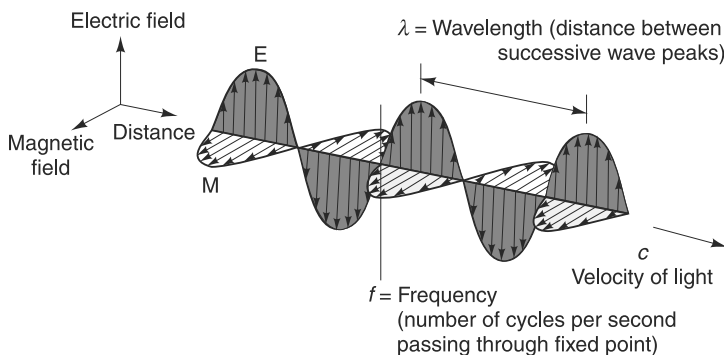


Fig. 6.2 *An electromagnetic wave*

The electromagnetic waves can be described in terms of the basic parameters such as velocity, wavelength, and frequency. Wavelength λ is the distance between any point on one wave to its same position on the next wave, usually expressed in micrometers (μm) or nanometers (nm). Frequency f is the number of cycles completed in one second, usually expressed in megahertz (MHz) or gigahertz (GHz), and c is the speed of the electromagnetic wave. The relation between these parameters is

$$c = \lambda f \quad (6.1)$$

Since c is a constant, frequency f and wavelength λ for any given wave are related inversely, and either term can be used to characterise a wave into a particular form. However, in remote sensing it is customary to categorise the electromagnetic waves by their wavelength location within the electromagnetic spectrum.

6.3 ELECTROMAGNETIC SPECTRUM

The range of different types of electromagnetic waves is shown in Fig. 6.3 and described in Table 6.1. To portray all the wavelengths, the horizontal scale is the log. Figure 6.4 shows the expanded spectrum part used in remote sensing.

Table 6.1 *Electromagnetic spectral regions*

<i>Region</i>	<i>Wavelength</i>	<i>Remarks</i>
Gamma ray	<0.03 nm	Incoming radiation is completely absorbed by the upper atmosphere and is not available for remote sensing.
X-ray	0.03 – 3.0 nm	Completely absorbed by atmosphere. Not employed in remote sensing.
Ultraviolet	0.03 – 0.4 μm	Incoming wavelengths less than 0.3 μm are completely absorbed by ozone in the upper atmosphere.
Photographic UV band	0.3 – 0.4 μm	Transmitted through atmosphere. Detectable with film and photodetectors, but atmospheric scattering is severe.
Visible	0.4 – 0.7 μm	Imaged with film and photodetectors. Includes reflected energy peak of earth at 0.5 μm .
Infrared	0.7 – 1.0 μm	Interaction with matter varies with wavelength. Atmospheric transmission windows are separated by absorption bands.
Reflected IR band	0.7 – 3.0 μm	Reflected solar radiation that contains no information about the thermal properties of materials. The band from 0.7 to 0.9 μm is detectable with film and is called the <i>photographic IR band</i> .
Thermal IR band	3 – 5 μm 8 – 14 μm	Principal atmospheric windows in the thermal region. Images at these wavelengths are acquired by optical-mechanical scanners and special vidicon systems but not by film.
Microwave	0.1 – 30 cm	Longer wavelengths can penetrate clouds, fog, and rain. Images may be acquired in the active or passive mode.
Radar	0.1 – 30 cm	Active form of microwave remote sensing. Radar images are acquired at various wavelength bands.
Radio	>30 cm	Longest wavelength portion of electromagnetic spectrum. Some classified radars with very long wavelength operate in this region.

There is no clear-cut dividing line between one nominal spectral region and the next. However for convenience, names (such as ultraviolet and microwave) are assigned to regions of the electromagnetic spectrum. Since Gamma-rays and X-rays are completely absorbed by the atmosphere, these cannot be registered with remote sensing techniques.

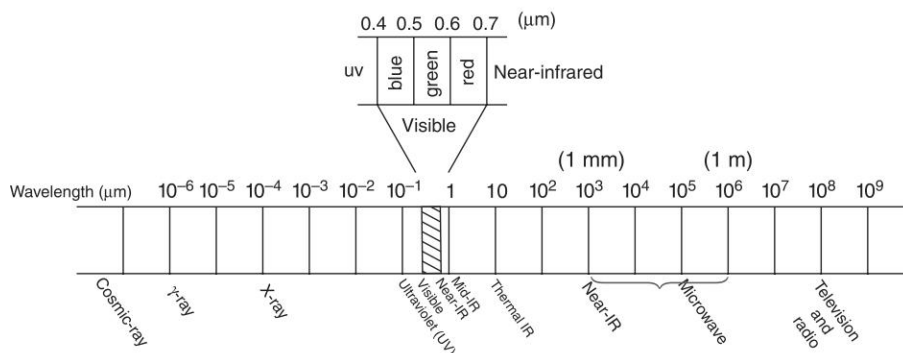


Fig. 6.3 The electromagnetic spectrum

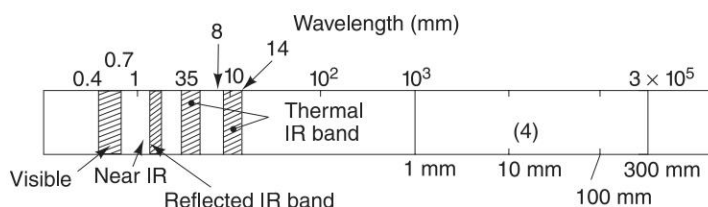


Fig. 6.4 Expanded spectrum part used in remote sensing

The various suitable regions of electromagnetic radiations are the visible region—wavelength ranging from 0.4 to 0.7 μm (of which 0.4–0.5 μm represents the blue, 0.5–0.6 μm the green, and 0.6–0.7 μm the red region), particularly suitable for photogrammetry; the reflected infrared region—ranging from 0.7 to 3 μm (of which 0.7–1.5 μm represents the near infrared, and 1.5–3 μm the shortwave infrared), which is invisible; the thermal infrared region—ranging from 3 to 14 μm (of which 3–8 μm represents the midwave infrared, and 8–14 μm the longwave infrared), which is invisible, and can be detected by crystal detectors, and such systems work day and night; the microwave region—wavelength ranging from 1 to 300 mm and which is used in radar.

The remote-sensing systems operate in one or more of the visible, reflected-infrared, thermal-infrared, and microwave portions of the electromagnetic spectrum described above. However, the existing gases and water vapours in the atmosphere, which absorb electromagnetic energy in specific wavelength bands, influence the selection of the spectrum for use in the remote-sensing system.

One of the most important aspect is the selection of those portions of the spectrum which have a high transmission of electromagnetic radiations, called *atmospheric windows*, and their corresponding wavelengths are very suitable for remote sensing since these produce good images. In simple words, the wavelengths which are able to pass through the atmosphere without loss are called the atmospheric windows. Most remote-sensing instruments acquire data from the discrete segments of an atmospheric window by making measurements with the detectors tuned to specific wavelengths.

Table 6.2 gives the wavelength regions along with the principal applications in remote sensing. Energy reflected from the earth during day time may be recorded as a function of wavelength. The maximum amount of energy is reflected at $0.5\ \mu\text{m}$, called the *reflected energy peak*. The earth also radiates energy, both during day and night time, with the maximum energy radiated at $9.7\ \mu\text{m}$, called the *radiant energy peak*.

Table 6.2 Wave length regions and their applications in remote sensing

<i>Region</i>	<i>Wavelength (μm)</i>	<i>Principal Applications</i>
1. Blue	0.45–0.52	Coastal morphology and sedimentation study, soil and vegetation differentiation, conifers and deciduous vegetation discrimination
2. Green	0.52–0.60	Vigor assessment of vegetation, rock and soil discrimination, turbidity and bathymetry studies
3. Red	0.63–0.69	Plant-species differentiation
4. Near infrared	0.70–1.50	Vegetation, biomass, delineation of water features, landforms/geomorphic studies
5. Reflected infrared	1.55–1.75	Vegetation, moisture content, soil moisture content, snow and cloud differentiation
	2.08–2.35	Differentiation of geological materials and soil
6. Thermal IR	3.0–5.0	For hot targets, i.e., fires and volcanoes
	10.4–12.5	Thermal sensing, vegetation discrimination, vegetation stress analysis, volcanic studies
7. Microwave/ Radar (0.1–30.0 cm)	2–6 cm	Suitable for sensing crop canopies and tree leaves Useful for determining ice types
	15–30 cm	Affords greater depth penetration measured in terms of metres; can penetrate 1–2 m into a dry material to reveal underlying bedrock structure. Useful for mapping total extent of ice and for sensing tree-trunks

6.4 INTERACTION OF ELECTROMAGNETIC ENERGY WITH MATTER

The wave theory discussed in Section 6.2 explains how electromagnetic energy propagates in the form of a wave and describes most of the characteristics of the electromagnetic radiations. However, this energy can only be detected when it interacts with matter. This interaction of electromagnetic energy with matter can be described by the particle theory which suggests that electromagnetic energy is composed of photons or quanta and describes as to how the electromagnetic energy interacts with matter. The photons have particle-like properties such as energy and momentum. These photons also move at the speed of light ($3 \times 10^8\ \text{m/s}$), exist as reflected or absorbed radiation, and the measurements of their varying energy levels forms the basis of remote sensing.

The energy of a quanta/photon is given as under:

$$Q = hf \quad (6.2)$$

where

Q = energy of quanta (J)

h = Planck's constant (6.626×10^{-34} Js)

f = frequency

From Eqs. (6.1) and (6.2)

$$Q = \frac{hc}{\lambda} \quad (6.3)$$

Equation (6.3) suggests that the energy of a photon is inversely proportional to its wavelength; the longer the wavelength, the lower is its energy content. This implies that for remote sensing, longer-wavelength emissions (such as microwaves) from terrain features are more difficult to sense than radiations of shorter wavelengths (such as emitted thermal infrared energy).

6.5 EFFECT OF ATMOSPHERE ON ELECTROMAGNETIC RADIATION

The electromagnetic radiations detected by remote sensors covers some distance, or path length, of the atmosphere. Since the atmospheric path length involved varies, the net effect of the atmosphere on radiation varies too. Some of the other factors, for example, the atmospheric conditions present, the wavelengths involved and the magnitude of the energy signal sensed, also contribute to the net effect of the atmosphere on radiation. The atmospheric effects are principally caused through the mechanism of atmospheric scattering and absorption.

6.5.1 Scattering

Atmospheric scattering is the unpredictable diffusion of radiation caused by the molecules of the gases, dust and smoke in the atmosphere. Scattering is basically classified as *selective* and *non-selective*, depending upon the size of particles with which the electromagnetic radiation interacts.

1. *Non-selective scatter*: The non-selective scatter occurs when the diameter of particles with which electromagnetic radiation interacts, is several times (about 10 times) the wavelength. Water droplets, pollen grains, ice and snow crystals, for example, cause non-selective scatter.

2. *Selective scatter*: The selective scatter is classified as Rayleigh scatter and Mie scatter.

(a) *Rayleigh scatter* In the upper layers of the atmosphere, scatter is mainly due to interaction of gas molecules with the radiation. Consequently, a haze results on the remotely sensed imagery. This is known as *Rayleigh scatter*. The effect of Rayleigh scatter is inversely proportional to the fourth power of the wavelength. So, the short wavelengths are scattered more by the scattering mechanism than the longer ones. A blue sky is a manifestation of Rayleigh scatter; with no scatter, the sky would appear black.

(b) *Mie scatter* In the lower layers of the atmosphere, the scatter is mainly because dust, water vapour and smoke particles interact with the radiation. This is known as *Mie scatter*. This type of scatter tends to influence the longer wavelengths. Although Rayleigh scatter dominates under most atmospheric conditions, Mie scatter is significant in overcast conditions.

6.5.2 Absorption

A part of electromagnetic radiation is absorbed by the molecules of ozone, carbon dioxide and water vapours. The absorption of radiation occurs in specific wavelength intervals called *absorption band* and governs the regions of the spectrum to be used in remote sensing. Wavelengths shorter than $0.3\ \mu\text{m}$ are completely absorbed by the ozone layer in the upper atmosphere (Fig. 6.5), which allows life on earth; a prolonged exposure to the intense energy of these wavelengths destroys living tissue. The aerosol-sized particles of liquid water in clouds absorb and scatter electromagnetic radiation at wavelengths less than about $0.3\ \mu\text{m}$. Only radiations of microwave and longer wavelengths are capable of penetrating clouds without being scattered, reflected, or absorbed.

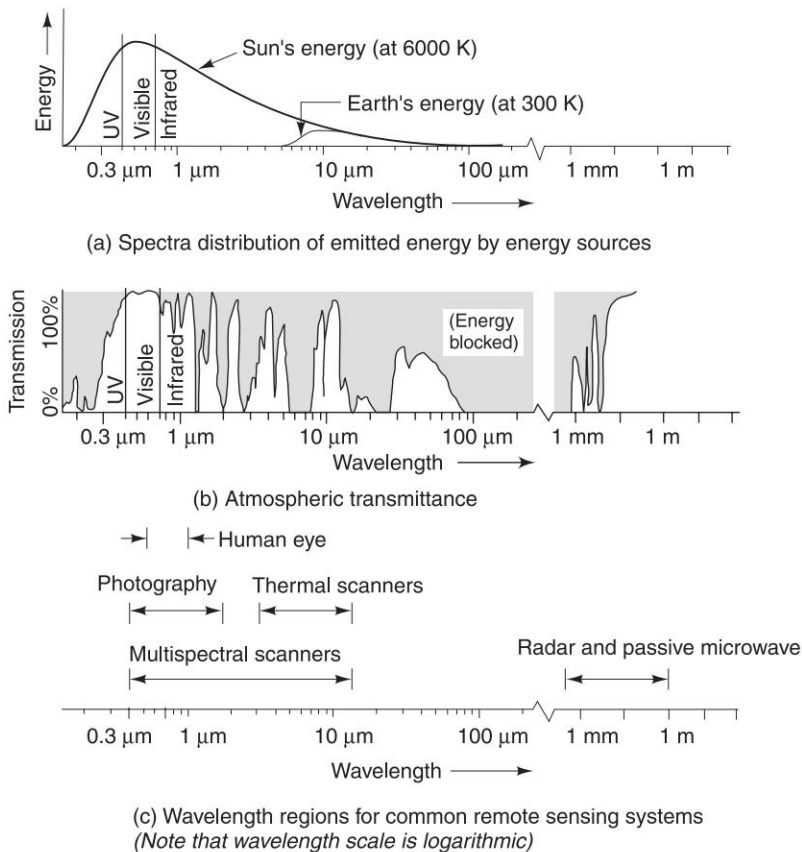


Fig. 6.5 Inter-relationship between energy sources and atmospheric absorption characteristics

Wavelength regions with high transmission are called *atmospheric windows* and are used to acquire remote-sensing images. Figure 6.5 shows these absorption bands, together with the applicability of a particular remote sensing system in the wavelength range. As can be seen, low wavelengths shorter than $0.3\ \mu\text{m}$ are completely absorbed by the ozone layer in the upper atmosphere and are responsible for life on earth. Thus, for selecting a sensor to be used in any given remote-sensing task, the following may be observed:

1. The spectral sensitivity of the sensors available,
2. The presence/absence of atmospheric windows in spectral range(s) and
3. The source, magnitude, and spectral composition of the energy available in these ranges.

Note

For passive remote-sensing tasks meant to study the earth's surface features using solar radiation, the most useful spectral windows in the visible and infrared portions are $0.3\text{--}0.7\ \mu\text{m}$, $3.0\text{--}4.5\ \mu\text{m}$ and $8.5\text{--}14\ \mu\text{m}$.

6.6 ENERGY INTERACTION WITH EARTH SURFACE FEATURES

Electromagnetic energy which comes in contact with matter is called *incident radiation*. Figure 6.6 shows the results of interaction of electromagnetic energy with matter; the incident energy is reflected (for smooth surfaces), transmitted (passed through the substance), absorbed (energy is given up—the matter is heated), scattered (deflected in all directions—caused by surfaces with roughness, relief, etc.) and emitted (transmitted) and their variations enables the interpreter to distinguish the features. The properties of incident radiation that can change are intensity, direction, wavelength, polarisation, and phase. The science of remote sensing consists in detecting and recording these changes.

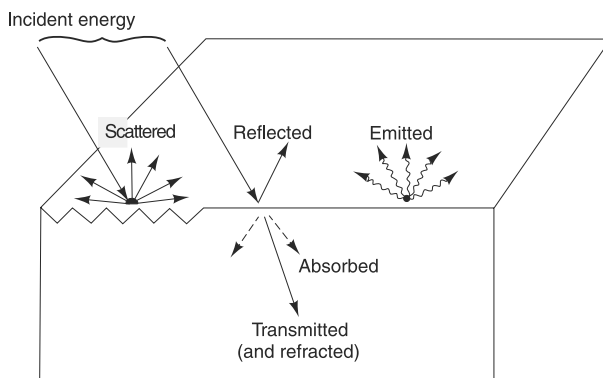


Fig. 6.6 Interaction between electromagnetic energy and matter

Emission, scattering and reflection are called *surface phenomena* and are determined primarily by the properties of the surface. Transmission and absorption are called *volume phenomena*, determined by the internal characteristics of the matter. These interactions are recorded on remote-sensing images, from which the characteristics of the matter can be interpreted.

The energy-balance equation at a particular wavelength will be as follows:

$$E_I = E_R + E_A + E_T \quad (6.4)$$

where E_I is the incidence energy, E_R is the reflected energy, E_A is the absorbed energy, and E_T is the transmitted energy.

The proportions of energy reflected, absorbed and transmitted will vary for different earth features. These differences enable us to distinguish different features on an image. It should be noted that two features may be indistinguishable in one spectral range and may appear different in another wavelength band. It is so because even within a given feature type, the proportions of reflected, absorbed, and transmitted energy will vary at different wavelengths. Equation (6.4) may be rearranged as

$$E_R = E_I - (E_A + E_T)$$

$$\text{or} \quad \frac{E_R}{E_I} = 1 - \left(\frac{E_A}{E_I} + \frac{E_T}{E_I} \right) \quad (6.5)$$

$$\text{or} \quad \rho = 1 - (\alpha + \gamma)$$

where ρ is the reflectance, α is the absorbance and γ is the transmittance.

Since almost all the earth surfaces are quite opaque, the transmittance (γ) can be neglected. Also, according to Kirchhoff's law, the absorbance (α) is taken as emissivity (ξ). Equation (6.5) reduces to

$$\rho = 1 - \xi \quad (6.6)$$

Equation (6.6) is the fundamental equation for the conceptual design of remote-sensing technology.

Notes

- i. For $\xi = 0$; $\rho = 1$: the total incident energy is reflected by the target and recorded by the sensor and the objects appear white (for example, snow).
- ii. For; $\xi = 1$; $\rho = 0$: the total incident energy is absorbed by the target and they appear black (for example, smoke).

Since most of the remote-sensing systems operate in the wavelength regions in which reflected energy predominates, the reflectance properties are of prime importance. The reflected energy is primarily a function of the surface roughness. Flat surfaces manifest mirror-like reflections (the angle of reflection is equal to the angle of incidence) and are called *specular reflectors* (Fig. 6.7(a)). Rough surfaces reflect energy uniformly in all directions and are called *diffuse* or *Lambertian reflectors* (Fig. 6.7(d)). Most earth features are neither perfectly specular nor Lambertian reflectors, and the characteristics of most earth features lie between these two extremes as shown in Fig. 6.7(b, c). The diffuse reflectance property of terrain features is of significant interest since it contains spectral information on the colour of the reflecting surface, whereas specular reflectors do not. That is why in remote sensing, the diffuse reflectance properties of terrain features are measured. To quantify the reflectance characteristics of the earth surface features, the reflected portion and incidence energy is measured. The reflected energy measured as a function of wavelength and called *spectral reflectance*, is expressed mathematically as

$$\rho_\lambda = \frac{E_R(\lambda)}{E_I(\lambda)} \times 100 \quad (6.7)$$

where ρ_λ is the spectral reflectance, $E_R(\lambda)$ is the energy of the wavelength reflected, and $E_I(\lambda)$ is the energy of the wavelength incident.

A graph of the spectral reflectance of an object as a function of the wavelength, called spectral reflectance curve, can be prepared in the laboratory using an instrument spectrometer. The configuration of this curve gives an insight into the spectral characteristics of the target. Several such curves for different targets may be prepared and used for comparing and discriminating various features in the images.

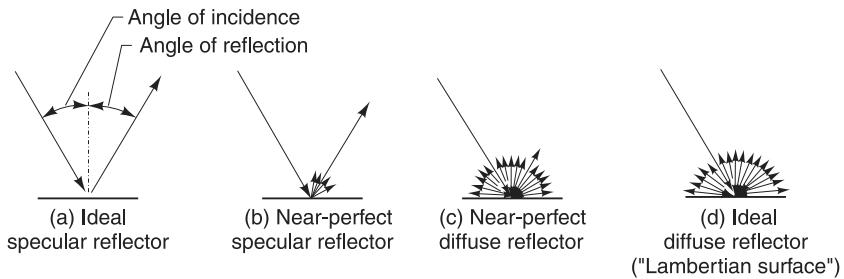


Fig. 6.7 Energy reflectance and surface characteristics of earth features

6.7 REMOTE-SENSING SENSOR SYSTEMS

The technology of remote sensing has developed rapidly during the recent years. Basically, remote sensing is concerned with the detection and recording of electromagnetic radiation. The instruments for recording electromagnetic radiations are called *sensors*. The sensor systems are required to be placed on suitable observation platforms, which can be lifted to a predefined altitude. These may be stationary (like a tripod for field observation) or mobile (aircraft and satellites). The aircrafts are mainly used for surveys of local or limited areas, whereas for a synoptic view of a large area, satellites are more useful.

6.7.1 Types of Remote-Sensing Sensor Systems

Remote-sensing sensor systems can be classified on the basis of the source of energy, either naturally occurring or man-made. Broadly speaking, remote sensing is concerned with detecting and recording electromagnetic radiation from the target area in view of the remote sensor—the instrument used to measure the electromagnetic radiation (reflected or emitted) leaving a surface under study. Sensors that sense natural radiations, either reflected or emitted from the earth, are called *passive sensors* while those which carry electromagnetic radiation of a specific wavelength or band of wavelengths to illuminate the earth's surface are called *active sensors*; the remote-sensing systems are named accordingly. The radiation may originate from separate components of the target area, which may be solar energy reflected from them or may be reflections of energy transmitted to the target area from the sensor itself. The radiation from any target can be expressed as a spectrum of which the different parts from short to longer wavelengths are ultraviolet, visible infrared and radio sectors. Each kind of sensor records radiation from a different sector of the electromagnetic spectrum.

Both the passive and active sensors could be either *imaging* or *non-imaging*. The imaging sensor measures the radiation received from specific points of the area being sensed through an image-producing device such as a camera, television (TV), and a computer monitor. These give a two-dimensional spatial distribution of the emitted or reflected intensity of electromagnetic radiation. Non-imaging sensors, on the other hand, measure the radiation received from all points within the area being sensed, integrate this radiation and produce the result in the form of electrical signals of varying strength. The imaging sensors stress upon spatial resolution whereas the non-imaging sensors stress upon time and wavelength resolution.

A sensor that measures the radiation from the entire scene instantaneously is known as a *framing system* and the one that measures the radiation from the scene point by point over a finite time is known as a *scanning system*.

Passive System A system that uses an existing source of energy, e.g., sun rays is called a *passive system*, such as taking photographs with a camera on a clear bright day. Passive remote-sensing systems sample emitted and reflected radiation from ground surface when the energy source is independent of the recording instrument. Without illumination from the sun, no photographs can be taken with a camera. The examples of passive systems are photographic cameras, television cameras, return beam vidicon (RBV), electro-optical scanners and imaging spectrometers.

Photographic cameras record images in the visible part of the spectrum and the immediately adjacent parts, i.e., ultraviolet and infrared portions. In this system, the focused image is usually recorded by photographic emulsion on a flexible film base. Television cameras also operate in the same wave bands as photographic cameras. Unlike the aerial photographs, images of TV cameras are formed as patterns of electrical charges on an image plate which is scanned by an electron beam and converted to electrical signals. These can be relayed over long distances by radio links or stored as videotape recordings that can be later displayed. The RBV is a multiband version of the TV camera. This is a scanning system that records black and white visible information in a form similar to a television picture. The electro-optical scanner constructs a scene by recording a series of scan lines as is in the case of a television camera. Line scanners employ a rotating system to scan successive strips of ground along the track of the aircraft/satellite. The instantaneous field of view (IFOV) determines the ground cell resolution and the angular field of view determines the length of the scan lines. IFOV of the sensor and the altitude of the platform determines the ground resolution cell viewed, and thus the spatial resolution. The ground width corresponding to the field of view of the scanning mirror is called *swath width*. Scanners can record in a wide range of spectral wave bands from the ultraviolet to the microwave. Unlike a camera, a scanner system builds up the picture from the spectral information directly from the target area.

The reception system of the scanner comprises of a mirror and photoelectronic detectors. The oscillations of the scanning mirror across the scene direct the reflected energy from the earth's surface to the detectors where it is converted into electrical signals. These signals are converted into numerical values and recorded on a magnetic tape. When played through a television monitor, pictures of the

earth's surface are displayed in varying patterns of brightness. Imaging sensor systems are classified as multispectral imaging sensor systems, thermal imaging sensor systems, and microwave imaging sensor systems. A multispectral scanner (MSS), measures more than one spectral band or wavelength regions. The scanning mirror after each sweep receives the radiation signals reflected by a specified number of parallel lines of the landscape. Every continuous line of signals received by the scanning mirror is divided into separable signals. The detector translates the single signal into an electrical signal, referred as digital number (DN) value, which is transferred to the ground for further processing and analysis of the image acquired. The MSS acquires the images by across-track (Whisk-broom scanner) and along-track (Push-broom scanner). These are shown in Fig. 6.8.

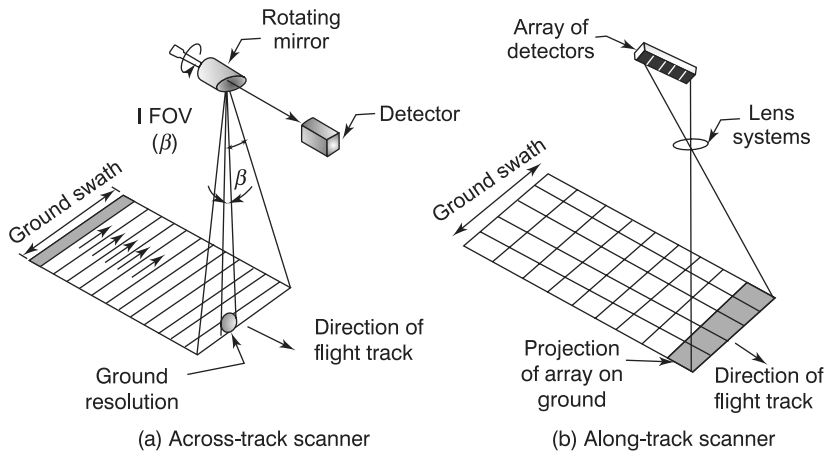


Fig. 6.8 Acquisition of multispectral scanner images

The across-track scanner consists of single detector and a rotating mirror. The detector beam sweeps a straight line over the earth's surface when mirror completes one rotation across the track of the satellite. These scan the earth in a series of lines. The lines are oriented perpendicular to the direction of motion of the sensor platform (i.e., across the swath). Each line is scanned from one side of sensor to the other, using a rotating mirror. As the platform moves forward over the earth, successive scan build up a two-dimensional image of the Earth's surface.

Along-track scanners also use the forward motion of the platform to record successive scan lines and build up two-dimensional image, perpendicular to flight direction. However, instead of scanning mirror, they use a linear array of detectors (thousands of equally spaced detectors acting as a linear camera recording one complete line) located at the focal plane of image formed by the lens systems, which are pushed along in flight track direction (i.e., along-track). Each individual detector measures the energy for single ground resolution cell and thus the size and IFOV of the detectors determines the spatial resolution of the system. Since it has no moving parts it is better than the across-track scanner.

The image spectrometer, also known as a *hyperspectral sensor*, can acquire an image in numerous, very narrow, contiguous infrared, and in thermal-infrared portions of the spectrum. This makes possible the construction of continuous

spectra of every pixel in the scene. The data so produced has a very high spectral resolution.

Active System A system which utilises man-made sources of energy for data collection is called an *active system*, e.g., taking photographs in dark places with the help of a flash bulb. The other examples of active systems are radar and laser scanner. In an active sensing system waves are propagated near the sensor and are bounced on the earth's surface to be recorded on their return. This is the essence of the RADAR (Radio Detection and Ranging) system. Images of landscapes, derived from airborne small side-looking radars (SLR) or side-looking air-borne radars (SLAR) resemble aerial photographs with low-angle solar illumination; in that, the shadow effects are produced. But the geometry of the radar pictures is essentially a ranging device which consequently produces lateral distortion of elevated objects.

The laser scanner emits pulses generally in the infrared portion of the electromagnetic spectrum. These are transmitted to the target and the reflected energy is recorded on the receiver. The laser scanner measures distances, the intensity of reflected energy and its amplitude parameters and finally a digital surface model (DSM) can be produced.

Laser scanning when integrated with other sophisticated and advanced technologies can be used to provide *Digital Elevation Models (DEMs)*. *LIDAR* (Light detection and ranging) is an example of such integration; the laser scanning is integrated with GPS for precise positioning of targets. LIDAR systems that can record multiple returns from the same pulse are used to determine the elevation of the ground as well as the elevation of the top of the targets. The data from LIDAR system can be directly integrated with GIS (Geographic Information System).

6.8 PLATFORMS

The remote-sensing sensors are operated or mounted on a vehicle, called a *platform*, and lifted to a predefined altitude. For an imaging system, the spatial resolution becomes poorer as the platform height increases, but the area coverage increases. The platforms may be a ground platform—a ladder, a scaffold or a truck; an aerial platform—a balloon, a helicopter or an aircraft; or a space platform—a rocket, a space-shuttle, or a satellite depending upon the needs on the observation mission and the constraints. Nowadays, use of specialised remote-sensing satellites are becoming increasingly common. Even India has launched its own IRS-series of remote sensing satellites.

Of these platforms, the most extensively used ones are aircrafts and satellites. Aircrafts are mainly useful for surveys of local or limited areas. The satellite platforms, on the other hand, are most suitable for global coverage and have enabled surveyors to gather information of the inaccessible areas. Besides this, there are several other advantages like synoptic view, monitoring the dynamic features such as change of crops, etc., and savings in time and cost.

Satellites may be classified as natural (moon) and artificial (IRS, etc.) the artificial satellites are grouped as geostationary and sun-synchronous satellites.

A satellite which is stationary with respect to a given position of the earth is known as a *geostationary satellite*. A satellite launched on the equatorial plane

and travelling along the same angular velocity as that at which the earth rotates and in the same direction will always remain above the same point on earth at all the times. This is possible only when the orbital altitude is 36,000 km. The orbital plane of geostationary satellites coincides with the equatorial plane and the orbit is circular. These satellites are deployed for communications and metrological observation. Examples are satellite of the INSAT series. However, when the orbit plane of a satellite is nearly polar and rotates at the same rate as the mean rotation rate of the earth around the sun, the satellite is known as *sun-synchronous satellite*. A satellite in a sun-synchronous orbit always passes over the same point on the earth at a given local solar time. These satellites are placed on relatively low altitudes (300–1000 km) and hence facilitate a good resolution. However, as these satellites are subjected to deceleration by terrestrial atmosphere, the orbital parameters need to be corrected periodically. Similarly, the altitude of the satellite is to be permanently controlled. The examples are IRS, etc.

6.9 IDEAL AND REAL REMOTE-SENSING SYSTEMS

An ideal remote-sensing system is illustrated in Fig. 6.9, wherein electromagnetic energy of all wavelengths and of known uniform intensity is produced by an ideal source; the energy propagates from the source without loss to a homogeneous target; and the energy of various wavelengths selectively interacts with the target, resulting in a return signal of reflected and emitted energy.

The returned signal propagates without loss to a sensor that responds linearly to energy of all wavelengths of any intensity. In real time, an intensity versus wavelength response is recorded, processed into an interpretable format, and recognised as being unique to the particular target in its particular chemical state. The information obtained without the particular target is made readily available in a useful form to the users.

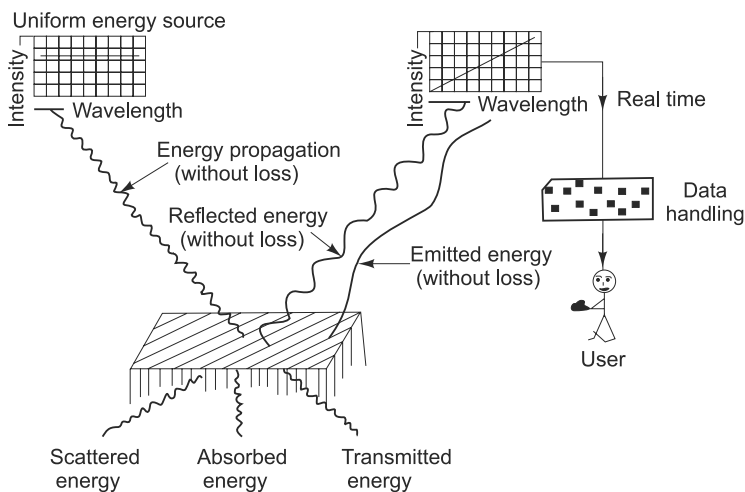


Fig. 6.9 Idealised remote sensing system

The basic components of an ideal remote sensing system include the following.

1. *Uniform energy source:* This source would provide energy over all wavelengths, at a constant, known, high level of output, irrespective of time and place.
2. *Non-interfering atmosphere:* This would be an atmosphere that would not modify the energy from the source in any manner, whether that energy were on its way to the earth's surface or coming from it. Again, ideally, this would hold irrespective of wavelength, time, place, and sensing altitude involved.
3. *Series of unique energy/matter interactions at the earth's surface:* These interactions would generate reflected and/or emitted signals that are not only selective with respect to wavelength, but are also known, invariant, and unique to each and every earth surface features type and subtype of interest.
4. *Super sensor:* This would be a sensor, highly sensitive to all wavelengths, yielding spatially detailed data on the absolute brightness (or radiance) from a scene as a function of wavelength, throughout the spectrum. This super sensor would be simple and reliable, require virtually no power or space, and be accurate and economical to operate.
5. *Real-time data handling system:* In this system, the instant the radiance versus wavelength response over a terrain element were generated, it would be processed into an interpretable format and recognised as being unique to the particular terrain element from which it came. This processing would be performed nearly instantaneously ("real time"), providing timely information. Because of the consistent nature of the energy/matter interactions, there would be no need for reference data in the analysis procedure. The derived data would provide insight into the physical–chemical–biological state of each feature of interest.
6. *Multiple data users:* These people would have knowledge of great depth, both of their respective disciplines and of remote sensing data acquisition and analysis techniques. The same set of "data" would become various forms of "information" for different users.

Unfortunately, an ideal remote sensing system does not exist. The elements of ideal system discussed above have the following general shortcomings:

1. *Energy source:* All passive remote sensing systems rely on energy that is either reflected and/or emitted from earth surface features. The spectral distribution of reflected sunlight and self-emitted energy is far from uniform. Solar energy levels obviously vary with respect to time and location, and different earth surface materials emit energy to varying degree of efficiency. While there is some control over the nature of sources of energy for active systems, the source of energy used in all real systems are generally non uniform with respect to wavelength and their properties vary with time and location.
2. *Atmosphere:* To some extent, the atmosphere always modifies the strength and spectral distribution of energy received by the sensor.

3. *Energy/matter interactions at the earth surface:* Remote sensing would be simple if every material reflected and/or emitted energy in a unique, known way. Although spectral response patterns (signatures) play a central role in detecting, identifying and analyzing earth surface materials, the spectral world is full of ambiguity.

4. *Sensor:* No single sensor is sensitive to all wavelengths. All real sensors have fixed limit of spectral sensitivity. They also have a limit on how small an object on the earth's surface can be "seen" by a sensor being separate from its surroundings.

5. *Data handling system:* Processing sensor data into an interpretable format can be—and often is—an effort entailing considerable thought, instrumentation, time, experience and reference data.

6. *Multiple data users:* Central to the successful application to any remote sensing system is the person (or persons) using the remote sensor data from that system. The "data" generated by remote sensing procedures become "information" only if and when someone understands their generation, knows how to interpret them, and knows how best to use them. A thorough understanding of the problem at hand is paramount to the productive application of any remote sensing methodology. Also, no single combination of data acquisition and analysis procedures will satisfy the needs of all data users.

6.10 DATA ACQUISITION AND INTERPRETATION

Data Acquisition Remote sensing images are recorded in digital forms and then processed with computers to produce images for interpretation. The images produced may be in the form of either photographic film or digital; the latter is supplied to the user in the form of computer readable tapes or CD-ROM. In a photographic film, variations in the scene characteristics are represented by the variations in brightness on the film. The part of scene reflecting more energy appears bright while the part of scene that reflects less energy appears black.

The digital image acquisition is the process of image generation by the combination of an illuminated source and the reflection or absorption of energy from that source by the elements of the scene being imaged. Imaging sensors are used to transform the illumination energy into digital images. The principal source for the images is the electromagnetic (EM) spectrum. The spectral bands are grouped from gamma rays (highest energy) to the radio waves (lowest energy).

The electromagnetic energy can be detected either photographically (for cameras) or electronically (for scanners). The energy variations within a scene on a photograph are detected by chemical reactions on the surface of a light-sensitive film in the process of photography whereas the same is detected with electronic sensors generating different electric signals. Photographic systems are relatively simple, inexpensive and provide a high degree of spatial detail and geometric integrity. Electronic sensors are complex and expensive but offer the advantage of a broader spectral range of sensitivity, improved calibration potential,

and the ability to electronically transmit data. A record of detected signals is obtained by developing a photograph and the film acts as both the detecting and recording medium. On the other hand, electronic sensor signals are recorded on some magnetic medium and then converted to an image on photographic film; the photographic film serves only a recording medium.

Digital Image Remote sensing images, acquired by satellites, are recorded in digital forms and then processed with computers to produce images for interpretation. Digital image is an array of numbers represented by a finite number of bits. Images acquired by satellites are useful in tracking of earth resources, monitoring of crop growths, control of flood and fire, and for environmental studies.

An *image* is a two-dimensional function $f(x, y)$ where x and y are the spatial coordinates, and the amplitude of f at any pair of coordinates (x, y) is called the intensity of image at that level. If x , y and the amplitude value of f are discrete and finite, the image is called a *digital image*. Digital image is composed of a finite number of elements each of which has a particular location and value. These elements are referred to as *picture elements*, *image elements*, *pels* and *pixels*¹. Pixel is the most widely used term to denote the element of a digital image. Thus a digital image consists of two dimensional arrays of discrete picture elements; pixels.

Each pixel has coordinates of (x, y) in the discrete space representing a continuous sampling of earth's surface. The origin of pixel coordinate system is considered to be the upper left corner of the image. Each pixel is associated with a number represented as digital number (DN); the digitally quantised brightness value. A digital image is therefore an array of DNs depicting distribution of field parameters such as reflectivity of EM radiation, topographic elevation, etc. Generally the digital number depicts the average radiance or brightness of relatively small area within a scene. Typically the DN values constituting a digital image are recorded over numerical ranges as 0 to 63 ($2^6 = 64$), 0 to 127 ($2^7 = 128$), 0 to 255 ($2^8 = 256$), 0 to 511 ($2^9 = 512$), or 0 to 1023 ($2^{10} = 1024$). These ranges represent the set of integers that can be recorded using 6-, 7-, 8-, 9-, and 10-bit binary computer coding scales, respectively. The size of this area under consideration affects the reproduction of details within the scene; smaller the pixel size more is the scene detail preserved in digital representation. The address of the pixel in two-dimensional image spaces is given in distances along rows and down columns. Some of the other terms used for pixel value are image intensity, image brightness, or grey value.

A digital image can also be obtained by converting a continuous/analog image in a digital form by *sampling* and *quantisation*. Given a continuous image $f(x, y)$, digitizing the coordinate values is called sampling and digitising the amplitude (intensity) values is called quantisation.

¹Pixels are the smallest individual elements in an image holding quantised values that represent brightness of a given colour at any specific point. Typically, pixels are stored in computer memory as a raster image or raster map, a two-dimensional array of small integers. These values are often transmitted or stored in a compressed form.

Data Interpretation The data interpretation aspects of remote sensing can involve analysis of photographs (images) and/or digital data. This can be performed by visual interpretation or with the help of computer assisted analysis techniques. Success in visual interpretation varies with the training and experience of the interpreter. Although the computer-assisted technique is better, both are complementary and a combination of the two techniques may yield better results. Image interpretation, however, can be best learned through the experience of viewing hundreds of remotely sensed images.

Photo-interpretation, already discussed in Chapter 5, can at the best be defined as the primitive form of remote sensing. It involves identifying and recognising objects in the aerial photographs and then judging their significance. Success depends upon the ability of the engineer to interpret the air photographs accurately. Knowledge of the local area, skill, and ingenuity contribute towards correct interpretation. Photo-interpretation, like remote sensing, is used for identifying objects and judging their significance without physically touching them, but a clear demarcation can be made between the two. Whereas in photo-interpretation, study is done of images recorded by energy near the visible portion of the electromagnetic spectrum, remote-sensing studies sensor records of energy over a broader range of the electromagnetic spectrum. Remote sensing is qualitative in nature, whereas photogrammetry results in quantitative as well as qualitative information about the terrain.

Some of the more important characteristics and aids with regard to identification of features in remote sensing imageries are reference data, signatures and False Colour Composites (FCC). These are discussed in the following sub-sections.

6.10.1 Reference Data

The supplementary data regarding measurements or observations about objects, areas or phenomena, sensed remotely, in any form and derived from any source is known as *reference data*. These data may be used to verify information extracted from remote-sensing data besides serving as an aid in the analysis and interpretation of remotely sensed data. Although it is expensive and time consuming to collect the reference data, but for reliability and accuracy it is customary to employ remote sensing with some or the other form of reference data. The reference data may also be used to calibrate the sensor. However, the remotely sensed data is of little value until validated by the ground reality.

6.10.2 Signatures

The word *signature*, in general, refers to a distinguishing character or a feature which is characteristic of an object. The word signature has been taken in complete analogy from human signature, which is supposed to be a well-defined characteristic of an individual. For example, the fingerprints of an individual human being is supposed to be unique for every individual person. The electromagnetic radiation when incident on a surface, either gets reflected, absorbed, radiated or transmitted through the material. In remote sensing, the term signature specifies the basic property, which directly or indirectly leads to identification of an object. From this, the characteristic expression of the object that distinguishes it from its surroundings can be known. The four major characteristics of targets facilitating discrimination

are spectral (variation in reflectance), spatial (spatial arrangement of terrain features providing attributes such as size, shape, etc., of the object), temporal (change in the reflectivity or emissivity with time), and polarisation (change in the polarisation of the radiation reflected or emitted by an object) variations. Of these characteristics, spectral variations are the most important. However, the other variations may also be useful, but for particular applications only.

Since spectral responses of features such as water, soil, vegetation, etc., are separable, these permit an assessment of the type and/or condition of the features. These responses are often referred to as *spectral signatures*. Spectral signature is used in a set of values which may be reflectance or radiance from the object averaged over different well-defined wavelength intervals. Since most of the remote-sensing systems operate in the wavelength regions in which reflected energy predominates, the reflectance properties are of prime importance. The colour of the objects is the manifestation of spectral variation in reflectance in the visible region. Table 6.3 gives the percentage reflectance for some of the common objects for the visible and the infrared part of the spectrum.

Table 6.3 Percentage of reflectance

	0.5–0.6 μm	0.6–0.7 μm	0.7–0.8 μm	0.8–1.1 μm
Sand	5.19	4.32	3.46	6.71
Clay	14.34	14.4	11.99	—
Ice	18.3	16.1	12.2	11.0
Snow	19.1	15.0	10.9	9.2
Water	3.75	2.24	1.20	1.89

Notes

- i. The term signature tends to imply a response pattern that is absolute and unique. However, the spectral response patterns may be distinctive but not necessarily unique.
- ii. Since spectral signatures are variable, these may sometimes cause severe problems in remote sensing data analyses.

6.10.3 False Colour Composite (FCC)

The remote-sensing images of a spectral band are on grey scale. However, it is more convenient to interpret various features on a colour image. In case, the images generated by remote-sensing measurements in blue (0.4–0.5 μm), green (0.5–0.6 μm) and red (0.6–0.7 μm) bands are combined by superposing the transmission through blue, green, and red filters respectively, the resultant print/transparency is known as *True Colour Composite*. In other words, for a true colour composite, an image obtained in blue, green and red spectral regions must be assigned blue, green and red colours to the image. However, while displaying the different bands of a multispectral data set, images obtained in different bands are displayed in different colours, other than that of their own bands; the colour composites obtained so are known as *False Colour Composites* (FCC). Reference may also be made to Section 7.2.3 (Chapter 7).

By assigning colours complementary to the observation band, one could improve the visual perception and this is achieved in standard FCC by assigning blue, green

and red colour to observations in green (0.52–0.59 μm), red (0.62–0.68 μm) and near infrared (0.77–0.86 μm) spectral bands respectively of IRS LISS sensors.

The vegetation normally reflects predominantly in near-IR region as compared to the green or red spectral band. Hence, vegetation appears red in standard FCC due to assignment of infrared band to red colour. Agricultural land and forest would appear in pink to deep red depending on leaf greenness (as a green band has been assigned to the blue colour, the blue-red composition will lead to a pink colour). In case of clear deep water, very less signals would be received by the satellite in green and red bands (the green band provides little more details than the red band) and practically nil in the near-IR band. Thus, deep clear water would appear dark blue, while shallow and turbid water would appear as a mixture of blue and green giving a sky-blue appearance depending on its depth and concentration of sediments.

The cultural features such as human settlements, cities would have similar reflectance in all the bands making all the three colours getting mixed in the same proportion to give it a grey colour in FCC. Snow, ice and cloud would reflect most of the energy in a similar manner for the bands and thus would appear white in the FCC.

The most typical colour signatures of some of the earth features in standard FCC are given in Table 6.4.

Table 6.4 *Signature of earth features on standard FCC*

<i>Earth feature</i>	<i>Colour signature</i>
Sand	Yellow/White
Red soil	Yellow
Snow/Clouds	White
Water	Dark blue
City/Town	Blue
Vegetation	Red/Magenta

6.11 RESOLUTION CONCEPT IN REMOTE SENSING

A working knowledge of resolution is essential for understanding both practical and conceptual details of remote sensing. Resolution refers to the ability of a remote sensing system to record and display fine details. Remote sensing image systems, whether analogue or digital, possess four major resolution characteristics viz. spatial resolution, spectral resolution, radiometric resolution, and temporal resolution. These characteristics are the most basic requirements of any sensor system that determine the type of geographic data that can be detected. However, there are no unique acceptable values for them as they depend on specific applications.

6.11.1 Spatial Resolution

Spatial resolution is by far the most important characteristic of a remote-sensing system. It determines the ability of a remote sensing system in recording spatial

detail. In analog photography, the spatial resolution of photograph refers to the sharpness of the image. The resolution depends on many factors, some of which are the resolving power of the film and camera lens, the image motion during exposure, the atmospheric condition at the time of image exposure, etc., Of these, power of a film and lens is quantifiable. The resolving power of the camera lens can be determined by means of a resolution test pattern made up of numerous sets of parallel black lines of varying thickness separated by white spaces of the same thickness, known as *line pairs* (l_p). The resolution target designed by the USA air force is generally referred to as standard and is shown in Fig. 6.10.

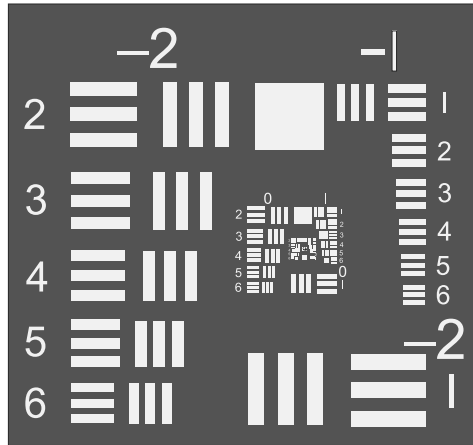


Fig. 6.10 Resolving power test chart

The number of line pairs per millimeter that can be detected visually from the photographed target under magnification is the resolution of the camera lens. The standard aerial photogrammetric camera with 23 cm $\times$ 23 cm (9" $\times$ 9") format is capable of producing aerial photographs with a resolution ranging between 20 and 40 lp/mm for the low-contrast target (1.6:1 contrast ratio). Modern photogrammetric cameras can produce aerial photographs with a resolution as high as 60 lp/mm.

The number of line pairs per millimetre does not take into account the scale of the aerial photograph. The ground resolution (GR), which combines the effects of the scale and resolution together, can be computed by using the following formula.

$$GR = w \times S.F.$$

where w is the width in millimetres for one line pair of the photographic system and $S.F.$ is the scale factor of the aerial photograph.

The spatial resolution of an imaging system is defined through various criteria, the geometric properties of the imaging system, the ability to distinguish between point targets, the ability to measure the periodicity of repetitive targets and the ability to measure the spectral properties of small targets. For electro-optical scanning systems, which produce digital images, spatial resolution is usually

described as the Instantaneous Field of View (IFOV) or the solid angle through which the detector is sensitive to the electromagnetic energy, and it is measured in terms of milliradians, a linear measure of angles. Obviously, IFOV is affected by the height of the imaging platform, the size of the detector element, and the focal length of the optical system. To calculate the spatial resolution of the image, the following formula is used.

$$D = H \beta$$

where D is the ground dimension of the detector element (in metres), H is the flying height (in metres), and β is the IFOV (in milliradians). This is the most commonly used definition based on geometrical properties of the imaging system. The other alternative measures of the IFOV based on the ability of device to distinguish between specified targets, periodicity of repetitive targets, and spectral properties of small targets.

One should note that spatial resolution of digital images is not the same as that of the analog aerial photograph. It was determined that for electro-optical remote sensing systems, two pixels are required to present the same content as one line pair on an aerial photograph. Therefore, for the standard aerial photographs with a spatial resolution of 20–40 lp/mm to be converted to digital form, the equivalent pixel size is between 25 μm (1016 dpi²) and 12.5 μm (2032 dpi).

The effects of resolution on scale can be combined to express image quality in terms of ground resolution distance (GRD).

$$\text{GRD} = \frac{\text{Reciprocal of image scale}}{\text{System resolution}}$$

6.11.2 Radiometric Resolution

Radiometric resolution is the smallest difference in radiant energy that can be detected by a sensor, and it is applicable to both photographs and digital images. In photography, radiometric resolution is inversely proportional to the contrast of the film, so that a higher contrast film will be able to resolve smaller differences in exposure. In other words, the film will be able to resolve subtle changes in grey tones. For digital images, radiometric resolution refers to the number of discrete levels into which a signal may be divided during the analog-to-digital conversion. This is also known as *quantisation* level. Radiometric resolution is commonly expressed as number of bits (binary digits) required storing the maximum level. For example the LANDSAT thematic mapper (TM) data are quantised to 256 levels (equivalent to 8 bits; 8 bit integer represents $2^8 = 256$ discrete values) and that FOR IRS LISS-III data are quantised to 128 levels (equivalent to 7 bits; 7 bit integer represents $2^7 = 128$ discrete values). The step size should not be less than the noise level in the data. A low-quality instrument with a high noise level would necessarily, therefore, have a lower radiometric resolution as compared with a high-quality, high signal-to-noise-ratio (S/N)

²dpi: Dots per inch is a measure of resolution of scanners.

instrument. Moreover, higher radiometric resolution may conflict with data storage and transmission rates as well.

6.11.3 Temporal Resolution

Temporal resolution refers to the frequency with which images of a given geographic location can be acquired. Therefore, it is the capability to view the same target, under similar conditions, at regular intervals. Obviously, more frequent remote sensing data acquisition will capture changes in environmental phenomena that occur daily, monthly, seasonally, and yearly. These environmental phenomena can be physical and cultural in nature. Remote sensing of vegetation and crops can benefit from a high temporal resolution. Vegetation shows phenological changes, while crop cultivation follows the seasons. In these cases, temporal resolutions, which use time as the discriminating factor, can compensate poor spatial and spectral resolutions to help identify vegetation and crop types. For satellite remote sensing, temporal resolution refers to repeat coverage at the Equator, or the number of days required for an orbiting satellite to return to the same point on earth. Satellites not only offer the best chances of frequent data coverage but also of regular coverage. The temporal resolution is determined by orbital characteristics and swath width, the width of the image area. Swath width is given by $2h \tan \times (FOV/2)$ where h is the altitude of the sensor, and FOV is the angular field of view of the sensor.

6.11.4 Spectral Resolution

Different materials on the earth surface exhibit different spectral reflectance and emissivities. These spectral characteristics define the spectral position and spectral sensitivity in order to distinguish materials. There is tradeoff between resolution and signal to noise. The use of well-chosen and sufficiently numerous spectral bands is a necessity, therefore, if different targets are to be successfully identified on remotely sensed images. Spectral resolution refers to the electromagnetic radiation wavelengths to which a remote sensing system is sensitive. There are two components—the number of wavelength bands (or channels) used and the width of each wave band. A larger number of bands and a narrower bandwidth for each band will give rise to a higher spectral resolution. An aerial photograph obtained by an aerial camera with a film sensitised to $0.4\text{--}0.7\ \mu\text{m}$ (the visible spectrum) has only one band, which gives the name panchromatic to the film. The spectral resolution is therefore very low.

The use of narrower bandwidths allows more unique spectral signatures of objects (such as crops and vegetation) to be recorded, thus helping to discriminate more subtle differences among these objects. As in spatial resolution, it is important to select the correct spectral resolution for the type of information to be extracted from the image. This involves deciding which wave bands are to be used.

6.12 APPLICATIONS OF REMOTE SENSING

Remote sensing affords a practical means for accurate and continuous monitoring of the earth's natural and other resources and of determining the impact of man's activities on air, water and land. Data obtained from remote

sensors had provided information necessary for making sound decisions and in formulating policy regarding land use and resource development. Detecting forest fires before the smoke appears, detecting diseased plants before these can be spotted by the naked eye, detecting the presence of pollutants and locating icebergs are some of the applications of remote sensing. Remote sensing was practised in India as early as in 1920, when black and white aerial photographs were used extensively for survey and geological exploration work. In 1926, the aerial photographs were used for flood assessment of the Indus river. Remote sensing as practised in the present day was first applied in the field of agriculture by the Indian Space Research Organization (ISRO) in 1970, for detecting the 'root wilt' disease in coconut plantation in Kerala. Some specific applications of remote sensing have been discussed below.

6.12.1 Environmental Applications

Remote sensing is a vital source of environmental information. It is the best means of acquiring basic information, particularly on the regional scale and on a repetitive schedule. Environmental satellites have been launched specifically for this purpose. The Indian INSAT series of satellites are currently being used for weather forecasting. Many aspects of the oceans which were poorly understood (such as circulation, seastate, etc.) are becoming better known through remote sensing techniques. Pollution in the form of oil spills and thermal plumes can easily be monitored. In fact, scientists are keeping a continuous watch on such important phenomena as the ozone layer depletion and global warming, by using remote sensors.

6.12.2 Mineral Exploration

Remote sensing techniques have great promise both for reconnaissance and detailed exploration of non-renewable resources like minerals and fossil fuels. Remote sensing has been proved to be indispensable in this field and has been used in the following:

1. Mapping regional lineaments along which an individual group of mining districts may occur.
2. Mapping local fracture patterns that may control individual ore deposits.
3. Detecting hydrothermally altered rocks associated with ore deposits.
4. Providing basic geological data.

6.12.3 Land Use/Land Cover Analysis

With the rapidly growing population of the world, land use and land cover patterns are also undergoing rapid changes. Agricultural land is being used for urban purposes and the forests are being stripped to provide land for cultivation. Many of these changes are harmful to human interests and may, in the long run, prove to be disastrous. To prevent this damage to the environment, it is necessary to plan and regulate these changes. Planning in turn requires extensive data regarding the existing patterns and this is precisely where remote sensing comes in. This application is covered in detail in Section 6.13.

6.12.4 Natural Hazards

Earthquakes, landslides, volcanic eruptions and floods are the natural hazards that cause colossal losses—both material as well as human. Remote sensing can by no means prevent any of these hazards, but it can definitely minimise the effects of hazards in the following ways:

1. By analysing the geological formation of the area, thereby identifying the risk-prone areas.
2. By providing specific warning of certain natural hazards, e.g., volcanic eruptions in Hawaii have been successfully predicted on the basis of the ground movements revealed by remote sensing techniques.
3. Assess the damage caused and thereby help in the rescue and aid operations.

6.12.5 Archaeology

Another field that has found a widespread use for remote sensing techniques is archaeology. Archaeological patterns of prehistoric land use may be recognised in remote sensing images. Many roads and towns of the old era are now covered by agricultural fields. But due to change in moisture content and other characteristics of the underlying old objects and the upper new layers, remote sensors are able to recognise the buried archaeologically important sites.

Remote sensing, as a technique for surveying, has opened up new frontiers and has provided means for investigating those regions and phenomena which were earlier out of our bounds due to the limitations of ground surveying. The technology of space remote sensing is undergoing steady growth and a continuous improvement.

6.12.6 Application in GIS

Remote sensing consists in data acquisition and data analysis. The most important uses of the remotely sensed data are for the analysis of geographical data acquired by digital images and detecting temporal changes. The acquired remotely sensed images, digital data as well as the information extracted from such data is the primary data sources of modern Geographic Information System (GIS). The GIS environment permits the synthesis, analysis, and communication of virtually unlimited sources and types of data. The raw data can be incorporated directly as layers in the raster based GIS. The integration of the remotely sensed data with both spatial and non-spatial data stemming from a range of sources provides a real-time spatial analysis. The user may then use the analysis made to their decision making process. However, it may be noted that interaction between remote sensing and GIS is two way in nature; the GIS information may be used to aid the interpretation of the remotely acquired imageries.

6.13 LAND USE/LAND COVER ANALYSIS

Land use describes how a parcel of land is used (such as for agriculture, residences, or industry), whereas land cover describes the materials (such as vegetation, rocks, or buildings) that are present on the surface. The land cover of an area may be

evergreen forest, but the land use may be lumbering, recreation, oil extraction, or various combinations of activities. Accurate, current information on land use/land cover is essential for many planning activities.

Remote sensing methods are becoming increasingly important for mapping land use/land land cover for following reasons:

1. Images of large areas can be acquired rapidly.
2. Images can be acquired with a spatial resolution that matches the degree of detail required for the survey.
3. Remote sensing images eliminate the problems of surface access that often hamper ground surveys.
4. Images provide a perspective that is lacking for ground surveys.
5. Image interpretation is faster and less expensive than conducting ground surveys.
6. Images provide an unbiased, permanent data set that may be interpreted for a wide range of specific land use/land cover, such as forestry, agriculture, and urban growth.

There are some disadvantages to remote sensing surveys:

1. Different types of land use may not be distinguishable on images.
2. Most images lack the horizontal perspective that is valuable for identifying many categories of land use.
3. For surveys of small areas, the cost of mobilising a remote sensing mission may be uneconomical.

Remote sensing interpretations should be supplemented by ground checks of areas that represent various categories of land cover.

6.13.1 Classification Principles and Systems

Classification systems should recognise both activities (land use) and resources (land cover). Such a classification system that utilises orbital and aircraft remote sensing data should meet the following criteria:

1. The minimum level of accuracy in identifying land use/land cover categories from remote sensing data should be at least 85 per cent.
2. The accuracy of interpretation for all categories should be approximately equal.
3. Repeatable results should be obtainable from one interpreter to another and from one time of sensing to another.
4. The system should be applicable over extensive areas.
5. The system should be usable for remote sensing data obtained at different times of the year.
6. The system should allow use of subcategories that can be derived from ground surveys or from larger scale remote sensing data.
7. Aggregation of categories should be possible.
8. Comparison with future land use data should be possible.
9. Multiple land uses should be recognisable.

6.13.2 Use of the Multilevel Classification System

A multilevel classification system (MCS) that would enable the user to select the type and scale of image that suits the objectives of the survey is needed. Table 6.5 lists the image systems and image scales employed for each of the four classification levels. For example, the Level I classification is suitable for an entire state, whereas Level III is suitable for a municipality. For more details about classification, reference may be made to Chapter 7 Digital Image Processing to follow.

Table 6.5 *Image systems and scale*

<i>Level</i>	<i>System</i>	<i>Image scale</i>
I	Landsat MSS images	1:250 000 and smaller
II	Landsat TM images and High-altitude aerial Photographs, IRS-1C/1D PAN and LISS-III images	1:80 000 and smaller
III	Medium-altitude aerial photographs	1:20 000 to 1:80 000
IV	Low-altitude aerial photographs	Larger than 1:20 000

As per Anderson's multilevel classification there are nine land use/land cover classes. The land use/land cover analysis consists in constructing an appropriate classification system. Although different approaches to classification of land use can be applied, the system proposed by Anderson (1976) provides one of the best and most widely applicable outlines for structuring the classification. Their system (Table 6.6) provides the general categories, but the interpreter must become involved in the design of the more detailed categories at Levels II and III.

Table 6.6 *Land-use/land-cover classification system*

<i>Level I</i>	<i>Level II</i>	<i>Level III</i>
100 Urban or built-up	110 Residential	111 Single unit, low density (less than 2 DUPA*) 112 Single unit, medium density (2 to 6 DUPA) 113 Single unit, high density (greater than 6 DUPA) 114 Mobile homes 115 Multiple dwelling, low rise (2 stories or less) 116 Multiple dwelling, high rise (3 stories or less) 117 Mixed residential
	120 Commercial and services	121 Retail sales and services 122 Wholesale sales and services (including trucking and warehousing) 123 Offices and professional services 124 Hotels and motels 125 Cultural and entertainment 126 Mixed commercial and services

(Contd.)

*Development units per acre.

(Contd.)

	130 Industrial	131 Light industrial 132 Heavy industrial 133 Extractive 134 Industrial under construction
	140 Transportation	141 Airports, including runways, parking areas, hangars, and terminals 142 Railroads, including yards and terminals 143 Bus and truck terminals 144 Major roads and highways 145 Port facilities 146 Auto parking facilities (where not directly related to another land use)
	150 Communications and utilities	151 Energy facilities (electrical and gas) 152 Water supply plants (including pumping stations) 153 Sewage-treatment facilities 154 Solid-waste disposal sites
	160 Institutional	161 Educational facilities, including colleges, universities, high schools, and elementary schools 162 Religious facilities excluding schools 163 Medical and health-care facilities 164 Correctional facilities 165 Military facilities 166 Governmental, administrative, and service facilities 167 Cemeteries
	170 Recreational	171 Golf courses 172 Parks and zoos 173 Marinas 174 Stadiums, fairgrounds and race tracks
	180 Mixed	
	190 Open land and other	191 Undeveloped land within urban areas 192 Land being developed; intended use not known
200 Agriculture	210 Cropland and pasture	211 Row crops 212 Field crops 213 Pasture
	220 Orchards, groves, vineyards, nurseries, and ornamental horticultural areas	221 Row crops 222 Non-citrus orchards 223 Nurseries 224 Ornamental horticultural 225 Vineyards
	230 Confined feeding operations	231 Cattle 232 Poultry 233 Hogs
	240 Other agriculture	241 Inactive agricultural land 242 Other
300 Rangeland	310 Grassland	

(Contd.)

(Contd.)

	320 Shrub and brushland	321 Sagebrush prairies 322 Coastal scrub 323 Chaparral 324 Second-growth brushland
	330 Mixed rangeland	
400 Forest land	410 Evergreen forest	411 Pine 412 Redwood 413 Other
	420 Deciduous forest	421 Oak 422 Other hardwood
	430 Mixed forest	431 Mixed forest
	440 Clearcut areas	
	450 Burned areas	
500 Water	510 Streams and canals	
	520 Lakes and ponds	
	530 Reservoirs	
	540 Bays and estuaries	
	550 Open marine waters	
600 Wetlands	610 Vegetated wetlands, forested	611 Evergreen 612 Deciduous 613 Mangrove
	620 Vegetated wetlands, non-forested	621 Herbaceous vegetation 622 Freshwater marsh 623 Saltwater marsh
	630 Non-vegetated wetlands	631 Tidal flats 632 Other non-vegetated wetlands
700 Barren land	710 Dry lake beds	
	720 Beaches	
	730 Sand and gravel other than beaches	
	740 Exposed rock	
800 Tundra		
900 Perennial snow or ice	910 Perennial snowfields	
	920 Glaciers	

1. *Urban or built-up (100)*: This category comprises areas of intensive land use where much of the land is covered by structures and streets. Included in this category are cities, towns, villages, highways and alongside developments, etc.
2. *Agriculture (200)*: In some parts of the world, one can readily recognise land used for crops and orchards on satellite images by rectangular or circular patterns. It also includes cropland and pasture defined by a wide variety of crops, which are subdivided at Level III into three categories for example pasture (category 213) refers to relatively small areas of grazing land commonly interspersed with croplands.
3. *Rangeland (300)*: Rangeland is land covered by natural grasses, shrubs, and forbs, which include non-woody plants such as weeds and flowers. Rangeland is capable of supporting native or domesticated grazing animals.
4. *Forest land (400)*: Forest lands have a crown density (also called the crown closure percentage) of 10 per cent or more and support trees capable of producing timber or other wood products. The 'forest' category also includes lands from which trees have been removed to a crown density of less than 10 per cent but have not been developed for other uses.
5. *Water (500)*: Lakes and ponds are natural water bodies, whereas reservoirs are artificially impounded. Shallow water areas where aquatic vegetation is submerged are also classified as water.
6. *Wetlands (600)*: These are the areas where the water table is at, near, or above the land surface for a significant part of most years. Most wetlands are located adjacent to water bodies and include marshes, swamps, tidal flats, and mudflats. Areas that are only seasonally flooded and do not support typical wetland vegetation are assigned to other categories.
7. *Barren land (700)*: Barren land has a limited ability to support life, and less than one-third of the area has vegetation cover. The surface is predominantly thin soil, sand, or rocks.
8. *Tundra (800)*: 'Tundra' refers to treeless regions beyond the limit of the boreal forest. Regions above the timber line in high mountain ranges are also classed as tundra. Tundra vegetation is low and dwarfed and it commonly forms a complete mat. The presence of permafrost and the prevalence of subfreezing temperatures most of the year are characteristics of tundra.
9. *Perennial snow or ice (900)*: This category is used for areas where snow and ice accumulations persist throughout the year as snowfields or glaciers.

6.14 METHODOLOGY FOR LAND USE/LAND COVER MAPPING

To map the land use and land cover of any region, IRS satellites LISS sensor False Colour Composite (FCC) at desired scale (1:250 000/1:50 000/1:25 000) of single season or different seasons can be used. The land use/land cover maps so

prepared will depict the details (spatial as well as area) of all land use/land cover classes for all the seasons.

As a case study, to find the spatial coverage and area statistics of different crops under Rabi and Kharif season, a methodology has been given in Fig. 6.11 by adopting visual interpretation techniques.

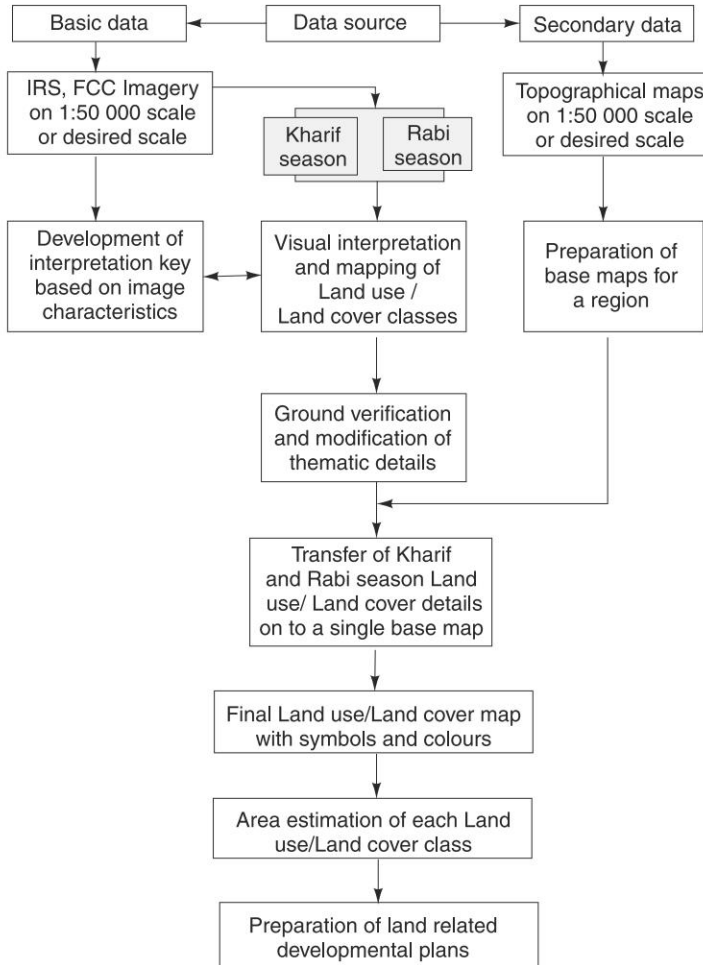


Fig. 6.11 Flowchart showing methodology for mapping land use/land cover through visual interpretation techniques

Satellite imagery contains a detailed record of features on the ground at the time of satellite overpass. An image interpreter systematically examines the image for generating the information required by him. Other supporting materials such as published maps and reports of various departments and field observations will increase the accuracy of interpretation.

There are certain fundamental photo elements or image characteristics seen on image which aid in the visual interpretation of satellite imagery. These are

colour, size, shape, texture, pattern, location, association, resolution and season. By understanding the image characteristics of each of the thematic class, an image interpretation key can be evolved. This will enable the image interpreter having the minimum of knowledge and interpretation skill to identify different features on satellite imagery. In other words, it simplifies and facilitates images interpretation of the imagery in a systematic way. However, it may be mentioned here that visual interpretation is subjective and differs from individual to individual and also depends upon the season, scale, spectral bands, spatial resolution, overall image contrast and quality of the data. The procedure for mapping land use/land cover from satellite imagery of different seasons and scale is well established and is fundamental for various advanced remote sensing applications.

6.15 REMOTE SENSING IN INDIA

The Satellite for Earth Observations (SEO), now called Bhaskara, was the first Indian remote sensing satellite. Construction of the satellite began in 1973 by ISRO and was launched by a Soviet launch vehicle from USSR into a near circular orbit in June 1979. The satellite weighing about 440 kg was polyhedral in shape with 26 flat faces and measured approximately 1.4 m from end to end. The sensors on Bhaskara satellite are two television cameras, three microwave radiometers and a Data Collection Platform (DCP) system. The TV cameras were intended for use in disciplines like forestry, snow, hydrology, geology and coastal phenomena. Data from microwave radiometers and the DCP system were for applications in oceanography and meteorology.

As a follow up of Bhaskara, SEO-II, which is also a prototype of SEO-I was launched on 20th November 1981 from a Soviet cosmodrome and is called Bhaskara-II. TV cameras of Bhaskara-I had some problems in the beginning. But one of the cameras started functioning and large amount of imageries over the Indian subcontinent was obtained. Cameras of Bhaskara-II returned about 1000 imageries of the Indian land mass in visible and near infrared bands, providing a total coverage of the country more than twice. This data was mainly used in geological and land cover mapping on 1:2 000 000 scale, forest, semi-arid regions, flood and snow cover mapping. Bhaskara imageries having a very coarse resolution (1 km) were in no way comparable to Landsat imageries of 80 m resolution, but it gave ISRO experience in the total system capability level.

Following the successful demonstration flights of Bhaskara-I and Bhaskara-II launched in 1979 and 1981, respectively, India began development of an indigenous IRS (Indian Remote Sensing Satellite) program to support the national economy in the areas of agriculture, water resources, forestry and ecology, geology, water sheds, marine fisheries and coastal management. The Indian Remote Sensing satellites are the main-stay of National Natural Resources Management System (NNRMS), for which Department of Space (DOS) is the nodal agency, providing operational remote sensing data services. Data from the IRS satellites is received and disseminated by several countries all over the world. With the advent of high-resolution satellites, new applications in the areas of urban sprawl, infrastructure planning and other large-scale applications for mapping have been initiated.

The first two IRS spacecraft, IRS-1A (March 1988) and IRS-1B (August 1991) were launched by Russian Vostok boosters from the Baikonur Cosmodrome. IRS-1A failed in 1992, while IRS-1B continued to operate through 1999. From their 22-day repeating orbits of 905 km mean altitude and 99 degrees inclination, the two identical IRS spacecraft hosted a trio of Linear Imaging Self-Scanning (LISS) remote sensing linear Charge Coupled Device (CCD) instruments working in four spectral bands: 0.45–0.52 μm , 0.52–0.59 μm , 0.62–0.68 μm , and 0.77–0.86 μm . The 38.5 kg LISS-I images swath of 148 km with a resolution of 72.5 m while the 80.5 kg LISS-IIA and LISS-IIB exhibit a narrower field-of-view (74 km swath) but are aligned to provide a composite 145 km swath with a 3 km overlap and a resolution of 36.25 m.

Each IRS spacecraft is 975 kg at launch with a design life of 2.5–3 years. The three-axis stabilised spacecraft is essentially rectangular (1.1 m by 1.5 m by 1.6 m) with two narrow solar arrays producing less than 1 kW electrical power. Spacecraft data transmissions are effected via X-band and S-band antennas at the base of the spacecraft.

IRS-1A and IRS-1B were to be joined in 1993 with IRS-1E, the modified IRS-1A engineering model which had been equipped with the LISS-1 and a German monocular electro-optical stereo scanner. The spacecraft was lost, however, when its PSLV launch vehicle failed to reach the earth's orbit. 13 months later in October 1994, the PSLV functioned correctly, allowing IRS-P2 to assume an 820 km, sun-synchronous orbit. This spacecraft continued in operation until September 1997. With an 870 kg mass (slightly less than IRS-1A and IRS-1B), IRS-P2 carried the LISS-II system with a ground resolution of 32 m across-track and 37 m along-track. The total swath width is 131 km, and the CCD array is tuned to four spectral bands between 0.45 and 0.86 μm . The spacecraft's solar arrays provide up to 500 W and are linked to conventional nickel–cadmium storage batteries.

As of late 1999, five IRS satellites were operating, and more were scheduled for launch by the year 2004–05. IRS-1C, successfully launched on 28th December 1995, on-board a Molniya rocket of Russia, was the last Russian launch of the program (Molniya rather than Vostok, while IRS-1D was orbited by India's PSLV). IRS-P3 was launched by PSLV in 1996 with a German modular electro-optical scanner and an Indian visible-IR scanner.

The Indian Space Research Organisation and its commercial marketing arm, ANTRIX Corp. Ltd., successfully launched the IRS-1D earth imaging satellite on 29 September 1997 from Sriharikota, India. The satellite is an identical twin to the IRS-1C, launched in December 1995. The dual use of these satellites provides 5.8 m resolution images to customers twice as often as was possible with just the IRS-1C.

IRS-1C and IRS-1D introduced a heavier (1350 kg), more capable earth observation platform. The spacecraft bus is similar to those of IRS-1A and IRS-1B, but a slightly larger solar array generates more than 800 W. Both IRS-1C and 1D produce 5.8 m panchromatic (0.5–0.75 μm —black and white) imagery, which is resampled to 5 m pixel detail. This resolution, which as of early 1998 was the best of any civilian remote sensing satellites in the world, is superior to the 8 m resolution initially reported for the panchromatic imager. These satellites are also

equipped with two-band wide-field sensors (WiFS) that cover a 774 km² (481 square mile) area in a single image, as well as LISS-III four-band (0.52–0.59, 0.62–0.68, 0.77–0.86, and 1.55–1.70 μm) multispectral sensors that provide 23.5 m resolution multispectral coverage. The 23.5 m resolution imagery is resampled to produce 20 m pixel detail. The spacecraft also carry a two-channel (0.62–0.68 and 0.77–0.86 μm) wide-field sensor (190 m resolution).

The IRS-1C, IRS-1D PAN sensor sacrifices swath width for its higher resolution. However, it can be pointed off the orbit path which allows 2- to 4-day revisits to specific sites. IRS-1C and IRS-1D data can be received and procured from EOSAT (USA) or in India at the NRSA, Hyderabad.

IRS-P4 (OCEANSAT-1) has payloads, specifically tailored for the measurements of physical and biological oceanography parameters. An Ocean Colour Monitor (OCM) with eight spectral bands, Multifrequency Scanning Microwave Radiometer (MSMR) operating in four frequencies will provide valuable ocean-surface related observation capability. The OCEANSAT-1 was launched by PSLV in early 1998.

1. IRS-P6 (RESOURCESAT-1) is state-of-the-art satellite mainly for agricultural applications and has a three-band multispectral LISS-IV camera with a spatial resolution of 5.8 m and a swath of around 25 km with across-track steerability for selected area monitoring. The satellite was unlaunched by the indigenously built Polar Satellite Launch Vehicle on October 17, 2003. The orbit parameters of IRS-P6 are same as IRS-1C. The payload system of IRS-P6 consists of three solid state cameras:

1. A high resolution multispectral sensor—LISS-IV
2. A medium resolution multispectral sensor—LISS-III
3. An Advanced Wide Field Sensor—AWiFS

The LISS-IV camera is a multispectral high-resolution camera with a spatial resolution of 5.8 m at nadir. The sensor consists of three linear odd-even pairs of CCD arrays, each with 12000 pixels. The odd and even pixel rows are separated by 35 microns, which correspond to five scan lines. The LISS-III is a multispectral camera operating in four spectral bands, three in the visible and near infra-red and one in short wave infra-red region, as in the case of IRS-1C/1D. The new feature in LISS-III camera is the SWIR band (1.55 to 1.7 microns), which provides data with a spatial resolution of 23.5 m like IRS-1C/1D (the spatial resolution is 70 m). The AWiFS camera provides enhanced capabilities compared to the WiFS camera on-board IRS-1C/1D, in terms of spatial resolution (56 m *vs.* 188 m), radiometric resolution (10 bits *vs.* 7 bits) and Spectral bands (4 *vs.* 2) with the additional feature of on-board detector calibration using LEDs. The spectral bands of AWiFS are same as LISS-III.

The AWiFS camera is realised in two electro-optic modules, *viz.*, AWiFS-A and AWiFS-B, each containing four band assemblies. A combined swath of 740 km is realised by mounting the two modules on the Deck, with their optical axes tilted by 11.94° away from the +y axis in opposite direction.

2. IRS-P5 (CARTOSAT-1) satellite, launched in May 5, 2005 has two panchromatic cameras with 2.5 m spatial resolution, to acquire two images simultaneously—one

forward looking (FORE) at +26 degree and one aft of the satellite at -5 degrees for near instantaneous stereo data. The time difference between the acquisitions of the same scene by the two cameras is about 52 seconds. The spacecraft body is steerable to compensate the earth rotation effect and force both Fore and Aft cameras to look at the same ground strip when operated in stereo mode. The stereo pairs have a swath of 26 km. This mission caters to the needs of cartographers and terrain modelling applications. The satellite will provide cadastral level information up to 1:10000 scale and will be useful for making 2–5 m contour maps. Simultaneous stereo pair acquisitions are of great advantage since the radiometric parameters of the images will be identical.

3. IRS-P7 (CARTOSAT-2) was built and launched by the PSLV on January 10, 2007. Weighing around 680 kg at launch, its applications will mainly be towards cartography in India. CARTOSAT-2 carries a state-of-the-art panchromatic (PAN) camera that works in the visible region of the electromagnetic spectrum. The swath covered by these high-resolution PAN cameras is 9.6 km and their spatial resolution is less than 1 metre. The satellite can be steered up to 45 degree along as well as across the track. It is an advanced remote-sensing satellite capable of providing scene-specific spot imagery. The data from the satellite will be used for detailed mapping and other cartographic applications at cadastral level, urban and rural infrastructure development and management, as well as applications in Land Information System (LIS) and GIS. The details are given below.

Altitude	630 km
Inclination	97.91 degree
Orbital per day	14
Revisit	4 days
Repetivity	310 days
Payload	One Panchromatic camera
Spatial Resolution	Less than 1 metre
Swath	9.6 km
Spectral Band	0.5–0.85 micrometre

4. CARTOSAT 2A is an earth observation satellite in a sun-synchronous orbit. The satellite is the thirteenth satellite in the Indian Remote Sensing (IRS) satellite series. It was launched by the Polar Satellite Launch Vehicle—C9 on April 28, 2008 along with the 87 kg Indian Mini Satellite (IMS-1). This satellite belongs to the Ministry of Defence mission of the Government of India. It will be a dedicated satellite for the Indian Armed Forces which is in the process of establishing an Aerospace Command. The satellite carries a panchromatic camera (PAN) which has spatial resolution less than 1 m. The highly agile CARTOSAT-2A can be steered up to 45° along as well as across the direction of its movement to facilitate imaging of any area more frequently. The properties are given as follows.

Revisit	4 days
Repetivity	310 days
Payload	One Panchromatic camera
Spatial Resolution	Less than 1 metre
Spectral Band	0.5–0.85 micrometre

5. OCEANSAT-2 satellite mission is conceived to provide continuity of services to the Oceansat-1 data users. This satellite will have enhanced capabilities. It carries an Ocean Colour Monitor (OCM) and Wind Scatterometer. Inclusion of a thermal infrared radiometer is also under consideration. This will be launched in 2009–2010.

6. CARTOSAT-2B is an earth observation satellite in a sun-synchronous orbit, and was launched on 12.07.2010. It is the seventeenth satellite in the IRS satellite series. It was launched in a Polar Satellite Launch Vehicle (PSLV) rocket from the spaceport at Sriharzkota. The satellite carries a panchromatic (PAN) camera capable of taking black-and-white pictures in the visible region of electromagnetic spectrum. It can be steered up to 26° along as well as across the direction of its movement to facilitate imaging of the area more frequently.

7. RESOURCESAT-2 is a follow on mission to RESOURCESAT-1 and the eighteenth remote sensing satellite built by ISRO. It is intended to continue the remote sensing data service to global users provided by resourcesat-1, but with enhanced multispectral and spatial coverage as well. LISS-4 multispectral swath has been enhanced from 23 km to 70 km suitable changes including miniaturisation in payload electronics have been incorporated in RESOURCESAT-2. The satellite consists of three multispectral cameras—the advanced Wide-Field sensor (AWiFS) with 56 m spatial resolution; the linear imaging self-scanning sensor (LISS-III) with 23.5 m spatial resolution; and LISS-IV camera with 5.8 m spatial resolution—on-board.

Megha-Tropiques is a satellite mission to study the water cycle in the tropical atmosphere in the context of climate change and was launched on 12.10.2011. Megha-Tropiques was successfully deployed into orbit by PLSV-C 18 rocket. It is designed to understand tropical, metrological, and climatic processes by obtaining reliable statistics on the water and energy budget of the tropical atmosphere. It also seeks to describe the evolution of major tropical weather system, with a focuss on repetitive measurement of the tropics.

8. RISAT-1 satellite has been launched by India, on 26.4.2012 to operate in the microwave region. The satellite carries a multimode, agile, Synthetic Aperture RADAR (SAR) payload operating in ScanSAR Mode to provide images with coarse and medium resolution and Strip Map Mode (SMM) to provide images with fine and high spatial resolution irrespective of light and weather conditions of the area being imaged. The details are given in Tables 6.7 and 6.8.

Table 6.7 Properties of different modes

Parameter	Scan SAR mode (MRS, CRS)	Strip map mode (FRS)
Altitude	608.958 km	608.145 km
Inclination	97.837°	97.833°
Repetivity cycle	193 orbits in 13 days	1767 orbits in 19 days
Orbit period	96.995 min	96.978 min
Swath	240 km	30 km

Table 6.8 Resolution details of different modes

Mode	Look	Resolution (m)	Swath (km)	Polarisation
Coarse Resolution mode	2–4	50	240	Co and/or cross
Medium Resolution mode (MRS)	1–2	25	120	Co and/or cross
Fine Resolution Striping Single mode (FRS-2)	9–12	12	30	Quad
Fine Resolution Strip map (FRS-1)	Single	3–6	30	Co and/or cross
High Resolution Spot light mode (HRS)	Single	1–2	10 × 10	Co and/or cross

Future IRS Launches The remote sensing satellites to be launched to strengthen the fleet of IRS satellites are as follows.

1. **RESOURCESAT-3:** It will carry more advanced LISS-III-WS (wide swath) sensor having similar swath and re-visit capability as AWiFS, thus overcoming any spatial resolution limitation. Satellite will also carry atmospheric correction sensor (ACS) for quantitative interpretation and geophysical parameter retrieval.
2. **CARTOSAT-3:** A continuation of CARTOSAT series, it will have a resolution of 30 cm and 6 km swath suitable for cadastre and infrastructure mapping and analysis. It will also enhance disaster monitoring and damage assessment. It is slated to be launched in 2012.
3. **OCEANSAT-3:** It will carry thermal infrared sensor and 12 channel ocean colour monitor, scatterometer, and passive microwave radiometer. IR sensor and ocean colour monitor will be used to analyse for operational potential fishing zones. The satellite is mainly for ocean biology and sea state applications. It is slated to be launched in 2012–13.

Remote-sensing applications in the country, under the umbrella of NNRMS, now cover diverse fields such as crop acreage and yield estimation, drought warning and assessment, flood control and damage assessment, land use/land

cover information, agro-climatic planning, wasteland management, water resources management, underground water exploration, prediction of snowmelt run-off, management of water-sheds and command areas, fisheries development, mineral prospecting and forest resources survey. Active involvement of the user ministries/departments has ensured in an effective harnessing of the potential of space-based remote sensing. An important application of IRS data is in the Integrated Mission for Sustainable Development (IMSD) initiated in 1992. IMSD, under which 174 districts have been identified, aims at generating local-specific action plans for sustainable development.

6.16 SATELLITE DATA PRODUCTS

The National Remote Sensing Agency (NRSA), Hyderabad is the only agency in India which acquires and archives satellite data. These are usually available as photographic products—paper prints, films, negatives, diapositives of individual band images in black and white, colour composites and digital data—as computer compatible tape (CCT), floppy or CD.

Summary

This is a broad subject and can not be justified within the limited scope of the book. However an attempt has been made to introduce the subject concisely and objectively.

Application of remote sensing in surveying is not new. Aerial photogrammetry is considered to be a form of remote sensing and is in practise since long. Remote sensing may be defined as a science and art of obtaining information about the object of interest without being in touch of the object (target). The term remote sensing generally refers to the use of aerial sensor technologies to detect and classify objects on earth (both on surface, and in the air and oceans) by propagating signals (e.g., electromagnetic radiation) from aircraft/satellite.

The technique consists in collecting, processing and interpreting information without being in physical contact with the object. Collection of the information is done from space by the deployment of satellites. The data collection is accomplished by sensors (cameras, multispectral scanner) mounted on a platform (aircraft, satellite). The information received by the sensor is manipulated, modified, and sent back and stored in a storing device (magnetic tapes, CD ROM). This stored data is processed on ground to produce photographs and other products. These photographs are then interpreted visually or digitally for object identification.

Human eyes sense a small and somewhat variable window of frequencies of electromagnetic radiation (EMR) called the visible spectrum. EMR is a form of energy emitted and absorbed by charged particles, which exhibit wave like behaviour as it travels through space. The electromagnetic spectrum is introduced. Since human eye system works only under visible light, most remote sensing methods use the reflected infrared bands, thermal infrared bands, or microwave portions of the EM spectrum. The main source of electromagnetic energy is sun and since the energy propagates through the atmosphere, interaction and effects of EM energy with atmosphere and earth surface features is described. In order for an observing sensor to acquire data about remote object, there must be flow of information between the object and the observing sensor. There has to be a carrier of that information, which in the case of remote sensing is EMR.

Sensors for data acquisition may be active or passive. Sensors that sense natural radiations either reflected or emitted from earth are called passive sensors; examples are taking photograph on a bright day, radar, etc. On the other hand, sensors which carry EM radiation of a specific wavelength or band of wavelengths to illuminate the earth's surface are called active sensors; example is taking photograph in dark place with flashbulb. Sensors are mounted on platforms which may be an aircraft or satellites. Aircrafts are suitable for small (local) areas while satellites cover the area globally. The latter are also suitable to monitor the dynamic features such as crop change, growth, yield, etc. The satellites may be classed as natural (moon) or artificial (geostationary, sun-synchronous).

Interpretation of the acquired data is an integral part of remote sensing. The quality of remote sensing data consists of its spatial, spectral, radiometric, and temporal resolutions. The size of a pixel that is recorded in a raster image represents spatial resolution. Typically, pixels correspond to square areas with side lengths up to 1000 m. Spectral resolution is represented by the recorded wavelength band of the different frequency bands. Usually, this is related to the number of frequency bands, including several in the infrared spectrum, ranging from a spectral resolution of 0.07 to 2.1 μm . The number of different intensities of radiation, the sensor is able to distinguish, represents the radiometric resolution. Typically, this ranges from 8 to 14 bits corresponding to 256 levels of the grey scale, and up to 16384 intensities or shades of colour in each band. The frequency of flyovers by the aircraft/satellite is represented by temporal resolution. The importance of resolution, i.e., the ability of remote sensing system to record and display fine details is explained. Importance and significance of reference data, signature, and false colour composites in interpreting the features is discussed.

Much interpretation and identification of objects in remotely sensed image is performed manually or visually by the interpreter. Recognising objects is the key to interpretation and information extraction. Observing the differences between objects and their backgrounds involves comparing different objects based on any, or all the visual elements of tone, shape, size, pattern, texture, shadow, and association. Visual interpretation is facilitated by a stereoscope, which makes possible to see the three dimensional model of the area imaged. The visual interpretation techniques have certain disadvantages, require extensive training and are labour intensive. Moreover, the limited ability of human eye to discern tonal values and analyse spectral changes also puts a limit. For remotely sensed data available in digital format, the spectral patterns can be more fully examined using computer-assisted techniques; the digital processing and analysis and is described in the next chapter.

Applications of remote sensing are innumerable. Some of the more important ones such as environmental, land-use/land-cover analysis, GIS, etc., are presented. The chapter ends with the details of remote sensing programmes in India.

Exercises

- 6.1 Define remote sensing. State how it differs from photogrammetry.
- 6.2 Define photo-interpretation. How does it differ from remote sensing?
- 6.3 Name the systems of remote sensing. Give a brief description of them.
- 6.4 List the applications of remote sensing and describe any two in details.
- 6.5 Give the applications of remote sensing with respect to the natural hazards and that of archaeology.
- 6.6 Discuss briefly the development of remote sensing in India and its utility.
- 6.7 What are the basic processes and elements involved in electromagnetic remote sensing of earth resources?

- 6.8** Write short notes on the following:
- (a) False colour composites
 - (b) Geostationary satellites
 - (c) Remote sensing platforms
 - (d) Sensors
- 6.9** Write short notes on the following:
- (a) Elements of remote sensing process
 - (b) Remote sensing sensor systems
 - (c) Digital image
 - (d) Satellite data products
- 6.10** Briefly discuss the following:
- (a) Resolution of remote-sensing data
 - (b) Reference data
 - (c) Return beam vidicon
 - (d) Atmospheric windows
- 6.11** Discuss briefly electromagnetic energy and electromagnetic spectrum.
- 6.12** Describe the following:
- (a) Particle theory
 - (b) Wave theory
 - (c) Energy interaction with atmosphere and earth surface features
- 6.13** What is idealised remote sensing system?
- 6.14** What are the various forms of remotely collected data? Briefly describe data acquisition and its interpretation.
- 6.15** Differentiate between the following:
- (a) Active and passive systems of remote sensing
 - (b) Selective and non-selective scattering
 - (c) Geostationary and sun-synchronous satellites

Objective-type Questions

- 6.1** Remote sensing can be defined as collecting information about a target
- (a) without seeing it
 - (b) by touching it
 - (c) without a physical contact with it
 - (d) from a ground station
- 6.2** Which of the following are examples of remote sensing systems?
- (a) Human eye
 - (b) Bat's guidance system
 - (c) Remote sensing satellites
 - (d) All of the above
- 6.3** Which of the following statements best expresses the relationship between remote sensing (R) and photo-interpretation (P)?
- (a) R is a form of P and vice versa
 - (b) P is a form of R
 - (c) R is a form of P
 - (d) cannot be concluded
- 6.4** Which of the following factors does not contribute towards accurate photo-interpretation?
- (a) Knowledge of the area
 - (b) Skill and ingenuity
 - (c) Incorrect interpretation of size
 - (d) None of these

- 6.5** Which of the following remote sensing systems employs only one detector?
 (a) Scanning system (b) Framing system
 (c) Both (a) and (b) (d) None of these
- 6.6** Which of the following cannot be achieved by remote sensing?
 (a) Detection of forest fires (b) Detection of pollutants
 (c) Prevention of earthquakes (d) Land use pattern
- 6.7** Remote sensing is being used in land use analysis because
 (a) images of large areas can be acquired rapidly
 (b) problem of access is eliminated
 (c) it is cheaper as compared to the ground surveys
 (d) All of the above
- 6.8** The energy of a quantum is
 (a) directly proportional to its wavelength
 (b) inversely proportional to its wavelength
 (c) directly proportional to the square of its wavelength
 (d) inversely proportional to the square root of its wavelength
- 6.9** Stefan–Boltzmann Law is
 (a) $M = \sigma T^4$ (b) $M = \sigma T^2$
 (c) $\lambda_m = AT$ (d) $\lambda_m = A/T$
 where σ , T , and A have their usual meanings.
- 6.10** Which one of the following is not an example of surface phenomena?
 (a) Emission (b) Scattering
 (c) Reflection (d) Absorption
- 6.11** A passive sensor uses which of the following sources of energy?
 (a) Sun (b) Flash light
 (c) Its own source (d) Moon
- 6.12** Which one of these is a long wavelength radiation?
 (a) Ultra violet (b) X-ray
 (c) Infrared (d) Gamma ray
- 6.13** The wavelength region of $2.08 \mu\text{m}$ is generally used for
 (a) vegetation discrimination (b) soil moisture discrimination
 (c) geological mapping (d) water penetration
- 6.14** If the mean temperature of the sun's surface 6000 K and λ_m of its radiation of $0.5 \mu\text{m}$, what is the mean temperature of earth's surface for which λ_m is $10.0 \mu\text{m}$, according to Wien's displacement law in the remote sensing concept?
 (a) 25° (b) 28° (c) 27° (d) 30°

Answers to Objective-type Questions

6.1 (c)	6.2 (d)	6.3 (b)	6.4 (c)	6.5 (a)	6.6 (c)
6.7 (d)	6.8 (b)	6.9 (a)	6.10 (d)	6.11 (a)	6.12 (c)
6.13 (c)	6.14 (c)				

7

DIGITAL IMAGE PROCESSING AND ANALYSIS

Introduction

Digital image data are usually the remotely sensed data, and as such, the data processing in remote sensing is referred as *digital image processing* (DIP). The DIP consists in the application of algorithms on digital images to process, analyse and extract information of interest by manipulation and interpretation of the images. The digital image processing is usually done using raster data structure—each image is treated as an array of values. Since it is easy to find and locate pixels and their values, these can be manipulated easily by image processing system. Although, there are many possible forms of digital image manipulation image rectification (preprocessing), image enhancement, image transformation and image classification are some of the common features. After the completion of the preprocessing, the analyst may use feature extraction* to reduce the dimensionality of the data. Steps involved in digital image processing are presented in Fig. 7.1.

The raw data received from the imaging sensors, on the satellite platforms, contain flaws and deficiencies and, therefore, needs processing, to get the originality of the data. The steps involved may vary from image to image depending on the type of image format, initial condition of the image, composition of the image, and the information of interest. The image analyst examines the images to identify objects and judge their significance through feature extraction, segmentation and classification.

By studying the remotely sensed data, the analyst logically detects, identifies, classifies, measures and evaluates the significance of the objects, their patterns and spatial distributions. The ultimate aim of DIP is to extract information from a digital image that is not readily apparent or available in its original form. For effective DIP the data must be originally recorded digital data, stored in digital data storage device (e.g., hard disc, CD, DVD, etc.), and a computer, referred to as an image analysis system, with necessary hardware and software to process the data. The scope of the book does not permit to get into the complex mathematics involved in digital image processing. However, principles behind the process are explained for in-depth understanding. *Feature extraction* is the process of isolating the most useful components of the data, discarding the errors, for the further study. It reduces the number of variables to be examined thereby saving time and resources. In a multispectral data, the feature extraction helps to portray

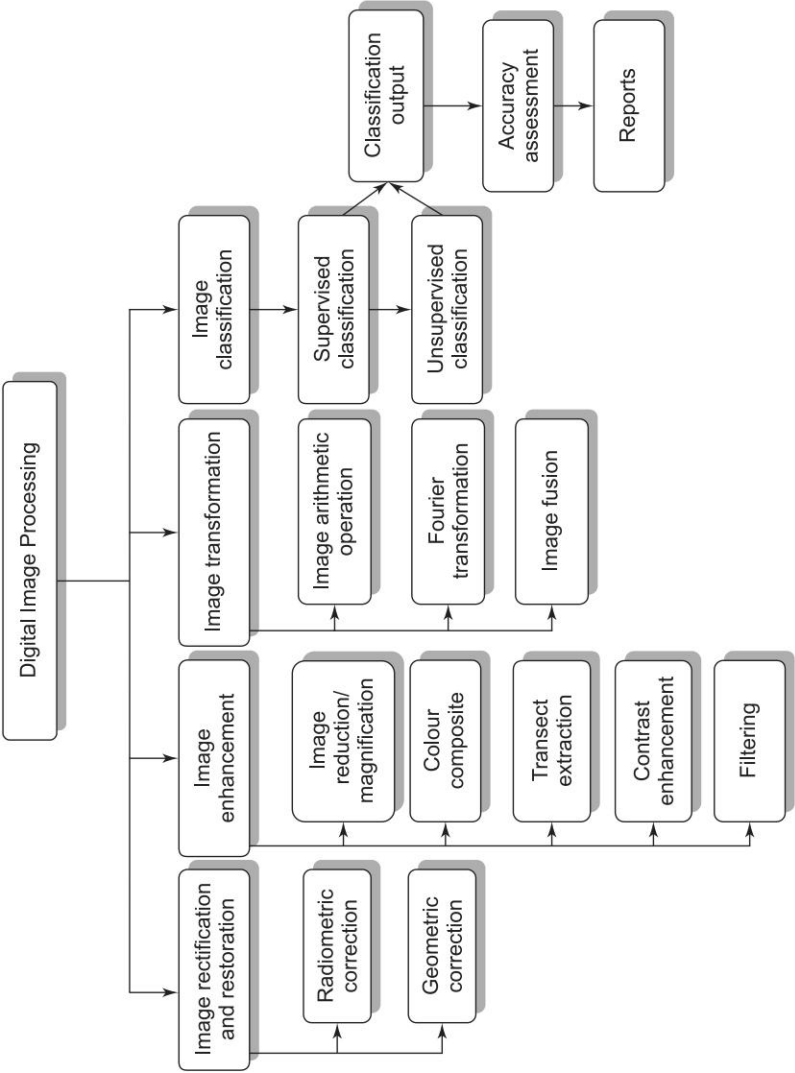


Fig. 7.1 Steps involved in digital image processing

the necessary elements of the image by enhancing the statistical characteristics of image data such as individual bands or combination of band values that carry information concerning systematic variation within the scene; feature extraction should not be got confused with the geographical features visible on the image. Once the feature extraction is over the analyst can work with the image comprising of desired channels or bands.

7.1 IMAGE RECTIFICATION AND RESTORATION

This operation is often termed as *preprocessing* since it normally precedes further manipulation and analysis of the image data to extract specific information. Distortions or degradations in the image stem from the image acquisition process. Therefore, the aim of image rectification and restoration is to correct the image data and is accomplished by offsetting problems with the band data by recalculating the DN values. The factors that affect are the *digital image acquisition type*—digital camera, along-track scanner, across-track scanner; *platform*—airborne vs. satellite; and *total field of view*. The preprocessing operation attempts to correct or compensate the remotely sensed data for errors such as *radiometric errors*—removal of sensor or atmospheric noise to represent ground conditions more accurately, i.e., to improve image fidelity by correcting for data loss, removing haze; and then *geometric distortions*—converting data to ground coordinates by removal of distortions from sensor geometry, enabling mapping relative to data layers. Once the errors are removed, the data is said to be *restored* to its original condition; the correct values are never known. Further, the attempt to correct the data may itself introduce errors. However, all these corrections might not be required in every case and will depend upon the nature of information to be extracted from the data.

7.1.1 Radiometric Corrections

When the image data contains errors in the measured brightness values of the pixels, it limits the ability of the analyst to visually interpret or quantitatively process and analyse the images. The potential sources of errors are: periodic drift or malfunctioning of a detector; electronic interference between sensor components; and intermittent disruptions in data transmission and recording. Consequently, the affects are: (i) a difference in the relative distribution of brightness over an image in a given band than that in a ground scene and (ii) distortion of the relative brightness of a single pixel from band to band compared with spectral reflectance character of the corresponding region of the ground.

It, therefore, becomes necessary to apply the radiometric corrections—detector response calibration, atmospheric correction, sun elevation correction, and earth-sun distance correction (the error due to the detector response influences the most)—by modifying DN values to account for noise. The correction involves a rearrangement of the DN in an image so that all areas have the same relationship between the DN and either radiance or back-scatter.

Detector Response Calibration Detector within a sensor has a curvilinear response to radiance and backscatter. Radiance is measure of the amount of EM radiation leaving or arriving at a point on surface, whereas backscatter is

reflection of signals back to the direction from which they came. If the form of detector response is known, it can be used to transform the detector output from a curvilinear to a linear response (Fig. 7.2). Further, since the original detector responses are measured in laboratory and do change with time, even after the application of correction the response remains slightly curvilinear and may require further treatment. This is known as *nominal correction*. The detector response calibration consists of noise correction, de-stripping, and line-drop outs. This is known as *supplemental correction*.

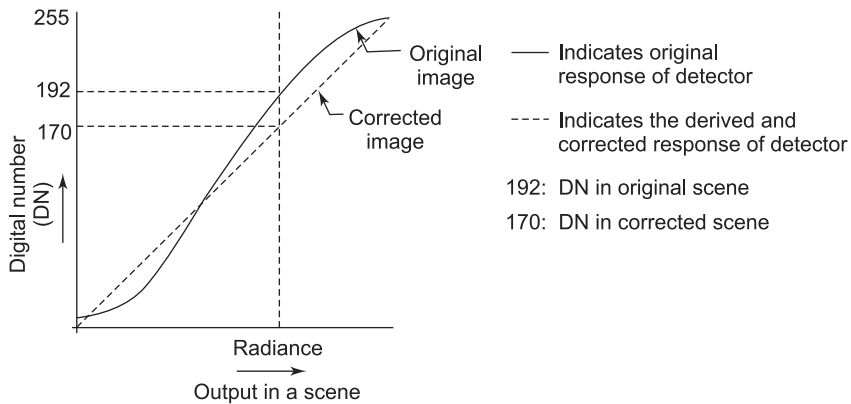


Fig. 7.2 Radiometric correction

1. **Noise correction:** Any unwanted disturbance in image data that is due to limitations in the sensing, signal digitisation, or data recording process is known as *image noise*. Noise can either degrade or totally mask the true radiometric information content of a digital image. Noise can be systematic (periodic), or random, or a combination of both. When odd pixels have a marked difference in their DN from those of adjacent pixels in the affected band, they are classed as *random noise*. The random noise problem is characterised by non-systematic variations in grey levels from pixel to pixel called *bit errors*. The noise values normally change much more abruptly than true image values. Pixel with spurious DN can be replaced with the average value of the neighbouring DNs. Moving neighbourhoods or windows of 3×3 or 5×5 pixels are typically used in such procedures (Fig. 7.3).

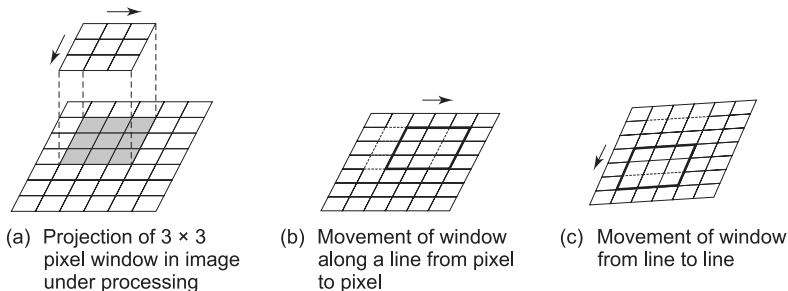


Fig. 7.3 Moving window concept

2. *De-stripping*: When a detector goes out of adjustment, *de-stripping*, also known as *line striping* or *banding* (Fig. 7.4), occurs. The error is due to periodic drift or malfunctioning of a detector. This was a common problem in early Landsat MSS data due to different drifts with time of the six MSS detectors resulting in different brightness by each detector. A systematic horizontal banding pattern on images produced by electro-mechanical scanners result in a repeated pattern of lines with consistently high or low DN's. Striping may also occur due to improper data recording and transmission.

The de-stripping process consists in constructing histograms of the problem band for each detector and calculating the mean and standard deviation for each histogram. In case, if these statistical parameters come out to be different, they are made equal to the mean and standard deviation value of the whole image. Application of the de-stripping correction enhances the visual appearance and interpretability of the image.

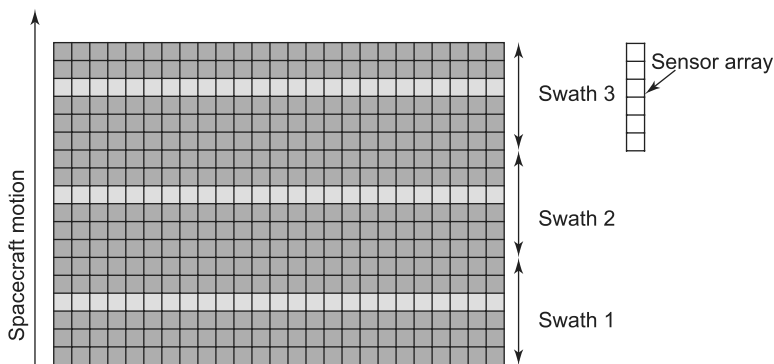


Fig. 7.4 *De-stripping (line striping)*

3. *Line-drop outs*: When a detector malfunctions completely or temporarily, a string of adjacent pixels in a scan line may contain defective data (spurious DN value) creating a horizontal streak, or miss a partial or complete line. The correction process consists in replacing the drop-out line (if only one line) by the average of the immediate preceding and next neighbouring lines. In case of two lost lines, the first lost line is replaced by the preceding line and the second lost line by the subsequent line. When three lines are lost, the outer two lines are treated as above and the middle line is recovered by averaging the two outer recovered lines. Recovery of three lost lines is the limit of the correction to be affected. Line-drop outs are also an example of systematic noise.

Atmospheric Corrections The ground surface receives the direct solar radiations as well as scattered radiation from the atmosphere. Further, a sensor receives the direct reflected or emitted radiation from a target as well as scattered radiations from the target and that from atmosphere, known as *path radiance*. This path radiance induces haze in the imagery and reduces image contrast. The effect of skylight and haze is negatively related to wavelength and positively related to the

distance from a point directly below the sensor. For the removal of the wavelength dependent effects it may be assumed that the longest wavelength recorded is unaffected by the effects of skylight and haze. This waveband then may be used to judge the effects in other wavebands. For off-vertical viewing effects, a correction based on the path length from the ground to the sensor may be used; else all data at the image edge may be omitted. However, the effect of atmosphere upon remotely sensed data is not considered to be an error, since they are part of the signal received by the sensing device.

Sun Elevation Correction The correction accounts for the seasonal position of the sun relative to the earth. The image data acquired under different solar illumination angles are normalised by calculating pixel brightness values assuming the sun was at the zenith on each date of sensing. The correction is applied by dividing each pixel value in a scene by the sine of the solar elevation angle for the particular timing and location of imaging (Fig. 7.5).

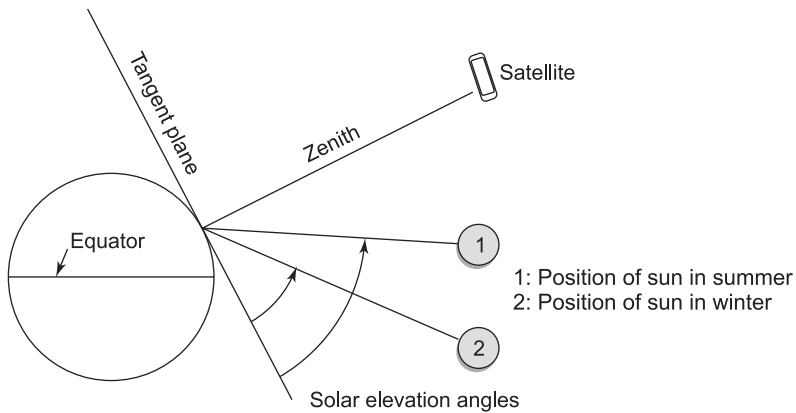


Fig. 7.5 Effects of seasonal change on solar elevation angle

Earth-Sun Distance Correction The correction is applied to normalise for the seasonal changes in the distance between the earth and the sun. The irradiance from the sun decreases as the square of the earth-sun distance. Neglecting the atmospheric effects, the combined influence of solar zenith angle and earth-sun distance on the irradiance incident on the surface of the earth can be expressed as

$$E = \frac{E' \cos \theta}{d^2}$$

where

- E = normalised solar irradiance
- E' = solar irradiance at mean earth-sun distance
- θ = angle of sun from zenith
- d = earth-sun distance, in astronomical units

Note

The solar elevation angle and the earth-sun distance for a given scene are ancillary data supplied with the digital data.

7.1.2 Geometric Corrections

It is the process of rectification of geometric errors introduced in the image during the process of acquisition. The aim is to transform the remotely sensed image to have the scale and projection properties of a map. The geometric distortions in the raw digital images arise from earth's rotation; panoramic distortion, further affected by earth curvature; scan skew; variations of platform's height, velocity, attitude (pitch, roll, and yaw); and aspect ratio distortion. The main source of geometric error in satellite data is satellite path orientation (non-polar). The distortions may be systematic distortions—the effects that are constant, can be predicted in advance; or non-systematic—caused due to variations in spacecraft variables, and atmospheric scatter. The systematic distortions (errors) are corrected by using ephemeris of the platform and the precisely known sensor distortion characteristics. The non-systematic distortions are corrected by matching the image coordinates of the physical features recorded in the image with the geographic coordinates of the same feature from a map or using GPS. The geometric correction process consists in first considering the systematic distortions and then the non-systematic ones.

Systematic Distortions The various types of systematic distortions are as follows.

1. *Scan skew*: It is caused by the forward motion of platform during the time required for each mirror sweep. The ground swath is not normal to the ground track but is slightly skewed, producing cross-scan geometric distortion (Fig. 7.6). The scanned lines do not remain exactly perpendicular to the ground track.

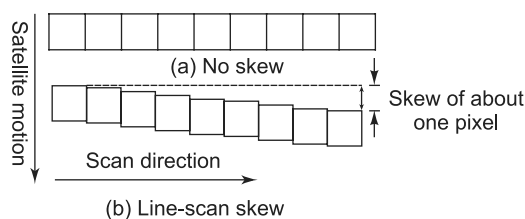


Fig. 7.6 Scan skew

2. *Scanner mirror velocity variance*: The mirror scanning rate is usually not constant across a given scan, producing along-scan geometric distortion. An oscillating mirror must stop at the end of each scan and reverse direction.

3. *Panoramic/scanner distortion*: The ground area imaged is proportional to the scan angle rather than to the angle itself (Fig. 7.7). Because data are sampled at regular intervals, this produces along-scan distortions.

4. *Spacecraft/platform velocity*: If the speed of the platform changes, the ground track covered by the successive mirror scan (IFOV) changes, producing along-track distortions in the form of oversampling—when forward platform velocity decreases (higher orbit), or undersampling—when forward platform velocity increases (lower orbit).

Table 7.1 *Geometric errors and their effects*

<i>Error type</i>	<i>Source</i>	<i>Effects</i>	<i>Direction</i>
<i>Non-systematic</i>			
Altitude	Platform	Deviation from nominal altitude of satellite	Along/across scan
Attitude	Platform	Deviation of sensor axis from normal to earth ellipsoid surface	Along/across scan
<i>Systematic</i>			
Scan skew	Platform	Scanned lines are not exactly perpendicular to ground track	Across scan
Space craft velocity	Platform	Change in along track IFOV (Instantaneous Field of View)	Across scan
Earth rotation	Scene	Westward shift of different scan lines of a scene	Along scan
Map projection	Scene	Geometric error in projecting image on 2D map plane	Along/across scan
Terrain relief	Scene	Relative planimetric error between objects imaged at different heights	Along/across scan
Earth curvature	Scene	Change in image pixel size than actual one	Along/across scan
Optical	Sensor	Barrel and pincushion distortions in image pixels	Along/across scan
Aspect ratio	Sensor	Image pixel size different in horizontal and vertical directions	Along/across scan
Mirror velocity	Sensor	Compression or stretching of image pixels at various points along scan line	Along scan
Detector geometry and scanning sequence	Sensor	Misalignment of different band scan lines of multispectral sensors	Along/across scan
Perspective projection	Scene and sensor	Enlargement and compression of image scene close and far off to nadir point respectively	Along scan
Panoramic	Scene and sensor	Introduces along scan distortions	Along scan

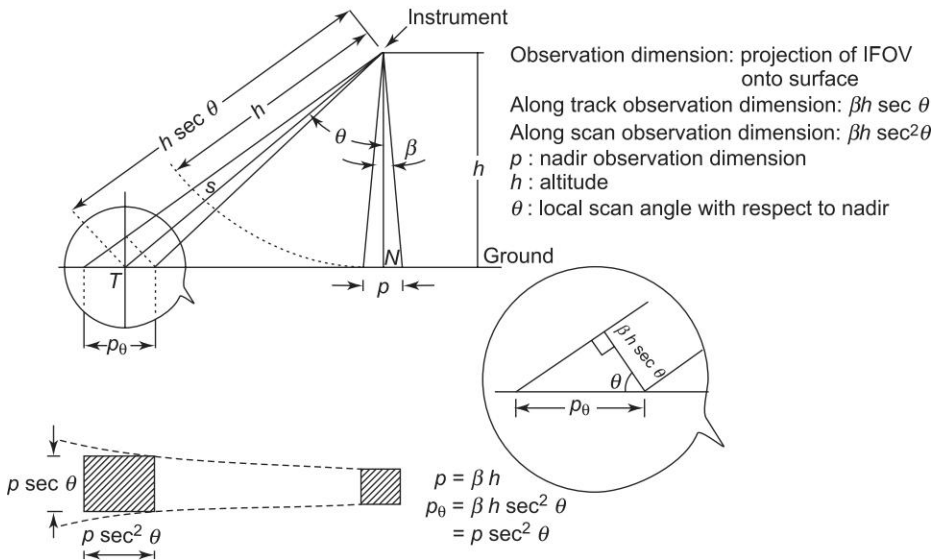


Fig. 7.7 Panoramic distortion

5. *Earth's rotation*: The earth rotates as the sensor scans the terrain. This results in a shift of the ground swath being scanned, causing along-scan distortion. Since between the time of first scan and the time of last scan, the earth rotates eastwards significantly relative to the resolution element, each optical sweep of the scanner covers an area slightly to the west of previous sweep. The amount of earth rotation during this period results in image distortion known as *skew distortion*. This distortion can be corrected by scanning successive groups of lines, offset towards the west, in proportion to the amount of movement of the ground during image acquisition. This results in the parallelogram outline of the restored image (Fig. 7.8(b)).

The systematic distortions are well understood and can be easily corrected by applying formulae derived by mathematically modelling the sources of distortions.

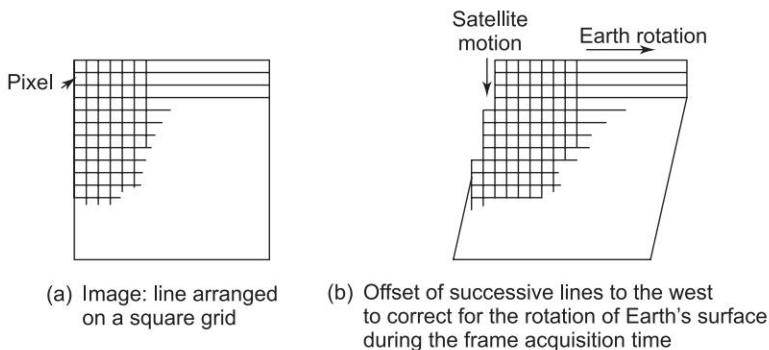


Fig. 7.8 Effect of earth rotation on scanner imagery

Non-symmetric Distortions The non-systematic geometric distortions occur because of followings. Platform motion affects the images with changes in altitude and attitude.

1. *Altitude variance:* If the sensor platform departs from the normal altitude or terrain increases in elevation, change in scale or pixel size occurs. With a change in altitude, the platform motion results in variations in scale (Fig. 7.9).

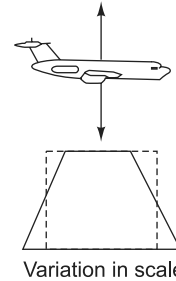


Fig. 7.9 Distortions due to altitude variance

2. *Platform attitude:* One sensor system axis is usually maintained normal to earth ellipsoid surface and other parallel to the spacecraft direction of travel. If the sensor departs from this attitude, the result is geometric distortion. Attitude change implies the change in platform orientation that is significant over the time required to scale a full scene. This is termed as *skew motion*—motion of aircraft/satellite perpendicular to the intended direction of motion. It may be yaw, pitch, or roll as shown in Fig. 7.10. The effect of attitude variations are shown in Fig. 7.11. *Pitch* is the vertical rotation of a sensor platform, in the direction of motion (nose-up plane), resulting in changes in the spacing of scan lines. *Roll* is the rotation of sensor platform around the velocity vector, and scale changes in the line direction resulting in lateral shift in scan line positions. *Yaw* is the rotation of a sensor platform in the horizontal plane, or about its vertical axis, hence in a nose-right direction. It causes rotation and skew distortion. Changes in yaw result in scan lines that are not parallel.

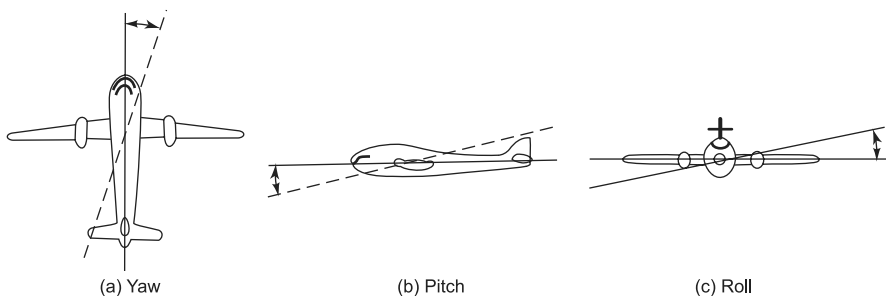


Fig. 7.10 Platform attitude change (skew motion)

Correction The non-systematic variations can be evaluated from tracking data, or from ground control. Also known as *random distortions*, these can be corrected by analysing well-distributed ground control points (GCPs) occurring in an image. Following are the two methods for correcting non-systematic distortions evaluated from ground control. The basic concept underlying is the transformation of satellite images into a standard map projection so that image features can be accurately located on earth's surface. Furthermore, the corrected image can be compared directly with other sources of geographic information, such as maps, etc.

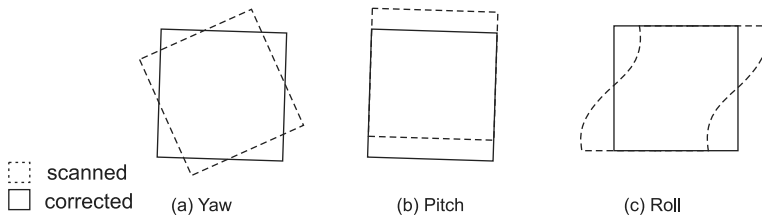


Fig. 7.11 Effect of attitude variance

Georeferencing The geographical features on the digital image, called *ground control points* (GCPs), whose positions are known, are identified. These points of known ground locations can be accurately located on digital image. Examples of GCPs are the intersection of streams, highways, runways, airports, etc. The latitudes and longitudes of GCPs are determined from base maps, if available; else GPS may be used for this purpose. It may be noted that the accuracy of fixation of GCPs has a direct bearing over the extent of rectification affected. GCPs should be reliably matched between source and reference, and should be widely dispersed throughout the source image.

The location of the output pixels are derived from the GCPs. Then these are used to establish the geometry of the output image and its relationship to the input image. The difference between the actual GCPs locations and their respective positions in the image are the required geometric transformations to restore the image.

Coordinate Transformation In this method the input pixels are rearranged on a new grid. The image correction is carried out using polynomial equations converting the source coordinates to rectified coordinates; the order of polynomials is decided depending upon the extent of geometric distortion. Usually, it is carried out with the help of 1st order and 2nd order transformations. For accuracy the number of control points must be more than the number of unknown parameters in polynomial equations. The accuracy should be within ± 1 pixel.

After carrying out the coordinate transformation of the image, a process called *resampling*, or *intensity interpolation* is used to determine the pixel values of the transformed image. Image resampling involves the reformation of an image on to a new grid. This is achieved by using features (GCPs) that are common to both the image and the new grid. Suppose that an image with pixel coordinates (x, y) undergoes geometric distortion. The process consists in first defining a geometrically correct geographical grid in terms of latitude and longitude. With computer the latitude and longitude values of each cell of the grid is transformed into x and y coordinates, which becomes the new address of an image pixel. The computer program scans this new address in the image and transfers the appropriate DN based on nearby DN values in original image to this address. The process is repeated until the geographical grid is full at which point the image has been geometrically corrected (Fig. 7.12).

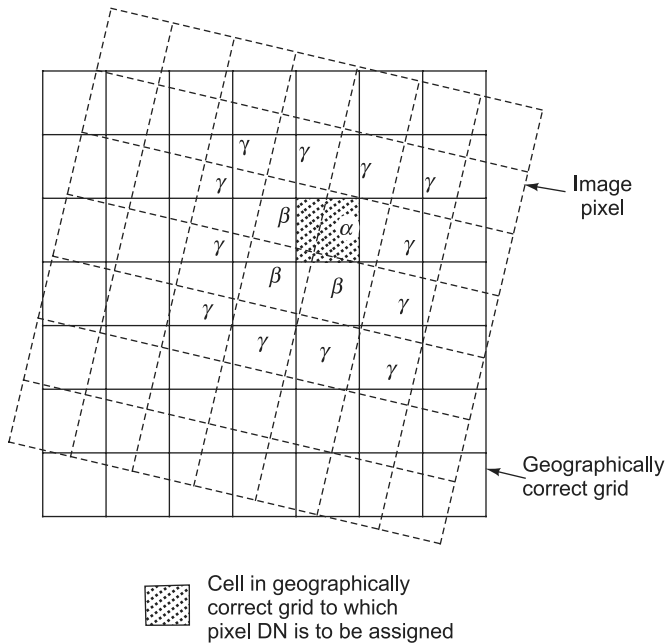


Fig. 7.12 Resampling procedure

There are generally three types of resampling methods, viz., the nearest neighbour method, the bilinear interpolation method and the cubic convolution method, to assign the appropriate digital number to an output cell or pixel. These methods are described as follows.

1. *Nearest neighbour method:* The method is also known as zero-order interpolation method. It consists in assigning each corrected pixel, the value of the nearest uncorrected pixel. The method is simple in application and has the advantage of preserving the original values in the altered scene. However it may introduce noticeable errors, especially in linear features, where the realignment of pixels is obvious. Some of the other disadvantages of the method are blocky picture appearance and spatial shifts. The effects although are negligible for most visual applications, but may be important for subsequent numerical analysis.

2. *Bilinear interpolation method:* In this method, the calculation of each output pixel value is based on a weighted average of values within a neighbourhood of (2×2) four adjacent input pixels. This process is simply the two dimensional equivalent to linear (first-order) interpolation. Since each output value is based on several input values, the output image is much smoother than nearest neighbour method. However, some changes such as loss of brightness values in the input image, a reduction in the spatial resolution of the image, and blurring of sharp boundaries in the picture do occur when bilinear interpolation creates a new pixel value.

3. *Cubic convolution method*: This method is also known as bi-cubic convolution or second-order interpolation method. It is supposed to be the most sophisticated and complex method of resampling. It uses a weighted average of values within a neighbourhood of (4×4) 16 adjacent input pixels. Though, the images produced are noticeably sharper than the previous two methods, but get drastically altered. This method, however, introduces some loss of high frequency information.

7.2 IMAGE ENHANCEMENT

Even after applying the radiometric corrections an image might not be visually interpretable and may require further treatment to enhance the image. The process consists in converting the image quality to a better understandable level for feature extraction or subsequent visual image interpretation; increasing the visual distinctions between the features in a scene is known as *image enhancement*. Thus, the main object of image enhancement is to make the satellite images more informative and suitable for a specific application. This goal can be achieved by *point application*, also referred to as *radiometric enhancement*—changing the values of individual pixels independent of the other pixels; and *local operations*, also referred to as *spectral enhancement*—changing the value of individual pixels with regards to the neighbouring pixel values. The image enhancing techniques are applied either to single-band image, or separately to the individual bands of a multiband image set, and are accordingly classed as *spectral enhancement techniques* and *multi-spectral enhancement techniques*. There are numerous image enhancement techniques, some of the most important are the image reduction, image enlargement, colour compositing, transect extraction, contrast enhancement, and digital filtering. These are described briefly as follows.

7.2.1 Image Reduction

Since the display monitor generally consists of 1024×768 screen resolution only, the computer display systems are unable to feature a full image at the normal image pixel scale (e.g., higher than 2500 rows and 2500 columns). Now, since the entire image cannot be viewed, location of exact coordinates of the area of interest (AOI) may not be possible. In such a case the original image data set is reduced to a smaller data set, and the analyst can view the AOI at one time on the screen by deleting systematically selected rows and columns. To accomplish this, n^{th} row and n^{th} column of the image may be selected and displayed, where n may be suitably chosen say 2, 3, 4,... etc. With a smaller n value, e.g., 2, there is a reduction of 25% pixels in the scene and it might not be enough to view the AOI defeating the effort. While a very large value of n , e.g., 10, will cause the image to retain only 1 % of the data, small enough to view the entire scene on the screen, but will not retain enough data for processing and interpretation. Therefore, a judicial selection of n value is an important key of the technique. Figure 7.13 shows an example of image reduction with $n = 2$.

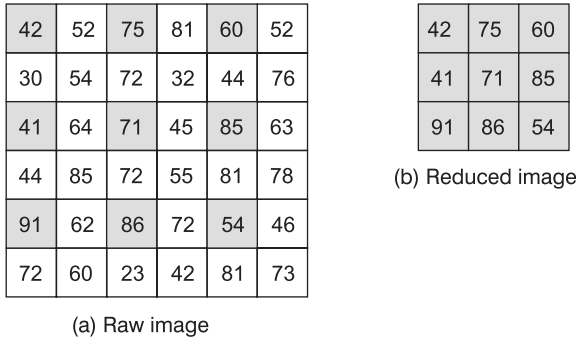


Fig. 7.13 Image reduction

7.2.2 Image Magnification

The process is also known as *zooming*. It is used to modify the scale of the image to improve the visual interpretation and to match the scale of another image. To accomplish this, an image is magnified by replication of rows and columns. For example, by using an integer factor $n = 2$ each pixel in the original image is replaced by $n \times n$ block of pixels as shown in Fig. 7.14; this will double the pixel size value. This technique is very useful when spectral reflectance or emitance of a very small area is of interest. Figure 7.14 shows an example of image enhancement with $n = 2$.

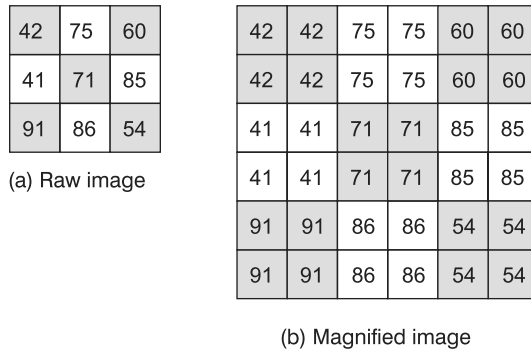


Fig. 7.14 Image magnification

7.2.3 Colour Compositing

Coloured images can be interpreted better than the black and white images. Images are usually coloured by pseudocolour on single images and normal or false colour on several images as discussed in Section 6.10.3 (Chapter 6). The pseudocolour involves replacement of each grey level in the image with a colour which separates the small grey scale differences that can not be distinguished by normal human eye.

Digital images are typically displayed as additive colour composite using the three primary colours red, green and blue denoted as RGB. The RGB colour cube (Fig. 7.15) is defined by the brightness level of each of the three primary colours. For a display with 8-bit-per-pixel data encoding (a common sample format), the range of possible DN's column component is 0 to 255. Hence, there are 256^3 (16,777,216) possible combinations of primary colours. The line from origin of the cube to the opposite corner is known as *grey line*, since DN's which lie on the line have equal components of primary colours. In DIP the RGB displays are used extensively to display normal colour, false colour infrared, and arbitrary colour composites. The procedure consists in contrast stretching and compositing the images and then colouring each waveband with a primary colour.

Colour compositing is a very useful technique to enhance visual interpretation. Combinations other than true colour combination produce false colour combinations resulting in identification of the features that were not visible in true colour composites. It is so because reflectance value of an object differs from band to band. For example, invisible infrared band is visible by obtaining red colour tones highlighting some invisible features.

For digital data three values corresponding to red, green and blue make various colour combinations. As an example, for a multispectral image of six bands, the number of colour composites that can be obtained in a combination of these would be

$$\frac{6!}{(6-3)!} = \frac{6!}{3!} = \frac{1 \times 2 \times 3 \times 4 \times 5 \times 6}{1 \times 2 \times 3} = 120$$

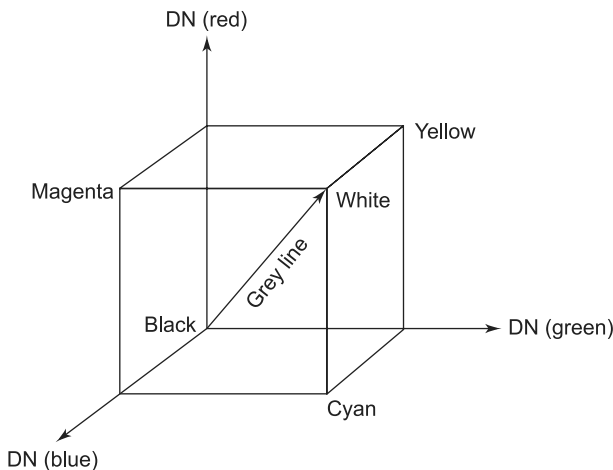


Fig. 7.15 The RGB colour cube

7.3 CONTRAST ENHANCEMENT

This is also referred to as *contrast manipulation*, or *radiometric enhancement*. In a raw image the useful data often populates only a small portion of the available

range of digital values (e.g., 8 bits or 256 levels). Materials on the earth's surface reflect and emit different levels of energy. A sensor might record a considerable amount of energy from any material in a certain wavelength, whereas another material is recorded at a much less energy on the same wavelength. The ranges of brightness values present on an image are referred as *contrast*.

Contrast enhancement is a process that makes the image features more clearly by making optimal use of the colours available on the display or output device. It involves changing the range of values in an image in order to increase the contrast. For example, an image having a range of brightness value between 35 and 80 when stretched to a range of 0 to 255 enhances the differences between the features. Before describing the two techniques (the linear technique and the non-linear technique) for enhancing the contrast of an image, it will be pertinent to discuss the key element of the enhancement techniques, the *histogram*.

Histogram is a graphical representation of brightness values that comprises an image. It may also be defined as a statistical graphic representation of an image of tones from dark to light and associated number of pixels for each tone of an image (Fig. 7.16(a)). For a single band of data, the histogram contains the numbers of all possible brightness values along their horizontal axis and the associated number of pixel values along the vertical axis. A broad histogram reflects a scene with significant contrast, while a narrow histogram may appear dull or flat. Contrast as defined above is measure of the differences between light and dark areas in a scene.

7.3.1 Linear Methods

The linear contrast enhancement technique is also referred to as *contrast stretching* technique. In this technique the original digital values of the remotely sensed data are linearly expanded into a new distribution. The method is most suited to Gaussian histogram, implying that all brightness values fall within a narrow range of the histogram. A linear stretch is used when equal weight is to be given to all DN regardless of their frequency of occurrence. The drawback of the linear stretch is that it assigns as many display levels to the rarely occurring image values as it does to the frequently occurring values. To improve upon the situation, a histogram-equalised stretch may be applied, where the image values are assigned to the display levels on the basis of their frequency of occurrence. Following are some of the methods which fall in this category, and are shown in Fig. 7.16.

Minimum-Maximum Stretch The technique utilises the stretch between minimum and maximum DN values, i.e., the range of the image values present in the scheme is expanded to fill the range of display values. Therefore, in this technique the original minimum and maximum values of the data are assigned a specified set of values to utilise full range of available brightness values of the display unit (Fig. 7.16(b)). The uniform expansion is called a *linear stretch*. The old intermediate values of the data are scaled proportionately between the new minimum and maximum values. This introduces appreciable variations in input image data values and the effect is reflected in output tones that can be more readily distinguishable by the interpreter.

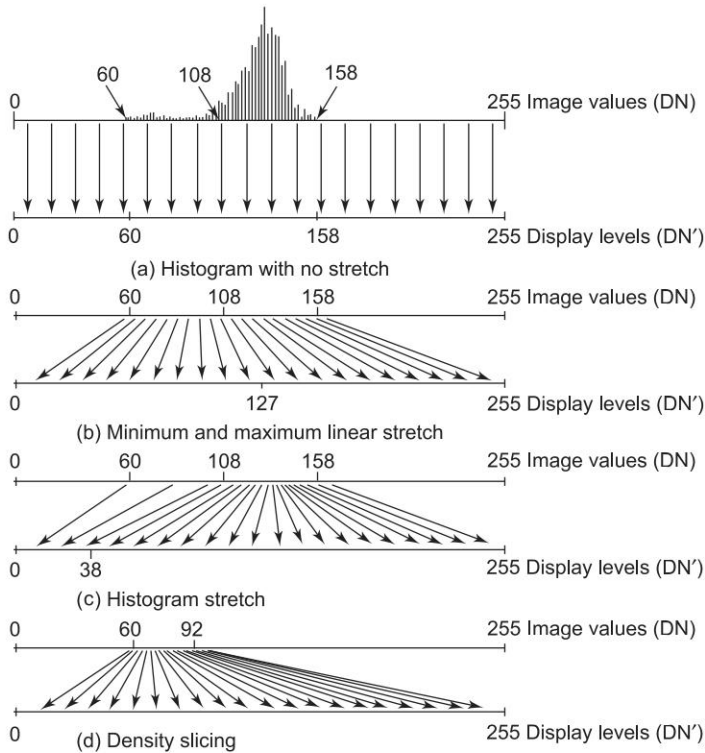


Fig. 7.16 Principle of contrast enhancement

Saturation Stretch This method is also called *percentage linear stretch* or *full trim*. The method is same as the above method, but with a difference that it utilises certain specified minimum and maximum values that lie in a certain percentage of pixels. Usually, the limits are 5th to 95th percentile on a histogram (Fig. 7.17).

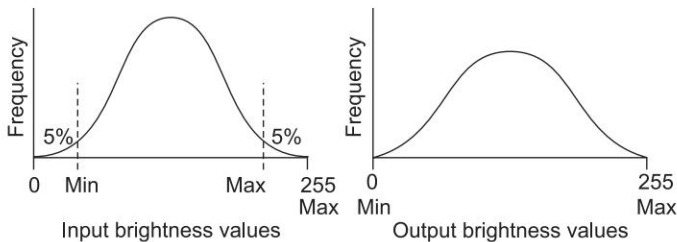


Fig. 7.17 Contrast enhancement: Saturation stretch

Average and Standard Deviation Stretch This method is same as the percentage linear method with a difference that a standard deviation from the mean is used to push the tails of the histogram beyond the original minimum and

maximum values. For example, it may be $\pm 2\%$ of the standard deviation from the mean of the histogram.

Piecewise Stretch For bi- or tri-linear modal histogram of an image, certain values of histogram may be stretched to increase enhancement of brightness values in selected regions.

7.3.2 Non-linear Methods

For asymmetric image-histogram the original histogram is converted to another histogram. However, the method suffers with the drawback that the objects lose their correct relative brightness values. Some of the techniques in use are as follows.

Histogram Equalisation In this method, histogram of any shape is changed to histogram having same frequency along the whole range of DN. It is assumed that each level in the displayed image contains an approximately equal number of pixel values, so that the histogram of these displayed values is almost uniform (though not all the 256 classes are necessarily occupied). A histogram equalisation stretch (Fig. 7.16(c)) is used when it is felt necessary to weight the DN by their frequency of occurrence. Contrast increases at the most populated range of brightness values of histogram. Although the resultant image would have greater overall contrast but the smaller features are distinguished with difficulty. The steps involved are as follows.

1. The pixel values of an image are assigned a certain number of bins, which are simply numbered sets of pixels.
2. The pixels are then given new values based upon the bins to which they are assigned.
3. The total number of pixels in the image is divided by the number of bins, equaling the number of pixels per bin
 $A = T/N$
 A = the equalised number of pixels per bin
 T = the total number of pixels in the image
 N = the number of bins
4. The pixels of each input value are assigned to bins, so that the number of pixels in each bin is as close to A as possible.
5. To assign pixels to output bin the following equalisation may be used:

$$B_i = \text{int} \left[\frac{\left(\sum_{k=0}^{i-1} H_k \right) + \frac{H_i}{2}}{A} \right]$$

where B_i is the output bin number, H_k is number of pixels per bin in input image, H_i is the number of pixels with the value (input), A is equalised number of pixels per bin, int is the integer function (truncating real numbers to integer, e.g., int (2.7) = 2), i = a particular brightness value (considered as input), and k = a particular bin number.

Histogram Normalisation Normal distribution, also known as *Gaussian distribution*, is a bell shaped distribution with most values at or near middle of the curve and rare values near the tails (Fig. 7.18). The method based upon the histogram of the pixel value is also called *Gaussian stretch*, because it consists in conversion of the observed histogram to a Gaussian histogram. Generally, a normal distribution of the density in an image creates a natural image. The histogram of the original image may, therefore, be normalised by fitting of the observed histogram to the Gaussian histogram. To achieve this transformation, the pixels with same grey scales are reallocated to other pixels with a different grey scales. This technique is most suited to unfocussed images with low dynamic range.

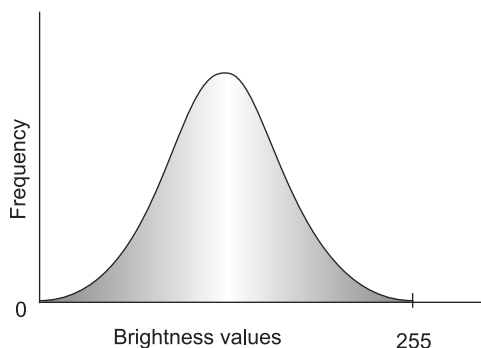


Fig. 7.18 Contrast enhancement: Histogram normalization

Reference Stretch The technique is also known as *histogram matching* or *histogram specification*. It consists in determining a look up table that converts the histogram of the image to resemble the histogram of another. The technique is very useful for matching the data of the same scene or adjacent scenes that were scanned on different days, or are different because of the sun angle or atmospheric effects.

Density Slicing The method, also called *level slicing*, is best suited to single band images as well as for a surface feature having a unique or narrow set of DN values. In this technique the DNs distributed along the x-axis of an image histogram are divided into a series of specified intervals or *slices*. All the DNs falling within a given interval in the input image are then displayed at a single DN in the output image. Consequently, if eight different slices are established, the output image contains only eight different grey levels. The output looks like a contour map, except that the areas between boundaries are occupied by pixels displayed at the same DN. Each level can also be shown as a single colour. Figure 7.16(d) shows a density slice.

Grey-level Thresholding This technique segments the image DNs into two distinct values (black = 0, white = 255) separated by a threshold DN as shown in Fig. 7.19.

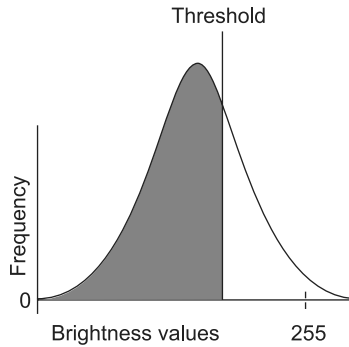


Fig. 7.19 Thresholding on a histogram

7.4 FILTERING

Spatial filtering can be described as selectively highlighting or suppressing specific features in an image based on their spatial frequency. The term spatial frequency is related to the image texture and refers to the frequency of the roughness of the tonal variations occurring in an image. Image areas of high spatial frequency are tonally rough; the grey levels in these areas change abruptly over a relatively small number of pixels, e.g., for roads, boundaries, etc. The normal row-column coordinate system—the (x, y) coordinate space of image, in which images are expressed as spatial location, is called *spatial domain* (changes in pixel positions correspond to changes in the scene). Filtering techniques can be implemented through the Fourier transform in the frequency domain (changes in image positions correspond to changes in the spatial frequency) by reconstruction with reconstruction filters or in the spatial domain by convolution (enhancement) filters.

7.4.1 Convolution Filtering

This technique is based upon the transformation of the image into its scale or spatial frequency components using Fourier transform. In this method a window, known as *convolution kernel*, of a set of pixels in dimension (3×3 , 5×5 , etc.), is moved over each pixel in the image, applying a mathematical calculation using the pixel values under that window, and replacing the central pixel with the new pixel. The kernel is moved along the row and column dimensions with one pixel at a time until the entire image is filtered and a new image is generated. The convolution filters are classed as low pass or high pass filters.

Low Pass Filter Also known as *smoothing* or *defocusing filter* is used for large portion of similar tone and to reduce smaller details in an image; they serve to smooth the appearance of an image. It enhances low spatial frequencies. Low pass filters may be of average, or median, or mode type depending upon the respective statistical parameter used.

1. *Averaging filter*: A typical low pass averaging kernel is shown in Fig. 7.20 (a) and Fig. 7.20 (b) shows the original data of a digital image. The output DN is found by the sum of the products of corresponding convolution kernel and image

elements. The smoothening effect of averaging filter ($(1/9) (4 + 12 + 10 + 10 + 3 + 9 + 7 + 6 + 13) = 8.22 \approx 8$) is as shown in Fig. 7.20(c).

2. *Median filter*: This filter operates by rearranging all DN values in sequential order within the defined window. The pixel of interest is replaced by the value in the centre of this distribution, the *median*. This filter is used for removing pulse or spike noise. When the values of digital image data of Fig. 7.20(d) are arranged in ascending order (3, 4, 5, 7, 8, 9, 10, 11, 12) the median will be 8 and the smoothening effect is as shown in Fig. 7.20(e).

3. *Mode filter*: This filter computes the mode of the grey level values (the most frequently occurring grey level value) within the filter window surrounding each pixel. As an example, the original digital data of an image is given in Fig. 7.20(f). When arranged as 8, 8, 9, 9, 9, 9, 11, 11, 11 the data results in a value nine that occurs maximum four times. It is the mode value and will replace the central value (11) of Fig. 7.20(f) as shown in Fig. 7.20(g).

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

(a) Typical low pass kernel

4	12	10
10	3	9
7	6	13

(b) The original data

4	12	10
10	8	9
7	6	13

(c) Smoothening effect of low pass filter

4	5	10
11	3	8
9	7	12

(d) The original data

4	5	10
11	8	8
9	7	12

(e) The enhanced image (the central pixel value 8 is the median of the original data)

11	9	9
9	11	8
8	9	11

(f) The original data

11	9	9
9	9	8
8	9	11

(g) The enhanced image (the central pixel 9 is the mode of original data)

-1	-1	-1
-1	5	-1
-1	-1	-1

(h) Typical high pass kernel

4	12	10
10	22	9
7	6	13

(i) The original data

4	12	10
10	39	9
7	6	13

(j) Smoothening effect high pass filter

Fig. 7.20 Convolution filtering

High Pass Filter These filters serve to sharpen the appearance of fine detail in an image by enhancing high spatial frequencies. It increases the spatial frequency by subtracting the low frequency image (resulting from a low pass filter) from the original image. Sharper and deblurred image can be obtained by adding back to the image some multiple of the high frequency components, extracted by the filter amplifying the high frequency detail. Figure 7.20 (h, i and j) shows the enhancement effect of a high pass filter. A typical high pass filter is shown in Fig. 7.20(h). The sharpening effect of high pass filter is achieved by replacing the original middle pixel value (22) by (39).

$$\left[\begin{array}{c} (-1 \times 4) + (-1 \times 12) + (-1 \times 10) + (-1 \times 10) + (+5 \times 22) + \\ (-1 \times 9) + (-1 \times 7) + (-1 \times 6) + (-1 \times 13) = 39 \end{array} \right]$$

Edge Detection Filter These are either high or low pass filters used to highlight linear features such as roads, canals, field boundaries, etc., or for features that are oriented in specific directions. Most common edge detection kernel is zero-sum kernel; the sum of all coefficients in the kernel equals zero. Sum of the coefficients is not used in the convolution equation, since division by zero is not defined. This generally makes the output values to be zero in areas where all input values are equal. The resulting image often contains only edges and zeros. The edge enhanced images attempt to preserve both local contrast and low frequency brightness information.

Adaptive Filter This filter has kernel coefficients calculated for each window position based on the mean and variance of the original DN in the underlying image.

7.4.2 Frequency Domain Filtering

The filtering discussed in previous subsection is in the spatial domain. Frequency domain is an alternative to the coordinate space that can be used for image analysis purposes. In this approach, an image is separated into its various spatial frequency components or scale components through the Fourier transform. These filters are used to remove, attenuate or amplify the amplitudes in specified wavebands. Conceptually, this operation amounts to fitting a continuous function through the discrete DN values if they were plotted along each row and column in the image. After an image is transformed into discrete component spatial frequencies, it is possible to display these values in a two-dimensional scatter plot known as a *Fourier spectrum*. The filtering operation is carried out in following steps:

1. The original image is transformed to its frequency representation using Fourier transform.
2. Image processing—selecting an appropriate filter transfer function and multiplying it to the elements of the Fourier spectrum.
3. The inverse Fourier transform is performed to return to the spatial domain for display purposes.

7.5 IMAGE TRANSFORMATION

This involves manipulation of multiple bands of data from image(s) of the same area acquired at different times (multitemporal image data), or from images of the same area acquired at different spatial resolution (multiresolution). The aim is to generate new images from two or more sources so as to highlight particular features or properties of interest better than the original input images. There are various methods of image transformation, e.g., image arithmetic operation, Fourier transformation, image fusion, etc. The image arithmetic operations are described as follows. The others involve complex mathematical techniques and are beyond the scope of this book.

Image Arithmetic Operations The operations of addition, subtraction, multiplication, and division are performed on two or more co-registered images of the same geographic area. These images may be separate spectral bands from a single multispectral data set or they may be individual bands from image data sets collected at different dates.

1. *Image addition:* For multiple, co-registered images of a given region for the same time and date of imaging, the new DN values of a pixel in the output image can be determined by averaging the DN values of corresponding pixels of input images (Fig. 7.21(a)). Image averaging finds application in the field of astronomy where low light levels cause sensor noise rendering poor image quality.

2. *Image subtraction:* This operation is carried out on co-registered scenes of the same area acquired at different times or dates, for change detection. This involves simple subtraction operation on the pixel values of the two images which have been geometrically registered (Fig. 7.21(b)). Image subtraction finds application in medical imaging called *mask mode radiography*.

3. *Image multiplication:* This operation is performed rarely. The best use of the method is made when an image of interest is composed of many distinctive regions, but the analyst is interested in any one region only (Fig. 7. 21(c)).

4. *Band ratioing:* This technique is also known as *spectral ratioing*. By division of DN values of one spectral band by corresponding DN value of another band, the resultant image enhances variations in the shapes of the spectral reflectance curves between the two different spectral ranges that may otherwise be masked by the pixel brightness in each of the bands. A major advantage of ratio images is that they convey the spectral or colour characteristics of image features, regardless of variations in scene illumination conditions. Hence, a ratioed image of the scene effectively compensates for the brightness variation caused by the varying topography and emphasises the colour content of the data. This is the most common arithmetic operation used in geological, ecological, and agricultural applications.

7	11	51	16
54	48	32	21
52	70	22	43
58	62	60	33

+

51	60	21	62
43	62	83	73
32	24	32	60
60	88	21	49

=

29	36	36	39
49	55	58	47
42	47	27	52
59	75	41	41

Example: 1st Pixel $(7 + 51)/2 = 29$

(a) Image addition

22	25	24	26
23	26	27	28
26	27	29	25
28	30	26	28

-

27	26	23	22
23	25	28	29
26	30	29	23
28	26	22	27

=

-5	-1	1	4
0	1	-1	-1
0	-3	0	2
0	4	4	1

Example: 1st Pixel $22 - 27 = -5$

(b) Image subtraction

7	8	5	7
5	6	8	7
3	7	6	5
2	4	3	5

×

7	8	5	7
5	6	8	7
3	7	6	5
2	4	3	5

=

49	64	25	49
25	36	64	49
9	49	36	25
4	16	9	25

Example: 1st Pixel $7 \times 7 = 49$

(c) Image multiplication

Fig. 7.21 Image arithmetic operations

7.6 IMAGE CLASSIFICATION

A major task after feature extraction is to classify the object into one of several categories. Classification of remotely sensed data is achieved by assigning levels with respect to groups with homogeneous characteristics so that multiple objects in a scene can be discriminated. The level is called *class*. Classification may, therefore, be defined as the process of assigning the pixels in an image into a finite number of individual classes based on their DN values. The classification is usually based on the patterns of their DN, spatial relationship with neighbouring pixels, and relationships between data acquired on different dates. The term *pattern* is not geometric in character; rather it refers to the set of radiance measurements obtained in the various wavelength bands for each pixel. The objectives of image classification are to detect different kinds of features in an image, discriminate the distinctive shapes and patterns, and to identify temporal changes in the image. Classification transforms the image data into information. Classification is considered as the most important technique to extract information from the digital images.

Spectral information is represented by digital numbers in spectral bands. Digital image classification attempts to classify each individual pixel based on this spectral information. The object is to assign all pixels in the image to particular classes or themes, e.g., water, wheat, etc. The resulting classified image will comprise of an array of pixels, each of which belongs to a particular theme. Classification generally comprises *image classification scheme*—usually classes such as agriculture, forest, etc., *image processing*—geometric correction, atmospheric correction, noise suppression, image enhancement, etc., *training signatures*—selection of the particular features which best describe the pattern; *decision*—choice of suitable method for comparing the image patterns with the target patterns; and *accuracy assessment*—comparing classification results with the field studies.

Different landcover types in an image can be discriminated using some image classification algorithms, using spectral features, i.e., the brightness and colour information contained in each pixel. The procedure can be supervised or unsupervised, depending upon whether or not a set of prototype is available. In the *supervised classification* system, also called *supervised learning*, each pixel is supervised for the categorisation of the data by specifying to the computer algorithm, numerical descriptors of various class types. The steps involved in typical supervised classification are *training stage*—identification of the training areas (representative sample sites of known landcover type) and development of a numerical description of the spectral attributes of the class or landcover type; *classification stage*—categorisation of each pixel, in the whole image, into landcover class to which it closely resembles; and *numerical mathematical approaches*—for application to spectral pattern recognition to form the class signature. These are shown in Fig. 7.22.

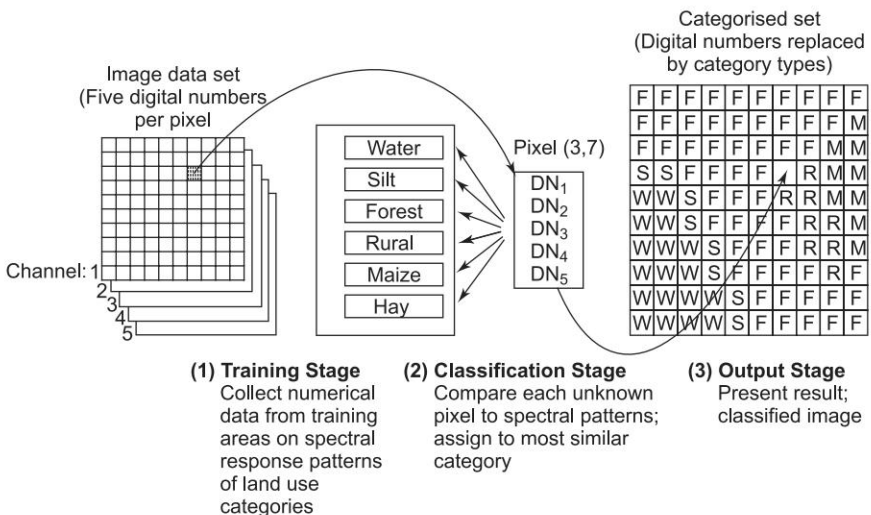


Fig. 7.22 Steps in supervised classification

The *unsupervised classification* (*unsupervised learning*) system does not utilise the training data as the basis of classification. It involves algorithms that examine

the unknown pixels in the image, and aggregate them into a number of classes based on the natural groupings (cluster) present in the image. The classes that result from this type of classification are spectral classes. Because they are based solely on the natural groupings in the image values, the identity of the spectral classes will not be initially known. Unsupervised classification consists in identification, labelling and mapping of these natural classes. The analyst then assigns a landcover type to each cluster. This method is used when less information about the data is available before classification.

Note *Each class of landcover is referred to as a theme and the product of classification as thematicmap.*

The utility of image classification depends on the requirement of the end user. The output product may contain the interpreted information in virtually unlimited forms; the commonly used forms are graphic, tabular, and digital information files.

7.7 SELECTIVE DIP SOFTWARE

Digital image processing consists of a number of steps that vary from image to image depending upon the format and initial condition of the image, the information of interest, the composition of scene elements, etc. DIP software requires specialised programs tailored for manipulation of processing for digital images. The available software have unlimited capabilities of manipulation and display of the output in desired format. They consist of predefined algorithms in executable command modules. Some of the software are ERDAS Imagine, ENVI, ELAS, GRASS, IDIMS, ERIPS, EMIPS, and EDRISI. These commercial software are mainly menu-driven, interactive, and can be used to their full capabilities. However, since there can be a number of different combination of techniques to affect the process, the software must satisfy the needs of the analyst. Further, the analyst has to be careful while selecting and using software for a particular application; his judgment in this concern is of utmost importance as this will influence the accuracy of the output, the efficiency of processing, and the cost involved.

SOLVED EXAMPLES

Example 7.1 Apply a suitable kernel to sharpen the image with pixels as given below:

204	200	196
201	107	210
198	202	211

Solution Since for the given image, the middle pixel with given DN value is surrounded by pixels of relatively higher DN values. The image can be sharpened by decreasing the pixel value of middle pixel. This can be achieved by using a

(3×3) kernel as shown in Fig. 7.23, and applying the high pass filter so that the middle pixel value is transformed from 107 to 90.

Note

It may be noted that a kernel with any other middle coefficient may also be selected so as to transform the middle pixel value lower than the original value. It depends on the experience and skill of the analyst and the image quality desired.

-1	-1	-1
-1	16	-1
-1	-1	-1

Fig. 7.23

$$204 \times (-1) + 200 \times (-1) + 196 \times (-1) + 201 \times (-1) + 107 \times 16 + 210 \times (-1) + 198 \times (-1) + 202 \times (-1) + 211 \times (-1) = 90$$

Example 7.2 Develop the image histogram of pixel values of an image given below. Also, suggest a method for improving the brightness values of the image.

51	123	212	130
42	32	120	123
130	51	51	120
110	110	120	123

Solution Image histogram is plotted between pixel values on x-axis and their frequency on y-axis (Fig. 7.24). Since number of pixels for different pixel value is not distributed uniformly, the image can be improved by contrast improvement through histogram equalisation.

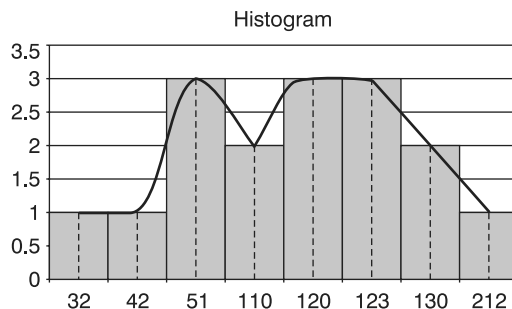


Fig. 7.24

Example 7.3 Apply histogram equalisation to a satellite image whose details are given below.

<i>Pixel value</i>	<i>Number of pixels</i>
0	5
1	4
2	10
3	14
4	61
5	59
6	41
7	29
8	11
9	6

Solution The histogram of the input data is shown in Fig. 7.25.

Total number of pixels in image $T = 240$

Number of bins $N = 10$ (for pixel value for range 0–9)

Equalised number of pixels per bin $A = T/N = 24$.

Output bin B_i for histogram equalisation is given by

$$B_i = \text{int} \left[\frac{\left(\sum_{K=1}^{i-1} H_K \right) + \frac{H_i}{2}}{A} \right]$$

The calculations for B_i are given in the following table.

<i>Input bin</i>	<i>Number of pixels H_K</i>	$\sum_{K=1}^i H_K$	$\sum_{K=1}^{i-1} H_K$	<i>Output bin value (B_i)</i>
0	5	5	0	0
1	4	9	5	0
2	10	19	9	0
3	14	33	19	1
4	61	94	33	2
5	59	153	94	5

6	41	194	153	7
7	29	223	194	8
8	11	234	223	9
9	6	240	234	9

After applying histogram equalisation, output pixels and their corresponding count is given below. The equalised histogram is shown in Fig. 7.26.

<i>Output pixel value</i>	<i>Number of pixels</i>	<i>Grouped input pixel values</i>
0	19	0, 1, 2
1	14	3
2	61	4
3	0	—
4	0	—
5	59	5
6	0	—
7	41	6
8	29	7
9	17	8, 9

Further after applying histogram equalisation, input ranges of pixels 3–7 is stretched to the range of 1 to 8.

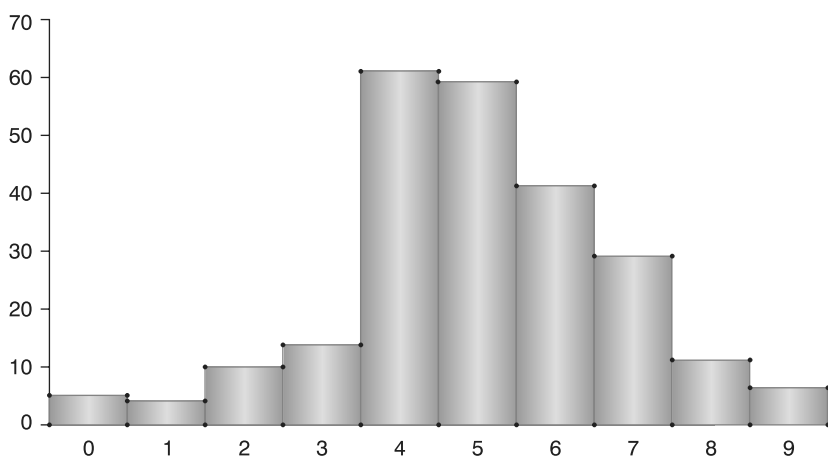


Fig. 7.25 *Input histogram*

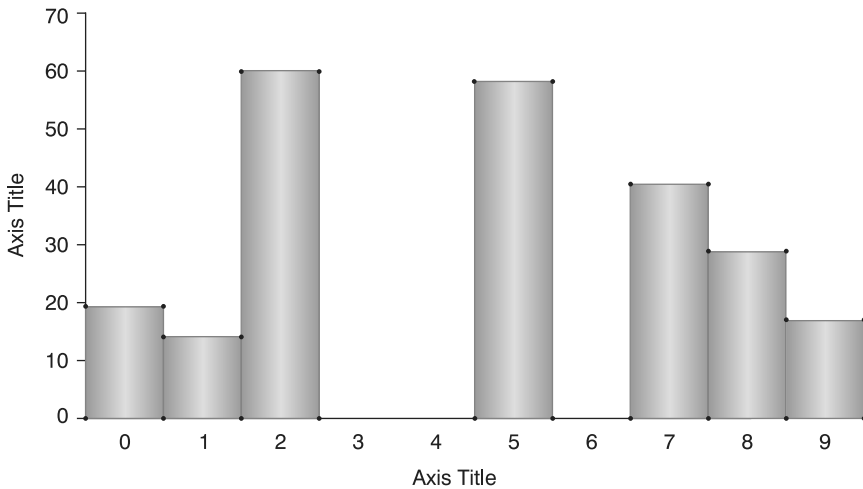


Fig. 7.26 Output histogram (after equalisation)

Summary

A brief and general introduction to digital image processing is presented in this chapter. Although the subject is extremely broad and the procedures involve complex mathematics, the matter presented provides a foundation to the readers interested in taking a full course in DIP. The underlying principles and fundamentals have been outlined to form a base for those who wish to get into depth of the subject and make it a profession.

There are two methods of image processing, analogue and digital. Analogue or visual techniques of image processing are used for hard copies like printouts, photographs etc. Such analysis is a combination of knowledge of the analyst and collateral data to image processing. Digital processing techniques help in manipulation of the digital images by computers. Raw data from imaging sensors from satellite platform contain deficiencies. The raw digital images carrying inherent flaws and deficiencies need processing to get originality of information. The steps involved in the processing are image rectification and restoration (preprocessing), image enhancement, image transformation, image classification, and object recognition.

Prior to data analysis, it is usual to carry out initial processing on the raw data. Preprocessing refers to removal or minimisation of known degradation in an image due to limitations of a sensor, non-linearities due to sensors, noise present in the data, calibration of data radiometrically, and corrections of geometric distortions due to earth's rotation and other imaging conditions.

In order to carry out visual interpretation, visual appearance of the object in the image can be improved by image enhancement techniques. The goal of image enhancement is to accentuate certain features for subsequent analysis or for better clarity of the image scene. The techniques consist in enhancing contrast, colour compositing, noise filtering, image magnification, etc. The image enhancement procedures play a vital role in feature extraction, image analysis, and visual information display. The process emphasises certain specified image characteristics without increasing the inherent information in the data.

Arithmetic operators, statistics or Fourier transforms are used to generate a new image, once the digital image is enhanced using filtering. The image is analysed by making

quantitative measurements from the image to produce a description of it. The technique requires extraction of certain features that aid in the identification of the object. Quantitative measurements of features allow classification and description of images. The chapter ends with an introduction of the software available for processing and analysis of remotely sensed data.

Exercises

- 7.1 What is a digital image? Describe the functions of elements of digital image processing system.
- 7.2 What are the various steps involved in digital image processing?
- 7.3 (a) Specify the elements of digital image processing.
(b) Specify the objective of image enhancement technique.
- 7.4 What is image acquisition in digital image processing?
- 7.5 Define digital image, pixel, brightness, grey level.
- 7.6 Write short notes on the following:
 - (a) Categories of image enhancement
 - (b) Histogram equalisation technique
 - (c) Density slicing
 - (d) Image subtraction
- 7.7 Define the following:
 - (a) Noise
 - (b) Histogram
 - (c) Restoration
 - (d) Kernel
- 7.8 Differentiate between
 - (a) linear and non-linear contrast enhancement methods
 - (b) systematic and non systematic distortions
 - (c) image and digital image
 - (d) georeferencing and coordinate transformation
- 7.9 Differentiate between
 - (a) spatial domain and frequency domain filtering
 - (b) supervised and unsupervised classifications
 - (c) histogram equalisation and normalisation
 - (d) low pass and high pass filters
- 7.10 Briefly describe the following:
 - (a) Geometric corrections
 - (b) Radiometric corrections
 - (c) Density slicing
 - (d) Convolution filtering
- 7.11 Describe the following with examples:
 - (a) Resampling methods
 - (b) Frequency domain filtering
- 7.12 Write short notes on the following:
 - (a) De-stripping
 - (b) Line-drop outs
 - (c) Sun elevation correction
 - (d) Classification
 - (e) Preprocessing
- 7.13 Explain the following:
 - (a) Saturation stretch

- (b) Reference stretch
 (c) Edge detection filters
 (d) RGB colour cube
- 7.14** Consider two 8-bit images whose grey levels span the full range from 0 to 255.
 (a) Discuss the limiting effect of repeatedly subtracting image 2 from image 1.
 (b) Would reversing the order of images yield a different result?
- 7.15** Discuss the limiting effect of repeatedly applying a 3×3 low pass spatial filter to a digital image. Ignore border effects.
- 7.16** Consider the middle pixel of an image shown below. Calculate its new value if
 (a) 3×3 median filtering is applied and (b) 3×3 averaging filtering is applied.

5	3	11
12	4	9
8	6	24

[Ans. (a) 8 (b) 9]

- 7.17** Suggest and apply the suitable image transformation (arithmetic) operation on Image 1 and Image 2 of an area in order to reduce the overall noise contribution for Image 1.

Image 1		
98	84	124
151	131	135

Image 2		
86	64	35
122	36	158

Ans.	92	74	79
	136	83	146

- 7.18** Using the pixel values of row 1 and row 3 for the image given below carry the de-striping for row 2.

15	32	12
8	6	15

[Ans. (11, 19, 13)]

- 7.19** Carry out $2 \times$ image magnification operation for the image with pixels given below:

204	200
201	107

Ans.	204	204	200	200
	204	204	200	200
	201	201	107	107
	201	201	107	107

Objective-type Questions

- 7.1** Which of the following is not part of DIP?
(a) Enhancement (b) Filtering
(c) Squaring (d) Classification
- 7.2** Which of the following is not an error due to platform instability?
(a) Altitude of spacecraft (b) Attitude of spacecraft
(c) Scan skew (d) Mirror sweep
- 7.3** Which of the following is not a radiometric correction?
(a) Earth's rotation (b) Sun elevation
(c) Atmospheric (d) Detector response calibration
- 7.4** Which of the following is not a platform attitude?
(a) Crab (b) Yaw
(c) Pitch (d) Roll
- 7.5** Which of the following is not a method of contrast enhancement?
(a) Density slicing (b) Band rationing
(c) Histogram normalisation (d) Reference stretch
- 7.6** Pick up the correct statement.
(a) A digital image is composed of two dimensional array of pixels.
(b) A digital image is defined as a row of DNs.
(c) The pixel value shows the spatial resolution of remote sensed data.
(d) One of this range of DNs for recording digital image is 0–500.

Answers to Objective-type Questions

7.1 (c)	7.2 (d)	7.3 (a)	7.4 (a)	7.5 (b)	7.6 (a)

8

GEOGRAPHIC INFORMATION SYSTEM (GIS)

Introduction

The development of technology and the geographical world is taking place at a very fast pace. This necessitates having a system capable to analyse and manipulate the spatially referenced data or remotely sensed information, and give the desired output (information) in a very short time. A geographic information system helps to better understand the world around us and enables development of spatial intelligence for logical decision-making. Several definitions of GIS exist in literature, some of which are as follows.

GIS is a computer-based information system which attempts to capture, store, manipulate, analyse and display spatially referenced and associated attribute data for solving complex research, planning and management problems.

GIS is a system of hardware, software, data, and people organising, collecting, storing, analysing and disseminating information about the areas of the earth.

GIS is an information technology which stores, analyses and displays both spatial and non-spatial data.

GISs are specialised data bases that preserve locational identities of the information that they record.

The word *graphic* in GIS carries two meanings: *earth* and *geographic space*. By *earth*, it implies that all data in the system are pertinent to earth's features and resources, including human activities based on or associated with these features and resources. By *geographic space*, it means that the commonality of both the data and the problems that the systems are developed to solve, is geography, i.e., location, distribution, pattern, and relationship within a specific geographical reference framework. The word *information* implies that data in a GIS are organised to yield useful knowledge—often as coloured maps and images, statistical graphics, tables and various on-screen activities. The word *system* implies that a GIS is made up from several inter-related and linked components with different functions.

The term 'Geographic Information System' first appeared in published literature in 1970. Although it sounds to be relatively a new term, many of its concepts have been in existence for centuries. For example, to illustrate fundamentals that still comprise the basis of GIS, consider the base plan of a locality of a city wherein it is required to show public utilities such as water supply network, fire hydrants,

drainage system, network of sewers, arrangement of manholes, underground electrical cables, etc. Since incorporation of all these into the base plan probably will make the plan unreadable and confusing, the practice in conventional surveying is to draw the utility plans on tracing papers separately and overlay them over the base plan as per the requirements. In addition, the data related to such utilities, for example, diameter, make, type of pipes, etc., are maintained in a separate register to be referred for complete information. Of course, maintenance of these records and updating the changes effected from time to time is a cumbersome process. All this information and more can be stored, manipulated and retrieved with the help of computer and softwares within no time, which is the essence of GIS. The use of GIS removes the need of paper plans and associated documents and speeds up the production of information in the form of maps, tables, etc., (the GIS products) by rapidly updating and editing the data in computers.

GIS is capable of acquiring spatially indexed data from a variety of sources, changing the data into useful formats, storing the data, retrieving and manipulating the data for analysis, and then generating the output required by the user. The acquired special indexed data are known as *layers*. Each layer represents a thematic approach to a particular purpose. For urban planning, a set of the layers may represent, for example, property lines of areas, land use, drainage and contour information, soil type, etc. Any selected layer or a combination of layers can be depicted on a map at any desired scale. Figure 8.1 shows the layer-based concept of data organisation in GIS.

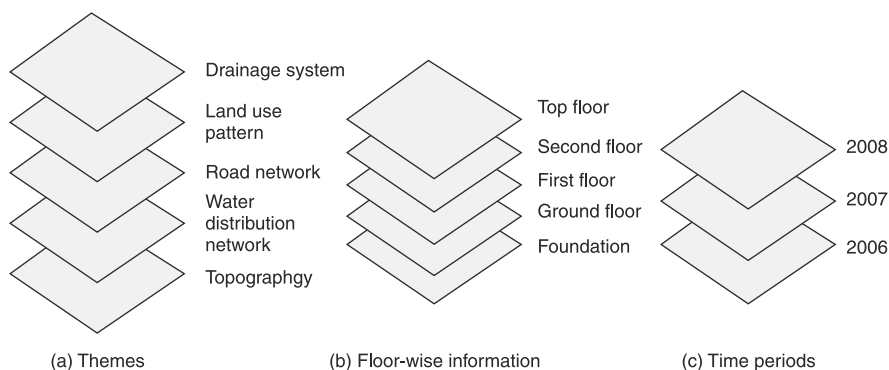


Fig. 8.1 A layered database concept

The great strength of GIS is the ability to handle a large multilayered, heterogeneous database and to answer queries about the existence, location and properties of a wide range of spatial objects in an interactive way. All the layers need to have a common reference system to enable two or more layers to be integrated. Also, maps can be created from the merged data and overlaid, if desired. However, the accuracy of the spatial analyses and hence the validity of decisions reached (map) as a result of these analyses depend directly on the quality of the spatially related information in its database. A digital computer provides the basis for storage, manipulation and display of large amounts of data that have been encoded in digital form. GIS consists of a package of computer programs with a user interface that provides access to particular functions.

The objective of collecting geographic data and converting them into useful information (desired output, e.g., map, table, etc.) by means of a GIS transcend the traditional boundary of data processing and information management. The main advantage of GIS is rapid analysis and display of data, with flexibility not possible using manual methods. GIS does not hold maps or pictures—it holds a database. That is how it is different from computer mapping, which can produce only good graphic output. GIS, as we understand it today, is very different from its predecessor. GIS, earlier used for computer-based applications for map-data processing, is now an essential component of the information technology infrastructure of modern society. It is a multidisciplinary science. GIS practitioners may be geographers, surveyors, planners, or computer engineers. Despite the diversity in approaches, a special set of skills and knowledge is required by professionals to use GIS in all its forms and implementations. GIS, today, has become an indispensable tool to manage land and natural resources, monitor the environment, formulate economic and community-development strategies, enforce law and order, and deliver social services.

The main purposes of GIS are:

1. To support decision-making based on spatial data; for example, an engineering geologist may evaluate slope stability conditions through GIS for deciding the best new route
2. To support general research
3. To collect, manipulate and use spatial data in database management
4. To produce standardised and customised cartographic production

8.1 SUBSYSTEMS OF GIS

A GIS must include at least three main elements: (i) computer hardware, (ii) computer programs, and (iii) data. A GIS may be considered to have five major component subsystems (Fig. 8.2), and are as follows:

- 1. Input** : Which deals with creating an image-based GIS from multigeodatasets.
- 2. Management** : The purpose is efficient storage, retrieval and database management.
- 3. Processing** : Data manipulation, feature enhancement and classification, etc.
- 4. Display** : Display and product generation.
- 5. Output** : Provides thematic maps, images, etc., for application.

The basic forms of the data for GIS are spatial data—data that provide locations and shapes of features in a map; tabular data—data that are collected or compiled for a given area, a GIS links to the features in a map; and image data—such as aerial photographs and products, satellite images, and scanned data (photographic prints converted to digital format).

The captured data can be transformed from existing maps, detailed observation, satellites and aerial photography into digital or computer compatible format. The remote sensing images and digital data are the primary source of modern GIS.

Using digitiser/scanner, keyboard entry of attribute information, etc., these data can be stored in a computer. The various types of geographic data are stored in

two different formats in a GIS in the form of: the Cartesian coordinates and the raster format (e.g., grids). The graphical features are stored with their location (x, y) defined by latitude and longitude, and attribute data may be qualitative (e.g., land use), or quantitative. The location can be represented either by a raster (grid cell format), or a vector format (polygon). In raster format, a location is defined by the row and column position of the cell it occupies; the value corresponding to the location indicates the type of feature. In the vector format, the geographic space is continuous and the data structure is more representative of the dimensionality as in a map. For GIS data input, either of the two data formats can be used. The two data formats, their structure and conversion from one format to the other are discussed in the sections to follow.

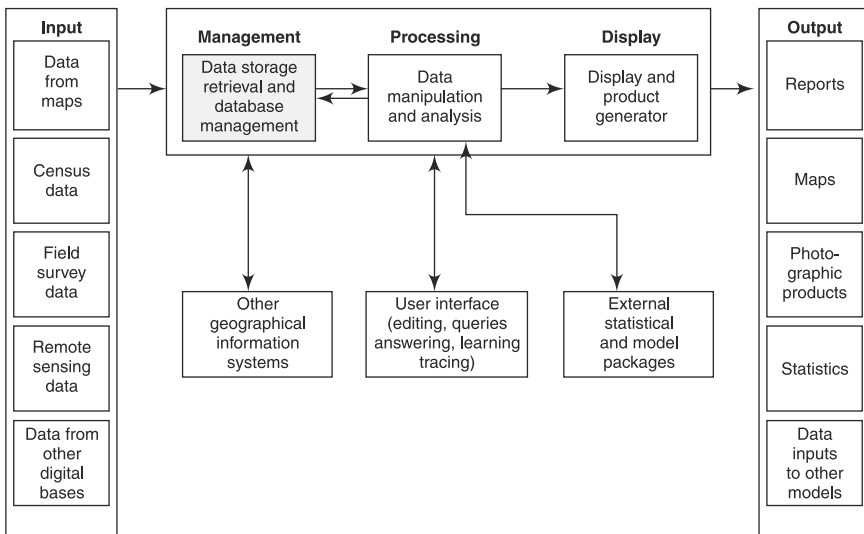


Fig. 8.2 *Component subsystems of GIS*

Once the data is transformed into the computer, this data has to be stored to create a permanent database for further data analysis and manipulation. The digital map file is stored on a magnetic or optical digital medium. The encoded spatial data are stored systematically in the form of layers, known as GIS layers. These layers are archived in the digital format as a geographically-referenced plane in the GIS database. The database files are stored in the central processing unit (CPU) memory and can be processed and manipulated.

The computer program that is employed to organise the database is commonly known as Database Management System (DBMS). The analysed and manipulated result of the data has to be displayed or presented to the user in a user-specified format for decision-making purpose. Either the data may be presented as maps, tables and figures on the screen, or recorded on magnetic media in digital format or as hardcopy output drawn on printer or plotter.

8.2 HARDWARE OF GIS

The hardware of GIS is made up of a configuration of core and peripheral equipment that is used for the acquisition, storage, analysis, and display of geographic information. At the heart of the GIS hardware architecture is the *central processing unit* (CPU) of the computer. The CPU performs all the data processing and analysis tasks and also controls the input/output connectivity with data acquisition, storage, and display systems. Depending on the data-processing power of the CPU, computers are classified as supercomputers, mainframes, minicomputers, workstations, and microcomputers or personal computers (PCs). All these classes of computers can be used as the hardware platforms for GIS. Conventionally, GIS were developed as stand-alone applications that ran on one of these classes of computers. Today's GIS are mostly implemented in a network environment using the client/server model of computing,

Client/server computing is based on the concept of division of work among different machines in local or distributed computer network. A server is the computer on which data and software are stored. A client, on the other hand, is the computer by which the users access the server. The application programs can be executed on either the server or the client computer. In the client/server environment, a client can access multiple servers, and similarly, a server can provide services to a number of clients at the same time. For GIS that are implemented on the client/server architecture, processor-intensive operations and data management are most commonly performed in the workstation class of servers, and PCs are used as the clients that provide the graphical interface to the system. Such a configuration, which combines the processing power of workstations and the economy of using PCs, has replaced the mainframes and minicomputers as the dominant hardware platforms for GIS.

Because of the very large amounts of data that are inherently required for a GIS, it is necessary to have access to tape drives and disk drives that permit the GIS to read the information transported on computer tape from other computers and other GISs. A *colour display* is often an important element of a GIS as a means of displaying several images or map patterns.

A GIS usually includes equipment for entering data into the system. Some data can be entered using a *video camera*, which is capable of recording an image in much the same way that a television camera does: a map or image that can be analysed in much the same way as any digital image. Although the video camera is useful for some applications, it often cannot produce the geometric accuracy or the fine detail that are required for GIS work. Therefore, other methods of data entry are required.

A *digitising table, or digitiser*, (Fig. 8.3) forms a third method of data entry. Digitiser is a traditional device for encoding the digital data from an existing map. The table consists of a fine grid of thin wires that have been encased in a dense, stable substance such as fibre glass. The wires are capable of sensing the x - y positions of a *cursor* that can be moved over the surface of the table. The more finely and accurately the wires are spaced, the more precise is the data generated by the digitiser. The analyst tapes a map or aerial photograph to the surface of the digitiser, and after establishing a coordinate system, can trace the outlines of areas,

or mark positions of points and lines with the cursor. As the cursor moves, the digitiser creates a digital record of its successive positions usually as a sequence of coordinate pairs.

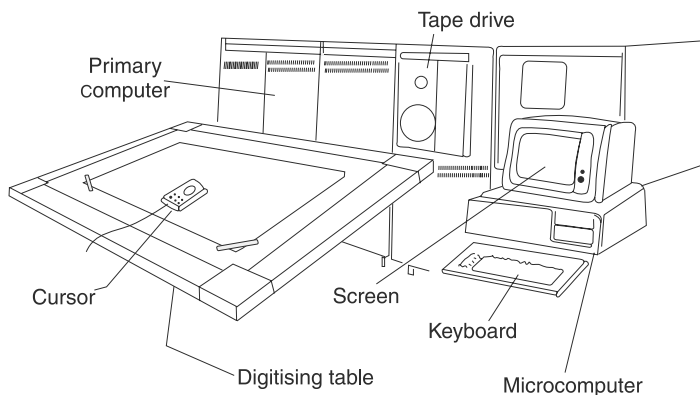


Fig. 8.3 *Equipment for a GIS*

Usually, the digitiser is supported by a small processor or computer that can allow the analyst to record data with separate codes (to identify lines that record streams, highways, or power lines, for example) and to perform geometric transformations (to correct positional errors). The best digitising systems may allow the analyst to see the data on a computer screen and to identify and correct errors as they occur.

8.3 DATA FOR GIS

The prime aspect in the construction of GIS is the acquisition of data. Data can be gathered directly in the field by an original survey specifically carried out for the purpose of GIS and is known as *captured data*. However, it will be very time consuming and costly and thus rarely resorted to. Alternatively, the data can be obtained/derived from a source that is already available, such as topographic maps, digitised maps and plans, aerial photographs, satellite imagery, or directly from GPS survey. Such data are called *encoded data*. Data that are obtained by human intervention, for example, from sketches of landscapes, a questionnaire, etc., are called the *interpreted data*. Those which are in a table or in GIS are called *structured* or *organised data*. The remotely sensed images and digital data, and the extracted information from these are the primary source of modern GIS. The digital remote sensing data are in raster format and are acquired by sensors through a scanning device. The device collects the data at successive instants of time by dividing the field of view in a grid pattern; each grid element is called pixel. A sensor records a series of radiometric values for each pixel of each band, it is sensing, and the image is built up by a combination of consecutive pixel lines (scan lines). Data in GIS may be classified as spatial and non-spatial.

1. *Spatial data* Also called *graphical data*, it consists of natural and cultural features that can be shown with lines or symbols on maps, or that can be seen

as images on photographs. The data in the different forms (maps, photographs, images, etc.) being in non-compatible formats create problems while integrating in GIS. In a GIS, these data must be represented and spatially located in digital form, by using a combination of fundamental elements called *simple spatial objects* (SSO). These SSO include points, lines and strings, areas or polygons, pixels, and grid cells. SSO can be represented by their respective symbols. In addition to these SSO, surface element is also there, which can represent most realistically the spatial objects we observe in the real world. A surface is an area or a polygon having a third dimension, i.e., height (elevation). Natural features such as hills, valleys, etc., and man-made features such as structures, etc., can be best described by the surface element. Modelling of surface is known as Digital Terrain Modelling (DTM) and is given in Appendix II.

2. *Non-spatial data*: Also called *attribute data*, it describes geographic regions or defines characteristics of spatial features within geographic regions. These data are usually alphanumeric and provide information such as colour, texture, quantity and quality. The non-spatial data are often derived from documents such as plans, files, reports, tables, etc. For example, for a highway, the data may be its number, pavement type, number of lanes, lane width, year of last resurfacing, etc.

In general, spatial data will have related non-spatial attributes. The linkage between the spatial and non-spatial data, therefore, must be established and is achieved with a common identifier, which is stored with both of these. The identifier may, for example, be a unique parcel identification number, a grid cell table, etc.

8.4 REPRESENTATION OF FEATURES

GISs are the information systems which offer the functionality and tools to collect, store, retrieve, analyse and display geographic information. Features, events and activities with spatial components are modeled, as points, lines, polygons, nets or links to form the geographic database under georelational schemes. The geometrical part is defined by a series of coordinates and is connected through the feature code with the attribute tables where the non-spatial information such as the properties, symbolism, etc., is stored. GIS can represent spatial data as the data which have physical dimension on the earth. In order to represent the complex three dimensional realities in a spatial database, the SSO used are described as follows.

1. *Point data*: Point data consist of observations that occur only at points or occupy very small areas in relation to the scale of the database. These define single geometric positions (Fig. 8.4(a)). Spatial location of points are given by their coordinates (x , y). Features such as wells, for example, illustrate data that occupy a single point even at the largest levels of detail. In contrast, features such as buildings sometimes occupy significant areas, even though they may be represented as a point in the database. Some of the other examples of point data are survey control points, monuments, mines, etc.

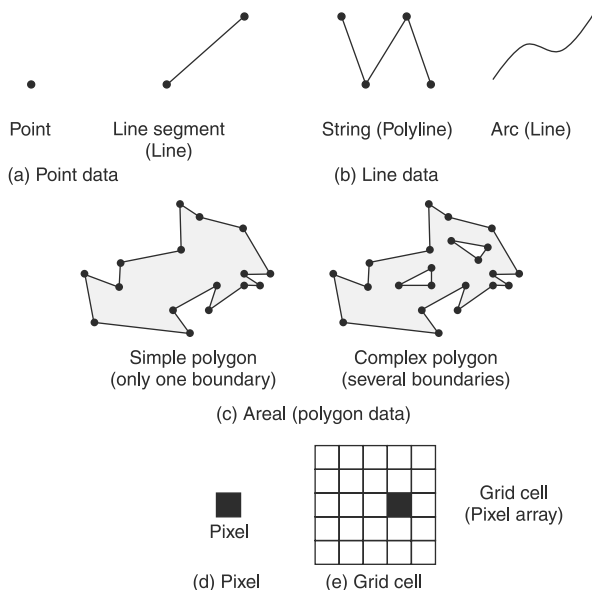


Fig. 8.4 Representation of features

2. Line and string data: Lines and strings are obtained by connecting points. A line connects two points (Fig. 8.4(b)), and a string is a sequence of two or more lines. Line and string data are formed by features such as highways, railways, canals, rivers, pipelines, power lines, etc. An arc, which is the locus of points, may be defined by a spline curve or polynomial mathematical function.

3. Areal data: An area or polygon consists of a continuous space within three or more connected lines. Examples of areal data (Fig. 8.4(c)) include distribution such as soil type, parcels (pockets) of land ownership, different types of land cover, vegetation classes, and other patterns that occupy area at the scale of the GIS.

4. Pixels: These are usually tiny squares that represent the smallest elements into which a digital image is divided (Fig. 8.4(d)). Continuous arrays of pixels, arranged in rows and columns, are used to enter data from aerial photos, satellite images, orthophotos, etc. The distributions of colours or tones throughout the image are specified by assigning a numerical value to each pixel. Pixel size can be varied and is specified either in terms of image or object scale. At the image scale, pixel size may be specified directly (e.g., 0.025×0.025 mm) or as a number of pixels per unit distance (e.g., 10 dots per cm, where a dot corresponds to a pixel). At the object scale, pixel size is usually expressed directly by dimension (e.g., 10 m pixel size).

5. Grid cells: These are single elements, usually square, within a continuous geographic variable. Similar to pixels, their sizes can be varied, with smaller cells yielding improved resolution. Grid cells (Fig. 8.4(e)) may be used to represent terrain slopes, soil types, land cover, water-table depths, land values, population density, etc. The distribution of a given data type within an area is indicated by

assigning a numerical value to each cell. For example, to show soil types in an area, numerals 1, 2, and 3 may be used to represent sand, silt and clay respectively.

8.5 DATA STRUCTURE FOR GIS

Data for a GIS must be represented in a form that preserves locational identities of each unit of information, so that it is possible to retrieve data by location and therefore to depict and analyse geographic patterns. Because data are frequently derived from a ‘conventional’ (non-digital) map or image, it is necessary to convert them into digital form suitable for use by a GIS. This process, known as *geocoding*, records the pattern/features of a map in a form that can be accepted and manipulated by computers.

The simple spatial objects described in Section 8.4 may be coded in two different formats—vector and raster—for storing and manipulating these (spatial data) in a GIS. Both of these data structures, also called *data models* or sometimes *data formats*, offer contrasting advantages and disadvantages. When data are depicted in the vector format, a combination of points, lines and strings, an area is used, whereas the raster format uses pixels and grid cells. Figure 8.5 shows a representation of simple spatial objects in vector and raster models.

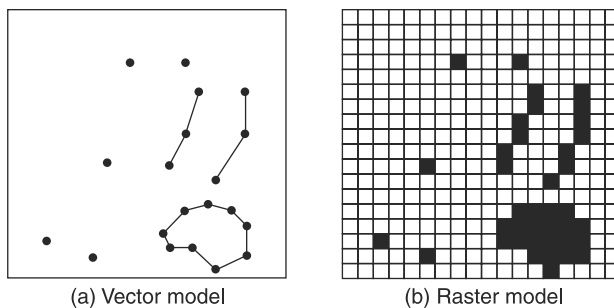


Fig. 8.5 *Spatial data models*

Usually, a GIS must be designed on either a raster or a vector format. Because of differences in equipment, computer programs, and expertise required for the two different approaches, the choice depends upon the facilities available, the kinds of data to be examined, and the purposes of establishing the GIS. Also, it is possible to convert from vector-to-raster format by applying relatively straightforward computer algorithms, but it is a little difficult to do raster-to-vector conversion. The data format conversions are described in Section 8.9.

8.6 VECTOR DATA STRUCTURE

Vector data depicts the real world by means of discrete points, lines and polygons and is stored as a collection of x, y coordinates. The vector format provides an accurate representation of spatially referenced data incorporating the topology (Appendix IV) and other spatial relationships between the individual entities. The three forms of *basic graphical elements* are the *point*—represented as a single pair of coordinates, the simplest type of vector data; the *line* or *arc*—represented as a string of coordinates, which begin and end with a *node*; and the *area* or

polygon—represented as a closed loop of coordinates. When graphical elements representing an individually identifiable real-world feature are logically grouped together, a *graphical entity* is formed. For example, the different line segments that represent a railway are graphical elements. When these line segments/graphical elements are identified in the database and logically joined together, a graphical entity railway is formed.

In addition to the formatting of the basic graphical elements, it is also necessary for the vector data to be properly linked to the descriptive data in geographic databases. This is usually achieved by the use of a unique *feature identifier* (FID) that is assigned to individual spatial objects. By using common FIDs, the graphical and descriptive elements of vector data can be correctly cross-referenced during database creation and spatial data processing. Usually, the assignment of the feature identifier is an automated procedure during the topology building process, but the linkage to descriptive data is normally a manual process that can be only partially automated.

A number of instruments are available to input a vector data into a GIS, but manual digitising is used commonly. Figure 8.6 illustrates the vector format for two adjacent parcels of land, designated as Parcel I and Parcel II. It consists of points, lines and areas. Vector representation of the data can be achieved by creating a set of tables that list these points, lines, and areas (Table 8.1). Data within the table are linked by using *identifiers* and are related spatially through the coordinates of points. As illustrated in column (i) of Table 8.1, all points in the area are indicated by an identifier. Similarly each line is described by its end-points, as shown in column (ii) of Table 8.1, and the end-point coordinates locate the various lines spatially. Areas in Fig. 8.6 are defined by lines that enclose them, as shown in column (iii) of Table 8.1. As before, coordinates of end-points locate the areas and enable the determination of their locations and magnitudes.

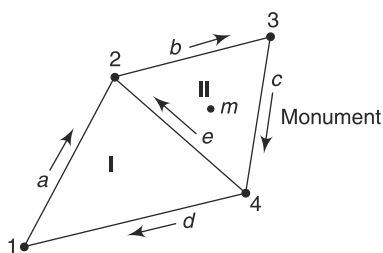


Fig. 8.6 Vector representation of simple graphical record

Table 8.1 Vector representation of graphical record of Fig. 8.6

(i)		(ii)		(iii)	
Point identifier	Coordinates	Line identifier	Points	Area identifier	Lines
1	(x_1, y_1)	a	1,2	I	a,e,d
2	(x_2, y_2)	b	2,3	II	b,c,e
3	(x_3, y_3)	c	3,4		
4	(x_4, y_4)	d	4,1		
Monument, m	(x_m, y_m)	e	4,2		

As another example, consider a land-cover map showing an orchard, a forest, a built-up area and a river with standard topographic symbols (Fig. 8.7(a)). A vector representation of the region is shown in Fig. 8.7(b). Here, the line and

string elements demarcate and locate boundaries of different regions. The river has been shown by string elements. Tables similar to 8.1 can be constructed and entered into GIS using the vector format.

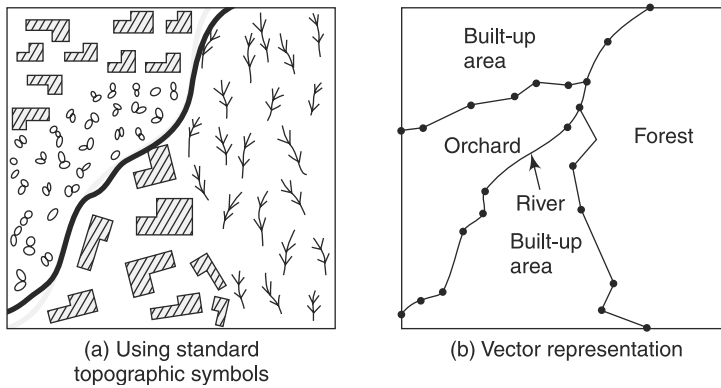


Fig. 8.7 Land cover map of a region

The representation of vector data is governed by the scale of the input data. For example, a building that is represented as a polygon on a large-scale map will become a point on a medium-scale map, and it will not be represented at all as an individual entity on a small-scale map (unless it is a very important landmark). The possibility of representing vector data differently at different scales is associated with two important concepts: (i) cartographic generalisation, whereby line and areal objects are represented by coordinates at a larger scale and (ii) cartographic symbolisation, whereby vector data are represented by different symbols that serve to visually distinguish them from one another when the data are displayed.

In the computer, vector data can be stored as integers or floating point numbers. In order to avoid the problem of rounding off errors that occur during data processing, most GIS software products store vector data by using double-precision floating point numbers. This creates the impression that vector data are accurate and precise representations of spatial objects in the real world. However, this is not necessarily true because the precision of data storage does not always mean accurate description of the data, and also, the boundaries of many spatial objects are *fuzzy* rather than exact entities. Thus, storing vector data by double-precision floating point numbers does not improve the quality of the data, but simply serves to avoid degradation of data quality due to rounding errors during data processing.

8.7 RASTER DATA STRUCTURE

Raster data structure, also called *cellular data structure* depicts the real world by pixels or grid cells. It is not as accurate or flexible as the vector format, as each coordinate may be represented by a cell and each line by an array of cells. Raster data can be positioned only to the nearest grid cell. Examples of data in the raster format are aerial photographs, satellite imagery and scanned maps or plans.

For the input of raster data, first the region of interest is subdivided into a network of cells of uniform size and shape (regular, square or rectangular). The linear dimensions of each cell define the spatial resolution of data or the precision with which the data is represented. Thus, the size of an individual pixel or cell is determined by the size of the smallest object in the geographic space to be represented. The size is also known as the *minimum mapping unit* (MMU). A general rule is that the grid size should be less than half the size of the MMU.

Once the grid cell size has been decided, each grid cell is assigned a value, which can be an integer, a floating point number, or a character (a code value). A raster data along with characteristics is shown in Fig. 8.8(a). The values marked indicate the quantity, or characteristics of the spatial object, or phenomenon that is found at the location of the cell. The input of the vector counterparts of this raster data is also shown in Fig. 8.8(b). The value 3 has been used to classify the raster cells according to land use—the road—at the given location. The remaining cells are filled with 0 indicating that no identity is present at that location. There are four methods for the input of the vector counterparts of the raster data—the dominant method, the precedence method, the presence/absence method, the per cent occurrence method (see Section 8.9)—available in the literature.

In a raster database, values pertaining to different characteristics at the same cell location are stored in separate files (*map layers*). For example, a road and forest cover for the same area are stored as separate road and forest data layers. When the data are used for processing, the appropriate layers are retrieved. This means that raster data processing always involves the use of multiple raster files, in the same way different layers are used in vector data processing.

When a specific raster layer is displayed, it is shown as a two-dimensional matrix of grid cells. In computer storage, the raster data are stored as a linear array of attribute values. Since the dimension of the data (the number of rows and columns) is known, the location of each cell is implicitly defined by its row and column numbers. There is no need to store the coordinates of the cell in the data file. The locations of the cells can be computed when the data are used for display and analysis.

In order to translate linear array storage to a two-dimensional display, enough information must be stored in the header section of the data file as well. In general, the file header contains information about the number of bits used to represent the value in each cell, the number of rows and columns, the type of image, the legend, the name or the colour palette (if the file uses one), and the name of the look-up table (if the file uses one). Some file headers also contain parameters for coordinate transformation so that raster data in the files can be georeferenced. This is, however, a system-dependent feature. The cells in each line of the image (Fig. 8.8(a)) are mirrored by an equivalent row of numbers in the file structure (Fig. 8.8(c)). The first line of the file structure indicates the computer that the file consists of 6 rows and 6 columns and that the maximum cell value is 3.

Raster data files are stored in different file formats. The differences between these file formats are due mainly to the different algorithms used to compress the raster data files. In order to minimise the data-storage requirements, raster data are often stored in compressed form. The data are decompressed 'on-the-fly'

when they are used by an application program. The raster model (the geometrical arrangement of the figures covering the surface) is best employed to represent geographic phenomenon that are continuous over a large area.

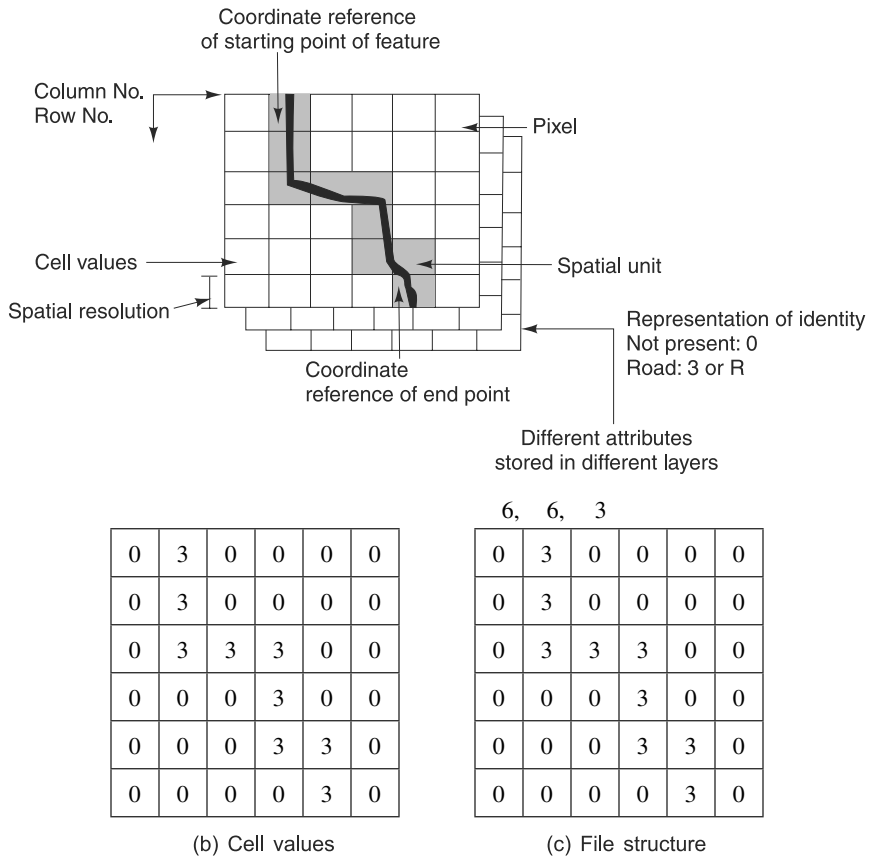


Fig. 8.8 Characteristics of raster data

8.8 VECTOR VS RASTER DATA STRUCTURES

The relative advantages/disadvantages and limitations of vector and raster data structures are presented with suitable examples.

Advantages of Vector Data over Raster Data

1. A vector database can depict point data as points which can be positioned accurately. However, a raster database can depict point data only at the level of the detail of a single cell. This leads to loss of accuracy, for example, a cell can show the presence of a tower, within a cell but cannot show its placement within this cell. Of course, the cell size would affect the depiction but a raster database cannot be as accurate as a vector database. It may further be noted that some points may represent the quantitative characteristics, e.g., amount of rainfall or elevation but these information cannot be included in the raster format.

2. A vector database can show a line data in exact and fine detail, whereas a raster database can show the same line as a zigzag or a comparatively more smooth line depending upon the resolution of the cell.
3. A vector database provides details and exact/fine boundaries among aerial patterns, e.g., land cover. However, in case of a raster database, the accuracy is lost for the reason explained earlier.
4. Discrete quantitative data such as population, which are grouped/associated with an area, are best depicted in finer detail by polygons (vector format). However, continuous data such as topographic elevation/contours, represented by a network of equally spaced observations, can probably be most directly presented by a raster format.
5. A vector database is best suited to represent various natural/artificial features and also these can be presented mathematically (coordinates). This makes the vector format conceptually more complex than the raster format.
6. A vector database require less storage space on the computer as compared to a raster database for the same information. Also, the vector formats are more accurate and present a finer detail of shapes and sizes as compared to the raster format.

Some of the other notable disadvantages of raster data are coding of a cell with a single value (category) whereas many features may be present in it; recording of spatial objects only to the nearest cell which may not correspond/match in reality and also may not exist in reality, e.g., watershed; and coarser resolution of spatial features leading to inaccurate representation.

Disadvantages of Vector Data over Raster Data

1. Vector formats sometimes prove costlier because of higher data-encoding cost. Also, the programs for data manipulation are more complex as compared to the raster format.
2. In case of the vector format, the superimposition/overlying different layers of data may be difficult because some polygons in different layers may not match exactly due to minor digitisation errors, forming small *slivers* or *strips* (Fig. 8.9).

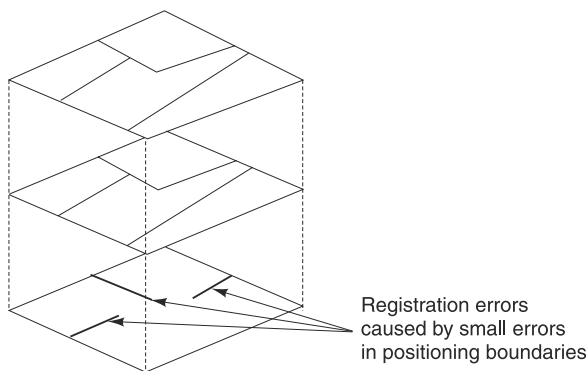


Fig. 8.9 *Overlays of polygon files may not register exactly*

Some of the notable advantages of raster format are use of simpler computer programs for data manipulation; ideally suitable for a variety of spatial analysis functions, for example, overlay, buffering and network analysis; direct use of remote-sensing data which are in this format; use of available image processing software for refining raster images; and in some data types (soil, boundary, wet land, built-up area, etc.) which are relatively vague, its use does not significantly degrade the inherent accuracy of the data.

8.9 DATA FORMAT CONVERSIONS

The usage of GIS data often requires integration of vector and raster data which needs conversion from one form to the other. The various procedures for converting vector data to raster format and vice versa are described below. It should be noted that interpretation/classification of boundary/mixed data varies with the method used and also that none of the methods is perfect in presenting the true feature.

1. *Vector to raster conversion:* Vector to raster conversion is also referred to as *coding*. Figure 8.10(a) shows a vector data/format (map) of the vector representation of the land-cover map shown in Fig. 8.7, overlaid/superimposed on a raster of grid cells. The size of these cells will depend upon the accuracy desired, time and computing facility available. Figure 8.10(b) illustrates the raster representation of vector counterparts of raster data using a coarse-resolution grid cell of the region of the land-cover map of Fig. 8.7(a), and that using a finer-resolution grid cell is shown in Fig. 8.10(c). The finer-resolution grid renders depiction/storage of areas with greater precision. Of course, the fine-resolution grid cell size will yield better results but will demand more time and computing facility and results in higher costs.

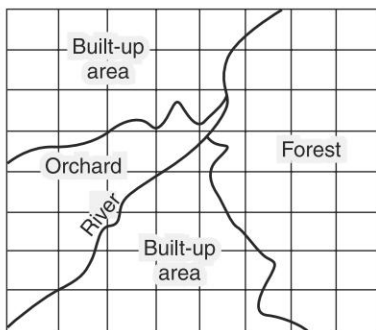
As discussed earlier, the vector counterparts of the raster data can be input by the four methods. Of these, the *presence/absence* and *dominant* types are the best method and are used depending upon the importance felt by the user. In the presence/absence method, for each grid cell, a decision is made as to whether the selected entity exists at the centre of the given grid cell or not.

In the presence/absence method, the cell is assigned a value corresponding to the characteristics (vector location) of its centre. Thus, for example, the centre of the cell (3, 4) is occupied by an orchard and hence this cell is assigned the value 0 or 2, as shown in Fig. 8.11(b). If it does not, it is ignored.

The cell in the precedence method is allotted a value corresponding to the most important characteristics or precedence with respect to the other characteristics present in it. For example, in the cell (6, 3), there are three different characteristics, viz., orchard, river and built-up area. Though the river in this cell occupies the least area; it is the most important among the three features, and also the river cannot be discontinuous. Thus, among the three characteristics of this cell, the river is given precedence over others and accordingly this cell is coded as R or 3, as shown in Fig. 8.11(c).

In the dominant type of conversion or coding, each cell of a grid is assigned a value corresponding to predominant characteristics of the area within the cell. For example, two different type of areas, viz., built-up area and orchard occupy the cell located in the third row and the third column (3, 3). This cell is assigned

the value B or 1, as shown in Fig. 8.11(d), because the larger part (more than 50 per cent) of the cell is occupied by buildings.



Built-up area I, B
Orchard O, 2
River R, 3
Forest F, 4

(a) Vector representation of map over grid cell

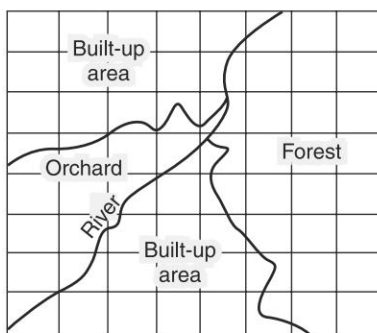
B	B	B	B	R	R	F	F
B	B	B	B	R	F	F	F
B	B	B	O	R	F	F	F
B	O	O	R	R	F	F	F
O	O	R	B	B	F	F	F
O	R	R	B	B	F	F	F
O	R	B	B	B	B	F	F
R	B	B	B	B	B	F	F

(b) Coarse-resolution grid cell

B	B	B	B	B	B	B	B	B	B	B	R	F	F	F	F
B	B	B	B	B	B	B	B	B	B	B	R	F	F	F	F
B	B	B	B	B	B	B	B	B	B	B	R	F	F	F	F
B	B	B	B	B	B	B	B	B	B	B	R	F	F	F	F
B	B	B	B	B	B	B	B	B	B	B	R	F	F	F	F
B	B	B	B	O	O	O	O	O	R	F	F	F	F	F	F
B	B	B	B	O	O	O	R	R	F	F	F	F	F	F	F
O	O	O	O	O	O	R	B	B	F	F	F	F	F	F	F
O	O	O	O	R	B	B	B	B	F	F	F	F	F	F	F
O	O	O	R	B	B	B	B	B	F	F	F	F	F	F	F
O	O	O	R	B	B	B	B	B	F	F	F	F	F	F	F
O	O	R	B	B	B	B	B	B	F	F	F	F	F	F	F
O	O	R	B	B	B	B	B	B	F	F	F	F	F	F	F
R	R	B	B	B	B	B	B	B	B	F	F	F	F	F	F
B	B	B	B	B	B	B	B	B	B	B	F	F	F	F	F

(c) Fine-resolution grid cell

Fig. 8.10 Raster representation of land cover map of Fig. 8.7



(a) Vector representation of map overlaid on a coarse grid



(b) Presence / absence method (3, 4)



(c) Precedence method (6, 3)



(d) Dominant method (3, 3)

Fig. 8.11 Data conversion from vector to raster format

It should be obvious by now that the coding of a particular cell, in which multiple features/characteristics are present, depends upon the conversion method used. For example, the cell (4, 5) would be assigned F or 4 in the predominate

method; R or 3 in the precedence method; and B or 1 in the presence/absence method. Further, the cell size would also affect the coding or conversion from vector to raster, with a larger cell resulting in a relatively inaccurate presentation of vector data.

In the above coding methods, alphabets or numerals may be used depending upon convenience. However, use of numbers is more common. These grid cell values can be used directly for computation or indirectly as code numbers referenced to an associated table.

2. Raster to vector conversion: In this conversion, the cells of the raster through which the vector line (e.g., river, boundaries of buildings, roads, etc.) passes are identified. Then the line (vector form) connecting these cells is drawn. One way is to connect the centres of cells with straight line segments. Obviously, this will produce a zigzag line, whereas in nature this line would be a smooth one. So, either curve-fitting is required or the cell size is to be reduced to extract/draw a smooth line passing through the cells. The first method involves complicated mathematical calculations, which does not necessarily give a unique solution. Either of the methods would require large capacity computers as well as time. Even after this, the line drawn, using raster data, may not exactly match with the actual feature present in nature. This conversion of data from raster to vector model is illustrated in Fig. 8.12.

Note *In case if the vector data is converted to raster data and then again converted back to the vector data, the resultant data set will not likely match with the original one. This is specifically true in case of boundaries of different common data or linear features.*

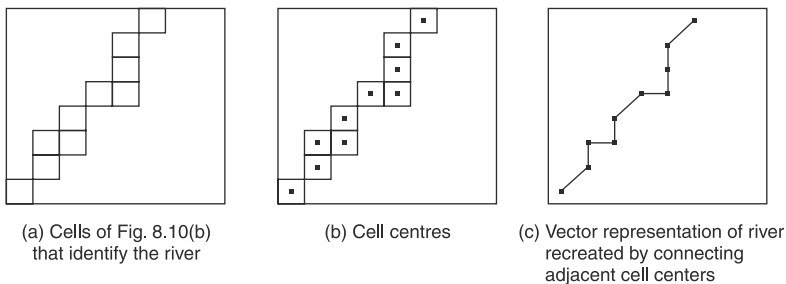


Fig. 8.12 Conversion of data from raster to vector format

8.10 CAPABILITIES/FUNCTIONALITIES OF GIS

Analysis in GIS basically refers to the processes of drawing inferences from data. The analysis is carried out on data available either in tabular form or maps. In general, the capabilities or functionalities of GIS are innumerable and beyond the scope of the book. However, queries, reclassification, buffering and overlay need special mention and are described in detail. Since map overlay analysis is the most important function, it is discussed in sections to follow.

1. Organisation: The importance of data organisation can be revealed by a person who has collected a large mass of data for any particular purpose. Data can be

arranged in many ways, and unless the organisation scheme is suitable for the application at hand, useful information cannot be easily extracted. Schemes for organising data are sometimes called data models (structures)—the vector and the raster—already discussed in the previous sections. Data models organise observations both by spatial and non-spatial attributes. Thus, data organisation has a fundamental importance.

2. *Visualisation:* Visualisation is achieved in GIS with colour and by specialised methods using perspective, shadowing and other means. The graphical capabilities of computers are exploited by transforming a table of data, for example, into a visual display through which the spatial associations can be visualised. Complex relationship probably can be better understood by visual display rather than from a table of data. Further a visual display can be manipulated to give alternative views/representation of the data, thereby enhancing the capability to analyse the anomalies and patterns through GIS. Visual display is obtained either on the video monitor or other output devices such as colour printers.

3. *Combination:* The ability to merge spatial data sets from quite different sources their manipulation and subsequent display can often lead to an understanding and interpretation of spatial phenomena that are simply not apparent when individual spatial data types are considered in isolation. The data measuring activity combines image data for a certain geographic area with other reference data of the same area. The GIS operator may overlay multiple images of this area at different dates—a technique used for identifying changes over time, for example, monitoring of forest fire or spreading of disease in tree species. The process of combining layers of spatial data is sometimes called *data integration* and can be carried out either by visualising composite displays of various kinds, or with integration models that effectively create a new map from two or more existing maps.

4. *Prediction:* Prediction is one of the purposes of GIS. For example, a number of data layers indicating population data in different regions of a city along with the growth patterns and civic facilities might be combined together to predict the future population at the desired time in different parts of the city. Such a map may then be used as a basis for making city development decisions. Prediction may sometimes also be a research exercise to explore the outcome of making a particular set of assumptions, often with the purpose of examining the performance of a model.

5. *Queries:* Since GIS is a decision support system, performing queries on a GIS database to retrieve information (data) is its essential part. Queries offer a method of data retrieval, and can be performed on data that are part of the GIS database, or on new data produced as a result of data analysis. These are useful at all stages of GIS analysis for checking the quality of data and the results obtained. A GIS typically stores spatial and non-spatial (also called aspatial or attribute) data in two separate files. The GIS has capability to search and display spatial data based on attribute criteria and vice versa. Accordingly, there are two general types of query that can be performed with GIS: spatial and aspatial. Aspatial queries are the questions about the attributes of features. ‘How many nursing homes are there?’

is an aspatial query since neither the question nor the answer involves analysis of the spatial component of data. This query could be performed by database software alone. A question requiring information about 'where' is a spatial query. This requires linking the data sets using location as the common key. A GIS has the capacity to satisfy the following queries:

- (a) About location: What exists at a particular location? The location of the particular region can be described in many ways using place name, post or pin code, or geographic reference, such as latitude and longitude.
- (b) Condition: This query requires spatial analysis to give an answer. Instead of identifying what exists at a certain location, one seeks to find a location where certain conditions are satisfied.
- (c) Pattern: This query is more sophisticated and important as one might want to know how many anomalies are there within an area over a time.
- (d) Trend: This query might involve both location and conditions and seeks to find differences within an area over a period of time.
- (e) Modelling: This query is posed to determine what happens if some addition or changes are done in the existing network, e.g., to determine the extent and level of contamination in an area if some toxic substances seeps into the ground water and thence to the local water supply. For answering these queries, both geographic and other information and possibly even scientific laws may be required. These queries require efficient search of data items and capability for deriving their geometric and topological attributes.

6. *Reclassification*: Although query is the most widely used function to retrieve data from a GIS database, irrespective of the vector or raster model, reclassification can also be used in place of query in the raster model. Consider a land-use image from which we require to extract information on areas of schools. The answer to this query could be obtained by creating a new coverage that eliminates all unnecessary data. Reclassification would result in a new image. For example, in a raster image, if cells representing schools in the original image had a value of 30, a set of rules for the reclassification could be

- (a) Cells with values 30 (schools) should take the new value of 1.
- (b) Cells with values other than 30 should take the new value of 0.

Such a reclassification will generate a new image with all schools coded with 1, and all the rest coded with 0. The resulting reclassified image is very useful for land use/land cover and environmental studies.

8.11 NEIGHBOURHOOD FUNCTIONS

There is a range of functions available in GIS that allow a spatial entity to influence its neighbours, or the neighbours to influence the character of an entity. The most common examples are buffering, proximity analysis and filtering.

1. *Buffer operation*: Buffering is the creation of a zone of interest around an entity. Buffering is possible in both vector and raster GIS. In the vector case, the result

is a new set of objects, while the result in the raster case is the classification of cells according to whether they lie inside or outside the buffer. Buffers are very useful for analysing landscapes, highway alignments, water supply networks and drainage studies.

In most GIS data analysis, there is more than one method of achieving an answer to a question. The trick is to find the most efficient method, and the most appropriate analysis. For example, the question, 'Which nursing homes are within 300 m of a main road?' could be approached in a number of ways. One option would be , first, to produce a buffer zone identifying all land up to 300 m from the main road; and then, to find out which nursing homes fall within this buffer zone using a point-in-polygon overlay. Then another query can be made to find the names of the nursing homes. An alternative approach would be used to measure the distance from each nursing home to a main road, and then to identify those which are less than 300 m away. Repeated measurement of distances from nursing homes to roads could be time consuming and prone to human error. Thus, the first approach using buffering would be more appropriate.

Conceptually, buffering is very simple but involves complex computational operation. If a point is buffered, a circular zone is created. Buffering lines and areas creates new areas (Fig. 8.13). Creating buffer zones around point features is the easiest operation; a circle of the required radius is simply drawn around each point. However, creating buffer zones around line and area features are a little more complicated. Some GIS do this by placing a circle of the required radius at one end of the line or area boundary to be buffered. This circle is then moved along the length of the segment. The path that the edge of the circle tangential to the link makes is used to define the boundary to the buffer zone. Sometimes, there may be a need for another buffer around a buffer. This is called a *doughnut* buffer.

2. *Proximity analysis*: While buffer zones are often created with the use of one command or option in vector GIS, a different approach is used in many raster GISs. Here, *proximity* is calculated which results in a new raster data layer where the attributes of each cell is a measure of distance. This is known as *proximity analysis*.

3. *Filtering*: Data filtering involves the recalculation/reallotment of cells in a raster image based on the characteristics of neighbours. *Filtering* is a technique used for the processing of remotely sensed imagery. Filtering will change the value of a cell based on the attributes of neighbouring cells. The filter is defined as a group of cells around a target cell. The size and shape of the filter are determined by the operator. Common filter shapes are squares and circles, and the dimensions of the filter determine the number of neighbouring cells used in the filtering process. The filter is passed across the raster data set (Fig. 8.14) and used to recalculate the value of the target cell that lies at its centre. The new value assigned to the target cell is calculated using one of a number of algorithms. Examples include the maximum cell value within the filter and the most frequent value. The raster data obtained from a classified satellite image may require filtering to 'smooth'

the noisy (erratic/fuzzy) data caused by high spatial variability in vegetation cover or problems with the data collection device.

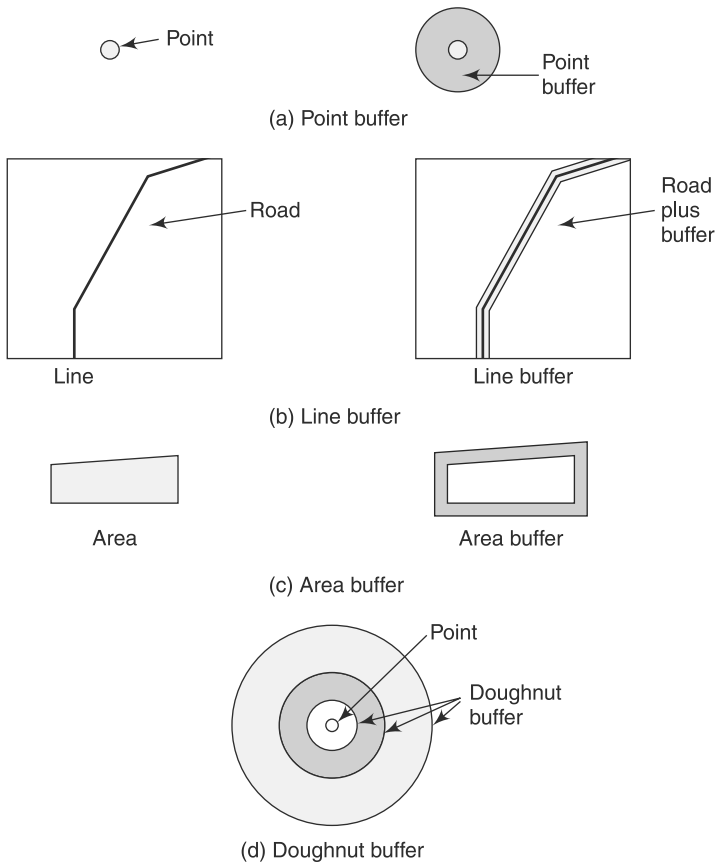


Fig. 8.13 *Types of buffers*

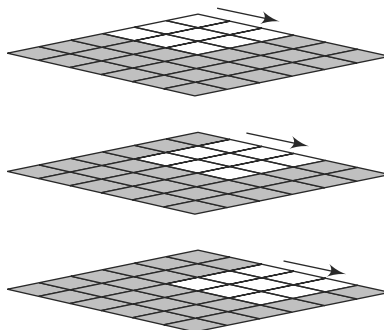


Fig. 8.14 *Filtering technique*

8.12 MAP OVERLAY ANALYSIS

The map overlay technique to integrate data from various sources is perhaps the key GIS analysis function. Using GIS, it is possible to take two different thematic map layers of the same area and overlay them one on top of the other to form a new layer through a common reference network or coordinate system. These individual layers must be spatially registered. For example, obtaining an answer to the question, 'Which nursing homes are within 300 m of a main road?' requires the use of several operations. First, a buffering operation must be applied to find all the area of land within 300 m of a main road, and then overlay function used to combine this buffer zone with the nursing home data layer. This will allow the identification of nursing homes within the buffer zone.

As with many other operations and analyses in GIS, there are differences in the way map overlay are performed between the raster and vector worlds. In vector-based systems, map overlay is time consuming, complex and computationally expensive whereas in raster-based systems, it is quick, straightforward and efficient.

The techniques of GIS map overlay are analogous to sieve-mapping in conventional methods of surveying; the overlaying of tracing of paper maps on a light table. The concept of map overlay is illustrated through a case study of soil erosion in a watershed as follows.

Figure 8.15 illustrates georeferenced data from a number of sources and is used to study the soil erosion potential. In this illustration, the data maps (a) are computer coded with respect to a grid (b). The data maps are encoded by recording the information category most dominant in each cell in the grid. That is, each cell is assigned a single soil type in the soil data file, a single cover type in the land cover file, and an average elevation in the topographic file. The activity of making the data type compatible is inherently accomplished by encoding the

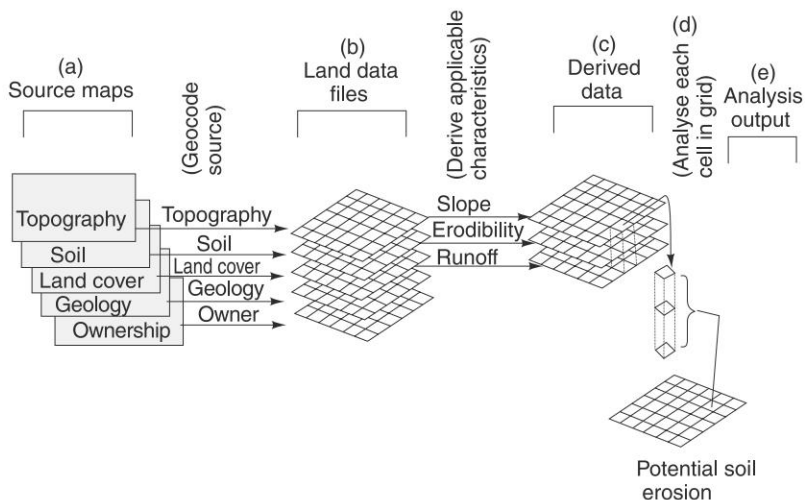


Fig. 8.15

Analysis procedure using geocoded data

maps on a common grid. The job of interpreting applicable characteristics (slope, erodibility, and run off) from the original data is a simple one for the computer. The slope information can be derived from the elevations in the topographic files. The erodibility—can be derived from the data-based management system; and the run off potential—an attribute associated with each land cover-type can also be calculated. The three sources of data can be interrelated by the analyst to identify the sites prone to soil erosion. The overlay, also called a *composite analysis*, consists of evaluating the data values within each cell in the combined grid matrix. Complex weighting schemes may be applied to increase the importance of the more critical variables. The resulting output grid can be displayed as a matrix or printed characters. The output can also be generated on a line plotter, a colour monitor, or a precision film recorder.

8.12.1 Vector Overlay Capabilities

Vector map overlay relies heavily on the two associated disciplines—geometry and topology. The overlaid data layers need to be topologically correct so that lines meet at nodes and all polygon boundaries are closed. To create topology for a new data layer produced as a result of the overlay process, the intersections of lines and polygons from the input layers need to be calculated using geometry. The three main types of vector overlay—the point-in-polygon, the line-in-polygon and the polygon-on-polygon are as shown in Fig. 8.16. The overlay of two or more data layers representing simple spatial features results in a more complex output layer. This will contain more polygons, more intersections and more line segments than either of the input layers.

The point-in-polygon overlay is used to find out the polygon in which a point falls. For example, using the point-in-polygon overlay, it is possible to find out in which land-use polygon are each of the fire stations located. Figure 8.16(a) illustrates this overlay process. On the output map, a new set of fire station points is created with additional attributes describing land use.

The line-in-polygon overlay is more complicated. Suppose that it is required to know the parts of the roads passing through the new and old city areas. To do this, we need to overlay the road data on a data layer containing city polygon. The output map will contain roads split into smaller segments representing roads in new city areas and those in the old city areas. Topological information must be retained in the output map (Fig. 8.16(b)); therefore this is more complex than either of the two input maps. The output map will contain a database record of each new road segment.

The polygon-on-polygon overlay of Fig. 8.16(c) could be used to examine the area of market in new/old city. Two input data layers—a market data layer contained in city polygons and the market boundary layer—are required. Three different outputs could be obtained which are shown in Fig. 8.16(c) and are presented below:

1. The output data layer could contain all the polygons from both the input maps. In this case, the question posed is ‘Where are areas of market or areas which are within the new/old city?’ This corresponds to the Boolean **OR** operation, or in mathematical set terms, UNION.

- The output data layer could contain the whole of the market area, and the city area within this. The boundary of the market would be used as the edge of the output map, and city areas would be cut away if they fall outside it. This operation is referred to as 'cookie cutting'. It is equivalent to the mathematical **IDENTITY** operation. The questions being answered are 'Where is the market boundary, and which areas of city are within this?' This overlay might be used for calculation of the percentage of the area of the city covered by the market.
- The output data layer could contain areas that meet both the criteria; that is, area that is both market and within the new city. An output map would be produced showing the whole of the new city polygon that are entirely covered by the market, and 'cut' away the new city polygon which crosses the market boundary. This is the mathematical **INTERSECT** operation, and the output map shows where the two input layers intersect. 'Where are market areas within the new city area?' is the question being answered.

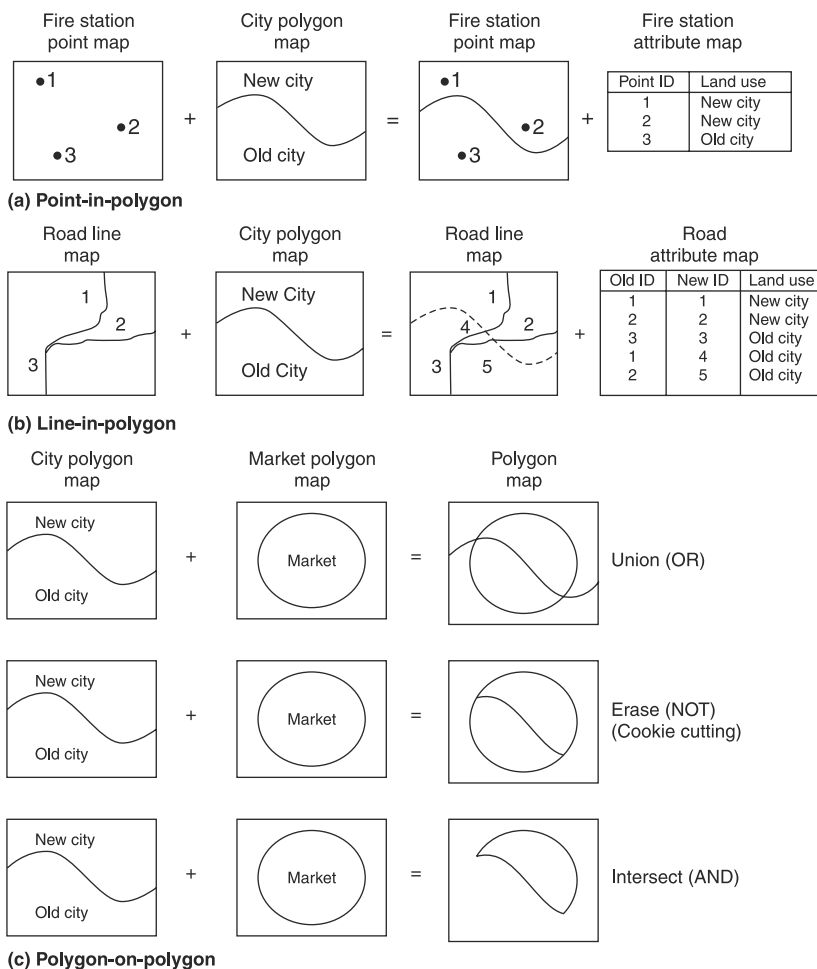


Fig. 8.16

Vector overlays

8.12.2 Raster Overlay Capabilities

In the raster data structure, everything is represented by grid cells—a point is represented by a single cell, a line by a string of cells and an area by a group of cells. A raster map overlay introduces the idea of *map algebra* or *mapematics*. Using map algebra, input data layers may be added, subtracted, multiplied or divided to produce output. Mathematical operations are performed on individual cell values from two or more input layers to produce an output value. Thus, the most important consideration in raster overlay is the appropriate coding of point, line and area features in the input data layers.

Consider five of the data layers of a hill station that have been registered and are as follows.

Layer	Code
1. Location of nursing home	1
2. Road	2
3. Agriculture land	3
4. Land use	
(i) Habitat	1
(ii) Water	2
(iii) Agriculture land	4
(iv) Forest	5
5. Hill Station	10

Note On all data layers, '0' is the value given to cells that do not contain features of interest.

To find out how many nursing homes are contained within the hill station, an operation equivalent to the vector point-in-polygon overlay is required. The two data layers may be added as shown in Fig. 8.17(a). The output map would contain cells with the following values:

- 0 for cells outside the hill station boundary and without nursing homes
- 1 for cells containing nursing homes, but outside the hill station boundary
- 10 for cells inside the hill station boundary, but without nursing homes
- 11 for cells inside the hill station boundary and containing nursing homes

To know about the sections/parts of roads that pass through forest area, an operation equivalent to the vector line-in-polygon method (Fig. 8.17(b)) is required. This would require the roads data layer, and reclassified version of the land use map that contain only forest area. The two data layers will be added.

The output map would contain cells with the following values:

- 0 for cells with neither roads nor forest present;
- 2 for cells with roads, but outside forest areas;
- 5 for cells with forest present, but roads absent;
- 7 for cells with both forest and roads present.

Note If the value '2' for a road was added to land-use codes, the new value for a cell could be the same as that for another land use type (for example, a road value of 2 + water value of 2 = 4 (which is the same as the value here for an agriculture land).

Thus, the coding of raster images used in overlay is very important, and frequently users employ Boolean images (using only codes 1 and 0) so that algebraic equations will produce a meaningful answer.

Nursing home	Hill stations	Output																												
<table> <tr><td>1</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>0</td></tr> </table>	1	0	0	0	0	1	0	0	0	<table> <tr><td>10</td><td>10</td><td>10</td></tr> <tr><td>10</td><td>10</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td></tr> </table>	10	10	10	10	10	0	0	0	0	<table> <tr><td>11</td><td>10</td><td>10</td></tr> <tr><td>10</td><td>10</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>0</td></tr> </table>	11	10	10	10	10	1	0	0	0	Neither hill stations nor Nursing home = 0 Nursing home not in hill station = 1 Hill station, no Nursing home = 10 Nursing home in hill station = 11
1	0	0																												
0	0	1																												
0	0	0																												
10	10	10																												
10	10	0																												
0	0	0																												
11	10	10																												
10	10	1																												
0	0	0																												

Nursing home = 1 Hill stations = 10

(a) Point-in-polygon (using add)

Road		Forest		Output																													
<table><tr><td>2</td><td>0</td><td>0</td></tr><tr><td>0</td><td>2</td><td>0</td></tr><tr><td>0</td><td>0</td><td>2</td></tr></table>	2	0	0	0	2	0	0	0	2	+	<table><tr><td>0</td><td>5</td><td>5</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>5</td><td>5</td><td>5</td></tr></table>	0	5	5	0	0	0	5	5	5	=	<table><tr><td>2</td><td>5</td><td>5</td></tr><tr><td>0</td><td>2</td><td>0</td></tr><tr><td>5</td><td>5</td><td>7</td></tr></table>	2	5	5	0	2	0	5	5	7	Neither road nor forest	= 0
2	0	0																															
0	2	0																															
0	0	2																															
0	5	5																															
0	0	0																															
5	5	5																															
2	5	5																															
0	2	0																															
5	5	7																															
					Road not in forest	= 2																											
					Forest, no road	= 5																											
					Road in forest area	= 7																											
Road = 2		Forest = 5																															
Other areas = 0		Other areas = 0																															

(b) Line-in-polygon (using add)

Forest		Hill station		Output																												
<table><tr><td>0</td><td>5</td><td>5</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>5</td><td>5</td><td>5</td></tr></table>	0	5	5	0	0	0	5	5	5	+	<table><tr><td>10</td><td>10</td><td>10</td></tr><tr><td>10</td><td>10</td><td>0</td></tr><tr><td>0</td><td>0</td><td>0</td></tr></table>	10	10	10	10	10	0	0	0	0	=	<table><tr><td>10</td><td>15</td><td>15</td></tr><tr><td>10</td><td>10</td><td>0</td></tr><tr><td>5</td><td>5</td><td>5</td></tr></table>	10	15	15	10	10	0	5	5	5	Neither road nor hill station = 0 Forest not in hill station = 5 Hill station, no forest = 10 Forest in hill station = 15
0	5	5																														
0	0	0																														
5	5	5																														
10	10	10																														
10	10	0																														
0	0	0																														
10	15	15																														
10	10	0																														
5	5	5																														
Forest = 5		Hill stations = 10																														
Other areas = 0		Other areas = 0																														

(c) Polygon-on-polygon (using add)

Forest		Hill station		Output																												
<table><tr><td>0</td><td>1</td><td>1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>1</td><td>1</td><td>1</td></tr></table>	0	1	1	0	0	0	1	1	1	+	<table><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>1</td><td>1</td><td>0</td></tr><tr><td>0</td><td>0</td><td>0</td></tr></table>	1	1	1	1	1	0	0	0	0	=	<table><tr><td>1</td><td>2</td><td>2</td></tr><tr><td>1</td><td>1</td><td>0</td></tr><tr><td>1</td><td>1</td><td>1</td></tr></table>	1	2	2	1	1	0	1	1	1	
0	1	1																														
0	0	0																														
1	1	1																														
1	1	1																														
1	1	0																														
0	0	0																														
1	2	2																														
1	1	0																														
1	1	1																														
Neither forest nor hill station					= 0																											
Forest or hill station					= 1																											
Forest and hill station					= 2																											
Forest = 1		Hill stations = 1																														
Other areas = 0		Other areas = 0																														

Forest		Hill station		Output																													
<table><tr><td>0</td><td>1</td><td>1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>1</td><td>1</td><td>1</td></tr></table>	0	1	1	0	0	0	1	1	1	+	<table><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>1</td><td>1</td><td>0</td></tr><tr><td>0</td><td>0</td><td>0</td></tr></table>	1	1	1	1	1	0	0	0	0	=	<table><tr><td>0</td><td>1</td><td>1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td><td>0</td></tr></table>	0	1	1	0	0	0	0	0	0	Forest and hill station	= 1
0	1	1																															
0	0	0																															
1	1	1																															
1	1	1																															
1	1	0																															
0	0	0																															
0	1	1																															
0	0	0																															
0	0	0																															
					Other areas	= 0																											
Forest = 1		Hill stations = 1																															
Other areas = 0		Other areas = 0																															

(d) Polygon-on-polygon (Boolean alternatives)

Fig. 8.17 Raster overlays

The polygon-on-polygon analysis is conducted in just the same way (Fig. 8.17(c)). For example, adding the forests layer and the hill station boundary would produce an output layer with the following codes:

- 0 for cells outside the hill station boundary and with forest absent
- 5 for cells outside the hill station boundary and with forest present
- 10 for cells inside the hill station boundary and with forest absent
- 15 for cells inside the hill station boundary and with forest present

The output map is equivalent to a union polygon-on-polygon overlay in vector GIS.

Reclassification will produce variants of this, and other overlay operations are available, by multiplying, subtracting or dividing the data layers. The algebraic manipulation of images in raster GIS is a powerful and flexible way of combining data and organising analysis. Equations can be written with maps as variables to allow the development of spatial models. Figure 8.17(d) represents polygon operation using Boolean alternative.

8.13 DATA QUALITY

The success of any GIS application depends on the quality of the geographic data used. Collecting high-quality geographic data for input to GIS is therefore an important activity. In GIS, *data quality* is used to give an indication of how good the data are. It describes the overall fitness or suitability of data for a specific purpose or is used to indicate data free from errors and other problems. Some pointers for gauging the overall quality of GIS database are *error*, *accuracy*, *precision* and *bias*. In addition, the *resolution* and *generalisation* of source data, and the data model used, may influence the portrayal of features of interest. Data sets used for analysis need to be complete, compatible and consistent, and applicable for the analysis being performed. These concepts are explained below.

1. *Flaws* in data are usually referred to as errors. Error is the physical difference between the real world and the GIS facsimile. Errors may be single, definable departures from reality, or may be persistent, widespread deviations throughout a whole database.

2. *Accuracy* is the extent to which an estimated data value approaches its true value. If a GIS database is accurate, it is a true representation of reality. It is impossible for a GIS database to be 100% accurate, though it is possible to have data that are accurate to within specified tolerances.

3. *Precision* is the recorded level of detail of the data. A coordinate in metres to the nearest ten decimal places is more precise than one specified to the nearest three decimal places. Computers store data with a high level of precision, though a high level of precision does not imply a high level of accuracy.

4. *Bias* in GIS data is the systematic variation of data from reality. Bias is a consistent error throughout a data set. A consistent overshoot in digitised data caused by a badly calibrated digitiser, or the consistent truncation of the decimal points from data values by a software program, are possible examples. These examples have a technical source. However, human sources of bias also exist. An aerial photograph interpreter may have a consistent tendency to ignore all features below a certain size. Although such consistent errors should be easy to rectify, they are often very difficult to spot.

5. *Resolution* is the term used to describe the smallest feature in a data set that can be displayed or mapped. In raster GIS, resolution is determined by cell size. For example, for a raster data set with a 20 m cell size, only those features that

are 20 m × 20 m or larger can be distinguished. At this resolution it is possible to map large features such as fields, lakes and urban areas but not individual trees or telegraph poles. Vector data can also have resolution, although this is described in different terms. Resolution is dependent on the scale of the original map, the point size and line width of the features represented thereon, and the precision of digitising.

6. *Generalisation* is the process of simplifying the complexities of the real world to produce scale models and maps. Cartographic generalisation is a subject in itself and is the cause of many errors in GIS data derived from maps. It is the subjective process by which the cartographer selectively removes the enormous detail of the real world in order to make it understandable and attractive in map form.

7. *Completeness*: A complete data set covers the study area and the time period of interest in its entirety. The data should be complete spatially and temporally, and should have a complete set of attribute information.

8. *Compatibility* data sets can be used together sensibly. With GIS it is possible to overlay two maps, one originally mapped at scale of 1:500 000 and the other at 1:25 000. The result, however, is largely worthless because of incompatibility between the scales of the source documents. Maps containing data measured in different scales of measurement cannot be combined easily. To ensure compatibility, ideally data sets should be developed using similar methods of data capture, storage, manipulation and editing.

9. *Consistency* applies not only to separate data sets but also within individual data sets. Inconsistencies can occur within data sets where sections have come from different source documents or have been digitised by different people. This will cause spatial variation in the error characteristics of the final data layer. One area of the final data set may contain more error than another. Problems of inconsistency also come from the manner in which the data were collected.

10. *Applicability* is a term used to describe the appropriateness or suitability of data for a set of commands, operations or analyses.

8.14 SOURCES OF ERRORS IN GIS

Spatial and attribute errors can occur at any stage in a GIS project. They may arise during the definition of spatial entities, from the representation of these entities in the computer, or from the use of data in analysis. In addition, they may be present in source data, arise during conversion of data to digital format, occur during data manipulations and processing, or even be produced during the presentation of results.

8.14.1 Errors Arising from Understanding and Modelling of Reality

Errors can originate from the ways in which we perceive, study and model reality. These errors can be termed conceptual errors, since they are associated with the representation of the real world for study and communication.

The different ways in which people perceive reality can have effects on how they model the world using GIS. The perception of reality influences the definition of reality, and in turn the use of spatial data. This can create real errors and often gives rise to inconsistencies between data collected by different surveyors, maps drawn by different cartographers, and databases created by different GIS users.

In geography, and GIS, spatial models are used to reflect reality. The main models in use are raster, vector, object-oriented and layer based. All of these spatial models have limitations when it comes to portraying reality. For instance, the raster model assumes that all real-world features can be represented as individual cells. This is clearly not the case. The vector model assumes that all features can be given a single coordinate or a collection of Cartesian coordinates. The world is actually made up of physical and biological materials, which is, in turn, made up of molecular and submolecular matter grouped into complex systems linked by flows of energy and materials (solids, liquids and gases). Whatever GIS model we adopt, it is a simplification of this reality, and any simplification of reality will include errors of generalisation, completeness and consistency.

8.14.2 Errors in Source Data for GIS

The models of reality in GIS are built from a variety of data sources including survey data, remotely sensed and map data. All sources of spatial and attribute data for GIS are likely to include errors.

Survey data can contain errors due to mistakes made by people operating the equipments or recording the observations, or due to technical problems with equipments.

Remotely sensed and aerial photography data could have spatial errors if they were spatially referenced wrongly, and mistakes in classification and interpretation would create attribute errors.

Maps are probably the most frequently used sources of data for GIS. Maps contain both relatively straightforward spatial and attribute errors caused by human or equipment failings, and more subtle errors, introduced as a result of the cartographic techniques employed in the map-making process. Generalisation is one cartographic techniques that may introduce errors.

8.14.3 Errors in Data Encoding

Data encoding is the process by which data are transferred from some non-GIS source, such as the paper map, satellite image or survey, into a GIS format. The method of data encoding, and the conditions under which it is carried out, are perhaps the greatest source of error in most GIS. Digitising, both manual and automatic, is an important method of data entry. Despite the availability of hardware for automatic conversion of paper maps into digital form, much of the digitising of paper maps is still done using a manual digitising table. Manual digitising is recognised by researchers as one of the main sources of error in GIS; however, digitising error is often largely ignored.

Sources of error within the digitising process are many, but may be broken down into two main types: *source map error* and *operational error*. Operational

errors are those introduced and propagated during the digitising process. Human operations can compound errors present in an original map and add their own distinctive error signature.

Automatic digitising, like manual digitising, requires correct registration of the map document before digitising commences, but there the similarity ends. By far the most common method of automatic digitising is the use of a raster scanner. This input device suffers from the same problems regarding resolution as the raster data model.

8.14.4 Errors in Data Editing and Conversion

After data encoding is complete, cleaning and editing are almost always required. These procedures are the last line of defence against errors before the data are used for analysis. Of course, it is impossible to spot and remove all the errors, but many problems can be eliminated by careful scrutiny of the data.

A different problem occurs when automated techniques are used to clean raster data. The main problem requiring attention is ‘noise’—the misclassification of cells. Noise can be easy to spot where it produces a regular pattern, such as striping. At other times, it may be more difficult to identify as it occurs as randomly scattered cells. These noise errors can be rectified by filtering the raster data to reclassify single cell or small groups of cells by matching them with general trends in the data. The ‘noisy’ cells are given the same value as their neighbouring cells.

After cleaning and editing data it may be necessary to convert the data from vector to raster or vice versa. During vector-to-raster conversion both the size of the raster and the method of rasterisation used have important implications for positional error and, in some cases, attribute uncertainty. The smaller the cell size, the greater is the precision of the resulting data. Finer raster sizes can trace the path of a line more precisely and therefore help to reduce classification error—a form of attribute error. Positional and attribute errors as a result of generalisation are seen as classification error in cells along the vector polygon boundary. The conversion of data from raster to vector format is largely a question of geometric conversion; however, certain topological ambiguities can occur, such as where differently coded raster cells join at corners.

8.14.5 Errors in Data Processing and Analysis

Errors may be introduced during the manipulation and analysis of the GIS database. GIS users must ask themselves questions before initiating a GIS analysis. For example: Are the data suitable for this analysis? Are they in a suitable format? Are the data sets compatible? Are the data relevant? Will the output mean anything? Is the proposed technique appropriate to the desired output? These questions may seem obvious but there are many examples of inappropriate analysis. These include the inappropriate phrasing of spatial queries, overlaying maps which have different coordinate systems, combining maps which have attributes measured in incompatible units, using maps together that have been derived from source data of widely different map scales, and using an exact and abrupt method of interpolation to interpolate approximate and gradual point data.

GIS operations that can introduce errors include the classification of data, aggregation or disaggregation of area data and the integration of data using overlay techniques.

Classification errors also affect raster data. Classified satellite images provide a reflectance value for each pixel within a specific wavelength range or spectral band (for example, red, near infrared or microwave). Raster maps of environmental variables, such as surface cover type, are derived by classifying each pixel in the image according to typical reflectance values for the range of individual cover types present in the image. Error can occur where different land cover types have similar reflectance values and where shadows cast by terrain, trees or buildings reduce the reflectance value of the surface. Careful choice of classification method can help to reduce this type of error.

Where a certain level of spatial resolution or a certain set of polygon boundaries are required, data sets that are not mapped with these may need to be *aggregated* or *disaggregated* to the required level. This is not a problem if the data need to be aggregated from smaller areas into larger areas, provided that the smaller areas nest hierarchically into the larger areas. Problems with error do occur, however, if we wish to disaggregate our data into smaller areas or aggregate into larger non-hierarchical units. The information required to decide how the attribute data associated with the available units aggregate into the larger but non-nested units or disaggregate into a set of smaller units, rarely exists.

Error arising from map overlay in GIS is a major concern and has correspondingly received much attention in the GIS literature. This is primarily because much of the analysis performed using GIS consists of the overlay of categorical maps (where the data are presented in a series of categories). GIS allows the quantitative treatment of these data (for example, surface interpolation or spatial autocorrelation), which may be inappropriate. Map overlay in GIS uses positional information to construct new zones from input map layers using Boolean logic or 'mapemantics'. Consequently, positional and attribute errors present in the input map layers will be transferred to the output map, together with additional error introduced by multiplicative effects and other internal sources. Data output from a map overlay procedure are only as good as the worst data input to the process.

Perhaps the most visual effect to positional error in vector map overlay is the generation of sliver polygons. Slivers (or 'weird' polygons) occur when two maps containing common boundaries are overlaid. If the common boundaries in the two separate maps have been digitised separately, the coordinates defining the boundaries may be slightly different as the result of digitising error. When the map are overlaid, a series of small, thin polygons will be formed where the common boundaries overlap (Fig. 8.18). Slivers may also be produced when maps from two different scales are overlaid. Of course, sliver polygons can and do occur by chance, but genuine sliver polygons are relatively easy to spot by their location along common boundaries and their physical arrangement as long thin polygonal chains.

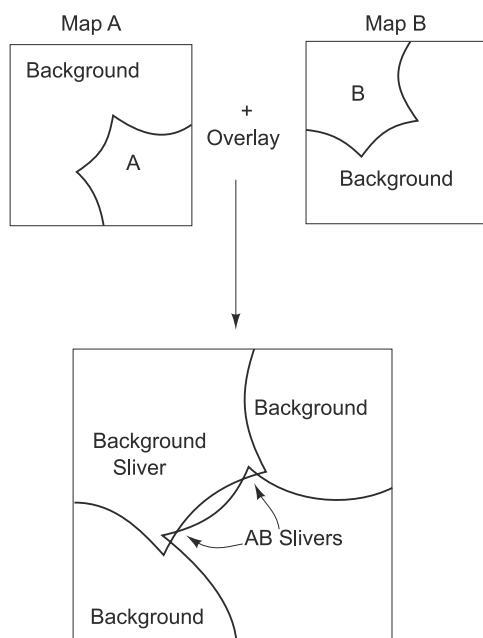


Fig. 8.18 *Generation of sliver polygons*

8.14.6 Error in Data Output

From the preceding discussion it should be clear that all GIS database will contain error. In addition, further errors will be introduced during manipulation and analysis of the data. Therefore, it is inevitable that all GIS output, whether in the form of a paper map or a digital database, will contain inaccuracies. The extent of these inaccuracies will depend on the care and attention paid during the construction, manipulation and analysis of the databases. It is also possible that errors can be introduced when preparing GIS output.

8.15 APPLICATIONS OF GIS

The concept of geographic information infrastructure has brought about a dramatic philosophical and technological revolution in the development of GIS. Instead of being used simply as a set of software tools for processing and analysing geographic data stored locally, GIS is now a gateway for accessing and integrating geographic data from different sources located locally and globally, increasingly used for interactive visualisation of scenarios resulting from different business decisions, as well as for the communication of spatial knowledge and intelligence among people all over the world. GIS has popularised the use of geographic information by empowering individuals and organisations to use such information in areas that earlier generations of GIS user could never have thought of even with their wildest imagination. It is now common place for ordinary people to use GIS to check the weather and traffic conditions before they leave home for work and

find information about the country or city they are about to visit. Increasingly, business people rely on GIS to identify locations where to set up their new shops and to determine the best routes to deliver their goods and services. At the same time, GIS has become an indispensable tool for government officials to manage land and natural resources, monitor the environment, formulate economic and community development strategies, enforce law and order and deliver social services. Major application areas in GIS are listed in Table 8.2.

Table 8.2 *Major application areas of GIS*

<i>Sectors</i>	<i>Application Areas</i>
Academic	Research in engineering, science and humanities. Primary and secondary schools—school district delineation, facilities, management, bus routing, spatial digital libraries.
Industry	Engineering—surveying and mapping, site and landscape development, pavement management. Transportation—route selection for goods delivery, public transit, vehicle tracking. Utilities and communications—electricity and gas distribution, pipelines, telecommunication networks. Forestry—forest resource inventory, harvest planning, wildlife management and conservation. Mining and mineral exploration. Systems consulting and integration.
Business	Banking and insurance. Real estate—development project planning and management, sales and renting services, building management. Retail and market analysis. Delivery of goods and services.
Government	Central government—national topographic mapping, resource and environmental management, weather services, public land management, population census, election, and voting. State government—surveying and mapping, land and resources management, highway planning and management. Local/municipal government—social and community development, land registration and property assessment, water and wastewater services. Public safety and law enforcement—crime analysis, deployment of human resources, community policing, emergency planning and management. Health care. International development and humanitarian relief.
Military	Training. Command and control. Intelligence gathering.

8.16 SELECTIVE GIS SOFTWARES

A GIS requires specialised programs tailored for manipulation of geographic data. Other kinds of databases may have very large volumes of data, but do not need to retain *locational* information for data. Therefore, GIS software must satisfy the special needs of the analyst who needs to reference data by geographic location. Furthermore, the GIS must provide the analyst with the capability to solve the special problems that arise whenever maps or images are examined—the problems of changing coordinate systems, matching images, bringing different images into registration, and so on. A GIS must be supported by the ability to perform certain operations related to the geographic character of the data. For example, it must be capable of identifying data by location or by specified areas in order to retrieve and to display data in a map-like image. Thus, the GIS permit the analyst to display data in a map-like format, so that geographic patterns and interrelationships are visible to the analyst.

Furthermore, the software for a GIS must be able to perform operations that relate values at one location to those at neighbouring locations. For example, to compile slope information from elevation data, it is necessary to examine not only specific elevation values, but also those at neighbouring locations, in order to calculate the magnitude and directions of the topographic gradient.

A GIS, of course, consists not only of a single data set, but also of many that together show several kinds of information for the same geographic area. Thus, a GIS may include data for topographic elevation, streams and rivers, land use, political and administrative boundaries, power line, and other variables. This combined data set is useful only if the several overlays register to one another exactly, and therefore the several kinds of data must share a common coordinate system, because separate variables are likely to be derived from quite different reference systems, and different cartographic projections. Thus, a GIS must have special programs to bring data into registration by changing scale and geometric qualities of the data.

The current trend of GIS software development is to move away from the proprietary development environment to open industry standards. It is now possible to build application software modules with programming languages, such as Visual Basic, Visual C++, and Power Builder, and to integrate them with the GIS functions originally supplied by the software vendor. The concepts and techniques of using generic computer languages to build GIS applications are based on the use of *component software*. This is a software engineering methodology that has been evolving since the early 1990s. There has been considerable success in using this approach to effectively address the integration of separate computer-based applications such as document imaging, optical character recognition, database query, and fax. GIS application can obviously be benefited from this new approach to software development.

8.16.1 Integrated Land and Water Information System (ILWIS) 3.1

ILWIS integrates image, vector and thematic data in one unique and powerful package on the desktop. ILWIS delivers a wide range of features including import/

export, digitising, editing, analysis and display of data as well as production of quality maps.

Key Features

1. Integrated raster and vector design
2. Import and export of widely used data formats
3. On-screen and table digitising
4. Comprehensive set of image processing tools
5. Orthophoto, image georeferencing, transformation and mosaicing
6. Advanced modelling and spatial data analysis
7. 3D visualisation with interactive editing for optimal view findings
8. Rich projection and coordinate system library
9. Geo-statistical analyses for improved interpolation
10. Production and visualisation of stereo image pairs

Application Areas

Earth sciences

Applied geomorphology and natural hazards	1. Hazard, vulnerability and risk analysis 2. Flood hazard analysis using multitemporal satellite 3. Modelling cyclone hazard 4. Modelling erosion potential of catchment 5. Statistical landslide hazard analysis 6. Deterministic landslide hazard zonation 7. Seismic landslide hazard zonation
Engineering geology	8. Creating an engineering geological database
Surface hydrology	9. Irrigation water requirement 10. Irrigation area characteristics 11. Determination of peak run off 12. Morgan approach for erosion modelling
Hydro-geology	13. Assessing aquifer vulnerability to pollution
Geology	14. Remote sensing and GIS techniques applied to geological survey 15. Geological data integration

Environment management

Neighbourhood modelling	16. Modelling with neighbourhood operators 17. Extracting topographic and terrain variables for distributed models
Data combination	18. Tools for map analysis applied to the selection of a waste disposal site

Land resource and urban surveys

Urban surveys	19. Updating a land use map with oblique air photos 20. Analysis of urban change and spatial pattern 21. Analysis of suitability for urban expansion

Land use planning
Soil surveys

22. Analysing the fuelwood demand
23. The erosion issue
24. Soil erosion modelling

Cartography

Cartography

25. Layer tinting and shading

8.16.2 Bentley Software Solutions for GIS Applications

1. *MicroStation GeoGraphics*: Full blown GIS solution, which provides tools to input, validate, manage, analyse, and visualise geographic information. MicroStation GeoGraphics also has the ability to manage a diverse combination of vector data, raster data, and non-spatial attributes.

2. *MicroStation descartes*: A comprehensive imaging solution, which provides high-performance raster image display and editing capabilities with functionality of semi-automatic digitisation of raster images. Supports engineering, mapping, and conversion projects that require the use of images, and may be used to create hybrid maps, plan sheets, and other visual documents.

3. *MicroStation GeoWater*: Intelligent geoengineering solution for creation and management of municipal water supply network. Provides a complete set of predefined water facilities like main supply and service pipes, valves, pumps, fire hydrants and other system components. Supports automatic service load calculations and burst-pipe isolation traces.

4. *MicroStation GeoWasteWater*: Intelligent geoengineering solution for creation and management of municipal wastewater supply network. Provides a complete set of predefined facilities like wastewater, storm and combined sewer pipes, manholes, pumps and other components. (User can also define additional component types via the tools and menu interfaces). The software supports the surface flow calculation for catchment areas and upstream/downstream traces.

5. *GeoPak Site*: A 3D GIS solution, which provides visual interactive environment for site design, utilising a set of very dynamic and flexible tools.

All the above solutions are based on the one common CAD base, i.e., MicroStation. This allows seamless interoperability between these solutions with the help of referencing files made in one software into the other. The powerful CAD features of MicroStation facilitate in faster and easier modelling and native support for drawing (AutoCAD) files allows interoperability with AutoCAD.

8.16.3 ArcGIS by ESRI, USA

ESRI ArcGIS is a family of software products that form a complete Geographic Information System (GIS) built on industry standards that provide exceptional, yet easy-to-use capabilities out of the box. ArcView, ArcEditor, and ArcInfo software comprise the ArcGIS desktop products, while ArcSDE and ArcIMS software make up the Application Services products.

Key features of ArcGIS are easy to use, extremely functional, scalable, web-enabled and developer-friendly.

The software follows standards for user interface design and interoperability. ArcGIS uses the following standard information technology (IT) components: Visual Basic for Application for customisation, commercial Database Management System (DBMS) for data storage, and TCP/IP and HTTP for networks.

1. *ArcView*: ArcView is an exceptional standalone desktop GIS product, as well as the entry point into ArcGIS, and provides core mapping and GIS functionality. ArcView provides geographic data visualisation, query, analysis, and integration capabilities along with the ability to create and edit geographic data.

2. *ArcEditor*: ArcEditor includes all the functionality of ArcView and adds the power to edit topologically integrated features in a geodatabase or coverage. Additional functionality includes support for multiuser editing, versioning, custom feature classes, feature-linked annotation, and dimensioning. ArcEditor allows to create and edit all ESRI-supported vector data formats.

3. *ArcInfo*: ArcInfo is the complete GIS data creation, update, query, mapping, and analysis system. Within the ArcGIS software family, ArcInfo is the top of the line. It includes all the functionality of ArcView and ArcEditor and adds the advanced geoprocessing capabilities that make ArcInfo the de Facto standard for GIS.

4. *ArcGIS Extensions* include ArcGIS Spatial Analyst, ArcGIS 3D Analyst, ArcGIS Geostatistical Analyst, ArcGIS StreetMap, ArcPress for ArcGIS, MrSID Encoder for ArcGIS.

8.16.4 Geomatica Software Solutions for GIS

Geomatica from PCI Geomatics, unites technologies that are dedicated to remote sensing image processing, GIS/spatial analysis, cartography, and desktop photogrammetry into a single integrated environment. The new product is greater than the sum of its parts. Geomatics professionals benefit from having all of these tools available from a single focal point: a new state-of-the-art viewing environment.

1. *Geomatica FreeView* is the new viewing environment for working with a variety of data, including imagery, vectors, and graphical bitmaps. Free View allows the viewing, enhancing, and examination of remotely sensed imagery such as LANDSAT, SPOT, RADARSAT, ERS-1, NOAA, AVHRR, and aerial photography. FreeView is used to overlay GIS data and view the associated attribute data.

2. *Geomatica GeoGateway* is the smart data translation tool that includes a viewing environment for working with a variety of data including imagery vectors, and graphical bitmaps.

3. *Geomatica Fundamentals* is the powerful, low-cost solution with tools for image geocorrection, data visualisation and analysis, and cartographic map production. Fundamentals include:

Geomatica, PCI's new visual environment for working with a variety of data, including imagery, vectors, and graphical bitmaps. It is an application for viewing, enhancing and examining remotely sensed imagery such as LANDSAT, SPOT,

RADARSAT, ERS-1, NOAA, AVHRR and aerial photography. It also enables users to overlay their GIS data and view the attribute information associated with such data. Focus includes many useful display tools, such as fast roam and zoom, image enhancements, numeric values display and attribute table display.

4. *Geomatica prime* contains the full ensemble of integrated tools for answering all of the geomatics needs. This includes the functionality from *Geomatica Fundamentals* plus new component architecture, access to PCI's wealth of over 250 processing algorithms, and scripting functionality using EASI, Visual Basic, and JAVA.

5. *PCI Modeler* the intuitive and powerful visual-scripting environment allowing users to construct powerful processing models via. a simple "point and click" interface.

6. *EASI* is an application-independent host environment that can handle all interactions between the user and the system. It can be used for scripting, process automation and command line interface.

7. *Fly!* is a powerful terrain visualisation tool which drapes airphoto and satellite imagery and vectors over DEM data to create 3D perspective scenes in near real time. The adverse weather conditions can be simulated with FLY! fogging feature.

Geomatica also includes a number of add-on modules, allowing clients to tailor their solution to exactly what they need. Add-on modules include: optical, radar, spatial analysis, cartography, hyperspectral analysis, orthoengine.

8.16.5 MapInfo

MapInfo developed by MapInfo Corporation, New York, comes under desktop mapping field and is used in windows environment. The various features of MapInfo include redistricting, raster image display, thematic mapping and layer control. It gives the processing power of the databases (including powerful SQL queries) a visual power of maps.

MapInfo is a potential tool for data analysis, sales and presentations. It can easily interact with already created files in dBASE, FoxBASE, delimited ASCII, Lotus 1-2-3 and Microsoft Excel. It enables multiple views of data in three formats: Map windows, Browser and Graph windows. It has raster underlay capability to add more dimensions to the map. The query capabilities range from simple selection of data from single file to complex SQL queries from one or more files. Also, it has the capability to change the projection of the map for display or digitisation.

8.16.6 Geomedia Professional

Based on Jupiter technology from Intergraph Corporation, this product is an enterprise GIS for the Windows 98 and Windows NT operating systems. This product is the perfect tool for collecting GIS data, populating an enterprise database, and turning information into precise finished maps for distribution and presentation.

As a viewing and analysis tool, this product allows the combining of geographic data from different sources, in different formats, and with different map projections, all into a single environment. Using this software, complex queries on spatial and attribute data from various sources can be performed and numerous views of highly sophisticated maps can be produced in a single workspace. Furthermore, integrating this software with SmartSketch gives the capability of printing those map views on a single sheet and adding borders, marginalia, and other finishing touches.

As a data capture and maintenance tool, this product allows to capture and to edit data more easily, faster and with more intelligence than other products. Its integrated vector and raster snaps allow capturing vector data from raster images, automatically identifying snap points to ensure accurate heads-up digitising. The software also provides table-top digitising and vector transformation for data requiring geometry transformation. Using the software, clean, accurate data can be captured the first time, thus minimising editing. Automatic vector breaking and coincident geometry digitising allow avoiding traditional data-capture problems. However, data-capture problems with automatic error detection can be located and then can be corrected with intelligent feature placement and editing tools. Furthermore, the data can be quickly annotated with powerful labelling and text-placement tools.

This product is also a software-development environment, and one can customise it with standard windows-development tools such as Microsoft Visual Basic and Visual C++.

8.17 REMOTE SENSING AND GIS

As our earth constantly changes, remote sensing from an orbital platform in space provides the easiest means to keep the GIS database up-to-date. GISs have evolved side by side with remote sensing systems, so that there are many close relationships between the two.

1. Remote sensing systems contribute data to GISs. Remotely sensed data can provide timely information at low cost and in a form that is compatible with requirements of GIS. It is easy to see some of the advantages for using remotely sensed data in a GIS. Satellite systems such as IRS-1C, IRS-1D, IKONOS, etc., acquire data for very large areas in a short time period, thereby providing essentially uniform coverage with respect to data and level of detail. Such data are already in digital form and are provided in more or less standard formats. Furthermore, these data are available for almost all the earth's land areas and are inexpensive relative to alternative sources. Although raw satellite data are not planimetrically correct, preprocessing can often bring data to acceptable levels of geometric accuracy with only modest effort. Images formed by analysis or interpretation of the raw data register to the original data. Thus, remotely sensed data have the potential to address some of the difficult problems encountered in the formation of GISs.
2. Both GISs and digital remote sensing systems use similar equipment and similar computer programs for analysis and display. Therefore, expertise in one field tends to form the foundation for work in the other.

3. The non-remote sensing data from a GIS can be used to assist in the analysis of remotely sensed images.

Thus, remote sensing and GISs have natural, mutually supporting relationships to each other.

Summary

Geographic information systems are computer based systems that can deal with any type of information about the features that can be referred by geographical location. These systems are capable of handling both locational data and attribute data about such features. For analysing and querying the computer stored large database, geographic information system has emerged as the most effective technique. Earlier the data in the database were connected by spatial relationship, usually coordinates, and each point in data base was connected to attribute data. But with the advancements in computer technology, the database can be augmented by the addition of digitised data from a wide variety of additional sources such as remote sensing data, photographic products, images, etc. The stored data can be assessed for query, or analysis, or can be manipulated stored and delivered in a variety of report formats using GIS software.

GIS have three main elements: computer hardware, software, and data. There are five major component sub-systems: input (deals with creating an image based GIS from multigeodata sets), management (efficient storage, retrieval, and database management), processing (data manipulation, feature enhancement, classification, etc.), display of geo-referenced data to solve complex problems regarding planning and management of resources, and output.

Introduction of topological techniques permit the data to be connected in the relational sense, in addition to the spatial connection. Consequently, making it possible to record and analyse descriptive characteristics about the features. In addition to the location of a point, line, or an area the features can also be analysed with respect to topology—adjacency, connectivity, and direction.

The ability of stored data on feature-unique layers permits production of spatial-feature map. Also, any selected layers or combination of layers can be depicted on a map at any desired scale. The ability to spatially interrelate multiple types of information stemming from a range of sources (the overlay analysis) makes GIS the most powerful tool. The neighbouring functions: buffer operation, proximity analysis and filtering used for facilitating analysis are described.

Most GIS use either raster (grid cell) or vector (polygon) format to represent the locational component of geographic information. In raster data approach, location of objects is defined by the row and column position of the cells they occupy. The value stored for each cell indicates the type of object. In vector format approach, the feature boundaries are converted to straight-sided polygons approximating the original regions and encoded. Both the formats have their own advantages and disadvantages. The conversion of the two data structures from one form to other is presented with examples. The vector and raster overlay capabilities are also discussed with examples.

The capabilities of integrating multi-sector, multilevel, and multi-period data base for decision making has made GIS an indispensable tool not only in surveying, but in any field one may think of. The success of any GIS application depends on the quality of the data used. The pointers reflecting the quality of the data are presented. Sources of errors in GIS are discussed. The chapter ends with an introduction to some commercial GIS software available in the market.

Exercises

- 8.1** Define Geographic Information System (GIS). What are the various sources from which data can be derived to be used for GIS?
- 8.2** Describe the raster and vector data structures. What are the relative advantages and disadvantages of the two data structures?
- 8.3** Describe briefly the following with regard to the representation of geographical features:
(a) Point data (b) Line data (c) Aerial data
- 8.4** (a) What are the component subsystems of GIS? Describe them briefly.
(b) What are the functionalities of GIS?
(c) The success of GIS application depends on the geographic data comment.
- 8.5** Write notes on the following:
(a) Buffering in GIS (b) Neighbourhood functions
(c) Map overlay (d) Feature identifier
- 8.6** (a) Describe the types of vector overlays with neat sketches.
(b) The algebraic manipulation of images in raster GIS is a powerful and flexible way of combining data and organising analysis. Explain with examples.
- 8.7** (a) Discuss the differences between vector and raster formats for storing data in GIS. For what types of data is the raster format best suited?
(b) Name and describe the simple spatial objects used for representing geographic data in digital form. Which objects are used in vector format representations?
- 8.8** Explain how data can be converted from
(a) vector to raster format (b) raster to vector format
- 8.9** Discuss the following GIS functions:
(a) Point buffering (b) Line buffering
(c) Polygon buffering
- 8.10** Briefly describe the applications of GIS. Discuss how closely GIS is related to remote sensing.
- 8.11** (a) What are the possible advantages and disadvantages of using a raster GIS as opposed to vector?
(b) What are the analytical capabilities and limitations of GIS?
- 8.12** What are filters? How are they caused for reclassification processes?
- 8.13** What are neighbourhood functions, buffers, and filters?
- 8.14** Compile a list of data layers and attributes that would likely be included in a GIS for
(a) selecting the optimum corridor for constructing an expressway connecting Allahabad and Lucknow.
(b) selecting the optimum location for a new airport in a metropolitan city.
(c) selecting the optimum site for hospital-waste disposal for a city of population 20 lakhs.

Objective-type Questions

- 8.1** Creation of image-based GIS from multigeodatasets is characterised by which subsystem of GIS?
(a) Management (b) Input (c) Processing (d) Display
- 8.2** Cell-like units are characteristic of
(a) raster data structures (b) cellular data structures
(c) Both (a) and (b) (d) vector data structures

- 8.3** The other name of vector format is
 (a) line format (b) point format
 (c) polygon format (d) All of these
- 8.4** The graphical elements of vector data structures are
 (a) point (b) arc (c) area (d) All of these
- 8.5** Concepts associated with the representation of vector data differently at different scales are
 (a) cartographic generalisation (b) cartographic symbolisation
 (c) unique feature identifier (d) topological data model
 (e) Both (a) and (b)
- 8.6** Observations occupying very small areas in relation to the scale of the database are represented by
 (a) point data (b) line data
 (c) areal data (d) continuous data
- 8.7** Line in polygon method is characteristic of
 (a) raster overlay (b) vector overlay
 (c) buffer operation (d) intersecting operation
- 8.8** In a raster overlay, a point is represented by a
 (a) string of cells (b) group of cells
 (c) single cell (d) All of these
- 8.9** Description of smallest features in data set that can be displayed or mapped is known as
 (a) accuracy (b) bias (c) resolution (d) generalisation
- 8.10** Microstation Geographics is related to
 (a) map overlay (b) GIS-based software
 (c) functionalities of GIS (d) data structure for GIS

Answers to Objective-type Questions

8.1 (b)	8.2 (c)	8.3 (d)	8.4 (d)	8.5 (a)	8.6 (a)
8.7 (b)	8.8 (c)	8.9 (c)	8.10 (b)		

9

GLOBAL POSITIONING SYSTEM (GPS)

Introduction

Some of the first known surveyors were Egyptians who used distant control points to replace property corners destroyed by floods. Later, the Greeks and Romans surveyed the settlements. French surveyors were probably the first to conduct surveys on large scales, by measuring the interior angles of a series of interconnecting triangles in combination with measured base lines, to determine the coordinates of selected points. Subsequently, triangulation technique was used by surveyors to determine accurate distances. The use of triangulation was limited by the line of sight. The series of triangles were generally fixed by astronomical points by observing selected stars to determine the position of that point on the surface of the earth. Since these astronomical positions could be in error by hundreds of meters, the interrelationship between the coordinates cannot be precisely fixed. To overcome this problem, optical global triangulation was developed.

The worldwide satellite triangulation program often called BC-4 program was carried out to establish interrelationships of the major world datums. The method involved photographing special reflective satellites against a star background with a metric camera fitted with a chopping shutter. The main problem in using this optical technique was that clear sky was required simultaneously at a minimum of two observing sites separated by some 4000 km, and the equipment was massive and expensive. Thus, optical technique was soon supported by the electromagnetic ranging technique because of all-weather capability and greater accuracy.

First attempts to (positionally) connect the continents by electromagnetic techniques was by the use of an electronic High Ranging (HIRAN) system developed during World War II to position aircrafts. A significant technology breakthrough occurred when scientists around the world experienced that the Doppler shift in the signal broadcast by a satellite could be used as an observable to determine the exact time of closest approach of a satellite. This knowledge, together with the ability to compute satellite ephemerides according to Kepler's laws, led to the instantaneous determination of the precise position anywhere in the world.

The immediate Predecessors of today's modern positioning system are Inertial Surveying System (ISS) and the Navy Navigational Satellite System

(NNSS), also called TRANSIT system developed by the United States (US), and TSIKADA system developed by Russia. All of these systems provide a two-dimensional high-accuracy positioning system. However, the frequency of obtaining a position fix varies with latitude. Theoretically, a TRANSIT user at the equator could obtain a position fix at the average of once every 10 min; whereas, at 80° latitudes, the fix rate would improve to an average of every 30 min.

The inertial surveying system (ISS) provides three-dimensional positioning. The system consists basically of an inertial platform with three accelerometers held in three mutually orthogonal axes by three similarly oriented precise gyroscopes. The ISS requires a vehicle (truck or helicopter) to occupy a point of known coordinates (X , Y , and elevation) and remain stationary for zero velocity update. As the vehicle moves, its location is constantly updated by the use of three computer-controlled accelerometers, each aligned to a north–south, east–west, and vertical axis. The accelerometer platform is oriented north–south, east–west, and plumb by means of three computer-controlled gyroscopes, each of which is aligned to one of the three axes (Fig. 9.1). The change in position of vehicle from the starting point is obtained by double integration of the acceleration with respect to time. Analysis of accelerometer data gives rectangular displacement factors (latitude and longitude) for horizontal movement, in addition to vertical displacement. The advantages of ISS are it is unaffected by atmospheric refraction, it does not require intervisibility between points; and it can be operated day and night. The drawback of the system is the high cost of both the ISS equipment and its operation.

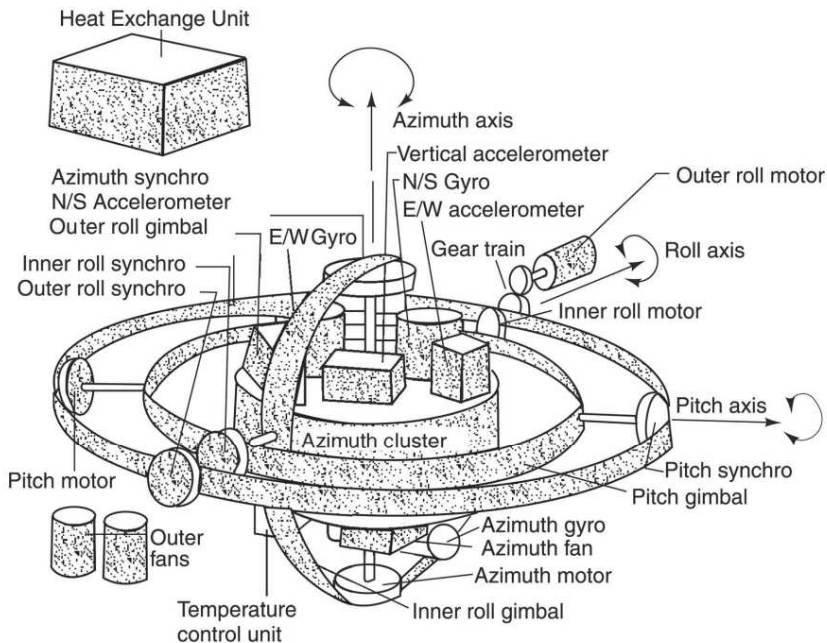


Fig. 9.1 *Inertial platform schematic*

The TRANSIT system was composed of six satellites orbiting at altitudes of about 1100 km with nearly circular polar orbits. TRANSIT system was developed primarily to determine the coordinates of vessels and aircrafts. Later on, this satellite system was used worldwide for navigation and surveying. The positioning analysis technique used in the TRANSIT system utilised a ground receiver capable of noting the change in satellite frequency transmission as the satellite first approached and receded from the observer. The change in frequency was affected by the velocity of the satellite itself. The change in velocity of transmissions from the approaching and then receding satellite, known as the *Doppler effect*, is directly proportional to the shift in frequency of the transmitted signals, and is thus proportional to the change of distance between the satellite and the receiver over a given time interval. With the precise knowledge of the satellite orbit and that of the satellite position in that orbit, the position of the receiving station could be computed. Some of the TRANSIT experiments showed that accuracies of about 1 metre could be obtained by occupying a point for several days.

The main problem with TRANSIT was the large time gaps in coverage. Since nominally a satellite would pass overhead every 90 minutes, users had to interpolate their position between *fixes* or *passes*. Unfortunately, the satellites that it uses are in very low orbit and there are not very many of them. So a user does not get a fix very often. Also, since the system is based on low frequency Doppler measurements, even small movements at the receiving end can cause significant errors in position. It was these shortcomings that led to the development of the US Global Positioning System (GPS), the European Satellite Based Navigation System (Galileo) and the Russian Global Navigation Satellite System (GLONASS). In these systems the satellites are high enough that they can avoid the problems encountered by land-based systems and they use technology accurate enough to really give pinpoint position anywhere in the world, 24 hours a day. The details of GLONASS system and Galileo system are presented in the following sections followed by the US Department of Defense's Global Positioning System (GPS) in details. The GLONASS and Galileo constellation systems are shown in Plate 3, and GPS constellation system in Plate 4.

9.1 GLONASS SYSTEM

GLONASS is a radio-based satellite navigation system, developed by Russian Aerospace Defence Forces for the Russian Government. It was made operational in 1996. The first GLONASS satellite was launched and placed in the orbit on 12th October, 1982. Thereafter, numerous rocket launchers added satellites to the system. By 2010, GLONASS had achieved 100% coverage of Russian territory. The full orbital constellation of 24 satellites was restored in October 2011, enabling full global coverage. GLONASS satellite designers have undergone several upgrades, having three generations, from GLONASS to GLONASS-M to GLONASS-K. The characteristics of the three generations of satellites are given in Table 9.1. In November 2011, four more GLONASS-M were placed into final orbit. Originally GLONASS was designed to have an accuracy of 65 m, but in reality, it had an accuracy of 20 m in the civilian signal and 10 m in the military signal. GLONASS uses a coordinate datum named FZ-90. Its satellites transmit two types of signals:

standard precision (SP) signal and high precision (HP) signal. GLONASS system is a counterpart and at par with the United States GPS system. Both the systems share the same principles in the transmission and positioning methods. The GLONASS system has both the precise positioning service and standard positioning service as in the GPS, but its datum and time reference system are different.

GLONASS like GPS consists of three segments: the space, the control, and the user segment. The *operational space segment* of GLONASS consists of 21 satellites in three orbital planes, with 3 on-orbit spares; making the total number of 24 satellites. The three orbital planes are separated by 120° and the satellites within the same orbit plane by 45° . Each orbital plane, therefore, has eight equally spaced satellites, operating at an altitude of 19,100 km at an inclination angle of 64.8° to the equator. Each satellite will complete an orbit in approximately 11 hr 15 min. The spacing of satellites is such that a minimum of five satellites are always in view round the globe. The geometric arrangement gives a considerable better coverage than GPS in Polar Regions above and below 50° latitude. The satellites work in GLONASS System Time, checked and updated twice daily, with a maximum time error of 15 ns. The ground control segment is entirely located within former Soviet Union Territory. The ground control station is located in Moscow. The user segment consists of antennas and receiver-processors that provide positioning, velocity and precise timing of the user.

Note As of today the total satellites in the constellation are 31 with 24 operational satellites.

Table 9.1 Characteristics of satellites of GLONASS system

Serial No.	Property	1st generation satellites (also called Urgan)	2nd generation satellites (GLONASS-M)	3rd generation satellites (GLONASS-K)
1.	Mass	1250 kg	1480 kg	750 kg
2.	Life	3 years	5 years	10 years
3.	Geometric properties	7.8 m tall, 7.2 m wide	3.7 m high, 2.4 m diameter	

9.2 GALILEO SYSTEM

Galileo, a satellite navigation system, is currently being built by the European Union and European Space Agency. The project is named after the Italian astronomer Galileo. The European satellite-based navigation system was conceived in 1999. The first test satellite (GLOVE-A) was launched from Kazakhstan on December 28, 2005 by the European Space Agency. A Russian Soyuz rocket launched the British built 600 kg satellite. The second test satellite (GLOVE-B) was launched in 2008. On 21st October 2011, the first two of four satellites were launched to validate the system. The next two are planned for 2012. Additional satellites are planned to be launched by the mid-decade. Full completion of 30 satellites Galileo system, with 27 operational and 3 active spares, are expected by 2019. This will be the first step towards a final constellation of 30 satellites.

The aim of Galileo is to provide a high-precision positioning system. It is intended to provide horizontal and vertical positions measurements within 1 m precision, and better positioning at high altitudes than GPS and GLONASS systems. The satellites will be equipped with transponder which will relay distress signals from them that their situation has been detected and that the help is on the way; a new feature not available with the other two systems.

9.3 GPS OVERVIEW

The GPS is a space-based all weather radio navigation system that provides quickly, accurately and inexpensively the time, position, and velocity of the object anywhere on the globe at any time. The current GPS is based on accurate ephemeris data on the real-time location of each satellite and on a precisely kept time. It uses the satellite signals, accurate time, and sophisticated algorithms to generate distances in order to triangulate positions anywhere on earth. The system provides accurate, continuous, worldwide, three-dimensional position and velocity information to users with the appropriate receiving equipment. Satellite transmissions are referenced to highly accurate atomic frequency standards on-board the satellites, which are in synchronisation with an internal GPS time base.

The Global Positioning System is comprised of three segments: satellite constellation called *space segment*, ground control/monitoring network called *operational control segment*, and user receiving equipment called *user equipment segment*.

The satellite constellation contains the satellites in orbit that provide the ranging signals and data messages to the user equipment. The operational control segment (OCS) tracks and maintains the satellite in space. The OCS monitors the satellite health and signal integrity and maintains the orbital configuration of the satellites. It also updates the satellite clock corrections and ephemerides as well as other parameters essential for the determination of user position, velocity and time. The equipment segment performs the navigation, timing or other related functions (e.g., surveying).

9.4 SPACE SEGMENT (SATELLITE CONSTELLATION)

As stated above, one of the main problems with the TRANSIT System was the fact that the six orbiting satellites were not able to provide continuous positioning. A sufficient number of satellites, to ensure that (at least) four are always electronically visible, are required for continuous global positioning capabilities. The first NAVASTAR (Navigation Satellite Timing and Range) satellite for the GPS system was launched in 1978. This full system has a total of 28 GPS satellites in orbit around earth, although four are prototypes not used in main constellation. The satellite constellation consists of the nominal 24-satellite constellation. The satellites are positioned in six earth-centred orbital planes with four satellites in each plane. The nominal orbital period of a GPS satellite is one-half of a sidereal day or $11^{\text{h}} 58^{\text{m}}$. The orbits are nearly circular and equally spaced about the equator at a 60° separation with an inclination relative to the equator of nominally 55° . The orbital radius (i.e., the nominal distance from the centre of the mass of the

earth to the satellite) is approximately 26 600 km. This disposition ensures that at least four satellites can be observed from any point on the earth, and in fact six are often visible. This satellite constellation provides a 24 h global navigation and time determination capability. Figure 9.2 depicts the GPS constellation.

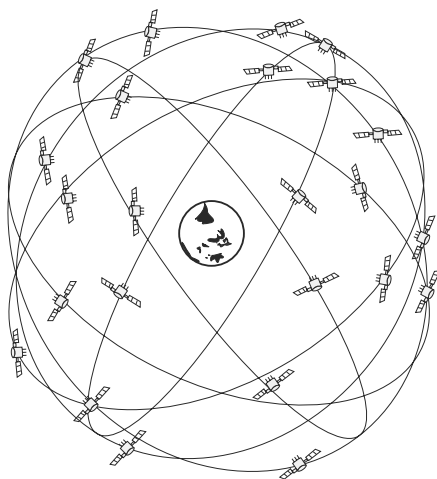


Fig. 9.2 GPS satellite constellation

The satellite orbits in a planar projection referenced to the epoch time of 0 hour, 1st July 1993 UTC are presented in Fig. 9.3. The slope of each orbit represents its inclination with respect to the earth's equatorial plane, which is nominally 55° . The orbital plane locations with respect to the earth are defined by the longitude of the ascending node, while the location of the satellite within the orbital plane is defined by the mean anomaly. The longitude of the ascending node is the point of intersection of each orbital plane with the equatorial plane. The Greenwich meridian is the reference point or point where the longitude of the ascending node has the value of zero. Mean anomaly is the angular position of each satellite within the orbit, with the earth's equator being the reference or point with a zero value of mean anomaly. It can be observed that relative phasing between most satellites in adjoining orbits is approximately 40° . In Fig. 9.3 *A, B, C, ...* are the orbital planes, and satellites within a plane are assigned the numbers 1, 2, 3, and 4. Thus, a satellite referenced as D_3 refers to satellite number 3 in orbital plane *D*. Another notation represents the configuration of the pseudo-random (PRN) code (Appendix V) generators on-board the satellite. These PRN code generators are configured uniquely on each satellite, thereby producing unique versions of both C/A-code and P-code. Thus, a satellite can be identified by the PRN codes that it generates.

Each satellite sends a full description of its own orbit data necessary to establish its location in space at any instant and an approximate guide to orbits of other satellites. The other information transmitted includes health of the satellite, parameters for propagation, error correction, etc.

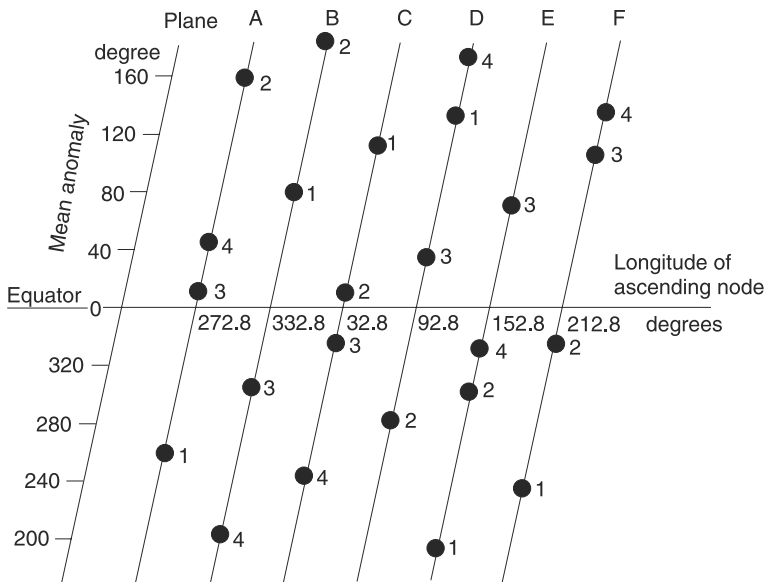


Fig. 9.3 GPS constellation planar projection

9.5 OPERATIONAL CONTROL SEGMENT (OCS)

The OCS has the responsibility for maintaining the satellites and their proper functioning. The control segment tracks all the satellites, ensures they are operating properly, and computes their position in space. The computed positions of satellites are used to derive parameters, which in turn are used to predict where the satellite will be later in time. These parameters are uploaded from the control segment to the satellites and are referred to as broadcast ephemeris, the prediction of current satellite position that are transmitted to the user in the data message. The OCS resolves satellite anomalies, makes pseudo-range and Delta-range measurements to determine satellite clock corrections, almanac, and ephemeris. To accomplish the above functions, the control segment is comprised of three different physical components: the master control station (MCS), the monitor stations (MS) and the ground antennas.

The monitor stations passively track the GPS satellites as they pass overhead by making pseudo-range and Delta-range measurements. These measurements are made using the two L-band frequencies: L_1 at 1575.42 MHz (wavelength of about 19 cm) and L_2 at 1227.6 MHz (wavelength at 24 cm). The L_1 signal is modulated with two codes and a navigation message: “short” Coarse Acquisition (C/A) code and “long” Precise (P) code. The C/A-code has a 1-ms period and repeats constantly, whereas the P-code satellite transmission is a 7-day sequence that repeats every midnight Saturday/Sunday. The C/A code is available to the public, whereas the P-code is designed for military use. Only the P-code is modulated on L_2 band. Each satellite transmits on these frequencies, but with different ranging codes than those employed by other satellites. These codes were selected because they have low cross-correlation properties with respect to one another. The

navigation data provides the means for the receiver to determine the location of the satellite at the time of signal transmission, whereas the ranging code enables the user's receiver to determine the transit time of the signal and thereby determine the user-to-satellite time.

The message contains clock corrections and predicted orbital parameters, which are used in computer programs to assist in positioning solutions.

9.6 USER EQUIPMENT SEGMENT

The user receiving equipment, referred to as a GPS receiver, receives and processes the L-band signals transmitted from the satellites to determine user position, velocity and time (PVT). Almost all GPS receivers have essentially the same basic components—an antenna, an RF (radio frequency) section, a microprocessor, a control and display unit (CDU), a recording device and a power supply. These components may be individual units, integrated into one unit, or partially integrated. Usually, all components, with the exception of the antenna, are grouped together and are referred to as receiver. A GPS receiver and antenna are shown in Plate 5. There has been a revolution in the technology of the GPS receiving sets, paralleling that of the electronics industry in general. The smallest sets today are those of wrist watch, while probably the largest is naval shipboard unit with a foot print of 1550 cm^2 and a weight of 31.8 kg. Selection of the GPS receiver depends on user's application. A block diagram of the GPS receiving set is shown in Fig. 9.4.

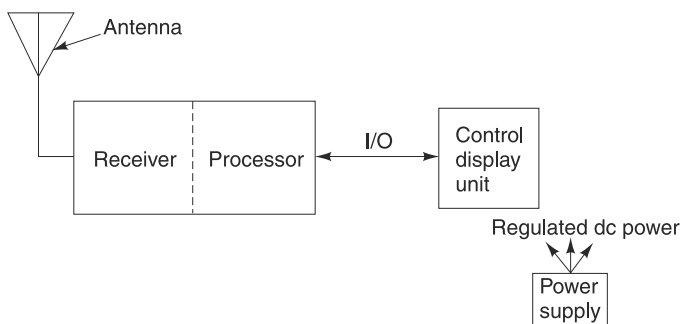


Fig. 9.4 A schematic representation of the GPS receiver

The basic functional subsystems of a GPS receiver consist of following:

1. *Antenna and preamplifier:* The task is to receive the electromagnetic radiation coming from the satellites and amplify them for further processing. The antennae should have broad beam characteristics so that they do not have to be pointed to source. There are several types of antennas used with GPS. Microstrip antenna is one of the widely used systems.
2. *Radiofrequency section and computer processor:* The RF section of a GPS receiver converts the incoming signal after preamplification to an intermediate frequency (IF) for further processing. The modulated signal extracted from the carrier is processed by a powerful microprocessor. There are different signal processing techniques—conventional and propriety—to generate the position fix.

3. *Input/Output device*: This enables operators to interact with the microprocessor, and the measured data is stored/displayed. It will also have a standard interface to transfer to digital data outside.

4. *Power supply*: Since transportable and handheld GPS are very common, internal battery packs form a part of the receiver.

Two fundamental methods can be used to establish the distance between the satellite and receiver: the pseudo-range technique and the carrier phase measurement technique. These are described in the following sections.

9.7 PRINCIPLE OF POSITION DETERMINATION VIA SATELLITE GENERATED RANGING SIGNALS

GPS employs Time of Arrival (TOA) ranging for user position determination. By making TOA measurements to multiple satellites, three-dimensional positioning is achieved. The satellite ranging signals travel at the speed of light (3×10^8 m/s). It is assumed that satellite ephemerides are accurate (i.e., the satellite locations are precisely known).

9.7.1 Three-Dimensional Position Location via Intersection of Multiple Spheres

Assume that there is a single satellite transmitting a ranging signal. A clock on-board the satellite controls the timings of the ranging signal broadcast. This clock and others on-board each of the satellites within the constellation are effectively synchronised to an internal system time scale denoted as GPS time. The user's receiver also contains a clock that (for the moment) we assume to be synchronised to the system time. Timing information is embedded within the satellite ranging signal that enables the receiver to calculate when the signal left the satellite. By noting the time when the signal was received, the satellite-to-user propagation time can be computed. The product of this propagation time and the speed of light yields the satellite-to-user range, R . As a result of this measurement process, the user would be located somewhere on the surface of a sphere centred about the satellite, as shown in Fig. 9.5(a).

If a measurement was simultaneously made using the ranging signal of a second satellite, the user would also be located on the surface of a second sphere that is concentric about the second satellite. Thus, the user would then be somewhere on the surface of both the spheres, which could be either on the perimeter of the shaded circle in Fig. 9.5(b) that denotes the plane of intersection of these spheres, or at a single point tangent to both spheres (i.e., where the spheres just touch). This latter case could only occur if the user was collinear with the satellites. The plane of intersection is perpendicular to a line connecting the satellites, as shown in Fig. 9.5(c).

Repeating the measurement process using a third satellite co-locates the user on the perimeter of the circle and the surface of the third sphere. This third sphere intersects the shaded circle perimeter at two points (Fig. 9.5(d)); however, only one of the points is the correct user position. With the third measurement user narrows down its position to just two points in all the universe.

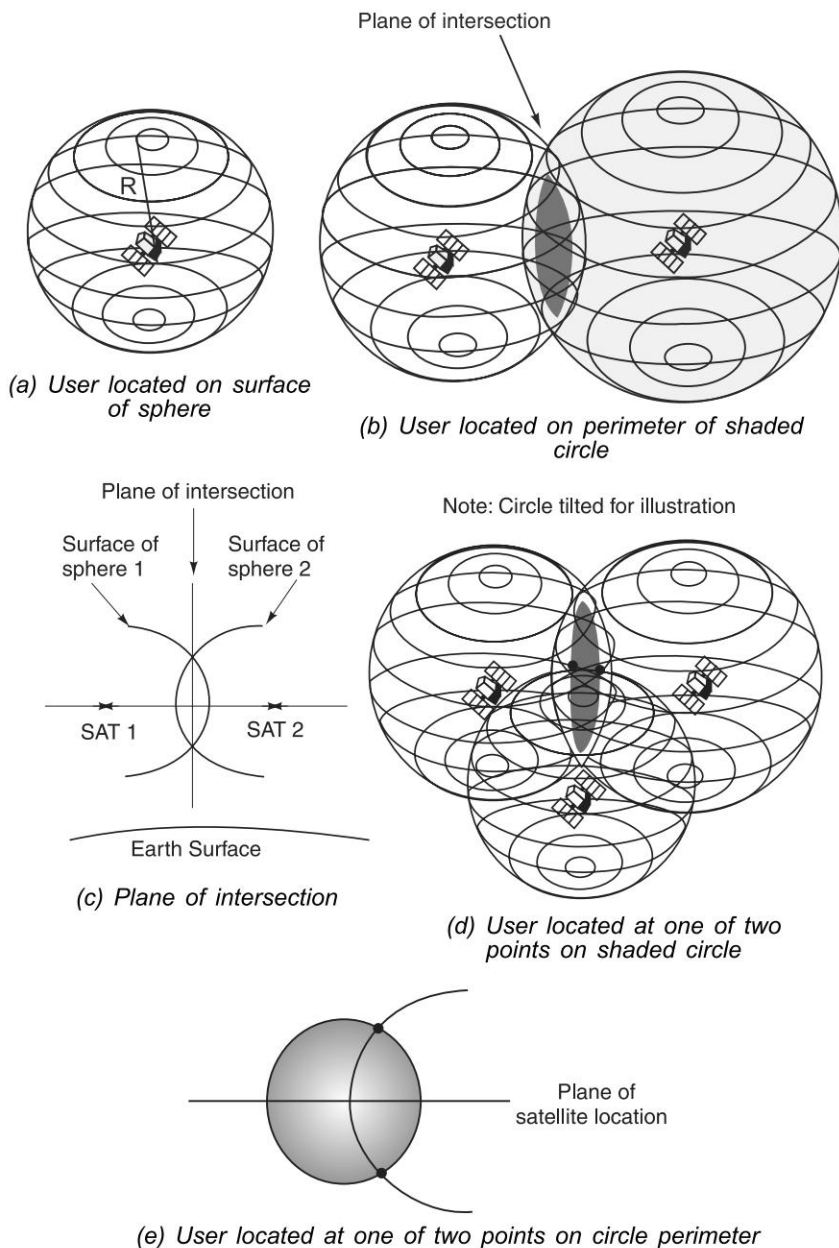


Fig. 9.5 User location

A view of the intersection is shown in Fig. 9.5(e). For a user on the earth's surface, it is apparent that the lower point will be the true position. However, users that are above the earth's surface may employ measurements from satellites at negative elevation angles. This complicates the determination of an unambiguous solution. Airborne/spaceborne receiver solutions may be above or below the

plane containing the satellites, and it may not be clear which point to select unless the user has ancillary information.

9.8 DETERMINING SATELLITE-TO-USER RANGE

The GPS satellites are configured, primarily to provide the user with the capability of determining his position, expressed, for example, by latitude, longitude, and elevation. This is accomplished by the simple resection process using the distances measured to the satellites. In Sec. 9.7, the theoretical aspects of using satellite ranging signals and multiple spheres to solve for user position in three dimensions have been discussed. It was presumed that the receiver clock was perfectly synchronised to the system time. In actuality, this is generally not the case.

The errors experienced from non-synchronised clocks are considerably large as compared to some of the other errors such as measurement noise, propagation delays, etc. Therefore, in the development of basic concepts, errors other than clock offset are omitted.

In Fig. 9.6, S denotes the position of a satellite in space at a given instant. It is required to determine the vector u , which represents a user receiver's position on earth's surface. The user's position coordinates x_u, y_u, z_u are considered unknown. Vector r represents the vector offset from the user to the satellite. x_s, y_s, z_s are coordinates of the satellite location. The position of the satellite with respect to the centre of the earth is represented by vector s .

Vector s is computed using ephemeris data broadcast by the satellite. The satellite-to-user vector r is given as

$$\mathbf{r} = \mathbf{s} - \mathbf{u}$$

The magnitude of vector r is

$$\|\mathbf{r}\| = \|\mathbf{s} - \mathbf{u}\|$$

or $R = \|\mathbf{s} - \mathbf{u}\|$ (9.1)

where R represents the magnitude of r . The range r to each satellite can be accurately measured by recording the time required for the satellite signal to reach the user receiver.

If the satellite clock and the receiver clock were perfectly synchronised, the correlation process would yield the true propagation time Δt . By multiplying this propagation time by the speed of light, the true (i.e., geometric) satellite-to-user distance can be computed. However, the satellite and receiver clocks are generally not synchronised.

The receiver clock will generally have a bias error from the system time. Further, satellite frequency generation and timing is based on highly accurate atomic clock, which is typically offset from the system time. Thus, the range determined by the correlation process is denoted as pseudorange ρ , as it contains

1. the geometric satellite-to-user range,

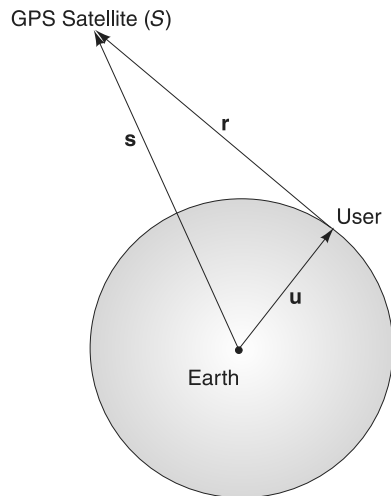


Fig. 9.6 User position vector representation

2. an attribute to the difference between the system time and the user clock, and
3. an offset between the system time and the satellite clock

Let,

T_s = system time at which the signal left the satellite

T_u = system time at which the signal reached the user receiver

δt = offset of the satellite clock from the system time (advance is positive; retardation (delay) is negative)

t_u = offset of the receiver clock from the system time

$T_s + \delta t$ = satellite clock reading at the time that the signal left the satellite

$T_u + t_u$ = user receiver clock reading at the time when the signal reached the user receiver

c = speed of light

$$\text{Geometric range, } R = c (T_u - T_s) = c \Delta t \quad (9.2)$$

$$\text{Pseudo range, } \rho = c [(T_u + t_u) - (T_s + \delta t)]$$

$$= c (T_u - T_s) + c (t_u - \delta t)$$

$$= R + c (t_u - \delta t)$$

$$\text{or} \quad R = \rho - c (t_u - \delta t) \quad (9.3)$$

From Eqs. (9.1) and (9.3)

$$\rho - c (t_u - \delta t) = \| \mathbf{s} - \mathbf{u} \| \quad (9.4)$$

where δt = the advance of the satellite clock with respect to the system time.

δt is composed of bias and drift contributions. The GPS ground monitoring network determines corrections for these offset contributions and transmits the corrections to the satellites for re-broadcast to the users in the navigation message. These corrections are applied within the user receiver to synchronise the transmission of each ranging signal to system time. Therefore, this offset can be assumed to be compensated for and δt therefore is not unknown. Equation (9.4) can be expressed as

$$\rho - ct_u = \| \mathbf{s} - \mathbf{u} \| \quad (9.5)$$

Notes

- i. The range error can be eliminated in advance by differencing the pseudoranges measured from one site to two satellites or two different positions of one satellite. In both cases, the Delta-range now defines a hyperboloid with its foci placed at the two satellite positions for the geometric location of the receiver.
- ii. Systematic errors in the satellite's position and eventual satellite clock biases in the pseudoranges can be reduced or eliminated by differencing the pseudoranges measured from two sites to the satellite.
- iii. One key dimension in positioning is the parameter of time. Time is kept on-board the satellites by atomic clocks, which have a precision of 1 nanosecond (10^{-9} second). The ground receivers are equipped with less precise quartz clocks. Uncertainties of these less precise clocks are resolved when observing the signals from four or more satellites instead of the basic three-satellite configuration required for point positioning.

9.9 CALCULATION OF USER POSITION

Since the intersection of three spheres yields the three unknowns (e.g., latitude, longitude, and height), hence, ranges to only three satellites would be needed. Since the clock of the ground GPS receiver is offset from the GPS time, the distance measured to the satellite is not the true range. Therefore, in order to determine the user position in three dimensions (x'_u, y'_u, z'_u) and the offset t'_u , pseudo-range measurements are made to four satellites resulting in the system of equations:

$$\rho_j = \| \mathbf{s}_j - \mathbf{u} \| + ct'_u \quad (9.6)$$

where j ranges from 1 to 4 and references the satellites. Equation (9.6) can be expanded into the following set of equations

$$\rho_1 = \sqrt{(x_1 - x'_u)^2 + (y_1 - y'_u)^2 + (z_1 - z'_u)^2} + ct'_u \quad (9.7)$$

$$\rho_2 = \sqrt{(x_2 - x'_u)^2 + (y_2 - y'_u)^2 + (z_2 - z'_u)^2} + ct'_u \quad (9.8)$$

$$\rho_3 = \sqrt{(x_3 - x'_u)^2 + (y_3 - y'_u)^2 + (z_3 - z'_u)^2} + ct'_u \quad (9.9)$$

$$\rho_4 = \sqrt{(x_4 - x'_u)^2 + (y_4 - y'_u)^2 + (z_4 - z'_u)^2} + ct'_u \quad (9.10)$$

where x_j, y_j and z_j denote the j th satellite's position in three dimensions.

These non-linear equations can be solved for the unknowns by employing either closed form solutions, or iterative techniques based on linearisation, or Kalman filtering. Linearisation is illustrated below.

If the approximate position of the receiver is known, then the offset of the true position (x'_u, y'_u, z'_u) from the approximate position (x'_u, y'_u, z'_u) can be denoted by a displacement ($\Delta x'_u, \Delta y'_u, \Delta z'_u$). By expanding Eqs. (9.7) to (9.10) in a Taylor series about the approximate position, the position offset ($\Delta x'_u, \Delta y'_u, \Delta z'_u$) as linear functions of the known coordinates and pseudo-range measurements can be obtained. This process is described below.

Let a single pseudorange be represented by

$$\begin{aligned} \rho_j &= \sqrt{(x_j - x'_u)^2 + (y_j - y'_u)^2 + (z_j - z'_u)^2} + ct'_u \\ &= f(x'_u, y'_u, z'_u, t'_u) \end{aligned} \quad (9.11)$$

Using the approximate position location (x'_u, y'_u, z'_u) and time-bias estimate t'_u , an approximate pseudorange can be calculated:

$$\begin{aligned} \rho'_j &= \sqrt{(x_j - x'_u)^2 + (y_j - y'_u)^2 + (z_j - z'_u)^2} + ct'_u \\ &= f(x'_u, y'_u, z'_u, t'_u) \end{aligned} \quad (9.12)$$

Since the unknown user position and receiver clock offset consist of an approximate component and an incremental component, these can be written as

$$\left. \begin{aligned} x_u &= x'_u + \Delta x'_u \\ y_u &= y'_u + \Delta y'_u \\ z_u &= z'_u + \Delta z'_u \\ t_u &= t'_u + \Delta t'_u \end{aligned} \right\} \quad (9.13)$$

Hence, $f(x_u, y_u, z_u, t_u) = f(x'_u + \Delta x_u, y'_u + \Delta y_u, z'_u + \Delta z_u, t'_u + \Delta t_u)$

Expanding the latter function about the approximate point and associated predicted receiver clock offset (x'_u, y'_u, z'_u, t'_u) using a Taylor series:

$$\begin{aligned} f(x'_u + \Delta x_u, y'_u + \Delta y_u, z'_u + \Delta z_u, t'_u + \Delta t_u) \\ = f(x'_u, y'_u, z'_u, t'_u) + \frac{\partial f(x'_u, y'_u, z'_u, t'_u)}{\partial x'_u} \Delta x_u + \\ \frac{\partial f(x'_u, y'_u, z'_u, t'_u)}{\partial y'_u} \Delta y_u + \frac{\partial f(x'_u, y'_u, z'_u, t'_u)}{\partial z'_u} \Delta z_u + \frac{\partial f(x'_u, y'_u, z'_u, t'_u)}{\partial t'_u} \Delta t_u + \dots \quad (9.14) \end{aligned}$$

To eliminate nonlinear terms the expansion has been truncated after the first order partial derivatives. The partial derivatives evaluate as follows:

$$\left. \begin{aligned} \frac{\partial f(x'_u, y'_u, z'_u, t'_u)}{\partial x'_u} &= -\frac{x_j - x'_u}{\hat{r}_j} \\ \frac{\partial f(x'_u, y'_u, z'_u, t'_u)}{\partial y'_u} &= -\frac{y_j - y'_u}{\hat{r}_j} \\ \frac{\partial f(x'_u, y'_u, z'_u, t'_u)}{\partial z'_u} &= -\frac{z_j - z'_u}{\hat{r}_j} \end{aligned} \right\} \quad (9.15)$$

$$\frac{\partial f(x'_u, y'_u, z'_u, t'_u)}{\partial t'_u} = c$$

where

$$r'_j = \sqrt{(x_j - x'_u)^2 + (y_j - y'_u)^2 + (z_j - z'_u)^2}$$

Substituting Eqs. (9.11), (9.12), (9.13) and (9.15) into Eq. (9.14) yields

$$\rho_j = \rho'_j - \frac{x_j - x'_u}{r'_j} \Delta x_u - \frac{y_j - y'_u}{r'_j} \Delta y_u - \frac{z_j - z'_u}{r'_j} \Delta z_u + c \Delta t_u \quad (9.16)$$

This completes the linearisation of Eq. (9.11) with respect to the unknowns Δx_u , Δy_u , Δz_u , and Δt_u .

Note In the above development, secondary error sources such as earth's rotation compensation, measurement noise, propagation delays, relativistic effects have been neglected.

Rearranging the above expression with the known quantities on the left and unknowns on right yield

$$\rho'_j - \rho_j = \frac{x_j - x'_u}{r'_j} \Delta x_u + \frac{y_j - y'_u}{r'_j} \Delta y_u + \frac{z_j - z'_u}{r'_j} \Delta z_u - c \Delta t_u \quad (9.17)$$

For convenience, the above equation may be simplified by introducing variables where

$$\begin{aligned}
\Delta\rho_j &= \rho'_j + \rho_j \\
a_{xj} &= \frac{x_j - x'_u}{r'_j} \\
a_{yj} &= \frac{y_j - y'_u}{r'_j} \\
a_{zj} &= \frac{z_j - z'_u}{r'_j}
\end{aligned} \tag{9.18}$$

The a_{xj} , a_{yj} and a_{zj} terms in Eq. (9.18) denote the direction cosines of the unit vector pointing from the approximate user position to the j th satellite. For the j th satellite, this unit vector is defined as

$$a_j = (a_{xj}, a_{yj}, a_{zj})$$

Equation (9.17) can be rewritten more simply as

$$\Delta\rho_j = a_{xj} \Delta x_u + a_{yj} \Delta y_u + a_{zj} \Delta z_u - c\Delta t_u$$

There are four unknowns: Δx_u , Δy_u , Δz_u , and Δt_u , which can be solved for by making ranging measurements to four satellites. The unknown quantities can be determined by solving the set of linear equations below:

$$\left. \begin{aligned}
\Delta\rho_1 &= a_{x1} \Delta x_u + a_{y1} \Delta y_u + a_{z1} \Delta z_u - c\Delta t_u \\
\Delta\rho_2 &= a_{x2} \Delta x_u + a_{y2} \Delta y_u + a_{z2} \Delta z_u - c\Delta t_u \\
\Delta\rho_3 &= a_{x3} \Delta x_u + a_{y3} \Delta y_u + a_{z3} \Delta z_u - c\Delta t_u \\
\Delta\rho_4 &= a_{x4} \Delta x_u + a_{y4} \Delta y_u + a_{z4} \Delta z_u - c\Delta t_u
\end{aligned} \right\} \tag{9.19}$$

These equations can be put in matrix form by making the definitions

$$\Delta\mathbf{p} = \begin{bmatrix} \Delta\rho_1 \\ \Delta\rho_2 \\ \Delta\rho_3 \\ \Delta\rho_4 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} a_{x1} & a_{y1} & a_{z1} & 1 \\ a_{x2} & a_{y2} & a_{z2} & 1 \\ a_{x3} & a_{y3} & a_{z3} & 1 \\ a_{x4} & a_{y4} & a_{z4} & 1 \end{bmatrix}, \quad \Delta\mathbf{x} = \begin{bmatrix} \Delta x_u \\ \Delta y_u \\ \Delta z_u \\ c\Delta t_u \end{bmatrix}$$

$$\Delta\mathbf{p} = \mathbf{H} \Delta\mathbf{x} \tag{9.20}$$

The solution of Eq. (9.20) is

$$\Delta\mathbf{x} = \mathbf{H}^{-1} \Delta\mathbf{p} \tag{9.21}$$

Once the unknowns are computed, the user's coordinates x_u , y_u , z_u and the receiver clock offset t_u can be calculated using Eq. (9.13).

Notes

- i. This linearisation scheme will work well as long as the displacement (Δx_u , Δy_u , Δz_u) is within close proximity of the linearisation point. The acceptable displacement is dictated by the user's accuracy requirements.
- ii. If the displacement does exceed the acceptable value, the above process is reiterated with $\mathbf{p}$ being replaced by a new estimate of pseudorange based on the calculated point coordinates x_u , y_u and z_u .

The true user-to-satellite measurements, however, are corrupted by uncommon (i.e., independent) errors such as measurement noise, deviation of the satellite path from the reported ephemeris, and multipath. These errors translate to errors in the components of vector $\Delta\mathbf{x}$, as below:

$$\boldsymbol{\varepsilon}_x = \mathbf{H}^{-1} \boldsymbol{\varepsilon}_{\text{meas}} \quad (9.22)$$

where $\boldsymbol{\varepsilon}_{\text{meas}}$ is the vector containing the pseudorange measurement errors and $\boldsymbol{\varepsilon}_x$ is the vector representing errors in the user position and receiver clock offset.

The error contribution $\boldsymbol{\varepsilon}_x$ can be minimized by making measurements to more than four satellites, which will result in an overdetermined solution set of equations similar to Eq. (9.19). Each of these redundant measurements will generally contain independent error contributions. Redundant measurements can be processed by least squares estimation techniques that obtain improved estimates of the unknowns. The receivers generally employ more than four user-to-satellite measurements to compute user position, velocity and time, and use the least squares estimation technique to obtain most probable values of the unknowns.

9.10 GPS SYSTEM TIME

A practical realisation of the dynamic time system is achieved by the atomic time scale. GPS time also belongs to this system. A GPS system time is a “paper”¹ time scale based on statistically processed readings from the atomic clocks in satellites and at various ground control segment components. GPS system time is a continuous time scale that is not adjusted for leap seconds. GPS system and Universal Time Coordinated (UTC)² system were coincident at 0 hour, 6th January 1980. As of writing this, UTC leads (GPS system) time by 19 sec. An epoch in GPS system time is distinguished by the number of seconds that have elapsed since Saturday/Sunday midnight and the GPS week number. GPS weeks are numbered sequentially and originate with week 0, which began at 0 hour, 6th January 1980.

9.11 CARRIER PHASE MEASUREMENT TECHNIQUE

In principle, this measurement technique is similar to the operation of EDM, although the signal is only transmitted one way, satellite to receiver, and thus the true phase shift is related to the accurate synchronisation of the two clocks. This problem can be overcome by differencing, a procedure of taking simultaneous reading on two satellites, and by observation onto a single satellite from two stations, as discussed below. A further difference from EDM relates to the method of establishing the number of complete wavelengths in the equation

$$R = (N\lambda + \phi)$$

¹A uniform time scale based on atomic second, which is defined as the fundamental unit of time in the International System of Units, is called International Atomic Time (IAT). IAT is referred to as a paper time scale, since it is not kept by a physical clock.

²UTC is a composite time scale. That is, UTC is comprised of inputs from a time scale derived from atomic clocks, a time scale referenced to the earth’s rotation rate.

where, R is the actual range between receiver and satellite, λ is the wavelength, N is the number of complete wavelengths between satellite and receiver, and ϕ is the part wave length.

For EDM the modulation wavelength on the carrier wave is changed to establish two or more data sets, allowing N to be eliminated from the calculations. With satellite observations this is not possible because the satellite is continuously moving, although two different carrier waves L_1 and L_2 are being transmitted simultaneously. In order to make use of the carrier waves it is necessary to strip off the binary code information, and there are various techniques for this, some requiring knowledge of *C/A* and/or *P* code, which is not possible with anti-spoofing (AS) active. The most common technique is *squaring* (no knowledge of ranging codes is required), which has the added advantage of effectively doubling the carrier frequency, hence halving the wavelength and allowing the millimetre accuracy on the measurement. Unfortunately, squaring destroys the data transmitted on the carrier that gives the satellite location and hence for accuracy it has to be used with calculated data for the satellite's positions.

9.12 GPS ERROR SOURCES

If the world was like a laboratory with perfect laboratory conditions, basic GPS would be a lot more accurate. Unfortunately, it is not so, and has plenty of opportunities for a radio-based system that spans the entire planet to get fouled up. Inaccuracies in GPS signals come from the variety of sources (Fig. 9.7), like satellite clocks, imperfect orbits and especially from the signal's trip through the earth's atmosphere. Since these inaccuracies are variable, it is hard to predict what they will be in order to correct for them. Although, these errors are small, but to get the kind of accuracies some critical positioning jobs require, all the errors, no matter how minor, are to be minimised. What is needed is a way to measure the actual errors as they happen. The error sources can be classified into three groups, namely satellite-related errors, propagation-medium related errors, and

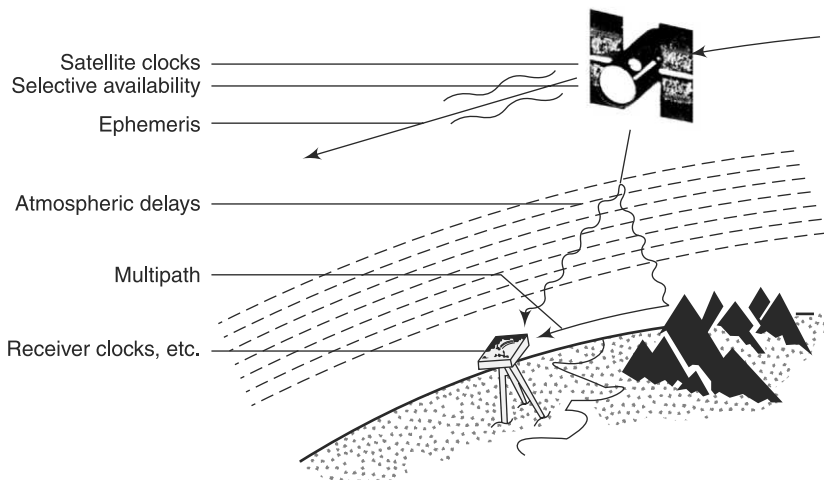


Fig. 9.7 Error sources

receiver-related errors. These are known as *systematic errors*. However, sometimes errors are introduced intentionally known as *selective availability*.

Note *Differential GPS works by canceling out most of the natural and man-made errors that creep into normal GPS measurements.*

9.12.1 Systematic Errors

Satellite Errors GPS satellites are equipped with very accurate atomic clocks. But as good as these clocks are, they are not perfect. Slight inaccuracies in their timekeeping ultimately lead to inaccuracies in our position measurements. The satellite's position in space is also important equally because it is the starting point for all of the GPS calculations. GPS satellites are placed into very high orbits and so are relatively free from the perturbing effects of the earth's upper atmosphere, but even so they still drift slightly from their predicted orbits contributing to the errors.

1. *Signal propagation errors:* GPS satellites transmit their timing information by radio, and that is another source of error because radio signals in the earth's atmosphere (ionosphere and troposphere) do not behave as predictably desired. It is assumed that radio signals travel at the speed of light, which is presumably a constant. However, the speed of light is not constant. It is only constant in vacuum. In the real world, light (or radio) slows down depending on what it is travelling through. As a GPS signal comes down through the charged particles in the ionosphere and then through the water vapor in the troposphere, it gets delayed a little. Since calculation of distance assumes a constant speed of light, this delay results into a miscalculation of the satellite's distance, which in turn translates into an error in position.

Good receivers add in a correction factor for a typical trip through the earth's atmosphere, which helps, but since the atmosphere varies from point to point and moment to moment, no correction factor or atmospheric model can accurately compensate for the delays that actually occur.

2. *Receiver errors:* The receivers are also not perfect. They can introduce their own errors which usually stem from their clocks or internal noise, multipath and antenna face centre variation.

(a) *Multipath:* As the GPS signal arrives at the surface of the earth it may get reflected by local obstructions and gets to the receiver's antenna via more than one path. This form of error is called *multipath error* because, in a sense, the signal is getting to the antenna by multiple paths. First, the antenna receives the direct signal it being the fastest, and then the reflected signals arrive a little later (Fig. 9.8). These delayed signals can interfere with the direct

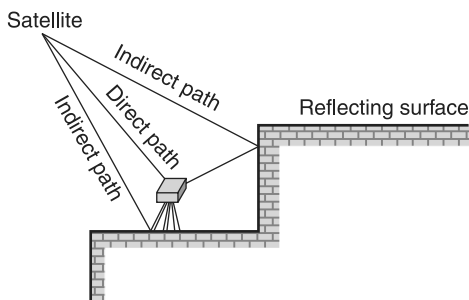


Fig. 9.8 *Multipath effect*

signal giving noisy results. Secondary effects are reflections at the satellite during signal transmission.

The influence of the multipath, however, can be estimated by using a combination of L_1 and L_2 codes, and carrier-phase measurement. The principle is based on the fact that troposphere, clock errors, and relativistic effects influence code and carrier phases by the same amount. This is not true for ionospheric refraction and multipath which are frequency dependent. Taking ionospheric-free code ranges and carrier phases, and forming corresponding differences, all aforementioned effects, except for multipath, cancel out. The most effective counter measure to multipath is to avoid sites where it could be a problem.

The elimination of multipath signals is also possible by selecting an antenna that takes advantage of the signal polarisation. GPS signals are right-handed circularly polarised, whereas the reflected signals are left-handed polarised. A reduction of multipath effect may also be achieved by digital filtering, wideband antennas, and radio frequency absorbent antenna ground planes. The absorbent antenna ground plane reduces the interference of satellite signals with low or even negative elevation angles which occur in case of multipath.

(b) Antenna phase centre offset and variation: The phase centre of the antenna is the point to which the radio signal measurement is referred and generally is not identical with the geometric antenna centre. The offset depends on the elevation, the azimuth, and the intensity of the satellite signal and is different for L_1 and L_2 codes. Also, the true antenna phase centre may be different from the manufacturer-indicated centre. This antenna offset may simply arise from inaccurate production series. Further, the antenna phase centre can vary with respect to the incoming satellite signals. The variation is systematic and may be investigated by test series.

Systematic effects can be eliminated by appropriate combinations of the observables. Differencing between receivers eliminates satellite-specific biases, and differencing between satellites eliminates receiver-specific biases. Thus, double-differenced pseudoranges are, to a high degree, free of systematic errors originating from the satellites and from the receivers. With respect to refraction, this is only true for short baseline where the measured ranges at both end points are affected equally. In addition, ionospheric refraction can be virtually eliminated by an adequate combination of dual frequency data.

Multipath is caused by multiple reflection of the signal. The interference between the direct and the reflected signal is largely not random; however, it may also appear as a noise. A similar effect is called *imaging*, where a reflecting obstacle generates an image of the real antenna which distorts the antenna pattern. Both effects, multipath and imaging, can be considerably reduced by selecting sites protected from reflections (buildings, vehicles, trees, etc.) and by appropriate antenna design. It should be noted that multipath is frequency dependent. Therefore, carrier phases are less affected than code ranges where multipath can amount to the meter level. The random noise mainly contains the actual observation noise plus random constituents of multipath (especially for kinematic applications).

The measurement noise, an estimation of the satellite biases, and the contributions from the wave propagation are combined in the User Equivalent Range Error (UERE). This UERE is transmitted via the navigation message. In combination with DOP factor, UERE allows for an estimation of the achievable point positioning precision.

Selective Availability (S/A) The errors described above are called the natural errors, systematic errors or biases¹. Sometimes noise² is introduced intentionally into the GPS satellite clocks (δ -process³), which reduces their accuracy. When implemented, it is the largest source of error in the GPS system. S/A is also known as *intentional error*. This error is created intentionally for *selective availability*, for example, for civil and military purposes. Also, the satellites may be given slightly erroneous orbital data, which will be transmitted back to the user as part of each satellite's status message (ϵ -process⁴). Those two factors can add up to a significant reduction in the accuracy of measurements made using the GPS.

Notes

- i. *If the user was to plot the output of a stationary receiver while S/A was in effect he/she would see its position solutions wander around within about a 100-metre circle.*
- ii. *Military receivers have a decrypting key that tells them what errors have been introduced so they can remove them. So for military users, basic GPS is a much more accurate system, probably capable of accuracies of the order of 15 metres.*

Differential GPS, as discussed in the section to follow, counteracts almost all of the above described errors. It tightens up the working accuracy of GPS to just a few metres, which is even better than the military encrypted signal. It also gives us a way of verifying the reliability of our measurements moment by moment. If some transient glitch in the system causes a satellite to send an erroneous signal, a differential GPS system will detect the problem and tell all the receivers using its corrections not to use that satellite's data.

Typical Error Budget (in metres) for a 100-km long base line are given in Table 9.2.

¹ A systematic error in a process that causes all measured values to deviate from the true value by a consistent amount in a consistent direction (higher or lower than).

² Error caused by noise is measured in terms of a distance which is the fraction of the time that the system output due to noise-corrupted signal differs from that due to signal alone.

³ The δ -process is achieved by introducing varying errors into the fundamental frequency of a satellite clock. Since the fundamental frequency is dithered, the code and pseudoranges are affected.

⁴ The ϵ -process is the truncation of the orbital information in the transmitted navigation message so that the coordinates of the satellites cannot accurately be computed. The error in the satellite position roughly translates to a like position error of the receiver.

Table 9.2 Summary of error sources

<i>Error sources</i>	<i>Standard GPS</i>	<i>Differential GPS</i>
<i>Per satellite accuracy</i>		
Satellite clocks	1.5	0
Orbit error	2.5	0
Ionosphere	5.0	0.4
Troposphere	0.5	0.2
Receiver noise	0.3	0.3
Multipath (reflections)	0.6	0.6
S/A	30	0
<i>Typical position accuracy</i>		
Horizontal	50	1.3
Vertical	78	2.0
3-D	93	2.8

9.13 DIFFERENTIAL GPS

Differential GPS (DGPS) is not just some technical pipe dream. The concept has been around for the quite a while and has been used extensively in scientific and industrial applications around the world. Differential GPS is a system in which differences between observed and computed coordinates or ranges called *differential corrections*, at a particular known point called the *reference station*, are transmitted to users to improve the accuracy of the user's receiver position. In fact, *DGPS is a way to make GPS even more accurate*. It can yield measurements good to a couple of meters in moving applications, and even better in stationary situations. That improved accuracy has a very profound effect on the importance of GPS as a resource and it becomes a universal measurement system, capable of positioning things on a very precise scale. DGPS is mainly a navigation method in which there is a fixed reference station and coordinates of moving platform or rover are determined in real time.

9.13.1 Principle of DGPS

DGPS is based on the concept that bias error in position of the location is similar to those for all locations in a given local area (say within 100 km). Unfortunately, a user can not just figure out the error once and use it to correct all the measurements made for the rest of the day, because the satellite errors are continuously changing. A user needs to have two receivers working simultaneously to do the job. The reference receiver stays put and continuously monitors the errors and then transmits or records corrections for those errors so that the second receiver (the one that is out roving around doing positioning work) can apply these corrections to its measurements, either as it is making them or some time later. Thus, by having a

reference receiver at a fixed location, the user can tune up the accuracy of a roving receiver, or for a whole fleet of roving receivers (Fig. 9.9).

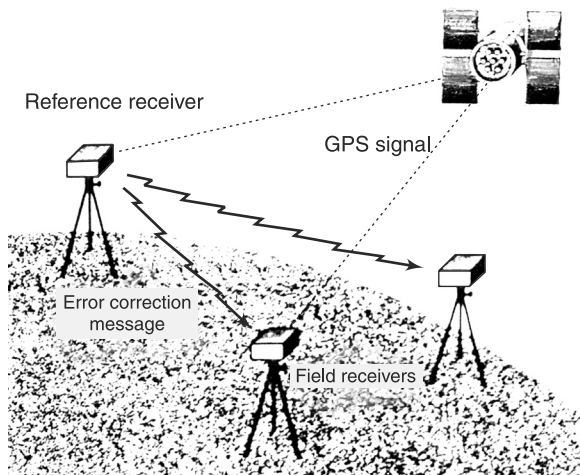


Fig. 9.9 *Differential GPS*

9.13.2 Working with DGPS

In basic GPS, a single receiver can wander around and make pretty good measurements by itself, using the GPS satellites as its reference. DGPS on the other hand, involves cooperation between two receivers, one that wanders around and another that is stationary. The stationary receiver is the key to the accuracy of DGPS. It ties all the satellite measurements into a solid local reference.

In GPS, the receivers use timing signals from at least three satellites to establish a position, and each of those timing signals has its own unique set of errors depending on what sort of perils have befallen it on its trip down to the user. As already discussed, the signals might be a little off even before they start their trip because the satellite clocks may be off and the satellites might be slightly out of place in their orbits. To that, selective availability adds a healthy dose of artificial clock and ephemeris (orbit) errors. Then as the signals travel down to earth they are refracted around in the atmosphere, which causes more delays. Finally those bent and delayed signals might bounce around in the user's local environment to give the user multipath errors. So by the time the receiver gets the signal it will have a pretty rough trip.

Here's where the sheer scale of the system comes to the user's rescue. The satellites are so far out in space that the little distances we travel here on earth are insignificant. That means that if two receivers are fairly close together, say within a few hundred kilometres or so, the signals that reach both of them will have traveled through virtually the same slice of atmosphere and will have virtually the same delays. In fact, most of the errors talked about, excluding multipath and receiver errors, will be common to both receivers. Since both receivers will have virtually the same errors, the user can have one receiver (reference receiver) measure those

errors and provide that error information to the other receiver (roving receiver) via telemetry (i.e., controlled radio link).

The reference receiver is put on a point that has been very accurately surveyed. This reference station receives the same GPS signals as the roving receiver but instead of working like a normal GPS receiver, i.e., instead of using timing signals to calculate a position, it uses its known position to calculate timing. Since the reference station knows where the satellites are supposed to be in space, and it knows exactly where it is, it can compute a theoretical distance between itself and each satellite. It divides that distance by the speed of light and gets a time, that is, how long the signals should have taken to reach it. It compares that theoretical time with the time they actually took. Any difference is the error (or delay) in the satellite's signal.

Now all the reference receiver has to do is give this error information to any roving receivers in the area so that they can use it to correct their measurements. Since the reference receiver has no way of knowing which of the many available satellites a roving receiver is using to calculate its position, it must go through all the visible satellites and compute their instantaneous errors. Then it encodes this information into a standard format and transmits it to the roving receivers. The roving receiver receives the complete list of errors and applies the corrections for the particular satellites they are using. There are a number of techniques to transmit correction factors from the reference receiver to the roving receivers, but the radio link technique is the most common.

One interesting drawback in the system has to do with the speed of the data transfer. The reference receiver cannot take its time waiting to send the corrections, because if it waits too long the corrections will no longer be accurate. Remember that GPS signal errors are constantly changing with the changing condition of the atmosphere and the effects of S/A, so the longer it waits the less applicable the corrections will be. Of course, all differential reference receivers are going to take some time to calculate the correction data. This period is called the *receiver's latency*. Latency also includes the time it takes to transmit the data, which can be significant. Some differential receiver links operate at transmission rates as low as 50 baud (bits per second), which means they could take over ten seconds to transmit the corrections for all the visible satellites. Experience has shown that an update rate of once every 5 seconds is much better, especially if S/A is in effect.

However, all DGPS applications do not need this radio link because some jobs do not require real-time corrections. It is one thing if the user is trying to position a drill bit over a particular spot on the ocean floor from a pitching ship, and quite another if the user just wants to record the track of a new road for inclusion on a map. For applications like the latter, the roving GPS receiver only needs to record all its measured positions and the exact time it made each measurement. Then later, this data can be merged with corrections recorded at the reference receiver for a final 'clean-up' of the data. This is known as *post-processed DGPS*.

There is also a little variation on this theme called *inverted DGPS* which may become important in fleet management situations. Say, the user has got a fleet of trucks who report their positions periodically to a base station. Instead of sending

differential corrections back to the truck, the user just corrects them at the base station. The trucks would only know their positions to 'raw' GPS accuracy, but the dispatcher would be able to pinpoint every truck right down to the side of the road it is on.

9.14 GPS SURVEYING TECHNIQUES

The position of stationary or moving object can be determined through GPS. When the position of a stationary or moving object is determined with respect to a well defined coordinate (x, y, z) by using a single GPS receiver and by making observations to four or more satellites, it is called *point positioning* or *absolute positioning* (Fig. 9.10). However, if the coordinates of an unknown point

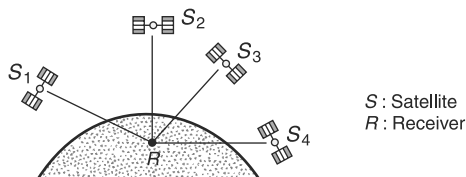


Fig. 9.10 Point positioning

are determined with respect to a known point (taking this as origin of a local coordinate system), it is called *relative positioning*. In other words, relative positioning aims at determination of the vector between the two points, by observations to four or more satellites by two receivers placed at the two points simultaneously (Fig. 9.11). In case the object to be positioned is stationary, it is termed as *static positioning*, while if the object is moving, it is called *kinematic positioning*. GPS surveying implies the precise measurement of the vector between two GPS instruments. The GPS surveys may be classed as static surveys—the traditional static surveying; dynamic surveys—the rapid-static surveying, pseudo-static surveying, and kinematic surveying.

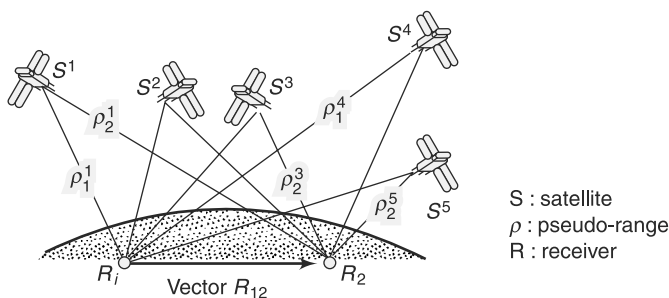


Fig. 9.11 Relative positioning

9.14.1 Static Survey Technique

Traditional Static Surveying This was the first high-precision method developed for GPS requiring at least two receivers. Both the receivers involved remain fixed in position. Accuracy obtained is 5–10 mm plus 1–2 ppm of the base line length. In this technique of GPS positioning, one base receiver is placed over a point of known coordinates, while others are placed over new permanent stations to be positioned. The limitation of the method is the need of hours of observation required; observation times depend on the receiver, the satellite's

geometric configuration, the length of base and atmospheric conditions. This technique is used for long lines in geodetic control, control densification, and photogrammetric control for aerial surveys.

9.14.2 Dynamic Survey Technique

Dynamic surveying implies some sort of motion. It allows user to move during surveying and to collect data on move. More often, the term *kinematic* is used in place of dynamic in the literature. However, the two terms are a little different. This can be explained by the example of modelling the movement of a vehicle in three-dimensional space which requires either the knowledge of the forces causing the motion or the measurement of the vehicle motion in a given three-dimensional coordinate system. The first type of modelling may be called dynamic, the second kinematic. The modelling of the orbit for GPS satellites is a *dynamic* procedure. As soon as the positions of the satellites are assumed to be known and given, positioning of a moving vehicle can be considered as kinematic procedure. Following are the various methods of this surveying technique.

1. *Rapid static surveying*: This technique, also known as fast-static technique, is much like static surveying except that the occupation times are shorter. Typically, the receivers need only occupy a base line for about 5–30 minutes to determine the components (Δx , Δy , and Δz) between the stations. Rapid static technique ideally requires one receiver to be positioned on a station of known coordinates while the others move from station to station. With this technique, there is no need to maintain lock on satellites while moving rover receivers. Accuracy of a few millimeters (10–30) are possible using this technique. Night observations take about half the time (5–15 min) than do day observations (15–30 min). The dual-frequency receivers involved in this technique are used over short (15 km) lines. The real power of this technique is the ability to solve the base lines in a very short period of time without having to maintain lock on a preset number of satellites.

2. *Pseudo-static surveying*: This is also known as *pseudo-kinematic* technique or *intermittent static* or *snapshot static*. It is more flexible than kinematic technique, since lock on satellites need not be maintained while travelling between stations, but it is not as quick. Also, it is faster than using static technique, but it is not as accurate. However, it is slower as well as less accurate than using the rapid static technique. This technique falls between static and kinematic techniques in terms of productivity. In this technique a pair of receivers occupies a pair of points for a small period (2–5 min) that are separated in time by 30–60 minutes. This technique is used when fewer than four satellites are available, or when the Geometric Dilution of Precision⁷ (GDOP) is weak (above a value of 7). The processing software will combine the satellite observations to provide a solution.

That is, if there are only three satellites available on the first occupation and the other three are available on the second occupation, the software will process

⁷The multiplying factor that modifies ranging error. It is caused solely by the geometry between the user and his set of satellites. It is also known as Dilution of Precision (DOP).

the data as if six satellites were available at once. This technique works well when a large number of sites have to be visited as this tends to reduce waiting periods between station re-occupations. This technique is sometimes also called *reoccupation technique*. The accuracy of coordinates obtained using this method varies between 10 and 30 m.

3. *Kinematic surveying*: Compared to static GPS surveying, the kinematic surveying greatly reduces the amount of time needed to determine the GPS vector. In this technique, a reference receiver is placed over a known point while a roving receiver moves around, stopping momentarily on various points. One of the receivers must always be stationary while the other is moving. In this method, both the base unit and roving unit occupy a base line and start from a point of known coordinates. The roving receiver then moves to new locations, keeping a lock on the satellite signals. If the lock is lost, the receiver is held stationary until the initial ambiguities are resolved so that the survey can continue. Observation times required are 10 sec (2 epochs) to compute accurate positions. Typically, relatively open areas are best suited for kinematic surveys; because locks on at least four satellites are needed all the time and a clear line of sight to the satellites must be available along the entire roving receiver route. A similar technique, called Real Time Kinematic (RTK), can commence without the roving receiver first occupying the base line. The base station transmits code and carrier-phase data to the roving receiver, which can use these data to help resolve ambiguities and to solve for change in coordinate differences between the reference and roving receivers. This technique can utilise either single-frequency or dual-frequency receivers. It is also known as On-the-Fly Ambiguity Resolution Technique.

4. *Traditional kinematic technique*: This technique computes a relative differential position at preset time intervals instead of at operator-selected points. This technique is used for road profiling, locating positions of ships and aircraft positioning.

Table 9.3 gives a summary of the relative accuracy and speed for each of the four GPS surveying methods: static, fast static, pseudostatic, and kinematic.

Table 9.3 *Relative accuracy of GPS methods*

<i>Method</i>	<i>Time</i>	<i>Accuracy</i>	<i>Restrictions</i>
Static	(For baselines 20 km or less) Slowest: 1 hour	High accuracy: Sub-centimeter, 1/100 000 to 1/5 000 000	None
Fast-static	Fast: 5–20+ minutes Depends upon baseline length and satellite geometry	Medium accuracy: 1/100 000 to 1/1 000 000	Need fast static capable receivers

(Contd.)

(Contd.)

Pseudostatic	Slow: 20 minutes per vector (Two 10-minute occupations)	Lower accuracy: 1/50000 to <1/500000	Requires two visitations to a point, at least 1 hour apart
Kinematic	Fastest: 2 minutes or less per vector	Medium accuracy: 1/100000 to <1/1000000	Requires survey initialisation and 4-satellite tracking at all times

9.15 INDIAN COORDINATE SYSTEM FOR USING GPS

GPS provides coordinates in a geocentric geodetic reference system known as World Geodetic System (WGS-84). This implies that the origin in WGS-84 system is centre of gravity of earth. However, for the Indian geodetic system Everest spheroid is considered as reference surface. The origin of this local geodetic datum, which best fits to the Indian subcontinent, has its origin away from the geocenter. Kalyanpur in central India is selected as initial point (origin) of the Indian geodetic coordinate system. Therefore, the positioning determined by GPS in India may result in difference of as much as 100 m. For the Indian geodetic datum based on Everest reference surface parameters of datum at Kalyanpur are as follows.

Latitude:	24°07'11.26"
Longitude:	77°39'17.57"
Meridional deflection of vertical:	-0.29"
Prime vertical deflection of vertical:	+2.89"
Semi-major axis of Everest spheroid:	6,377,301 m
Flattening of Everest spheroid:	1/300.8017
Geodetic undulations:	0 m
Azimuth to Surantal:	190°27'06.39"

Horizontal control over the survey is achieved by providing network of control stations all over the country by various surveying methods. To convert the coordinates of Indian geodetic reference system to WGS-84 geocentric system for precise positioning, the transformation parameters are required. Since reasonably accurate transformation parameters are not available for point positioning which can be applied in every part of the country, following procedure may be adopted in using GPS in India.

1. Two single frequency or double frequency precise receivers may be used while using GPS in relative positioning static mode.
2. The distances are determined and used to obtain the positions of points as in traditional surveying methods.
3. Slope distances are reduced to chord distances in space and then reduced to chord distances at spheroid level which is then further used in computations of other quantities.

9.16 USES AND APPLICATIONS OF GPS

GPS is a complex system that can be used in many ways. For basic point positioning using geodetic receivers with a computer and post-processing

software accuracy at the centimetre level is achievable. Although the accuracy is important, some surveyors feel that the main advantage of GPS is that it can be used in any weather conditions day or night. This enables GPS surveying to be carried out over extended periods at any time of the year without restrictions such as rain, fog and poor visibility. Another advantage when surveying with GPS is that intervisibility between stations or points surveyed is not necessary. This allows control stations to be placed conveniently. Further, with differential GPS accuracy of 100 m for navigation purposes, 1 m for mapping and a few millimetres for geodetic positioning is possible. However, the high cost of GPS surveying has restricted the realisation of the full potential of GPS till date. Added to this, there are difficulties in defining heights above datums such as mean sea level and with real time data processing and control. Despite these drawbacks, GPS has been very successfully used in surveying and other fields.

GPS is a tool that will provide the world a new “international standard” for defining locations and distances and it would allow nations to monitor and use natural resources more efficiently than ever before. The general overall uses of GPS are numerous. GPS itself, or in combination with Geographic Information System (GIS), or other spatially related data bases has emerged as a new, dynamic, spatially related information utility. This utility can process both spatial data and relational data, as well as it is capable of the real-time processing required for navigation and routing. The initial conception of GPS was military positioning, weapons aiming and navigation system. It was to replace TRANSIT and other navigation systems and to provide worldwide weather-independent guidance for military use. But because of its potential GPS will soon be part of the overall utility of technology. Some of the uses and applications global, regional and local are as follows.

9.16.1 Navigation

Marine and air navigation are perhaps the two most obvious applications of GPS. Both military and civilians wish to know their spatial locations as precisely as possible. The regional applications of GPS include exploration, transportation management, structural monitoring, and various types of automation. As a local application, GPS can aid in berthing and docking of large vessels.

For approaches to airports and harbours, Differential Global Positioning System (DGPS) has been developed. In order to achieve high accuracies in approach and landing, fixed receivers must broadcast range corrections to incoming aircrafts so that they can compute more accurate positions (by DGPS) as they approach the runway. GPS aided approach and landing will be an economical answer to airport control.

9.16.2 Surveying

The combination of real-time positioning, mobile data communications, and on-board data processing and applications software, all contribute to a new era in surveying.

GPS software has capabilities in mission planning with graphic displays at each step, loop closures, network adjustments, interface to data collectors and surveying

software. GPS can be applied to programs for topography, radial and linear stakeout, cut/fill; and for corrections to sea-level, curvature and refraction, datum transformation and geoid. Global applications of GPS provides a powerful geodetic tool. Global changes over time, a key to understand long-term geodynamical phenomena, can be monitored with GPS. Applications include measuring crustal deformations, post-glacial rebound, volcanic uplift, plate tectonics, and earth rotation. Geodists have long desired to measure crustal movements for various scientific purposes. One use would be to predict earthquakes by measuring certain precursor ground movements. The accurate positioning through GPS enables the development of precise seismic maps and location of drill sites with respect to geological structures. Also, in oil exploration activities, the risk of drilling dry wells will be the least. The availability of GPS and DGPS is a boon to these activities.

GPS techniques permit the collection of data on specified profile, cross section, and boundary locations; contours may be readily plotted from the collected data. The surveying data can be collected by a single surveyor and can be processed to the required accuracy using inexpensive computing facilities. Local use of GPS include local property and site survey, to perform topographic survey, to use GPS kinematic surveying to determine the coordinates of the photocentre during aerial mapping flights.

GPS is very useful for layout works. One base receiver supported by many rover receivers permits the instantaneous layout of boundaries, pipelines, roads and building locations.

GPS is also useful, when beginning the survey, in locating boundaries and control markers that may be covered by snow or other ground covers. Data captured using these techniques can be added to a mapping or GIS database or directly plotted to scale using digital plotter.

GPS techniques are particularly useful in open-pit mining where original, in-progress, and final surveys can easily be performed for quantity and payment purposes. Also, material stock-piles can be surveyed quickly and volumes can be computed directly.

9.16.3 Mapping

A major early implementation of GPS was in the provision of ground truthing, or orientation of aerial photogrammetry. Aircrafts or spacecrafts are used to photograph large areas of the earth's surface. Index marks are often surveyed on ground to provide reference locations on these photographs, which can be used in determining their scale and orientation. GPS can be used to survey these references. Further, the use of these references can be eliminated altogether if the position of the camera can be known accurately enough at the precise moment it took the picture. This technology has been developed using GPS augmented by accurate inertial navigation system (INS). The generation of roads map, or any other kind of feature map, is now extremely easy, achieved simply by recording a series of positions as a receiver is moved over the area to be mapped.

9.16.4 Remote Sensing

It is also possible to integrate GPS positioning into remote-sensing methods such as photogrammetry and aerial scanning, magnetometry, and video technology. Using DGPS or kinematic techniques, depending upon the accuracy required, real time or post-processing will provide positions for the sensor which can be projected to the ground, instead of having ground control projected to an image.

9.16.5 Geographic Information System

GPS are becoming very effective tools for GIS data capture. Anyone charged with the responsibility of managing a distributed inventory, such as might be the case with the utility, municipality, or goods yard, might appreciate the ability to locate and identify this inventory quickly and accurately. This is the role played by GPS in conjunction with GIS. The initial input of data and timely updating is a huge task using conventional means of data collection. With GPS, it is possible to capture position-reinforced data in the field with a simple hand-held computer. As an example, there may be a mix of fluorescent, sodium and mercury street lights, with several varieties of each in a city. The maintenance engineer, capable of recognising the types, can use GPS-based data collector to log the location of each type of installation. This information can be loaded into a central database, so that when maintenance is necessary, the appropriate replacements can be ordered, stocked, and dispatched.

A rapidly growing and highly visible endeavour is the management of natural resources. Another field where GIS is prevalent is the environmental impact assessment studies involving the collection of large amounts of position-related data. Here, GPS can be instrumental in data collection.

Common applications of GIS where GPS can provide three-dimensional information about the features such as engineering mapping, automated photogrammetry, subdivision design (cut/fill, street layout, parcel layout), cadastral mapping, highway mapping, utility/facility mapping and management, surface-water mapping, event mapping (accident, crime, fire, facility breakage, etc.), census and related statistical mapping, watershed prioritisation, land-use planning and management, and environmental impact studies.

9.16.6 Military and Space

GPS was conceived and constructed for the use of military. The use of GPS in spacecraft has been widespread. GPS has flown on several shuttle missions, and has been useful in providing better orbital positioning in much shorter time than had been previously possible. Incorporation of GPS altitude determination in space station design is currently mandated.

9.16.7 Agriculture

GPS is opening a new era of precision farming in the advance countries. A farmer can analyse the soil condition of every region of his farm and compile a fertiliser-demand map. This map is digitised and stored in the GPS system computer. As the chemical spreader moves through his fields, its GPS-measured position is

correlated with the stored demand map to determine the exact amount of fertiliser or pesticide to be applied at every point. The farmer profits from higher yields and the environment benefits from lower chemical usage.

This same accuracy is also being applied to aerial fertiliser and pesticide delivery. With a DGPS guidance system, pilots can design exact flight paths such as a grid of application swaths and then have the system guide them precisely through it. These systems can also record the actual flight path for reporting purposes.

9.17 GPS ACCURACY

Accuracy achievable with GPS depends on several conditions, e.g., single or multi-receiver operation; single or dual frequency data; receiver noise level; Selective Availability on or off (S/A has been put off since May 2nd 2000); P-code available or not; static or kinematic positioning; real-time or post-processed results; extent of data modelling; accuracy of orbits; and fiducial concept (fiducial points means previously well-determined coordinates). The ultimate accuracy of GPS is determined by the sum of several sources of error already described in the previous section. The contribution of each source may vary depending on atmospheric and equipment conditions. In addition, the accuracy of GPS can purposefully be degraded by Selective Availability.

9.17.1 Satellite Geometry and Accuracy Measures

The accuracy of GPS positioning depends on two factors:

1. The accuracy of a single pseudo-range measurement expressed by associated standard deviation σ_r , and
2. The geometric configuration of satellites used.

The relation between σ_r and the standard deviation of positioning σ^* is described by a scalar quantity, which is frequently used and called *Dilution of Precision (DOP)*:

$$\sigma^* = DOP \times \sigma_1$$

Different *DOP* terms are given by

$$\sigma_H = HDOP \times \sigma_1 \quad (\text{for horizontal positioning})$$

$$\sigma_V = VDOP \times \sigma_1 \quad (\text{for vertical positioning})$$

$$\sigma_P = PDOP \times \sigma_1 \quad (\text{for 3-D positioning})$$

$$\sigma_T = TDOP \times \sigma_1 \quad (\text{for time determination})$$

The combined effect of position and time is called

$$GDOP = [(PDOP)^2 + (TDOP)^2]^{1/2}$$

PDOP can be interpreted as the reciprocal of the volume *V* of a tetrahedron that is formed from the satellite and user positions and is given as

$$PDOP = 1/V$$

The best geometric solution exists when the volume is maximised.

The second criterion is that the ranges from satellites intersect nearly at right angles. *PDOP* should favourably be, 3 or less than 3.

Error Budget (Source: Trimble Navigation)*Error source (typical)*

<i>Satellite clock error</i>	<i>2 feet</i>
<i>Ephemeris error</i>	<i>2 feet</i>
<i>Receiver error</i>	<i>4 feet</i>
<i>Atmospheric/ionospheric</i>	<i>12 feet</i>
<i>Worst case S/A (if implemented)</i>	<i>25 feet</i>
<i>Total (root-square sum)</i>	<i>15 to 30 feet</i>
<i>(depending on S/A)</i>	

To calculate the predicted accuracy, multiply the total error by *PDOP* (Position Dilution of Precision). *PDOPs* under good conditions range from 4 to 6. So the position accuracy that can be expected would be

1. Typical-good receiver	60–100 feet
2. worst case	200 feet
3. if S/A implemented	350 feet

Summary

The static information from GIS when added to the global positioning system (GPS) provides a new information utility, which has a great impact on modern society. Land locations can be identified by their 3-D coordinates, moving vehicles can be traced as they move with direction and time continuously updating their positions in the database. GPS in conjunction with GIS and other spatially related databases has emerged as a new dynamic spatially-related information utility. The utility will be capable of real-time processing required for navigation and routing.

GPS is based on a constellation of 24 satellites orbiting the earth over twenty thousand kilometres up. These satellites act as reference points from which receivers on ground triangulate their positions. The satellite can act as reference points because their orbits are very accurately monitored by ground station here on earth. By measuring travel time of signal transmitted from satellites, a GPS receiver on the ground can determine its distance from each satellite. With distance measurements from four satellites, the receiver can calculate its, latitude, longitude, altitude, course, and speed. Accuracy of receivers can be up to 100 m. The basic functioning of GPS technique and its principle are explained.

The orbital motion of satellites is continuously monitored by ground stations so their precise instantaneous positions are always known. A receiver on ground can determine when a timing signal left the satellite and when it arrived on its own antenna. The difference is the travel time of the signal. The distance is given by the speed of light $\times$ travel time. If three such measurements are made, the intersection of three sphere theory gives two possible points of intersection, i.e., location positions of which only one will be on the earth. A fourth such distance measurement is added to make sure that receiver's clock is truly synchronised with universal time. The principle of 3-D position location via intersection of multiple spheres is explained. The mathematics involved in determining the user position is presented. GPS survey techniques—the static survey technique and dynamic survey technique are described. The best part of GPS is that it allows each signal of earth surface to have unique address.

Differential GPS and the techniques for surveying measurements are described. This method results in refined measurements by making corrections for most of the natural and human errors. Inaccuracies in GPS signal are because of satellite clocks, imperfect orbits

and especially from signal trips through the earth's atmosphere. The earth ionosphere and atmosphere cause delays in the GPS signal that can translate in to errors in position. It can yield measurements good to a couple of metres.

In DPGS, in addition to the basic receiver, a second receiver is placed on a spot of known position. It calculates its position from satellite data and then compares with its own position; the difference is the error in GPS signal. However, this cannot be treated as standard error (differential) of the whole day work as the error will keep on changing. This is the reason that two receivers working simultaneously are used. The reference receiver stays put and continuously monitors the errors, transmits or records correction, so that the second receiver can apply these corrections to its measurement. DPGS has a wide range of applications. Aviation, navigation, coast guarding, natural resource management, offshore exploration, geodetic, transportation and fleet management, agriculture, and social applications, etc., are but a few examples. The chapter ends with the use and application of GPS.

Exercises

- 9.1 Give a brief account of the origin of Global Positioning System.
- 9.2 Give a brief description of GPS. What are its main basic advantages over traditional method of surveying?
- 9.3 What do you understand by Global Positioning System? Give an overview of GPS.
- 9.4 What are the three segments of GPS? Describe them briefly.
- 9.5 Briefly describe the following:
(a) GPS (b) TRANSIT (c) ISS (d) UTC
- 9.6 Write short notes on the following:
(a) Satellite constellation
(b) Operational control segment
(c) Equipment segment
- 9.7 Write short notes on the following:
(a) GLONASS system
(b) Galileo system
(c) Applications of GPS
(d) Indian Coordinate System
- 9.8 Describe briefly the concept of three-dimensional position location via intersection of multiple spheres.
- 9.9 Describe briefly the various methods of GPS surveying. Also, give a brief account of the applicability and limitations of each technique.
- 9.10 (a) Discuss the method used for point positioning and relative positioning in GPS.
(b) Explain difference between relative positioning and differential GPS.
- 9.11 Describe in detail how the following can be determined:
(a) Satellite-to-user range (b) User position
- 9.12 What are the main basic observables in GPS? Write and explain the equation for single pseudo-range observation.
- 9.13 Describe sources of error in GPS. What is meant by Selective Availability?
- 9.14 What do you mean by GPS time?
- 9.15 What are the problems in using GPS data to determine positioning in India and how is it solved?

Objective-type Questions

- 9.1** The satellite constellation of GPS consists of
(a) 4 satellites (b) 6 satellites (c) 18 satellites (d) 24 satellites
- 9.2** What is the minimum number of satellites required from which signals can be recovered to enable a global positioning system receiver to determine latitude, longitude and altitude?
(a) One (b) Two (c) Three (d) Four
- 9.3** A TRANSIT user at the equator can obtain a position fix at the average of about once every
(a) 1 min (b) 10 min (c) 30 min (d) 2 h
- 9.4** The number of orbital planes and satellites in each orbit are respectively
(a) 6, 4 (b) 4, 6 (c) 4, 3 (d) 3, 4
- 9.5** In GPS, satellites are placed in orbits with orbital radius approximately of
(a) 1500 km (b) 15 200 km (c) 26 600 km (d) 18 400 km
- 9.6** The speed of light is
(a) 3×10^8 m/s (b) 2×10^8 m/s
(c) 3×10^9 m/s (d) 2×10^9 m/s
- 9.7** GPS weeks are numbered sequentially and originate with week
(a) 4 (b) 2 (c) 1 (d) 0
- 9.8** Which of the following system provides three-dimensional positioning system(s)?
(a) ISS (b) TRANSIT (c) GPS (d) all the above
- 9.9** Satellite on-board
(a) quartz clock (b) atomic clock
(c) stop watch (d) digital clock
- 9.10** The international atomic time is abbreviated as
(a) AIT (b) IAT (c) TAI (d) ATI
- 9.11** Selective Availability in GPS is a
(a) natural error (b) intentional error
(c) systematic error (d) bias
- 9.12** One of the carrier frequencies in a GPS is
(a) L_1 with frequency 1520 MHz
(b) L_1 with frequency 1575.42 MHz
(c) L_1 with frequency 1227.60 MHz
(d) L_1 with frequency 1230.43 MHz
- 9.13** The range measurements in GPS are made with
(a) L_3 band frequencies
(b) L_2 band frequencies
(c) C_2 band frequencies
(d) C_3 band frequencies
- 9.14** In the computation of satellite-to-user range, the predominant error(s) is/are due to
(a) measurement noise
(b) propagation delay
(c) non-synchronised clocks
(d) deviation of satellite path from the reported ephemeris
- 9.15** Geostationary satellites have
(a) same distance from earth's centre
(b) same speed as earth's rotation
(c) same mass as global weight
(d) same angle with geodetic stations

Answers to Objective-type Questions

9.1 (d)	9.2 (d)	9.3 (b)	9.4 (a)	9.5 (c)	9.6 (a)
9.7 (d)	9.8 (c)	9.9 (b)	9.10 (b)	9.11 (b)	9.12 (b)
9.13 (b)	9.14 (c)	9.15 (b)			

10

FIELD ASTRONOMY

Introduction

Astronomy is the science which deals with heavenly bodies. In general, all the methods of surveying provide the relative positions of points on the surface of earth. Whenever the absolute position of any point is desired, reference has to be made to one or more heavenly bodies, e.g., sun, polaris (north star), etc.

The importance of true meridian for large and lasting surveys has already been stressed in Chapter 3, Volume 1. If the survey is to be located as representing a portion of the surface of the earth, the true bearing of at least one line and the latitude, and the longitude of at least one of the stations must be obtained. Also, if the survey lines are to be retraced in future, the survey work has to be referred to the true meridian rather than to the magnetic meridian, since the former is invariable. In Chapter 5, Volume 1, while discussing the checks on an open traverse, the need for astronomical observations was mentioned. Checks can be made at intervals, on angular work by referring to the true meridian and astronomical observations.

A surveyor, therefore, has to deal with that branch of mathematical astronomy which relates to the determination of true meridian, latitudes, longitudes and time. The latitude and longitude of a survey station are obtained by observation of the sun or a star, to locate its absolute position, and also the true meridian—the direction from it of a true north and south line.

A surveyor is concerned with the angular directions of the sun or the stars relative to his horizon, viz. their angular elevation above it, and their horizontal direction compared with that of a fixed terrestrial mark or line. Since all these quantities change with time on account of the earth's rotation, it becomes essential to observe the time of observation. The branch of surveying which treats these matters is known as *field astronomy*.

Astronomical observations are occasionally needed in field survey work. A complete theoretical explanation of this subject, involving astronomy and spherical trigonometry, would require an entire volume. Fortunately, however, the following relatively limited treatment, coupled with a little practice, will enable an engineer to do all that he will ordinarily be called upon to accomplish.

The foundation upon which the theory of engineering astronomy is based is the relative motions of celestial bodies. On a clear night, the stars appear to be

studded on an imaginary invisible sphere—the *celestial sphere*—revolving about the fixed earth. Owing to this illusion, ancient astronomers, and sometimes even modern ones, found it to be a convenient fiction by which the movement of the stars and heavenly bodies could be represented. Therefore, to have a clear concept of astronomy, it is essential to have a clear concept of the earth, its shape and motions, and the motions of celestial bodies. However, to understand the real and apparent motions of these celestial bodies, a surveyor must be familiar with the geometry of spheres and spherical triangles.

Abbreviations

E.T.	Equation of Time
G.M.T.	Greenwich Mean Time
G.A.T.	Greenwich Apparent Time
G.S.T.	Greenwich Sidereal Time
G.M.N.	Greenwich Mean Noon
G.M.M.	Greenwich Mean Midnight
H.A.	Hour Angle
L.M.T.	Local Mean Time
L.A.T.	Local Apparent Time
L.S.T.	Local Sidereal Time
L.A.N.	Local Apparent Noon
L.M.M.	Local Mean Midnight
L.Std.T.	Local Standard Time
R.A.	Right Ascension
S.T.	Sidereal Time
S.I.	Sidereal Interval

10.1 THE EARTH

A spheroid is the figure traced by the revolution of an ellipse about one of its axes. When the axis of revolution is the major axis, a *prolate spheroid* is generated, whereas when the minor axis is the axis of revolution, an *oblate spheroid* is generated.

The earth rotates about its minor axis, known as the *polar axis*, on an average, once in 24 hours and is considered to be an oblate spheroid. The intersection of the polar axis with the surface of the earth gives two points, the *geographical north* and *south* or the *terrestrial poles*. The particular plane passing through the centre of the polar axis and at right angles to it is known as the *equatorial plane*. The circle that it traces with the surface of the earth is known as the *equator*.

The earth also revolves relative to the sun in a plane inclined at about $23^{\circ}27'$ to the plane of the equator, and a complete revolution takes an year.

The earth's orbit around the sun is elliptical with the sun at one of its foci (Fig. 10.1). On 4th January, the earth in its orbit is nearest to the sun and it is said to be in *perihelion*. On 6th July, the earth is at the most distant point of its orbit from the sun and is said to be in *aphelion*. These two points situated at the two ends of the major axis of the ellipse are termed as *apses of the orbit*.

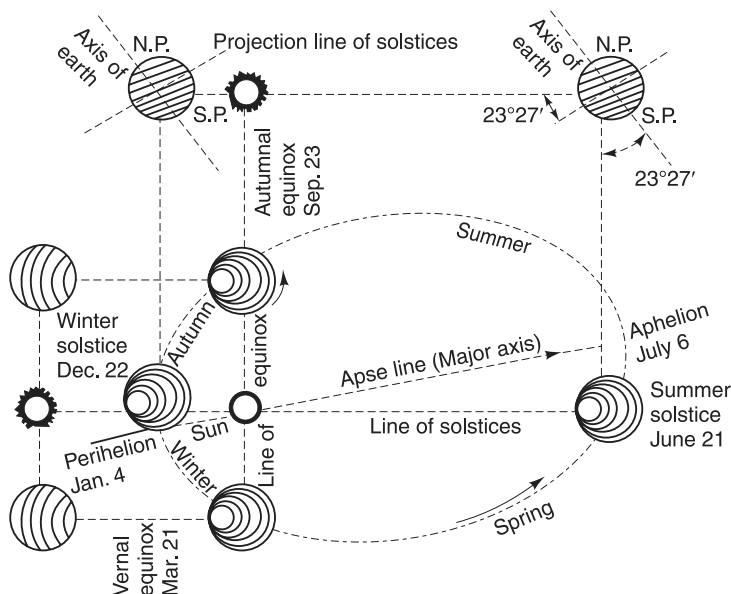


Fig. 10.1 Effects of earth's annual motion

The phenomenon of day and night results from the earth's rotation about its polar axis, whereas the seasons are due to the motion of the earth in its orbit around the sun.

The real motion of the earth is from west to east or clockwise. This can be established from the fact that the stars, sun and planets rise in the east and set in the west. It can also be verified experimentally. At night, in the northern hemisphere, point a camera towards north and expose the film for several hours. A careful study of a picture of the stars would show the apparent motion of the earth which will be counter clockwise. Therefore, the real motion of the earth is clockwise or from west to east. Careful observations on different days, but at the same time, show that the celestial bodies change positions, which indicates that the earth's location in space and its relation to other celestial bodies are continually changing.

10.2 CELESTIAL SPHERE

In making astronomical observations, a surveyor is concerned with the determination of angular positions of stars and not their linear distances from the earth. For this purpose, the universe is considered to be a sphere of infinite radius, with the earth at its centre. The celestial bodies are considered as points studded on the imaginary sphere known as *celestial sphere*. Its center is taken at the eye of the observer.

Since the earth rotates about its polar axis, the celestial sphere appears to rotate. In astronomical problems it is convenient to assume that the earth is stationary and consider the apparent movements of the celestial bodies as real movements relative to the earth. The points at which the polar axis, when produced, intersects the celestial sphere are known as *celestial poles* (P and P'). The celestial equator (Fig. 10.2) is the great circle of the celestial sphere as determined by the plane

passing through the centre of the earth and perpendicular to the polar axis. It passes through the east and west points of the horizon.

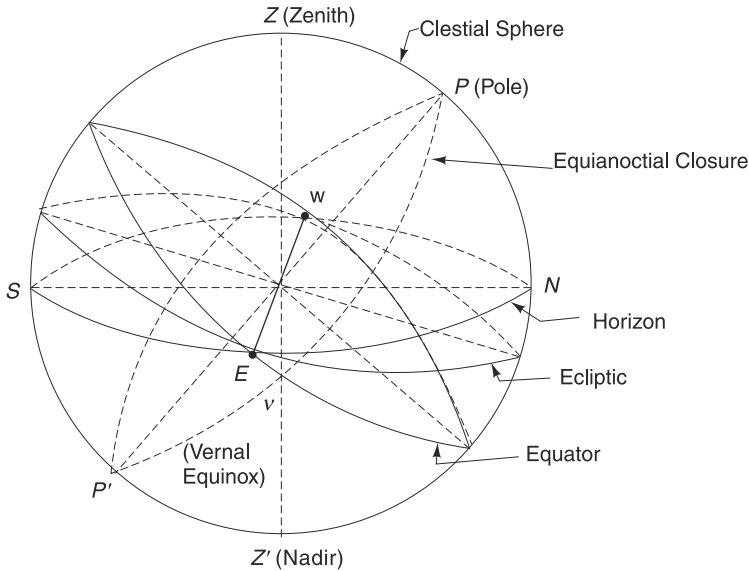


Fig. 10.2 The celestial sphere

10.3 SOLAR SYSTEM

The sun together with its planets and satellites is known as the solar system.

10.3.1 The Sun

The sun is at a distance of 149.6×10^6 km from the earth. Its diameter is about 109 times that of the earth and subtends an angle of $31'59''$ at the centre of the earth. Its temperature is about 6366 kelvin at surface and about 12 million degree kelvin at the centre. The sun is also known as a *yellow dwarf star* of G0¹ group.

The sun is considered to possess a rotary motion around its own axis, but with respect to the earth it is considered to be at rest. The sun has two apparent motions, one with respect to the earth and the other with respect to the stars. As regards the former, the apparent path of the sun is a result of the real motions of the earth and is in the plane of earth's orbit. Since this plane passes through the centre of the celestial sphere, it intersects the latter in a great circle known as *ecliptic* (Fig. 10.2). The plane of an ecliptic is inclined to that of the equator by about $23^\circ 27'$ with an annual diminution of $0.47'$ and is known as *obliquity* of ecliptic. Thus the axis of the earth is inclined to the plane of the ecliptic at an angle of $66^\circ 33'$.

On about 21st March, the axis of the earth is perpendicular to the line joining the earth and the sun. The sun is in the plane of the equator and passes it from south to north. This instant is known as *vernal equinox* (Fig. 10.1), also called *first*

¹ G0 in Morgan Keenan classification system of stars is abbreviated from where G stands for yellow stars and number 0 stands for hottest star.

point of aries (γ). The literal meaning of equinox is equal nights. Sun's H.A. will be 90° at both east (at rise) and west (at set). At this instant, the days and nights are of equal duration throughout the earth and the declination of sun is zero.

On about 21st June, the axis of the earth is in the plane perpendicular to the ecliptic. The sun is vertically on a point on the parallel of latitude $23^\circ 27' \text{N}$, called the *tropic of cancer*. This instant is known as *summer solstice*. The sun at rise and set has an H.A. of more than 90° and thus the day is longer than the night.

On about 23rd September, when again the sun is in the plane of the equator and passes it from north to south, it is called the *autumnal equinox*. This instant is known as *first point of libra* (Ω). The declination of the sun is zero at this point and the days and nights are of equal duration. At present, the vernal equinox does not lie at the first point of aries, but is in the beginning of the zodiac pisces. Also, the autumnal equinox is approximately the last point of vigo. This is because of the precession of equinoxes.

On about 22nd December, the sun is over the *tropic of capricorn* at $23^\circ 27' \text{S}$. The instant is known as *winter solstice*. The sun at rise and set has an H.A. of less than 90° and thus the day is shorter than the night.

10.3.2 Planets

The term planet is derived from the Greek word 'planets' which means 'wanderer'. Like the earth, the other celestial bodies also move in distinct orbits and in the same direction around the sun. The sun has nine planets. Mercury and Venus are nearer to the sun than the earth and are known as *inferior* planets. Mars, Jupiter, Saturn, Uranus, Neptune and Pluto being remote from the sun, in the order specified, are known as *superior* planets. The data about the various planets are listed in Table 10.1. Some additional information about the planets is listed in Appendices VI and VII.

The planets can be distinguished from the stars on the basis of the following. Through an ordinary telescope, planets appear as discs of light, while all the stars appear to be merely points of light. Also, the planets are non-luminous; they reflect the light of the sun, whereas stars are self-luminous and intensely hot. The stars twinkle whereas planets do not.

10.3.3 Satellites

The term satellite is derived from a Latin word 'satellite' which means 'attendants'. The celestial bodies revolving around the planets are known as *satellites*, e.g., the earth has one natural satellite—the moon, mars has two, jupiter has 47, saturn has 17, uranus has five, neptune has two, and pluto has one. The physical characteristics of the satellites of the planets are given in Appendix VIII.

10.3.4 Stars

These are the celestial bodies whose apparent relative positions remain sensibly unchanged over long periods of time. For the purpose of classifying and conveniently distinguishing particular stars, these have been grouped and are called *constellations*. Hence a group of stars, which forms a definite shape is termed as constellation, whereas one not forming a definite shape is known as *asterism*.

Table 10.1 *The solar system*

	<i>Mercury</i>	<i>Venus</i>	<i>Earth</i>	<i>Mars</i>	<i>Jupiter</i>	<i>Saturn</i>	<i>Uranus</i>	<i>Neptune</i>	<i>Pluto</i>
Distance from sun (in millions of km)	57.9	108.2	149.6	227.9	778.3	1427	2 871	4 497	5 914
Diameter (in km)	4 878	12 100	12 102	6 786	142 984	108 728	51 118	49 528	2 284
Mass (relative to earth)	0.055	0.82	1	0.11	317.94	95.18	14.53	17.14	0.0022
Density (water-1)	5.4	5.25	5.5	3.95	1.33	0.69	1.29	1.64	2.03
Time taken for the one orbit around sun (in earth units)	87.97 days	224.7 days	365.24 days	686.98 days	11.86 years	29.46 years	84.01 years	164.79 years	248.54 years
Time taken for one rotation (in earth units)	59 days	243 days	23 hrs 56 mins	24 hrs 37 mins	9 hrs 55 mins	10 hrs 40 mins	17 hrs 14 mins	18 hrs 30 mins	6 days 9 hrs 18 mins
Average temperature (max.) (atmosphere)	427°C	460 °C – 183 °C (min) (Surface)	10 °C	– 58.5 °C	– 140 °C	– 180 °C	– 220 °C	– 230 °C	– 240 °C
Approximate composition of atmosphere	No atmosphere	Dense Carbon dioxide	Oxygen, Nitrogen, and Carbon dioxide	Thin Carbon dioxide	Hydrogen, Methane, Ammonia, Helium	Same as Jupiter	Hydrogen and Hydrogen compounds	Same as Uranus	May be no atmosphere
Known satellites (natural)	0	0	1	2	47	21	15	8	1

There are about 88 constellations out of which only 20 are very prominent. Bayer, in 1603, named the brighter stars of each constellation by assigning them a Greek alphabet followed by the name of the constellation in the generic case. The Greek letters assigned being in order of brightness, the brightest being α , the next brightest β , and so on. Thus α -*tauri* (Aldebran-Rohini) is the most bright star of the constellation taurus, with the next bright star being β -*tauri* (EL-Nath). The letters indicate the brightness of the stars only in the constellation to which they belong and do not serve to compare brightness of stars of other constellations.

The circle of 12 constellations through which the ecliptic passes constitutes the *signs of the zodiac*. The constellations taken in order are aries, taurus, gemini, cancer, leo, virgo, libra, scorpio, sagittarius, capricorn, aquarius and pisces. The sun's greatest north declination is equal to the obliquity of the ecliptic and occurs when the sun's right ascension is 90° . The sun's greatest south declination is also equal to the obliquity of the ecliptic and occurs when the sun's right ascension is 270° .

10.3.5 The Moon

The moon is the only natural satellite of the earth. It revolves around the earth in an elliptical orbit, inclined at an angle of $5^\circ 8'$ to the plane of the ecliptic, which is intersected at points known as *nodes*. The nodes are not stationary, but have a retrograde motion of about 19° per year (Appendix (IX)). The moon has its motion from west to east, relative to the fixed stars. Approximately 59% of the moon is always seen by an observer on the earth because the rotation of the moon about its axis is equal to its revolution about earth.

10.4 DEFINITIONS

1. *Spherical triangle*: A triangle bounded by three arcs of great circles is known as *spherical triangle*. The astronomical triangle *ZPS* is an example of a spherical triangle wherein the zenith *Z*, the pole *P* and celestial body *S* form the three vertices (Fig. 10.3).

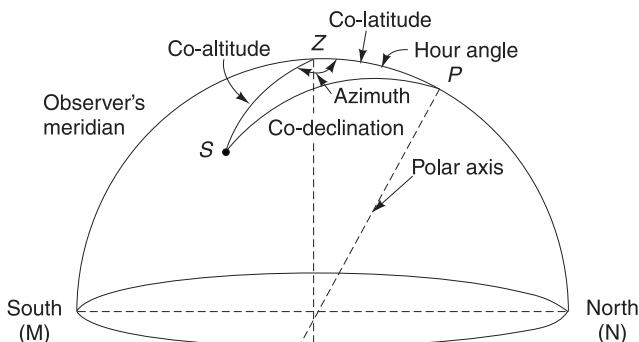


Fig. 10.3 Spherical triangle

2. *Spherical excess*: The quantity by which the sum of the three angles of a spherical triangle exceeds 180° is known as spherical excess of the triangle.

3. *Horizon*: A plane passing through an observer at right angles to the direction of gravity at the place of observation is called the *sensible horizon*. The line of sight of a level telescope lies in this plane. When this plane passes through the centre of the earth, it is known as the *true, rational, or geocentric horizon*. The distance between these planes is insignificant in comparison with the radius of the celestial sphere (infinity) and, therefore, are regarded as coincident.

A great circle on the celestial sphere obtained by the intersection of the true horizon is known as *horizon or celestial horizon* (Fig. 10.4).

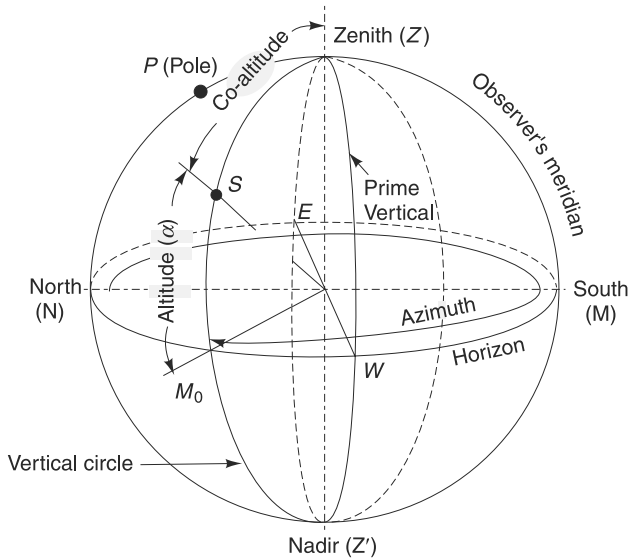


Fig. 10.4 Celestial horizon

4. *North and south points*: The projected points of the north and south poles on the horizon are known as north (N) and south (M) points, respectively.

5. *Zenith and nadir*: Zenith and nadir are the poles of the celestial horizon. These are the points on the celestial sphere obtained by producing the direction of gravity at any point on the earth's surface piercing the celestial sphere in either direction. The corresponding point, on the celestial sphere, above the observer is called the zenith (Z), and the nadir (Z') is the point on this sphere directly beneath the observer (Fig. 10.4).

6. *Meridian*: The *celestial meridian* is the great circle that passes through the poles, the zenith and the nadir (Fig. 10.2). It cuts both the horizon as well as the equator at right angles. The *terrestrial meridian* of an observer at any point on the earth's surface is the great circle passing through the terrestrial poles and the point of observation.

7. *Equinoctial coloure*: A great circle passing through the celestial poles and the vernal equinox is known as *equinoctial coloure* (Fig. 10.2).

8. *Declination circle*: A great circle passing through the celestial body and the poles is known as *declination circle* (Fig. 10.5).

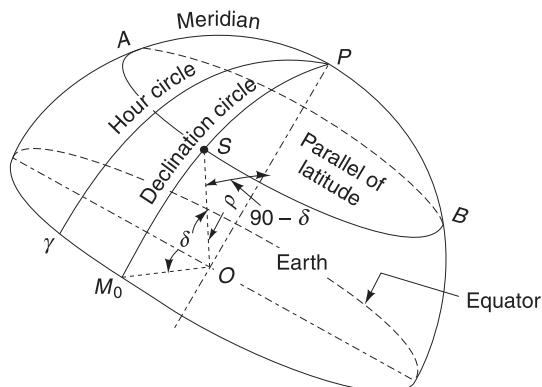


Fig. 10.5 Hour circle

9. *Hour circle*: Projections of the meridians or longitudinal circles in infinite numbers to intersections with the celestial sphere are called *hour circles* (Fig. 10.5). Thus the hour circle of a celestial body will be its declination circle.

10. *Parallel of latitudes*: These are the circles on the earth's surface parallel to the equator. The terrestrial equator is at 0° latitude and all the points on the earth lie on either the south or north latitudes. The terrestrial poles are at 90° latitude, north and south.

11. *Vertical circles*: These are the great circles of the celestial sphere through the zenith and the nadir (Fig. 10.4). The observer's meridian $ZNZ'M$ and the prime vertical $ZWZ'EZ$ are examples of vertical circles.

12. *Prime vertical*: A vertical circle which is at right angles to the meridian is known as a *prime vertical*. It intersects the horizon in the east and west points (Fig. 10.4).

13. *East and west points*: The points on which the prime vertical meets the horizon are known as the east (E) and west (W) points, respectively. These points may also be obtained by the intersections of the equator and horizon.

14. *Altitude (α)*: The altitude of a celestial body is its angular distance above the horizon, as measured on the vertical circle passing through the body (Fig. 10.4).

15. *Co-altitude*: The co-altitude is also known as *zenith distance* (Fig. 10.4). It is the complement of the altitude and is written as $90^\circ - \alpha$. It may be defined as the angular distance between the body and the zenith.

16. *Azimuth (A)*: The azimuth of a celestial body is the angle between the observer's meridian and the vertical circle through the celestial body. It is generally measured on the horizon from 0° to 360° from the north point towards the east point. Thus, the azimuth of stars in Fig. 10.4 is $270^\circ + WM_0$ (E). But sometimes it is measured

towards the west also, in which case the azimuth of stars will be NM_0 (W). When the azimuth of a star is 90°E or 90°W , it is on the prime vertical.

17. *Latitude (θ)*: The latitude of a place is the angle between the direction of a plumb line at the place and the plane of the equator (Fig. 10.6). It is marked plus or minus according to whether the place is north or south of the equator. It may also be defined as the angle between the zenith and the celestial equator.

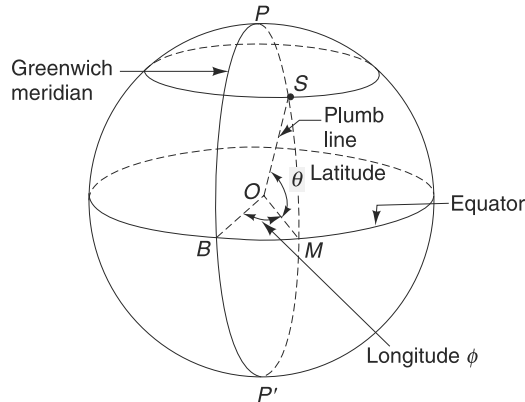


Fig. 10.6 Longitude of a place

18. *Co-latitude*: It is the complement of the latitude and is written as $90^\circ - \theta$ (Fig. 10.3). It is the angle between the zenith and the celestial pole.

19. *Longitude (ϕ)*: The longitude of a place is the angle between a fixed meridian called the prime meridian and the meridian of the place of the observer (Fig. 10.6). The prime meridian universally adopted for astronomical and geodetic work is that of Greenwich. It is measured on the plane of the equator and may be expressed in degrees or hours, minutes and seconds.

20. *Declination (δ)*: The declination of a celestial body is the angular distance measured on a star's meridian north or south of the celestial equator (Fig. 10.5). It can have any value from 0° to 90° and is marked plus or minus according to whether the body is north or south of the equator.

21. *Co-declination*: It is the angular distance between the body and the pole and is written as $90^\circ - \delta$ (Fig. 10.5). It is also known as *polar distance (p)* (Fig. 10.7). These are always positive. When the body is north of the equator, its declination north is positive and declination south is negative, so polar distance is $90^\circ - \delta$ or $90^\circ + \delta$ ($90^\circ - (-\delta)$), respectively. When the body is south of the equator, the polar distance is $90^\circ + \delta$ or $90^\circ - \delta$ as the declination is north or south, respectively.

22. *Right ascension (R.A.)*: The right ascension of a celestial body is the angle between the declination circle through the 'first points of aries' and the declination circle of the body (i.e., hour circle). It is measured eastwards along the equator from the first point of aries, γM_0 (Fig. 10.5) and varies from 0° to 360° . It is also equal to the spherical angle at the pole between the equinoctial colour and the hour circle through the celestial body. Since it is related to time measurement,

it is convenient to express it in hours, minutes and seconds; 24 hours being equal to 360° .

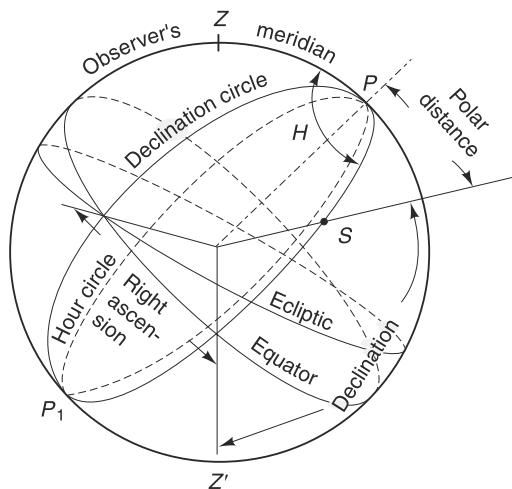
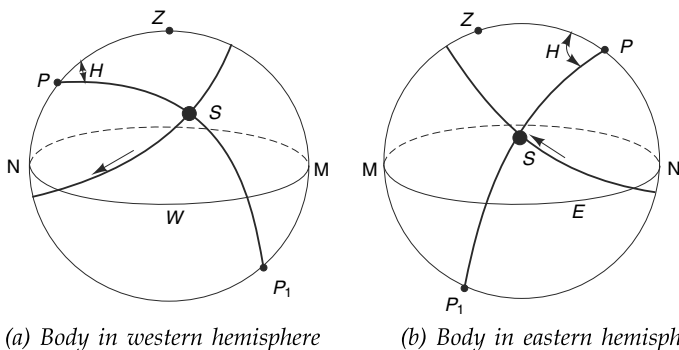


Fig. 10.7 Polar distance

23. Hour angle (H.A.): The hour angle of a celestial body is the angle between the meridian of the observer and the meridian of the declination circle of the celestial body (Fig. 10.3). It is reckoned westwards from 0° to 360° or from 0 to 24 h, each hour being equivalent to 15° . It is also defined as the time interval between the instant of observation of a celestial body and the culmination of that body on the observer's meridian. Owing to the apparent movement of the celestial body, the hour angle keeps on changing constantly.

As stated above, it is measured from 0^h to 24^h ; however, it is customary to quote the H.A. as less than 12^h . When a celestial body is in the western hemisphere, its azimuth is west and its H.A. is between 0^h and 12^h (0° to 180°) and is the spherical angle ZPS (Fig. 10.8(a)) and vice versa. On the other hand, when the celestial body is in the eastern hemisphere, its azimuth is east and the H.A. is between 12^h and 24^h (180° to 360°) and will be $360^\circ - \text{spherical angle } ZPS$ (Fig. 10.8 (b)) and vice versa.



(a) Body in western hemisphere

(b) Body in eastern hemisphere

Fig. 10.8 Hour angle

24. *Culmination*: When a celestial body crosses the observer's meridian, it is said to *culminate* or *transit*. The instant of meridian passage, when the celestial body is above the pole (altitude is maximum), is known as *upper culmination*; when below the pole (altitude is minimum), it is known as *lower culmination*. For example, for star S_2 in Fig. 10.9, L and L_1 are the upper and lower culminations, respectively.

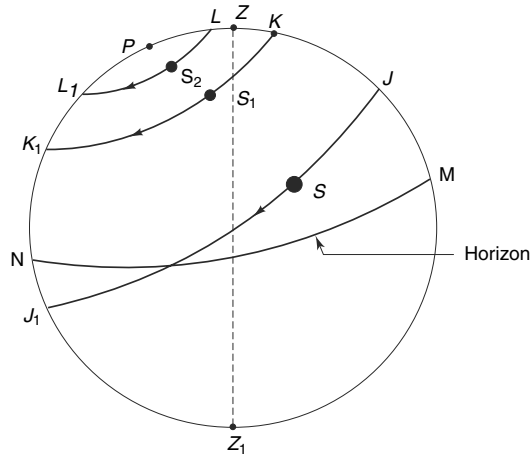


Fig. 10.9 Star at culmination

25. *Elongation*: When a celestial body appears to attain its maximum distance from the meridian, it is said to be at elongation. The elongation can be eastern or western.

26. *Circumpolar star*: The stars having polar distances less than the latitude of the place of observation are known as *circumpolar stars*. These always remain above the horizon and therefore do not set. In field astronomy, stars having polar distances less than 10° are only reckoned as circumpolar stars, e.g., star S_2 (Fig. 10.9).

If δ is declination of the star and θ the latitude of the place, then the polar distance of star S (Fig. 10.10),

$$PK_1 = 90^\circ - \delta$$

Latitude, $PN = \theta$

PK_1 must be less than PN , i.e., $(90^\circ - \delta) < \theta$ or $\delta > 90^\circ - \theta$ (the co-latitude).

Hence, for a circumpolar star, its declination must be greater than the co-latitude.

In Fig. 10.9, for the stars S , S_1 and S_2 , the upper culmination is at J , K and L , respectively, when their altitudes are maximum, and are at lower culmination at J_1 , K_1 and L_1 when their altitudes are minimum. Since the star S has its lower culmination below horizon, it sets during its lower culmination, it is not a circumpolar star. The upper culmination of star S_2 occurs north of the zenith and that of stars S and S_1 south of zenith. Zenith distance of star S_1 at K ,

$$ZK = PK - ZP = (90^\circ - \delta) - (90^\circ - \theta) = \theta - \delta$$

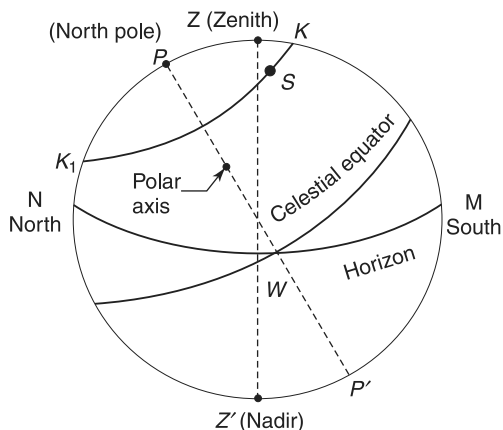


Fig. 10.10 Circumpolar star

Hence, when the declination (δ) of a star is less than the altitude (θ), the upper culmination occurs on the south of the zenith.

Zenith distance of star S_2 at L ,

$$LZ = ZP - PL = (90^\circ - \theta) - (90^\circ - \delta) = \delta - \theta$$

Hence, when the declination (δ) of a star is more than the altitude (θ), the upper culmination occurs on the north side of the zenith.

Also, when $\delta = \theta$, the zenith distance of the star is zero, the culmination of the star occurs in the zenith.

The zenith distance of star S_1 at lower culmination,

$$ZK_1 = ZP + PK_1 = (90^\circ - \theta) + (90^\circ - \delta) = 180^\circ - (\theta + \delta)$$

When a star is at its upper culmination, its hour angle = 0^h

When a star is at its lower culmination, its hour angle = 12^h

10.5 POSITION OF A CELESTIAL BODY

The location of a point on the surface of the earth can be completely specified by the latitude and the longitude. Similar means can be employed to designate the position of a body on a celestial sphere.

Several coordinate systems are available in field astronomy according to the particular reference plane or reference point adopted. In any of the coordinates systems the reference plane should essentially pass through the centre of the sphere. One of the two angular coordinates is measured upon the reference plane, the other being measured at right angles to it. The location of a heavenly body can be stated in one of three ways: the *altitude and azimuth system*, the *declination and right ascension system*, and the *declination and hour angle system*. In the first system of coordinates, the reference plane is the horizon, whereas in the second and third systems, the equator is the reference plane. In the three systems of coordinates, the values required are the azimuth, the altitude, the right ascension, the declination and the hour angle. The azimuth and the altitude can be observed directly with a theodolite but the right ascension, the declination and the hour angle have to be

computed from the solution of the astronomical triangle. Therefore, the azimuth and altitude system is used for field observations, and the other two systems are used for the preparation of star catalogues.

10.5.1 Altitude and Azimuth System

This system is also known as *horizon system*. In this system it is assumed that at the instant of observation, all the celestial bodies are static, so that the celestial body at that instant would be definitely at a fixed location on the celestial sphere.

The vertical axis of a theodolite set up and adjusted on the earth's surface, when extended by an infinite distance, would pierce the celestial sphere at the zenith and nadir. When the telescope is revolved about the vertical axis in a horizontal plane, and this plane is extended by an infinite distance, the great circle intersected on the celestial sphere will be the horizon. When the telescope is revolved about the horizontal axis and the vertical plane so generated is extended by an infinite distance, the great circle intersected on the celestial sphere would pass through the zenith and nadir. Such great circles are known as *vertical circles*. If this latter revolution of the line of sight is made with the telescope pointing towards the true north or south, the particular vertical circle on the celestial sphere will be the meridian of the place.

The horizontal angle measured from true north in a clockwise direction to the vertical circle through the celestial body is the azimuth of the body. The spherical angle at zenith between the vertical circles through the celestial body and the true north equals to the azimuth angle (Fig. 10.4).

These coordinates—the altitude and azimuth, depend upon the position of the observer and also constantly change owing to the apparent diurnal motion of the celestial body about the pole. They cannot be employed for permanently recording the position of heavenly bodies.

10.5.2 Right Ascension and Declination System

This is also known as *equatorial system*. Any celestial body may be located on the celestial sphere by the spherical coordinates, right ascension and declination (Fig. 10.7). Their values, for a celestial body, although nearly constant, are not absolutely so. This is mainly because of precession and nutation. The slow secular movement of the plane of the equator because of the disturbing couple due to the attractions of the sun, moon, and planets on the earth's surface (not spherical) is known as *precession*. On the other hand, *nutation* results because of the variation in the value of this couple, the effect being very small. Both of them cause a variation in the position of the equator and the 'first point of aries' and consequently the declination and right ascension of a star undergo slow alteration. A further source of error in the magnitude of declination or right ascension is from the fact that the stars do not occupy fixed positions on the celestial sphere. The variation is more for the sun than for the stars.

Because of the reasons stated above, the observations of field astronomy do not involve the measurement of declination and right ascension. Since this system is independent of the observer's position, these coordinates are generally used in the reduction of observations and recording the positions of celestial bodies, i.e., for tabulating the data for heavenly bodies in almanacs published annually.

10.5.3 Declination and Hour Angle System

This is also known as *hour angle system*. In this system of coordinates, the reference plane is the plane of the equator. The reference direction is the observer's meridian and the reference point is the point of intersection of the celestial equator by the meridian (Fig. 10.7).

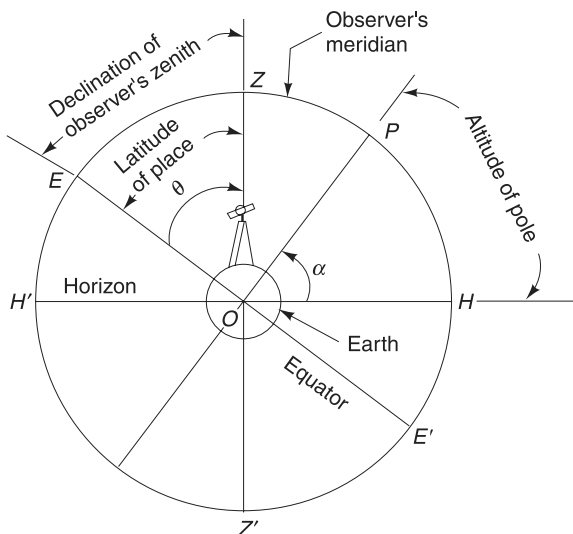


Fig. 10.11 Relation between altitude of pole and latitude

10.6 RELATIONSHIP BETWEEN COORDINATES

1. *The altitude of the pole is equal to the latitude of the observer:* In Fig. 10.11, let O be the centre of the earth, ZO the perpendicular to the horizon, and PO the perpendicular to the equator.

$$\text{Latitude of place} = \theta = \angle EOZ$$

$$\text{Altitude of pole} = \alpha = \angle HOP$$

$$\angle EOP = 90^\circ = \angle EOZ + \angle ZOP = \theta + \angle ZOP$$

$$\angle HOZ = 90^\circ = \angle HOP + \angle ZOP = \alpha + \angle ZOP$$

Equating the two equations

$$\theta + \angle ZOP = \alpha + \angle ZOP$$

$$\text{or} \quad \theta = \alpha$$

Hence, the altitude of the pole is equal to the latitude of the observer.

2. *Relation between latitude of place, declination and altitude of a point on meridian:* In Fig. 10.12, S is the star.

$$ES = \delta = \text{declination}$$

$$MS = \alpha = \text{meridian altitude of star}$$

$$SZ = z = \text{meridian zenith distance of star}$$

$$EZ = \theta = \text{latitude of place}$$

$$EZ = ES + SZ$$

$$\theta = \delta + z$$

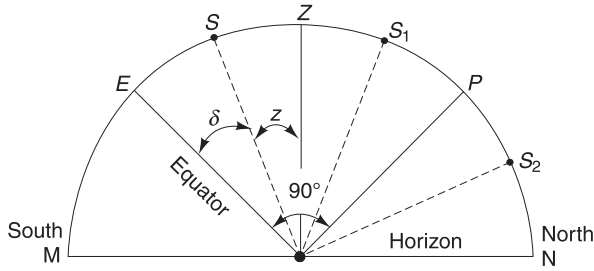


Fig. 10.12

If the star is below the equator, δ will be with negative sign. If the star is to north of the zenith, z will be with negative sign. For the star position S , which is north of zenith and above the pole,

$$\begin{aligned} ZP &= ZS_1 + S_1P \\ (90^\circ - \theta) &= (90^\circ - \alpha) + p \end{aligned}$$

or $\theta = \alpha - p$

when the star position S_2 is north of zenith and below the pole,

$$\begin{aligned} ZS_2 &= ZP + PS_2 \\ (90^\circ - \alpha) &= (90^\circ - \theta) + p \end{aligned}$$

$$-\alpha = -\theta + p$$

or $\theta = \alpha + p$

3. *Relation between right ascension and hour angle:* Figure 10.13 depicts the plan of celestial sphere on the plane of equator, and S is the star.

$$\angle SPM = \text{westerly hour angle}$$

$$\angle \gamma PS = \text{right ascension of star}$$

$$\text{Hour angle of equinox} = \text{H.A. of star} + \text{R.A. of star}$$

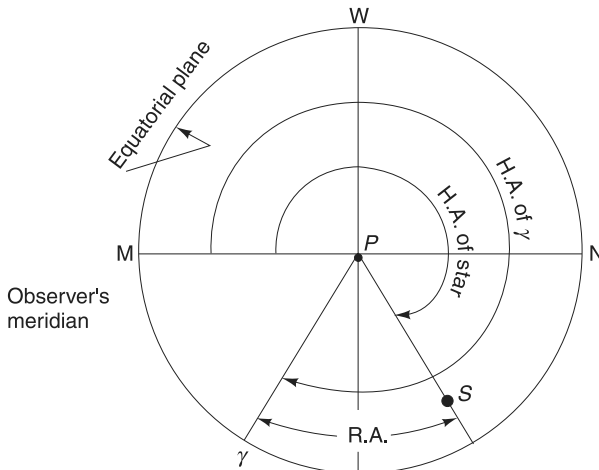


Fig. 10.13

10.7 SPHERICAL TRIGONOMETRY AND SPHERICAL TRIANGLE

Properties of Sphere The important properties of a sphere are as follows:

1. A section of a sphere by any plane is a circle whose radius is inversely proportional to the perpendicular distance of the plane from the centre of the sphere.
2. A section of a sphere is called a great circle if the section plane passes through the centre of the sphere.
3. A section of a sphere is called a small circle when the section plane does not pass through the centre of the sphere.
4. The axis of a circle is a diameter of a sphere perpendicular to a great circle. The two ends of this axis are called poles of the axis.
5. The area of small circles of the same sphere are proportional to the angle they subtend at the centre of the sphere.
10. The great circles which pass through the poles of any other great circle are called *secondaries*, and the given great circle is called the *primary*.
7. The angle between the planes of two secondaries to another great circle, is equal to the arc they intercept on their primary.
8. If the poles of one great circle lie on another great circle, then the poles of the latter will be on the former. The two circles are mutual secondaries.
10. The shortest distance between any two points on the surface of a sphere is along an arc of a great circle passing through the points.

The value of the shortest distance may be obtained by multiplying the radius of earth R by the angle, say α , subtended by the arc of great circle passing through the points at the centre of the earth in radians. Hence,

$$\text{shortest distance} = \frac{\pi R \alpha}{180^\circ}.$$

10. The distance of any point on a small circle from its near pole, is called the *angular radius* or the *polar distance* of the small circle.
11. The length of an arc of a great circle is equal to the angle which it subtends at the centre of the sphere of a unit radius.
12. The angular distance from the pole of a great circle to any point on that great circle is a right angle.
13. The arc of a small circle is equal to the corresponding arc of the great circle multiplied by either the cosine of the distance between the two circles, or the sine of the angular radius of the small circle.

Properties of Spherical Triangle The following are the properties of a spherical triangle:

1. Any angle is less than two right angles or π .
2. The sum of the three angles is less than six right angles or 3π and greater than two right angles or π .
3. The sum of any two sides is greater than the third.
4. If the sum of any two sides is equal to two right angles or π , the sum of the angles opposite to them is equal to two right angles or π .
5. The smaller angle is opposite to the smaller side, and *vice versa*.

Formulae in Spherical Trigonometry The six quantities involved in a spherical triangle are three angles A , B and C and the three sides a , b and c . Out of these, if three quantities are known, the other three can very easily be computed by the use of the following formulae in spherical trigonometry:

$$1. \text{ Sine formula: } \frac{\sin a}{\sin A} = \frac{\sin b}{\sin B} = \frac{\sin c}{\sin C} \quad (10.1)$$

$$2. \text{ Cosine formula: } \cos A = \frac{\cos a - \cos b \cos c}{\sin b \sin c} \quad (10.2)$$

$$\text{or} \quad \cos a = \cos b \cos c + \sin b \sin c \cos A \quad (10.2a)$$

$$\text{Also,} \quad \cos A = -\cos B \cos C + \sin B \sin C \cos a \quad (10.3)$$

3. For computation purposes:

$$\sin A/2 = \sqrt{\frac{\sin(s-b) \sin(s-c)}{\sin b \sin c}} \quad (10.4)$$

$$\cos A/2 = \sqrt{\frac{\sin s \sin(s-a)}{\sin b \sin c}} \quad (10.5)$$

$$\tan A/2 = \sqrt{\frac{\sin(s-b) \sin(s-c)}{\sin s \sin(s-a)}} \quad (10.6)$$

$$\text{where } s = \frac{1}{2}(a + b + c)$$

4. Similarly,

$$\sin \frac{a}{2} = \sqrt{\frac{-\cos S \cos(S-A)}{\sin B \sin C}} \quad (10.7)$$

$$\cos \frac{a}{2} = \sqrt{\frac{\cos(S-B) \cos(S-C)}{\sin B \sin C}} \quad (10.8)$$

$$\tan \frac{a}{2} = \sqrt{\frac{-\cos S \cos(S-A)}{\cos(S-B) \cos(S-C)}} \quad (10.9)$$

$$\text{where } S = \frac{1}{2}(A + B + C)$$

$$5. \quad \tan \frac{1}{2}(a + b) = \frac{\cos \frac{1}{2}(A - B)}{\cos \frac{1}{2}(A + B)} \tan \frac{1}{2}c \quad (10.10)$$

$$\tan \frac{1}{2}(a - b) = \frac{\sin \frac{1}{2}(A - B)}{\sin \frac{1}{2}(A + B)} \tan \frac{1}{2}c \quad (10.11)$$

$$6. \quad \tan \frac{1}{2}(A + B) = \frac{\cos \frac{1}{2}(a - b)}{\cos \frac{1}{2}(a + b)} \cot \frac{1}{2}C \quad (10.12)$$

$$\tan \frac{1}{2}(A - B) = \frac{\sin \frac{1}{2}(a - b)}{\sin \frac{1}{2}(a + b)} \cot \frac{1}{2}C \quad (10.13)$$

Example 10.1 Find the shortest distance between two places A and B , given that the latitudes of A and B are 14°N and $15^\circ 06'\text{N}$ and their longitudes are $70^\circ 10'\text{E}$ and 76°E , respectively. Radius of the earth is 6372 km.

Solution In the spherical angle ABP (Fig. 10.14),

$$\begin{aligned} b = AP &= 90^\circ - \text{latitude of } A \\ &= 90^\circ - 14^\circ = 76^\circ \end{aligned}$$

$$\begin{aligned} a = BP &= 90^\circ - \text{latitude of } B \\ &= 90^\circ - 15^\circ 06' = 74^\circ 54' \end{aligned}$$

$$\begin{aligned} \angle P = \angle APB &= \text{difference of longitudes of } A \text{ and } B \\ &= 76^\circ - 70^\circ 10' = 5^\circ 50' \end{aligned}$$

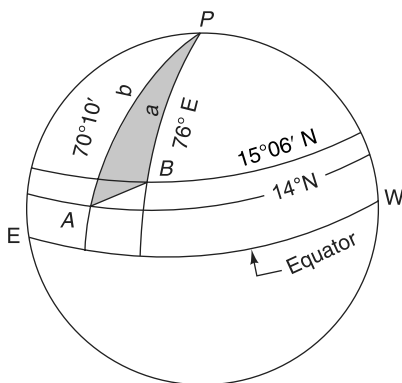


Fig. 10.14

Knowing the two sides a , b and the angle P , the third side p can be obtained by using the cosine rule.

$$\begin{aligned} \cos p &= \cos a \cos b + \sin a \sin b \cos P \\ &= \cos 74^\circ 54' \cos 76^\circ + \sin 74^\circ 54' \sin 76^\circ \cos 5^\circ 50' \\ &= 0.99496 \\ p &= 5^\circ 45' 7.74'' \end{aligned}$$

Knowing the side p , the value of arc p , i.e., the shortest distance will be

$$= \text{radius} \times \text{central angle} = \frac{6372 \times 5^\circ 45' 7.74'' \times \pi}{180^\circ} = 639.71 \text{ km}$$

10.8 ASTRONOMICAL TRIANGLE

Since the radius of the celestial sphere is infinite, the position of an observer on the earth's surface is assumed to be the earth's centre. Let S be the celestial body, the sun. Figure 10.15(a) shows the sun's apparent position above the equator. The declination of the sun is north. The observations have been taken in the morning. Figure 10.15(b) shows the conditions at the instant of afternoon sun observation during the season of the year, when the sun's declination is south. From the two figures, it is observed that the three points: the pole P , the zenith Z and the celestial body S form a spherical triangle known as *astronomical triangle*. To obtain the values of the latitude, azimuth and time, use of the astronomical triangle is made.

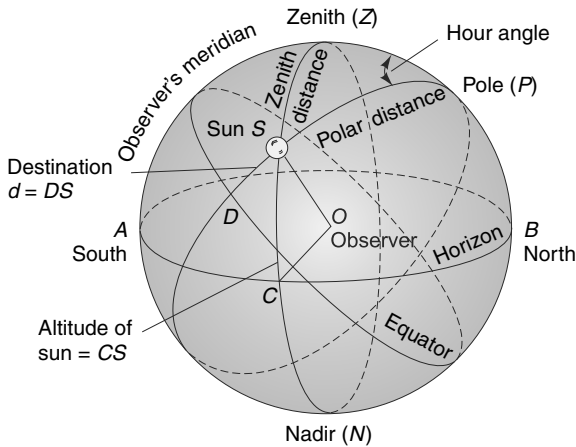


Fig. 10.15(a)

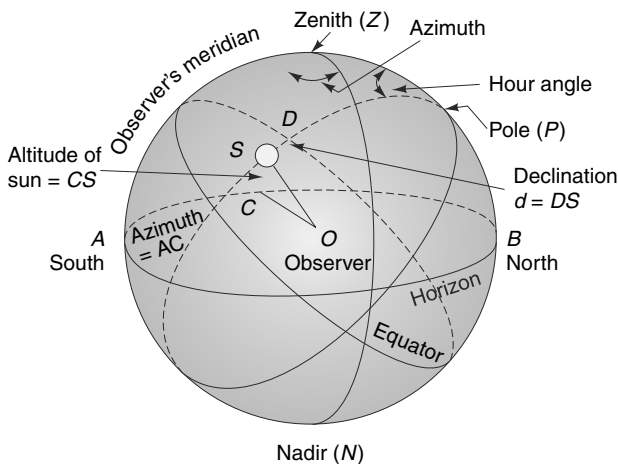


Fig. 10.15(b)

Let α = altitude of the celestial body S
 δ = declination of the celestial body S
 θ = latitude of the observer
 ZP = co-latitude of the observer = $90^\circ - \theta = c$
 SP = co-declination of the celestial body = $90^\circ - \delta = p$
 ZS = co-altitude of the celestial body = $90^\circ - \alpha = z$

The angle at $Z = \angle SZP$ = the azimuth A of the celestial body, the angle at $P = \angle ZPS$ = the hour angle H of the celestial body, and the angle at $S = \angle ZSP$ = the parallactic angle.

If the three sides c, p, z of the astronomical triangle are known, the angles A and H can be computed from the use of any of the following formulae:

$$1. \quad \cos A = \frac{\sin \delta}{\cos \alpha \cos \theta} - \tan \alpha \tan \theta \quad (10.14)$$

$$\cos H = \frac{\sin \alpha}{\cos \delta \cos \theta} - \tan \delta \tan \theta \quad (10.15)$$

$$2. \quad \tan A/2 = \sqrt{\frac{\sin(s-z) \sin(s-c)}{\sin s \sin(s-p)}} \quad (10.16)$$

$$\tan H/2 = \sqrt{\frac{\sin(s-c) \sin(s-p)}{\sin s \sin(s-z)}} \quad (10.17)$$

$$\sin A/2 = \sqrt{\frac{\sin(s-z) \sin(s-c)}{\sin z \sin c}} \quad (10.18)$$

$$\sin H/2 = \sqrt{\frac{\sin(s-c) \sin(s-p)}{\sin c \sin p}} \quad (10.19)$$

$$\cos A/2 = \sqrt{\frac{\sin s \sin(s-p)}{\sin z \sin c}} \quad (10.20)$$

$$\cos H/2 = \sqrt{\frac{\sin s \sin(s-z)}{\sin c \sin p}} \quad (10.21)$$

where $s = \frac{1}{2}(z + c + p)$

Application

Star at horizon: If a star is at horizon, its altitude will be zero and the zenith distance will be 90° . If the latitude and declination are known, the azimuth and hour angle can be calculated from the following equations with $\alpha = 0$:

$$\cos A = \frac{\sin \delta}{\cos \alpha \cos \theta} - \tan \alpha \tan \theta$$

$$\text{or} \quad \cos A = \frac{\sin \delta}{\cos \theta} \quad (\text{when } \alpha = 0) \quad (10.22)$$

$$\text{and} \quad \cos H = \frac{\sin \alpha}{\cos \delta \cos \theta} - \tan \delta \tan \theta$$

$$\text{or} \quad \cos H = -\tan \delta \tan \theta \quad (\text{when } \alpha = 0) \quad (10.23)$$

10.9 NAPIER'S RULES OF CIRCULAR PARTS

The relationships of right angled spherical triangles can be obtained by Napier's rules.

Let ZPS be a right angle, spherical triangle; the angle S being 90° . The five parts of the spherical triangle, the components adjacent to 90° , i.e., angle $(90^\circ - \alpha)$, $(90^\circ - \delta)$, the complement of spherical angle A , i.e., $(90^\circ - A)$, the complement of angle H , i.e., $(90^\circ - H)$ and the spherical angle θ can be arranged as they stand in order in the triangle around a circle as shown in Fig. 10.16. If any part is considered as the middle part, the two parts adjacent to it are termed as *adjacent parts* and the remaining two as *opposite parts*.

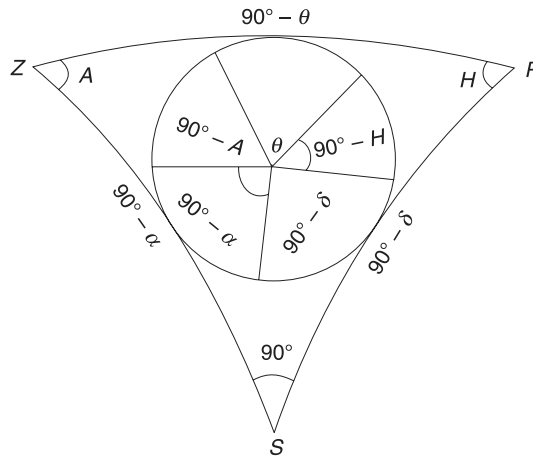


Fig. 10.16

Rule 1 *Sine of middle part = product of tangents of adjacent parts*

For example, let θ be the middle part, then

$$\sin \theta = \tan(90^\circ - A) \tan(90^\circ - H)$$

Rule 2 *Sine of middle part = product of cosine of opposite parts*

As before, if θ is the middle part, then

$$\sin \theta = \cos(90^\circ - \alpha) \cos(90^\circ - \delta)$$

Applications

1. *Star at elongation:* Figures 10.17(a) and (b) shows the star S at its eastern elongation. If the declination and the latitude of the place of observation is known,

the azimuth, the hour angle and the altitude of the celestial body can be found by the Napier's rule (Figs 10.17(c) and (d)).

From Fig. 10.17,

$$\sin (90^\circ - H) = \tan \theta \tan (90^\circ - \delta)$$

or

$$\cos H = \tan \theta \cot \delta \quad (10.24)$$

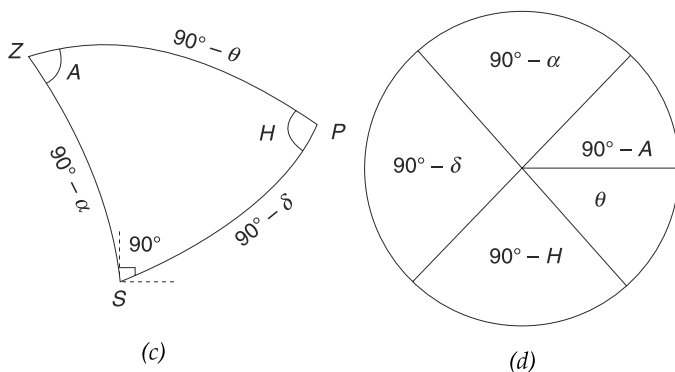
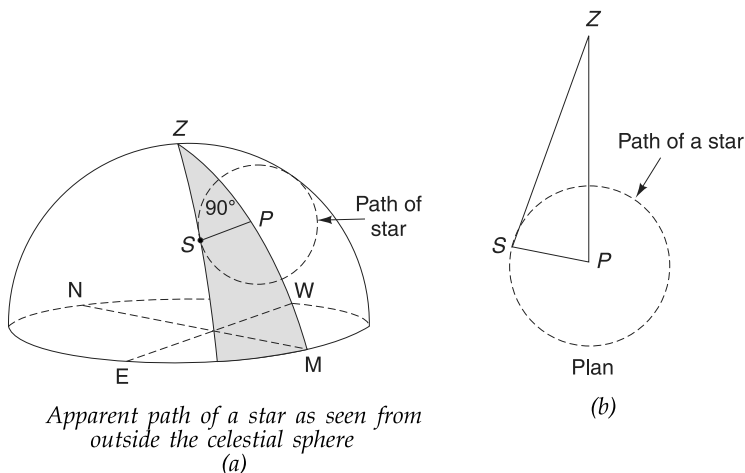


Fig. 10.17

From Fig. 10.17(d),

$$\sin \theta = \cos (90^\circ - \alpha) \cos (90^\circ - \delta)$$

or

$$\sin \theta = \sin \alpha \sin \delta \quad (10.25)$$

also

$$\sin (90^\circ - \delta) = \cos (90^\circ - A) \cos \theta$$

or

$$\cos (90^\circ - A) = \frac{\sin (90^\circ - \delta)}{\cos \theta}$$

or

$$\sin A = \cos \delta \sec \theta \quad (10.26)$$

2. *Star at prime vertical:* If the latitude and declination of the place of observation are known, the altitude and the hour angle can be calculated.

From Fig. 10.18,

$$\sin \delta = \cos (90^\circ - \theta) \cos (90^\circ - \alpha)$$

or $\sin \delta = \sin \theta \sin \alpha$

or $\sin \alpha = \frac{\sin \delta}{\sin \theta}$

or $\sin \alpha = \sin \delta \operatorname{cosec} \theta$ (10.27)

also $\sin (90^\circ - H) = \tan (90^\circ - \theta) \tan \delta$

or $\cos H = \cot \theta \tan \delta$ (10.28)

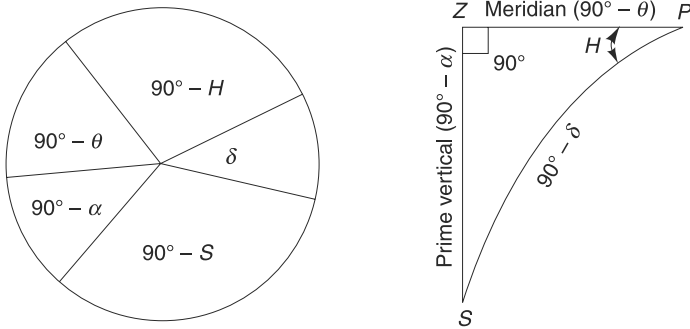


Fig. 10.18

Example 10.2 Find the zenith distance of a star in the following cases:

Star	S_1	S_2	S_3	S_4	S_5
Latitude (θ)	50°15'N	55°30'N	37°30'N	40°15'N	46°15'S
Declination (δ)	22°30'N	20°30'S	60°30'N	48°30'N	60°15'S
Condition of culmination	Upper	Upper	Upper	Lower	Lower

Which star is circumpolar star?

Solution

Star S_1 : $\delta < \theta$

$$\theta = 50^\circ 15', \quad \delta = 22^\circ 30'$$

Hence, the upper culmination occurs on the south of the zenith.

$$\text{Zenith distance} = \theta - \delta = 50^\circ 15' - 22^\circ 30' = 27^\circ 45'$$

Star S_2 : $\delta < \theta$

$$\theta = 55^\circ 30', \quad \delta = -20^\circ 30'$$

$$\text{Zenith distance} = \theta - \delta = 55^\circ 30' - (-20^\circ 30') = 76^\circ$$

Star S_3 : $\delta > \theta$

$$\theta = 37^\circ 30', \quad \delta = 62^\circ 30'$$

Hence, the upper culmination occurs on the north of the zenith.

$$\text{Zenith distance} = \delta - \theta = 62^\circ 30' - 37^\circ 30' = 25^\circ$$

Star S_4 : $\theta = 46^\circ 15'$, $\delta = 48^\circ 30'$

$$\text{Zenith distance} = 180^\circ - (\theta + \delta) = 180^\circ - (46^\circ 15' + 48^\circ 30') = 85^\circ 15'$$

Star S_5 : $\theta = 46^\circ 15'$, $\delta = 60^\circ 15'$

$$\text{Zenith distance} = 180^\circ - (\theta + \delta) = 180^\circ - (46^\circ 15' + 60^\circ 15') = 73^\circ 30'$$

Star	S_1	S_2	S_3	S_4	S_5
δ	$22^\circ 30'$	$20^\circ 30'$	$62^\circ 30'$	$48^\circ 30'$	$60^\circ 15'$
θ	$50^\circ 15'$	$55^\circ 30'$	$37^\circ 30'$	$40^\circ 15'$	$46^\circ 15'$
$90^\circ - \theta$	$39^\circ 45'$	$34^\circ 30'$	$52^\circ 30'$	$49^\circ 45'$	$43^\circ 45'$

A star is said to be circumpolar when its declination is greater than co-latitude ($90^\circ - \text{latitude}$) of the place, i.e., $\delta > 90^\circ - \theta$. Hence, star S_3 and S_5 are circumpolar stars.

Example 10.3 The altitude of a star at the upper culmination is $72^\circ 30'$ and that at the lower culmination is $20^\circ 30'$, both the culminations being to the north side of the zenith of the place. Determine the declination and the latitude of the observer.

Solution Refer to Fig. 10.19.

$$\alpha = 72^\circ 30', 20^\circ 30'$$

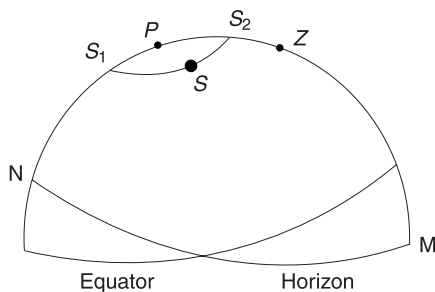


Fig. 10.19

At upper culmination

$$\text{Zenith distance} = \delta - \theta = 90^\circ - \alpha$$

$$\text{or } \delta - \theta = 90^\circ - 72^\circ 30' = 17^\circ 30' \quad (i)$$

At lower culmination

$$\text{Zenith distance} = 180^\circ - (\delta + \theta) = 90^\circ - \alpha$$

$$\text{or } 180^\circ - (\delta + \theta) = 90^\circ - 20^\circ 30' = 69^\circ 30'$$

$$\text{or } \delta + \theta = 110^\circ 30' \quad (ii)$$

From Eqs. (i) and (ii)

$$\delta = 64^\circ 00' \text{N} \quad \text{and} \quad \theta = 46^\circ 30' \text{N}$$

Example 10.4 A star has a declination of $50^{\circ}15'N$. Its upper (culmination) transit is in the zenith of the place. Find the altitude of the star at the lower transit (culmination).

Solution Refer to Fig. 10.20.

Since the star's upper transit is in zenith.

Polar distance = co-latitude

$$\text{or} \quad 90^{\circ} - \delta = 90^{\circ} - \theta$$

$$\text{or} \quad \delta = \theta$$

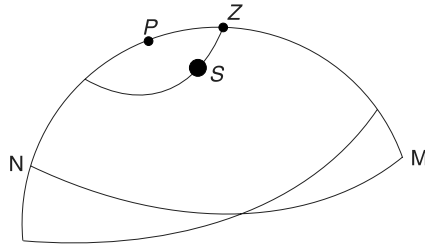


Fig. 10.20

At lower transit

$$\text{Zenith distance} = 180^{\circ} - (\theta + \delta)$$

$$= 180^{\circ} - (\delta + \delta) = 180^{\circ} - 2\delta = 180^{\circ} - 2 \times 50^{\circ}15' = 79^{\circ}30'$$

$$\text{Altitude of the star, } \alpha = 90^{\circ} - \text{zenith distance}$$

$$= 90^{\circ} - 79^{\circ}30' = 10^{\circ}30'$$

Example 10.5 Determine the azimuth and altitude of a star from the following data:

Latitude of observer = $48^{\circ}N$

Hour angle of star = 43°

Declination of star = $18^{\circ}20'N$

Solution Refer to Fig. 10.21. In triangle ZPS ,

$$ZP = 90^{\circ} - \theta = 90^{\circ} - 48^{\circ} = 42^{\circ}$$

$$PS = 90^{\circ} - \delta = 90^{\circ} - 18^{\circ}20' = 71^{\circ}40'$$

$$\angle ZPS = H = 43^{\circ}$$

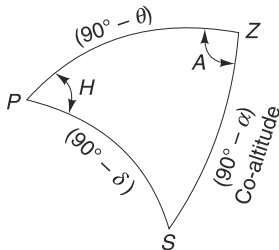


Fig. 10.21

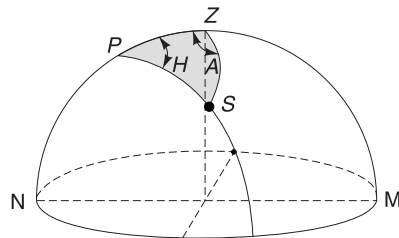


Fig. 10.22

Using cosine formula for H (Fig. 10.22)

$$\begin{aligned}\cos ZS &= \cos ZP \cos PS + \sin ZP \sin PS \cos H \\ &= \cos 42^\circ \times \cos 71^\circ 40' + \sin 42^\circ \times \sin 71^\circ 40' \times \cos 43^\circ\end{aligned}$$

$$\Rightarrow ZS = 43^\circ 51' 38''$$

$$\text{Since } ZS = \text{co-altitude} = 90^\circ - \alpha$$

$$\therefore \alpha = 90^\circ - ZS = 90^\circ - 43^\circ 51' 38'' = 46^\circ 8' 22''$$

Using cosine formula for A

$$\cos PS = \cos PZ \cos ZS + \sin PZ \sin ZS \cos A$$

$$\text{or } \cos A = \frac{\cos PS - \cos PZ \cos ZS}{\sin PZ \sin ZS}$$

$$\frac{\cos 71^\circ 40' - \cos 42^\circ \times \sin 43^\circ 51' 38''}{\sin 42^\circ \times \sin 43^\circ 51' 38''} = -0.4604575$$

$$\text{or } A = 117^\circ 24' \text{W}$$

Note

If the hour angle is less than 180° , the star will be in the western hemisphere. Therefore, the azimuth θ of the star will be western.

Example 10.6 Find the azimuth and altitude of a star from the following data.

Latitude of the place = 46°N

Hour angle of the star = $20^h 40^m$

Declination = $18^\circ 38' \text{S}$

Solution Refer to Fig. 10.23. In the triangle ZPS ,

$$ZP = 90^\circ - \theta = 90^\circ - 46^\circ = 44^\circ$$

$$PS = 90^\circ - \delta = 90^\circ - (-18^\circ 38') = 108^\circ 38'$$

Since, H.A. is more than 12^h (180°), therefore the star is in the eastern hemisphere. Therefore, hour angle will be

$$\begin{aligned}\angle ZPS &= H = 24^h - 20^h 40^m = 3^h 20^m \\ &\quad (\text{since } H \text{ is more than } 12^h)\end{aligned}$$

$$1^h = 15^\circ$$

$$\text{Hence, } 3^h 20^m = 50^\circ$$

$$H = 50^\circ$$

Using cosine formula for H

$$\begin{aligned}\cos ZS &= \cos PZ \cos PS + \sin PZ \sin PS \cos H \\ &= \cos 44^\circ \times \cos 108^\circ 38' + \sin 44^\circ \times \sin 108^\circ 38' \times \cos 50^\circ\end{aligned}$$

$$\Rightarrow ZS = 78^\circ 51' 21''$$

$$\text{But } ZS = 90^\circ - \alpha$$

$$\text{or } \alpha = 90^\circ - ZS = 90^\circ - 78^\circ 51' 21'' = 11^\circ 8' 39''$$

Hence, altitude of the star = $11^\circ 8' 39''$

Using cosine formula for A

$$\cos PS = \cos PZ \cos ZS + \sin PZ \sin ZS \cos A$$

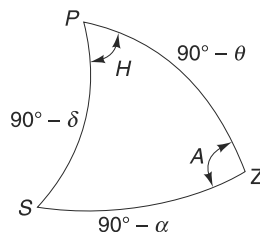


Fig. 10.23

$$\begin{aligned} \text{or} \quad \cos A &= \frac{\cos PS - \cos PZ \cos ZS}{\sin PZ \sin ZS} \\ &= \frac{\cos 108^\circ 30' - \cos 44^\circ \times \cos 78^\circ 51' 21''}{\sin 44^\circ \times \sin 78^\circ 51' 21''} = -0.6727852 \end{aligned}$$

or $A = 132^\circ 16'$
The azimuth of the star is $132^\circ 16'E$.

Note The azimuth of the star is eastern because the hour angle was more than 180° , i.e., the star was in the eastern hemisphere.

Example 10.7 Determine the declination and hour angle of stars for the following data.

Star	Altitude (α) of star	Azimuth (A) of star	Latitude (θ) of observer
S_1	22°	$50^\circ W$	$49^\circ 20' N$
S_2	22°	$150^\circ E$	$49^\circ 20' N$
S_3	22°	$50^\circ E$	$49^\circ 20' N$

Solution Refer to Fig. 10.24.

Star S_1

In triangle ZPS

$$ZP = 90^\circ - \theta = 90^\circ - 49^\circ 20' = 40^\circ 40'$$

$$ZS = 90^\circ - \alpha = 90^\circ - 22^\circ = 68^\circ$$

$$A = 50^\circ$$

Using cosine formula

$$\begin{aligned} \cos PS &= \cos ZP \cos ZS + \sin ZP \sin ZS \cos A \\ &= \cos 40^\circ 40' \times \cos 68^\circ + \sin 40^\circ 40' \times \sin 68^\circ \times \cos 50^\circ \end{aligned}$$

$$\text{or} \quad PS = 47^\circ 44' 17.21''$$

$$\text{Now,} \quad PS = 90^\circ - \delta = 47^\circ 44' 17.21''$$

$$\text{or} \quad \delta = 42^\circ 15' 42.79'' N$$

Hence, declination of star = $42^\circ 15' 42.79''$.

Using cosine formula

$$\cos ZS = \cos ZP \cos PS + \sin ZP \sin PS \cos H$$

$$\begin{aligned} \text{or} \quad \cos H &= \frac{\cos ZS - \cos ZP \cos PS}{\sin ZP \sin PS} \\ &= \frac{\cos 68^\circ - \cos 40^\circ 40' \times \cos 47^\circ 44' 17.21''}{\sin 40^\circ 40' \times \sin 47^\circ 44' 17.21''} = 0.28097 \end{aligned}$$

$$\text{or} \quad H = 73^\circ 40' 53.17''$$

Hence, hour angle is $73^\circ 40' 53.17''$.

Star S_2

$$A = 150^\circ$$

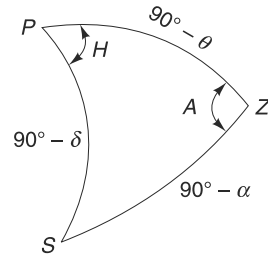


Fig. 10.24

Using cosine formula

$$\begin{aligned}\cos PS &= \cos ZP \cos ZS + \sin ZP \sin ZS \cos A \\ &= \cos 40^\circ 40' \times \cos 68^\circ + \sin 40^\circ 40' \times \sin 68^\circ \times \cos 150^\circ \\ &= -0.2391136\end{aligned}$$

or $PS = 103^\circ 50' 3.23''$

Now, $PS = 90^\circ - \delta = 103^\circ 50' 3.23''$

or $\delta = -13^\circ 50' 3.23'' = 13^\circ 50' 3.23''\text{S}$

Using cosine formula

$$\begin{aligned}\cos H_1 &= \frac{\cos ZS - \cos ZP \cos PS}{\sin ZP \sin PS} \\ &= \frac{\cos 68^\circ - \cos 40^\circ 40' \times \cos 103^\circ 50' 3.23''}{\sin 40^\circ 40' \times \sin 103^\circ 50' 3.23''}\end{aligned}$$

or $H_1 = 28^\circ 31' 07''$

Since the azimuth is eastern, the star S_2 will be in the eastern hemisphere.

Hence, hour angle, $H = 360^\circ - 28^\circ 31' 07'' = 331^\circ 28' 53.57''$.

Star S_3

For star S_3 the calculation for the declination and hour angle will be same as that of star S_1 . But since star S_3 is in the eastern hemisphere, therefore, the actual hour angle value will be obtained by subtracting the calculated H.A. value ($73^\circ 40' 53.17''$) from 360° .

Hence, the hour angle = $360^\circ - 73^\circ 40' 53.17'' = 286^\circ 19' 6.83''$.

Example 10.8 The observations were made for α -lyrae (vega) at Allahabad on 17th July 1992. The declination of star was $38^\circ 45' 48''\text{N}$ and the hour angle was $80^\circ 57' 48''$. Find the coordinates of star vega in horizon system, i.e., altitude azimuth system (latitude of Allahabad is $25^\circ 25' 25''$).

Solution Refer to Fig. 10.25.

Hour angle, $H = 80^\circ 57' 48''$; star is in the eastern hemisphere.

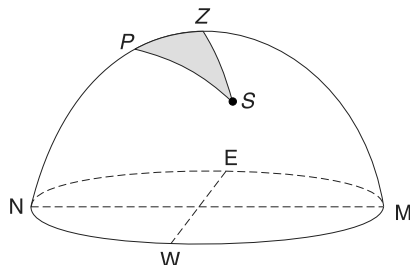


Fig. 10.25

In triangle ZPS ,

Co-latitude of the place = $ZP = 90^\circ - 25^\circ 25' 25'' = 64^\circ 34' 35''$

Co-declination of the star = $PS = 90^\circ - 38^\circ 45' 48'' = 51^\circ 14' 12''$

Hour angle, $H = \angle ZPS = 80^\circ 57' 48''$

Using cosine rule

$$\begin{aligned}\cos ZS &= \cos ZP \cos PS + \sin ZP \sin PS \cos H \\ &= \cos 64^\circ 34' 35'' \times \cos 51^\circ 14' 12'' \\ &\quad + \sin 64^\circ 34' 35'' \times \sin 51^\circ 14' 12'' \times \cos 80^\circ 57' 48''\end{aligned}$$

or $ZS = 67^\circ 42' 12.17''$

Altitude of the star $= 90^\circ - 67^\circ 42' 12.17'' = 22^\circ 17' 47.83''$.

$$\begin{aligned}\cos A &= \frac{\cos PS - \cos PZ \cos ZS}{\sin ZP \sin ZS} \\ &= \frac{\cos 51^\circ 14' 12'' - \cos 64^\circ 34' 35'' \times \cos 67^\circ 42' 12.17''}{\sin 64^\circ 34' 35'' \times \sin 67^\circ 42' 12.17''}\end{aligned}$$

or Azimuth, $A = 56^\circ 20' 46.6''$

Example 10.9 Calculate the hour angle and azimuth of the sun at sunrise when its declination is 18°N . The latitude of the place is 46°S .

Solution Refer to Fig. 10.26.

(a) consider triangle $Z'P_1S$, where P_1 is the south pole and S is the sun.

$\theta = 46^\circ$, $\delta = 18^\circ\text{N} = -18^\circ$

(Since sun's latitude is south)

Since it is sunrise, the sun is at horizon, its altitude will be zero and hence $SZ' = 90^\circ$

$Z'P_1 = 90^\circ - \theta = 90^\circ - 46^\circ = 44^\circ$,

$SP_1 = 90^\circ - \delta = 90^\circ - (-18^\circ) = 108^\circ$

Using cosine rule

$\cos SP_1 = \cos SZ' \cos Z'P_1 + \sin SZ' \sin Z'P_1 \cos A$

or $\cos 108^\circ = 0 + \sin 44^\circ \times \cos A$

or $\cos A = \frac{\cos 108^\circ}{\sin 44^\circ}$

or $A = 116^\circ 24'$

Hence, the azimuth of sun is $116^\circ 24'\text{E}$.

Using cosine rule

$$\begin{aligned}\cos Z'S &= \cos SP_1 \cos P_1Z' + \sin SP_1 \sin P_1Z' \cos H \\ \cos 90^\circ &= \cos 108^\circ \times \cos 44^\circ + \sin 108^\circ \times \sin 44^\circ \times \cos H\end{aligned}$$

or $\cos H = \frac{\cos 90^\circ - \cos 108^\circ \times \cos 44^\circ}{\sin 108^\circ \times \sin 44^\circ}$

or $H = 70^\circ 20' 18''$

Since the sun is setting, its hour angle is eastern.

Hence, westerly hour angle of sun $= 180^\circ - 70^\circ 20' 18'' = 109^\circ 39' 42''$.

Example 10.10 Calculate the sun's azimuth and hour angle at sunset, at a place of latitude 41°N if its declination is (a) 20°N and (b) 20°S .

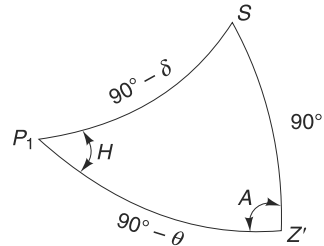


Fig. 10.26

Solution Refer to Fig. 10.27.

$$\theta = 41^\circ\text{N}, \quad \delta = 20^\circ\text{N} = 20^\circ$$

Since at sunset the sun will be at horizon, its altitude,

$$\alpha = 0.$$

$$SZ = 90^\circ$$

$$PS = 90^\circ - 20^\circ = 70^\circ$$

$$PZ = 90^\circ - 41^\circ = 49^\circ$$

Using cosine rule

$$\cos PS = \cos PZ \cos SZ + \sin PZ \sin SZ \cos SZP$$

$$\cos 70^\circ = \cos 49^\circ \times \cos 90^\circ + \sin 49^\circ \times \sin 90^\circ \times \cos A$$

$$\text{or} \quad \cos A = \frac{\cos 70^\circ - \cos 49^\circ \times \cos 90^\circ}{\sin 49^\circ \times \sin 90^\circ}$$

$$\text{or} \quad A = 63^\circ 3' 7.32''$$

Hence, azimuth of sun is $63^\circ 3' 7.32''\text{W}$.

Using the cosine rule again

$$\cos SZ = \cos SP \cos PZ + \sin SP \sin PZ \cos ZPS$$

$$\text{or} \quad \cos 90^\circ = \cos 70^\circ \times \cos 49^\circ + \sin 70^\circ \times \sin 49^\circ \times \cos H$$

$$\text{or} \quad \cos H = \frac{\cos 90^\circ - \cos 70^\circ \times \cos 49^\circ}{\sin 70^\circ \times \sin 49^\circ}$$

$$\text{or} \quad H = 108^\circ 26'$$

Hence, hour angle at sunset = $108^\circ 26'$.

$$(b) \quad \theta = 41^\circ\text{N}, \quad \delta = 20^\circ\text{S} = -20^\circ$$

$$PZ = 90^\circ - \theta = 90^\circ - 41^\circ = 49^\circ$$

$$PS = 90^\circ - \delta = 90^\circ - (-20^\circ) = 110^\circ$$

$$SZ = 90^\circ$$

Using cosine rule

$$\cos PS = \cos PZ \cos SZ + \sin PZ \sin SZ \cos SZP$$

$$\cos 110^\circ = \cos 49^\circ \times \cos 90^\circ + \sin 49^\circ \times \sin 90^\circ \times \cos A$$

$$\text{or} \quad \cos A = \frac{\cos 110^\circ - \cos 49^\circ \times \cos 90^\circ}{\sin 49^\circ \times \sin 90^\circ}$$

$$\text{or} \quad A = 116^\circ 56'$$

Hence, the azimuth of sun is $116^\circ 56'\text{W}$.

Using the cosine rule again

$$\cos SZ = \cos SP \cos PZ + \sin SP \sin PZ \cos ZPS$$

$$\cos 90^\circ = \cos 110^\circ \times \cos 49^\circ + \sin 110^\circ \times \sin 49^\circ \times \cos H$$

$$\text{or} \quad \cos H = \frac{\cos 90^\circ - \cos 110^\circ \times \cos 49^\circ}{\sin 110^\circ \times \sin 49^\circ}$$

$$\text{or} \quad H = 71^\circ 33' 17''.$$

Hence, the hour angle of sun is $71^\circ 33' 17''$.

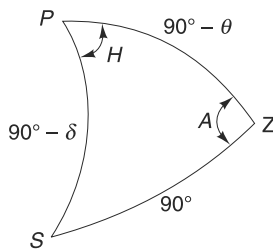


Fig. 10.27

10.10 TIME

The apparent motion of celestial bodies caused by two actual motions of the earth, one about its axis and the other in its orbit around the sun, introduces a time element into the astronomical work. Since the hour angle, altitude and azimuth of a celestial body are functions of time, a knowledge of the various aspects of time measurement is essential.

The orbit of the earth is an ellipse with the sun at one of its foci. This accounts for the variable speed of earth. Therefore, the sun's apparent variable motion is used as a method of measuring time; the construction of mechanical devices, e.g., clocks would be rather difficult. Sun's apparent motion is also not regular and this causes variable hour angles for equal distances along the ecliptic.

The time systems used for measuring time are sidereal time, apparent solar time, mean solar time, and standard time. The first two kinds of time are suitable for the astronomers while the latter two are convenient for everyday use.

1. *Sidereal time*: When the measurement of time is based upon the diurnal motion of a star or the 'first point of aries', it is known as sidereal time. The apparent position of a star is affected by refraction, precession, nutation, aberration, and parallax. After allowing for these, it is found that the stars have certain very small individual motions known as *proper motions*. Therefore, instead of a star, the 'first point of aries' is used to measure the sidereal time. It is well-known that the first point of aries moves backwards (opposite to the direction of the sun's movements) on the ecliptic at a rate of $50.2''$ per annum. We can, therefore, define a mean first point of aries moving with a uniform angular velocity equal to the average angular velocity of the true first point of aries.

The west hour angle of the mean first point of aries at any place and instant gives the *uniform sidereal time*, whereas the true first point of aries gives the *true sidereal time*.

2. *Sidereal day*: The time interval between two successive upper transits of the first point of aries over the same meridian is known as *sidereal day*. It is divided into 24 hours, each hour subdivided into 60 minutes, and each minute into 60 seconds.

3. *Sidereal year*: It is the time required by the sun to make a complete circuit of the ecliptic, with reference to a star on the ecliptic having no proper motion. The average interval between the two consecutive passages of the sun through the first point of aries, which is in motion, is known as *solar* or *tropical year*. Since the first point of aries moves backwards, the tropical year is slightly shorter than the sidereal year:

$$\text{tropical year} = 365.2422 \text{ mean solar days}$$

$$\text{sidereal year} = 365.2564 \text{ mean solar days}$$

4. *Local sidereal time (L.S.T.)*: The sidereal time of an instant is defined as the westward hour angle (when expressed in time) of the first point of aries (Fig. 10.28). The right ascension of a star is also the sidereal time of its upper transit.

Since the hour angle of a star is the sidereal time that has elapsed since its transit, we have

$$\text{L.S.T.} = \text{R.A. of a star} + \text{Westerly hour angle of a star}$$

Note

If L.S.T. so calculated is more than 24 hours, subtract 24 hours, while if it comes out to be negative, add 24 hours.

When the star is at the upper transit, H.A. = 0. Hence, the local sidereal time of transit of a star is equal to the R.A. of a star. A sidereal clock, therefore, records the right ascension of stars as they make their upper transits. The disadvantage of using sidereal time in everyday affairs is that L.S.T. = 0 hours, occurs at a different time of each day of the year.

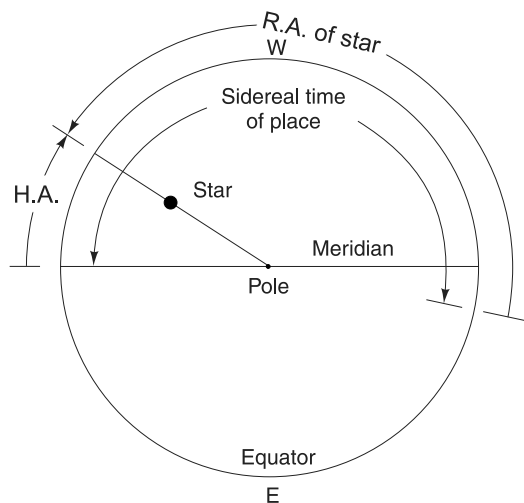


Fig. 10.28(a)

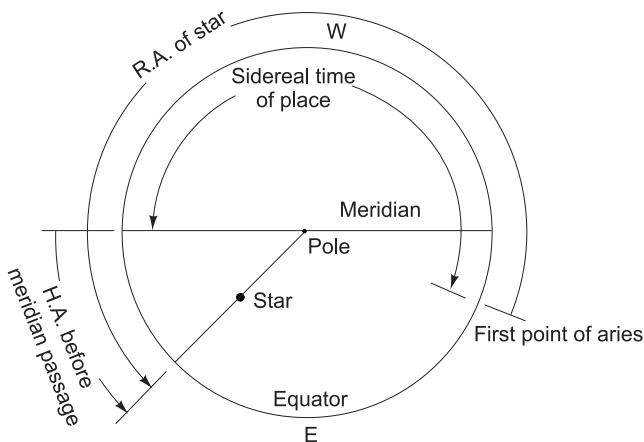


Fig. 10.28(b)

5. *Apparent Solar Time (A.S.T.):* The apparent solar time at any instant is the westerly hour angle (0^h to 24^h) of the sun's centre.

6. *Solar day:* A solar day is the interval of time that elapses between two successive lower transits of the sun's centre over the meridian of the place. The lower transit being chosen so that the date may change at midnight. The apparent solar day is not of constant length and thus modern clocks and chronometers cannot measure it. The apparent solar time is measured by a sundial. The variation in the length of solar days is because of two reasons:

- (a) The orbit of the earth round the sun is elliptical with the sun at one of its foci. The distance of the sun from the earth is thus variable. Due to this, the sun reaches the meridian sometimes earlier and sometimes later, resulting in a variation of length of the solar day.
- (b) The apparent diurnal path of the sun differs from those of the stars because it lies in the ecliptic. In consequence, even although the eastward progress of the sun in the ecliptic is uniform, the time elapsing between the departure of a meridian from the sun and its return thereto would vary because of the obliquity of the ecliptic.

A solar day is divided into 24 hours, each hour subdivided into 60 minutes, and each minute into 60 seconds. Solar noon is at upper transit, when local apparent time is 12 hours.

7. *Mean Solar Time (M.S.T.):* This is also known as *civil time* or *mean time*. Since the sun has a movement relative to the celestial sphere and since it is not uniform, it becomes necessary to assume a mean sun, a point travelling the celestial equator at a uniform rate so as to make solar days of uniform length. The motion of the mean sun is the average of that of the true sun in its right ascension. It differs from the actual sun by amounts varying up to $1/4$ hour. It starts and returns to the vernal equinox at the same time as the true sun. Since the longitude of the mean sun increases uniformly, the R.A. of the mean sun increases uniformly, hence its motion does give a uniform measurement of time.

8. *Greenwich Mean Time (G.M.T.):* It is measured from the lower transit of the Greenwich meridian by the mean sun. It is also known as *universal time* (U.T.).

9. *Mean solar day:* It is the interval between two successive transits of the mean sun over an observer's meridian. It is the average of all the apparent solar days of the year. It is divided in two periods, one from midnight to noon and the other from noon to midnight, each period being of 12 hours. Time of event occurring before mean noon is denoted by the letters a.m. (*ante meridian*) and that after mean noon as p.m. (*post meridian*). For convenience, zero hour of the mean day is reckoned at the local mean midnight (L.M.M.) and the local mean time (L.M.T.) is that reckoned from the L.M.M. But the astronomers divide the mean solar day into 24 hours as usual.

- (a) L.S.T. = R.A. of sun + H.A. of sun
- (b) L.S.T. = R.A. of mean sun + H.A. of mean sun
- (c) L.S.T. of apparent noon = R.A. of sun

- (d) L.S.T. of mean noon = R.A. of mean sun
- (e) H.A. of sun = 0 at upper transit
- (f) Solar time (apparent or mean) = 0 at lower transit
- (g) Apparent solar time = H.A. of sun + 12 hours
- (h) Mean solar time = H.A. of mean sun + 12 hours
- (i) One sidereal day = 1 solar day – ($3^m 55.91^s$) mean time
- (j) One solar day = 1.0027379 sidereal day
- (k) 24 hour sidereal time = ($24^h - 3^m 55.91^s$) mean solar time

10. *Lunar day*: It may be defined as the time for one complete rotation of the earth relative to the moon. One lunar day equals $24^h 51^m$. Since the moon shifts eastward, against the celestial background, much more rapidly than the sun, a lunar day is longer than a solar day. The moon is about 13° to the east and thus it rises 51^m late than the previous day. The length of a lunar day varies because the inclination of ecliptic to the horizon changes.

11. *Equation of time*: The difference between apparent and mean solar times at any instant is known as the equation of time (E.T.). The values of the E.T. at 0 hour at Greenwich are tabulated in the 'nautical almanac'. It varies from 0 to about 6 min at different seasons of the year.

The values of E.T. are sometimes prefixed with + or – sign indicating that they are to be added or subtracted from apparent time to give mean time.

$$\text{E.T.} = \text{R.A. of mean sun} - \text{R.A. of sun}$$

The E.T. vanishes four times during the year, on or about 6th April, 6th July, 6th October, and 4th January. On these dates the true and apparent sun are on the same meridians. The variation of E.T. is represented as shown in Fig. 10.29(a).

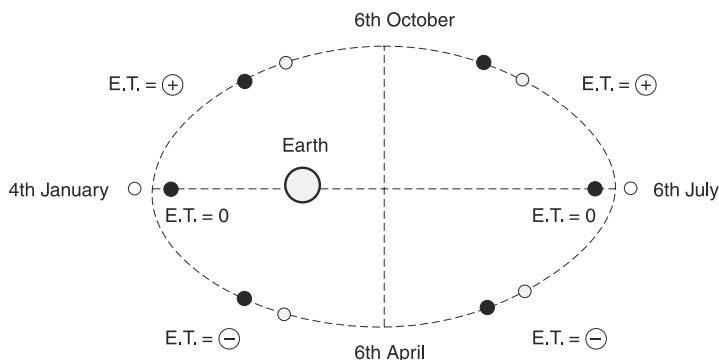
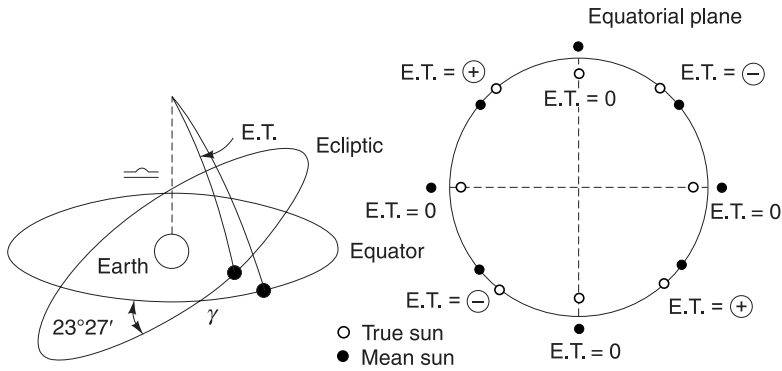


Fig. 10.29(a) Effect of ellipticity of the orbit

The reasons for the difference in the mean and apparent solar times is the non-uniform motion of the sun along the ecliptic (ellipticity of the orbit) and obliquity of the ecliptic. The two causes have been explained diagrammatically in Fig. 10.29(a) and Figs. 10.29(b, c), respectively.

**Fig. 10.29(b)** *Obliquity of ecliptic***Fig. 10.29(c)** *Effect of obliquity*

Because calculations are based on the fictitious mean sun, while observations are necessarily taken to the apparent sun, it is necessary to know the variation of the former with respect to the latter. The relationship is given by the formula

$$\text{L.M.T.} = \text{L.A.T.} \pm \text{E.T.}$$

12. *Standard time*: The use of different local times may create confusion and to avoid it, a particular meridian known as *standard meridian* is adopted for the whole of the country. This standard meridian is so chosen that it lies at an exact number of hours from Greenwich. Time referred to the standard meridian is known as *standard time* throughout the country. In India, the standard meridian has a longitude of $82^{\circ}30'E$ ($5^h 30^m E$). For Great Britain, Belgium and Spain, Greenwich is the standard meridian.

$$\text{Standard time} = \text{L.M.T.} \pm \text{difference of longitude in time W/E,} \\ \text{between the given place and the standard meridian}$$

Note Use plus sign if the place is to the west of the standard meridian and minus sign if it is to the east.

10.11 CONVERSION OF TIME

10.11.1 Local Time and Standard Time

The local mean time differs from the standard time because of the difference of longitude between the given place and the standard meridian. Thus,

$$\text{Standard time} = \text{L.M.T.} \pm \text{difference of longitude W/E}$$

Note Use plus sign if the place is to the west of standard meridian and minus if it is to the east.

10.11.2 Local Time and Greenwich Time

$$\text{Greenwich time} = \text{local time} \pm \text{longitude of the place W/E}$$

Note Use plus sign if the place is to the west of Greenwich meridian and minus if it is to the east.

10.11.3 Mean Solar Time and Sidereal Time

In one tropical year there are 365.2422 mean solar days, whereas the number of sidereal days is 366.2422. Thus,

$$1 \text{ mean solar day} = 1 + \frac{1}{365.2422} \text{ sidereal day}$$

$$= 24^h 3^m 56.56^s \text{ sidereal day}$$

Thus, the mean solar day is $3^m 56.56^s$ longer than the sidereal day.

Also, 1^h mean solar time = $1^h + 9.8565^s$ sidereal time

1^m mean solar time = $1^m + 0.1643^s$ sidereal time

1^s mean solar time = $1^s + 0.0027^s$ sidereal time

The correction of 9.8565^s per hour of mean time is known as *acceleration*. It is to be added to the mean solar time to obtain sidereal time. The correction of 9.8565^s per hour of sidereal time is known as *retardation* and is to be subtracted from the sidereal time to obtain the mean solar time.

10.11.4 Greenwich Sidereal Time and Local Sidereal Time

Since 24 hours of solar time is equal to $24^h 3^m 56.56^s$ of sidereal time, therefore the sidereal clock will be continually gaining on the mean clock at a rate of 9.8565^s per mean hour. The Greenwich sidereal time at Greenwich mean midnight is then the difference between the sidereal clock and the mean clock at that instant.

Local sidereal time at local mean midnight = Greenwich sidereal time at Greenwich mean midnight $\pm$ acceleration.

Note

Use plus sign if the longitude is to the west and minus sign if the longitude is to the east of Greenwich.

Example 10.11 Express the following angles in hours, minutes and seconds:
(a) $5^\circ 15' 25''$, (b) $102^\circ 18' 29''$, (c) $274^\circ 40' 18''$.

Solution

$$1^h = 15^\circ, 1^m = 15' \text{ and } 1^s = 15''$$

$$\text{Hence, (a) } 5^\circ 15' 25'' = \frac{5^\circ 15' 25''}{15} = 0^h 21^m 1.67^s$$

$$\text{(b) } 102^\circ 18' 29'' = \frac{102^\circ 18' 29''}{15} = 6^h 49^m 13.9^s$$

$$\text{(c) } 274^\circ 40' 18'' = \frac{274^\circ 40' 18''}{15} = 18^h 18^m 41^s$$

Example 10.12 Express the following hours, minutes and seconds into degrees:

(a) $2^h 02^m 42^s$, (b) $14^h 15^m 28^s$.

Solution

$$1^h = 15^\circ, 1^m = 15', 1^s = 15''$$

Hence,

$$\begin{aligned} \text{(a)} \quad 2^h 02^m 42^s &= 15 \times (2^h 02^m 42^s) = 30^\circ 40' 30'' \\ \text{(b)} \quad 14^h 15^m 28^s &= 15 \times (14^h 15^m 28^s) = 213^\circ 52' 0'' \end{aligned}$$

Example 10.13 Find the local mean time (L.M.T.) at a place having a longitude of $94^\circ 20'E$, when the standard time is $10^h 06^m 18^s$ and the standard meridian is $78^\circ 30'E$.

Solution

$$\text{Longitude of the place} = 94^\circ 20'E$$

$$\text{Longitude of the standard meridian} = 78^\circ 30'E$$

$$\begin{aligned} \text{Difference in longitude} &= 94^\circ 20' - 78^\circ 30' \\ &= 15^\circ 50' = 1^h 3^m 20^s \end{aligned}$$

The place is to the east of the meridian.

$$\text{Standard time} = \text{L.M.T.} - \text{Difference in longitude}$$

$$10^h 06^m 18^s = \text{L.M.T.} - 1^h 3^m 20^s$$

or

$$\text{L.M.T.} = 11^h 09^m 38^s$$

Example 10.14 The local mean time at a place in longitude $69^\circ 30'E$ is $8^h 20^m 16^s$. Find the standard time if the place is in India.

Solution The standard meridian for India is $82^\circ 30'E$

$$\text{Longitude of the place} = 69^\circ 30'E$$

$$\text{Longitude of the standard meridian} = 82^\circ 30'E$$

$$\begin{aligned} \text{Difference in longitude} &= 82^\circ 30' - 69^\circ 30' \\ &= 13^\circ = 0^h 52^m 0^s \end{aligned}$$

The place is to the west of the standard meridian.

$$\begin{aligned} \text{Hence, standard time} &= \text{L.M.T.} + \text{difference in longitude} \\ &= 8^h 20^m 16^s + 0^h 52^m 0^s = 9^h 12^m 16^s \end{aligned}$$

Example 10.15 In the example 10.14, if the longitude of the place is $69^\circ 30'W$, what will be the standard time?

$$\begin{aligned} \text{Difference in longitude} &= 82^\circ 30' - (-69^\circ 30') \\ &= 152^\circ 00' = 10^h 8^m 0^s \end{aligned}$$

The meridian of the place is to the west of the standard meridian.

$$\begin{aligned} \text{Standard time} &= \text{L.M.T.} + \text{difference in longitude} \\ &= 8^h 20^m 16^s + 10^h 8^m 0^s = 18^h 28^m 16^s \end{aligned}$$

Example 10.16 Find the G.M.T. corresponding to the local mean time of $10^h 20^m 6^s$ a.m. at places having longitudes: (a) $48^\circ 36'E$ and (b) $48^\circ 36'W$.

Solution L.M.T. = $10^h 20^m 6^s$

$$\text{Longitude of the place} = 48^\circ 36' = 3^h 14^m 24^s$$

(a) The place is towards east of Greenwich.

$$\begin{aligned} \text{Hence, G.M.T.} &= \text{L.M.T.} - \text{longitude of the place} \\ &= 10^h 20^m 6^s - 3^h 14^m 24^s = 7^h 5^m 42^s \end{aligned}$$

(b) The place is towards west of Greenwich.

Hence, $G.M.T. = L.M.T. + \text{longitude of the place}$

$$\begin{aligned} &= 10^h 20^m 6^s + 3^h 14^m 24^s \\ &= 13^h 34^m 30^s = 1^h 34^m 30^s \text{ p.m.} \end{aligned}$$

Example 10.17 The clock time at a given instant at Allahabad on 5th May 1992 was 10:00 p.m. Find L.M.T. and G.M.T. corresponding to that time. Given:

Spherical coordinates of Allahabad: $25^\circ 25' 25''\text{N}, 81^\circ 55'\text{E}$

Indian standard longitude for time: $82^\circ 30'\text{E}$

Solution Difference between longitude of Allahabad and I.S.T.

$$= 82^\circ 30' - 81^\circ 55' = 0^\circ 35' 00'' = 0^h 2^m 20^s$$

Allahabad is westward to I.S.T.

$$\text{Hence, } L.M.T. = 10^h 0^m 0^s - 0^h 2^m 20^s = 9^h 57^m 40^s$$

L.M.T. is $9^h 57^m 40^s$ p.m. on 5th May.

Difference between I.S.T. and G.M.T.

$$= 82^\circ 30' - 0^\circ 0' = 82^\circ 30' = 5^h 30^m 0^s$$

Greenwich is towards west of India.

$$G.M.T. = 10^h 0^m 0^s - 5^h 30^m 0^s = 4^h 30^m 0^s$$

G.M.T. is 4.30 p.m. on 5th May.

Example 10.18 Find the L.M.T. at the place having a longitude $62^\circ 30'\text{E}$ if the Greenwich mean time (G.M.T.) is

(a) $5^h 20^m 2^s$ a.m. on 5th May.

(b) $5^h 20^m 2^s$ p.m. on 5th May.

Solution Longitude of the place $= 62^\circ 30' = 4^h 10^m$

The place is towards the east of Greenwich.

(a) Hence, $G.M.T. = L.M.T. - \text{longitude of the place}$

$$9^h 2^m 20^s = L.M.T. - 4^h 10^m$$

or $L.M.T. = 13^h 12^m 20^s = 1^h 12^m 20^s$ p.m. on 5th May

(b) $G.M.T. = 9^h 2^m 20^s$ p.m. $= 21^h 2^m 20^s$ (past midnight)

$G.M.T. = L.M.T. - \text{longitude of the place}$

$$21^h 2^m 20^s = L.M.T. - 4^h 10^m$$

or

$$L.M.T. = 25^h 12^m 20^s$$

$$= 1^h 12^m 20^s \text{ a.m. on 6th May.}$$

Example 10.19 Find the sidereal time interval corresponding to a mean time interval of $6^h 6^m 20^s$.

Solution Correction (acceleration) for $6^h = 6 \times 9.8565 = 59.139^s$

$$6^m = 6 \times 0.1643 = 0.9858^s$$

$$20^s = 20 \times 0.0027 = 0.054^s$$

$$\text{Total correction} = 60.1788^s = 1^m 0.1788^s$$

Hence, sidereal time interval = mean time interval + acceleration

$$= 6^h 6^m 20^s + 1^m 0.1788^s = 6^h 7^m 20.1788^s$$

Example 10.20 Find the mean time interval corresponding to the sidereal time interval of $5^h 56^m 20^s$.

Solution Correction (retardation) for

$$5^h = 5 \times 9.8565 = 49.2825^s$$

$$56^m = 56 \times 0.1643 = 9.2008^s$$

$$20^s = 20 \times 0.0027 = 0.054^s$$

$$\text{Total correction} = 58.5373^s$$

$$\text{Hence, mean time interval} = 5^h 56^m 20^s - 58.5373^s = 5^h 55^m 21.4627^s$$

Example 10.21 Find the local apparent time of an observation at a place on longitude $78^\circ 10'E$ corresponding to the local mean time $10^h 20^m 20^s$. The equation of time at G.M.N. being $2^m 6.46^s$ subtractive from apparent time and decreasing at 0.27^s per hour.

Solution Longitude of the place = $78^\circ 10' = 5^h 12^m 40^s$

The place is towards east of Greenwich.

$$\begin{aligned} \text{G.M.T.} &= \text{L.M.T.} - \text{longitude of the place} \\ &= 10^h 20^m 20^s - 5^h 12^m 40^s = 5^h 7^m 40^s \end{aligned}$$

$$\text{Mean time interval before G.M.N.} = 12^h - 5^h 7^m 40^s = 6^h 52^m 20^s$$

Increase $6^h 52^m 20^s$ at 0.27^s per hour

$$= 6^h 52^m 20^s \times 0.27 = 1.855^s$$

$$\text{Equation of time at G.M.N.} = 2^m 6.46^s$$

$$\text{Increase} = 1.855^s$$

$$\therefore \text{equation of time at observation} = 2^m 8.315^s$$

$$\text{G.M.T. of observation} = \text{G.A.T. of observation} - \text{Equation of time}$$

$$5^h 7^m 40^s = \text{G.A.T. of observation} - 2^m 8.315^s$$

$$\text{G.A.T. of observation} = 5^h 9^m 48.314^s$$

$$\text{Now, G.A.T. of observation} = \text{L.A.T. of observation} - \text{longitude of place.}$$

$$\Rightarrow 5^h 9^m 48.315^s = \text{L.A.T. of observation} - 5^h 12^m 40^s$$

$$\text{Hence, L.A.T. of observation} = 10^h 22^m 28.315^s$$

10.11.5 Conversion of L.S.T. to L.M.T.

The following rules may be observed:

1. L.S.T. of L.M.N./L.M.M. = G.S.T. of G.M.N./G.M.M. $\pm$ 9.856^s per hour of longitude W/E
2. L.S.T. = L.S.T. of L.M.N./L.M.M. + S.I. from L.M.N./L.M.M.
3. L.M.T. of transit of γ = M.T. of transit of γ at Greenwich $\pm$ 9.8296^s per hour longitude E/W

Data given L.S.T., G.S.T. of G.N.N./G.M.M. and longitude

Procedure

1. Express longitude in hours, etc.
2. Find L.S.T. of L.M.N./L.M.M.
3. Obtain the sidereal interval.
4. Deduct 9.8296^s per hour of sidereal interval from sidereal interval, thus obtaining the local mean time.

Example 10.22 Find the L.M.T. at a place of longitude $84^\circ 30' 10^s$. L.S.T. at the place being $17^h 32^m 12^s$ and G.S.T. of G.M.N. being $6^h 40^m$.

Solution Longitude of the place = $84^\circ 30' 10^s = 5^h 38^m 0.67^s$

Acceleration for longitude west at the rate of 9.856^s per hour
 $= (5^h 38^m 0.67^s) \times 9.856 = 55.52^s$

L.S.T. of L.M.N. = G.S.T. of G.M.N. + acceleration
 $= 6^h 40^m + 55.52^s = 6^h 40^m 55.52^s$

Now, L.S.T. = L.S.T. of L.M.N. + S.I. from L.M.N.

$17^h 32^m 12^s = 6^h 40^m 55.52^s + \text{S.I. from L.M.N.}$

or S.I. from L.M.N. = $10^h 51^m 16.08^s$

Retardation for conversion of sidereal time to mean time is 9.8296^s per hour of sidereal time.

Retardation = $10^h 51^m 16.08^s \times 9.8296 = 106.69^s = 1^m 56.69^s$

Mean time interval from L.M.N. = S.I. from L.M.N. – retardation
 $= 10^h 51^m 16.08^s - 1^m 56.69^s = 10^h 49^m 19^s$

Hence, L.M.T. = $10^h 49^m 19^s$ p.m.

Data given G.M.T. of transit of first point of aries, L.S.T. and longitude.

Procedure

1. Express the longitude in hours, etc.
2. Find the local mean time of transit of γ .
3. Find the mean time interval corresponding to the local sidereal time by deducting 9.8296^s per hour of sidereal time.
4. Find the local mean time by adding (ii) + (iii).

Example 10.23 Given L.S.T. = $5^h 32^m 10^s$, longitude of the place $74^\circ 30'E$, G.M.T. of transit of $\gamma = 15^h 41^m 24.38^s$, find the L.S.T.

Solution Longitude of the place = $74^\circ 30' = 4^h 58^m$

Correction for longitude at 9.8296^s per hour of longitude
 $= 4^h 58^m \times 9.8296 = 48.82^s$

L.M.T. of transit of γ = G.M.T. of transit of γ + 48.82^s
 $= 15^h 42^m 13.20^s$

L.S.T. = $5^h 32^m 10^s$

Retardation = $5^h 32^m 10^s \times 9.8296 = 54.41^s$

Mean time interval since transit of γ = L.S.T. – retardation

$= 5^h 32^m 10^s - 54.41^s = 5^h 31^m 15.59^s$

$$\begin{aligned}
 \text{L.M.T.} &= \text{Mean time interval since transit of } \gamma + \text{L.M.T. of transit of } \gamma \\
 &= 5^h 31^m 15.59^s + 15^h 42^m 13.20^s \\
 &= 21^h 13^m 29.19^s = 9^h 13^m 29.19^s \text{ p.m.}
 \end{aligned}$$

10.11.6 Conversion of L.M.T. to L.S.T.

Data given L.M.T., G.S.T. of G.M.N., and longitude

Procedure

1. Express the longitude in hours, etc.
2. Find L.S.T. of L.M.N./L.M.M.
3. Find S.I. from L.M.N./L.M.M.
4. Find L.S.T. by adding S.I. to the L.S.T. of L.M.N./L.M.M.

Example 10.24 Find the L.S.T. corresponding to 5.05 a.m. on 20th May 1990 at a place on longitude $54^\circ 12' \text{W}$. The G.S.T. of G.M.M. being $7^h 12^m 54^s$.

Solution Longitude of the place = $68^\circ 12' \text{W} = 4^h 32^m 8^s$
Acceleration (for west longitude) = $4^h 32^m 8^s \times 9.8565 = 44.70^s$
L.S.T. of L.M.M. = G.S.T. of G.M.M. + acceleration
 $= 7^h 12^m 54^s + 44.70^s = 7^h 13^m 38.70^s$
L.M.T. = $5^h 05^m 0^s$

Convert M.T. interval ($5^h 05^m$) since L.M.M. to S.I. at 9.8565^s per hour of M.T. interval.

$$\begin{aligned}
 \text{Acceleration} &= 50.10^s \\
 \text{S.I. since L.M.M.} &= 5^h 05^m 0^s + 50.10^s = 5^h 05^m 50.10^s \\
 \text{L.S.T.} &= \text{L.S.T. of L.M.M.} + \text{S.I. since L.M.M.} \\
 &= 7^h 12^m 54^s + 5^h 05^m 50.10^s = 12^h 18^m 44.10^s
 \end{aligned}$$

Data given G.M.T. of transit of γ , L.M.T. and longitude

Procedure

1. Express the given longitude in hours, etc.
2. Find the L.M.T. of transit of γ .
3. Find the mean time interval since transit of γ .
4. Convert this mean time interval into sidereal time by adding 9.8565^s per hour.

Example 10.25 Find the L.S.T. corresponding to 4.30 a.m. at a place on longitude $62^\circ 10' \text{E}$ and G.M.T. of transit of γ being $13^h 20^m 30^s$.

Solution Longitude of the place = $62^\circ 10' \text{E} = 4^h 8^m 40^s$
Correction for longitude = $4^h 8^m 40^s \times 9.8565 = 40.85^s$
L.M.T. of transit of γ = G.M.T. of transit of γ + acceleration
 $= 13^h 20^m 30^s + 40.85^s = 13^h 21^m 10.85^s$
L.S.T. = $4^h 30^m$
Retardation = $4^h 30^m \times 9.8565 = 44.35^s$
Mean time interval = L.S.T. – retardation

$$= 4^h 30^m - 44.35^s = 4^h 29^m 15.65^s$$

L.M.T. = L.M.T. at transit of γ + mean time interval

$$= 13^h 21^m 10.85^s + 4^h 29^m 15.65^s = 17^h 50^m 26.50^s$$

Example 10.26 Find the L.M.T. of elongation of a star of right ascension $14^h 45^m 27^s$ and declination $75^\circ 04' 20''\text{N}$, at a place in latitude $52^\circ 21' 10''\text{N}$, longitude $110^\circ 45' 05''\text{E}$, on the night of 20th July, given that the G.S.T. of the previous G.M.M. will be $8^h 29^m 58^s$. State whether the elongation will be eastern or western.

Solution The star is at elongation; therefore, the angle ZSP is 90° , but it is not known at this stage whether the star is at eastern or western elongation (Fig. 10.30(a) and (b)).

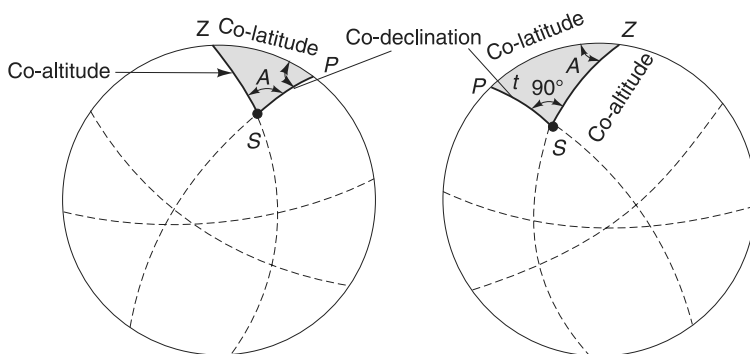


Fig. 10.30(a) Star at eastern elongation **Fig. 10.30(b)** Star at western elongation

Using Napier's rule (Fig. 10.30 (c)) or by deduction from the sine rule:

$$\cos t = \frac{\tan \text{latitude}}{\tan \text{declination}} = \frac{\tan 51^\circ 21' 10''}{\tan 75^\circ 04' 20''}$$

or

$$t = 60^\circ 46' 54.6''$$

$$= 4^h 39^m 7.6^s$$

If star is at eastern elongation (Fig. 10.30(d)):

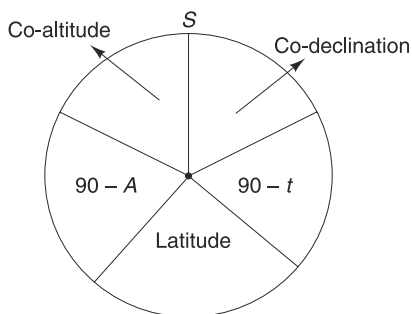


Fig. 10.30(c)

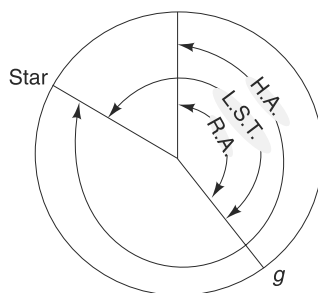


Fig. 10.30(d)

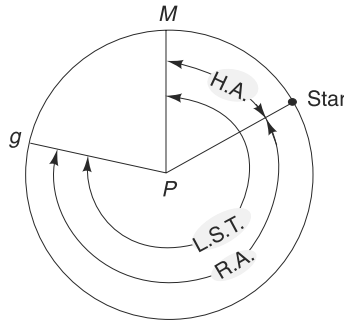


Fig. 10.30(e)

$$\text{hour angle} = 24^h - t = 19^h 20^m 52.4^s$$

If star is at western elongation (Fig. 10.30(e)):

$$\text{hour angle} = 4^h 39^m 7.6^s$$

Now L.S.T. = R.A. + H.A. (deducting 24^h if the sum is more than 24^h). R.A. is given as $14^h 45^m 27^s$.

If star is at eastern elongation:

$$\begin{aligned} \text{L.S.T.} &= 14^h 45^m 27^s + 19^h 20^m 52.4^s - 24^h \\ &= 10^h 6^m 19.4^s \end{aligned}$$

If star is at western elongation:

$$\begin{aligned} \text{L.S.T.} &= 14^h 45^m 27^s + 4^h 39^m 7.6^s \\ &= 19^h 24^m 34.6^s \end{aligned}$$

Now G.S.T. of previous G.M.M. = $8^h 29^m 58^s$

In longitude $110^\circ 45' 05''\text{E}$ a negative correction is necessary to obtain L.S.T. of L.M.M. This amounts to $\frac{110^\circ 45' 05''}{360^\circ} \times 3^m 56.56^s$, the last figure being the daily increase in sidereal time. The correction works out to be $1^m 12.8^s$ and, as already stated, is negative.

$$\therefore \text{L.S.T. of previous L.M.M.} = 8^h 28^m 45.2^s$$

Comparing the L.S.T. of L.M.M. and the L.S.T. of eastern and western elongations, it is seen that the sidereal time interval from midnight is about 2^h to the eastern elongation and 11^h to the western elongation. Ignoring the slight difference between a sidereal time interval and a mean time interval, it is clear that the western elongation, occurring 11^h after midnight, will be invisible because of daylight. The star must therefore be at eastern elongation.

The exact sidereal time interval from local midnight

$$= 10^h 6^m 19.4^s - 8^h 28^m 45.2^s = 1^h 37^m 34.2^s$$

Convert this sidereal time interval to a mean time interval.

$$\text{Mean time interval after midnight} = \frac{365.2422}{366.2422} \times 1^h 37^m 34.5^s = 1^h 37^m 18.2^s$$

i.e., L.M.T. = $1^h 37^m 18.2^s$ at eastern elongation.

Example 10.27 A star of right ascension $10^h 05^m 05^s$ and declination $12^\circ 20' 02''\text{N}$, is observed in the east from a station in latitude $42^\circ 29' 45''\text{N}$ at a corrected altitude of 43° at U.T. (G.M.T.) $22^h 49^m 31^s$.

Find the longitude of the observation if at the time of observation $R = 8^h 47^m 56.8^s$, R being defined as the difference between G.H.A. aries and U.T.

Solution Refer to Fig. 10.31.

$$\text{Latitude} = 42^\circ 29' 45''$$

$$\therefore \text{co-latitude} = 47^\circ 30' 15'' = ZP$$

$$\text{Declination} = 12^\circ 20' 02''\text{N}$$

$$\therefore \text{co-declination} = 77^\circ 39' 58'' = PS$$

$$\text{Altitude} = 43^\circ 00' 00''$$

$$\therefore \text{co-altitude} = 47^\circ 00' 00'' = ZS$$

The three sides of the astronomical triangle PZS (Fig. 10.31) being obtained from the data.

$$\tan \frac{A}{2} = \sqrt{\frac{\sin(s - ZP) \sin(s - PS)}{\sin s \sin(s - ZS)}}$$

where

$$\begin{aligned} s &= \frac{1}{2} (PS + ZP + ZS) \\ &= \frac{1}{2} \times 172^\circ 10' 13'' = 86^\circ 05' 06.5'' \end{aligned}$$

$$\tan \frac{A}{2} = \sqrt{\frac{\sin 38^\circ 34' 51.5'' \times \sin 8^\circ 25' 8.5''}{\sin 86^\circ 5' 6.5'' \times \sin 39^\circ 5' 6.5''}}$$

or

$$A = 41^\circ 42' 48'' = 2^h 46^m 51.2^s$$

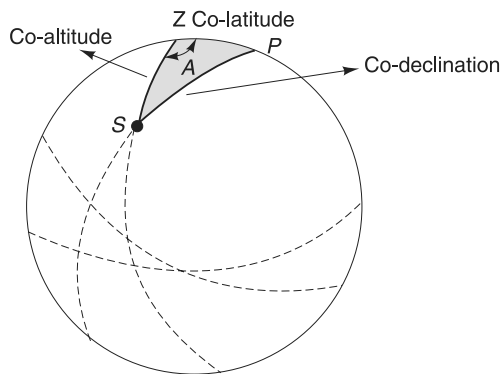


Fig. 10.31

Since the star is observed in the east:

$$\text{L.H.A.} = 24^h - 2^h 46^m 51.2^s = 21^h 13^m 8.8^s$$

Now, G.H.A. aries (G.S.T.)

$$= \text{U.T.} + R \text{ (deducting } 24^h \text{ if sum exceeds } 24^h)$$

$$\text{U.T.} = 22^h 49^m 31^s + 8^h 47^m 56.8^s$$

$$= 31^h 37^m 27.8^s$$

$$\therefore \text{G.H.A. aries} = 31^h 37^m 27.8^s - 24^h = 7^h 37^m 27.8^s$$

$$\text{G.H.A. star} = \text{G.H.A. aries} - \text{R.A. of star}$$

(add 24^h to G.H.A. aries if value less than R.A.)

$$\text{G.H.A. aries} = 7^h 37^m 27.8^s + 24^h = 31^h 37^m 27.8^s$$

$$\therefore \text{G.H.A. star} = 31^h 37^m 27.8^s - 10^h 05^m 05^s = 21^h 32^m 22.8^s$$

$$\text{Longitude} = \text{G.H.A. star} - \text{L.H.A. star},$$

in which longitude is west if between 0^h and 12^h . If greater than 12^h , it must be subtracted from 24^h to give east longitude.

$$\text{G.H.A. star} = 21^h 32^m 22.8^s$$

$$\text{L.H.A. star} = 21^h 13^m 8.8^s$$

$$\therefore \text{Longitude} = 19^m 14.0^s \text{ west}$$

Converting to arc by multiplying by 15

$$\text{Longitude} = 4^\circ 48' 30'' \text{ W}$$

10.12 ASTRONOMICAL CORRECTIONS

In the solution of many astronomical problems it is necessary to know the true altitude of the centre of a celestial body. The observed altitude (measured vertical angle) of a celestial body is not its true altitude. Before using it in the solution of a *ZPS* triangle, the corrections to be applied to the value of the altitude are corrections for parallax, refraction, semi-diameter and dip.

10.12.1 Correction for Parallax

Parallax is defined as the difference in direction of a celestial body caused by the fact that the observations are made from the surface of the earth and not from the centre of the earth. Correction for parallax is required when the observations are made for celestial bodies like sun, moon, etc., which are at a finite distance from earth. The correction is zero for the celestial bodies infinitely distant from earth. This is because of the fact that for finite distances, the parallax is significant, since the observer is on the earth's surface and not at the earth's centre. But for infinitely distant celestial bodies the parallax value will be practically zero. The parallax correction is *always additive* and is very small. For the sun its value is $8.8''$ and for the moon it is 1° .

In Fig. 10.32, the angle $S'AS$ (α') is the observed altitude, whereas α , the angle $S'OH$, is the true altitude, corrected for parallax. p_a is the parallax correction and p_h is called sun's horizontal parallax.

$$\text{From triangle } SAO, \sin p_h = \frac{R}{SO}$$

The sun's horizontal parallax varies inversely as its distance from the earth's centre and its average value is $8.8''$.

$$\begin{aligned}\text{The altitude, } \alpha &= S'OH = S'BS \\ &= BS'A + BAS' \\ &= p_a + \alpha'\end{aligned}$$

Hence, parallax correction, $p_a = \alpha - \alpha'$

In triangles $AS'O$, applying sine rule

$$\frac{AO}{\sin AS'O} = \frac{S'O}{\sin S'AO}$$

or

$$\frac{AO}{\sin p_a} = \frac{S'O}{\sin (90^\circ + \alpha')}$$

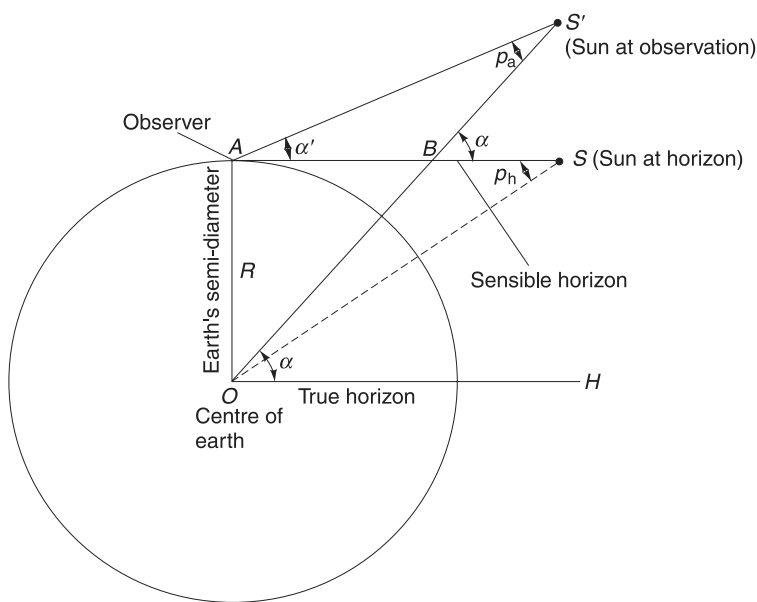


Fig. 10.32 Parallax correction

$$\begin{aligned}\sin p_a &= \frac{AO}{S'O} \sin(90^\circ + \alpha') \\ &= \frac{AO}{SO} \cos \alpha' \quad (\text{since } S'O = SO) \\ &= \sin p_h \cos \alpha' = 8.8'' \cos \alpha'\end{aligned}$$

The correction is always positive and is maximum when the sun is at the horizon.

10.12.2 Correction for Refraction

The earth is surrounded by layers of atmospheric air. The layers are dense near the surface and get thinner and thinner with distance from the surface of the earth.

This causes the bending of light rays as they pass through the earth's atmosphere and is termed as *refraction*. In Fig. 10.33, the observation has been made to a star S . The apparent direction AS' of the body from the earth's surface being tangent to the curve AS of the refracted ray from the position of the observer A results in the observed altitude $S'AH$, whereas the correct altitude of the body is SAH . Therefore, the refraction correction is the angle SAS' and is always subtractive. The value of this correction varies from 0 when the celestial body is at zenith, to $11'$ when the body is a few degrees above the horizon. It also decreases slightly as the temperature increases. At a pressure of 76 cm of mercury and a temperature of 10°C , the correction for refraction is given by

$$C_R = 58'' \cot \alpha'$$

where α' is the apparent altitude of the celestial body. The combined correction for parallax and refraction is subtractive.

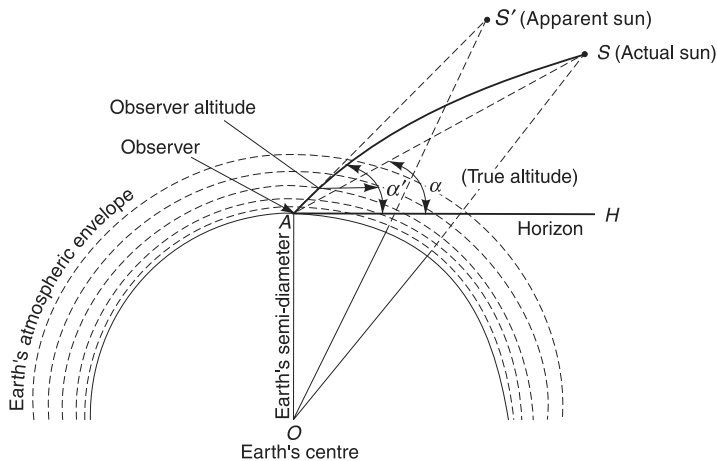


Fig. 10.33 Correction for refraction

10.12.3 Correction for Semi-diameter

When the celestial body, on which observations are made, has a large diameter, e.g., sun and moon, it is customary to obtain the altitude by measuring vertical angles to either the upper or lower limb (edge). The observed value so obtained must be corrected for the semi-diameter of the celestial body. The semi-diameter of the body is half the angle subtended at the centre of the earth by the diameter of the body. Since the distance of the sun from the earth varies throughout the year, the correction is not constant and varies from $15'46''$ to $16'18''$. Its mean value is $16'1.18''$.

If α is the correct altitude (Fig. 10.34), then

$$\alpha = \alpha_1 + \frac{\gamma}{2}$$

also
$$\alpha = \alpha_2 - \frac{\gamma}{2}$$

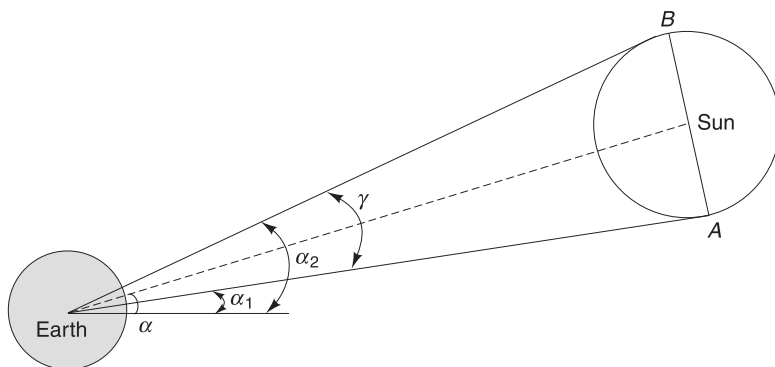


Fig. 10.34 Correction for semi-diameter

Hence, the correction for semi-diameter is additive if the observation is made for the upper limb of the sun and negative for its lower limb.

10.12.4 Correction for Dip

When the observation is made with a sextant at sea, the altitude is observed from the *visible* or *sea* horizon. The angle between the *sensible* and *visible* horizon is known as *dip*, and is shown in Fig. 10.35 as β .

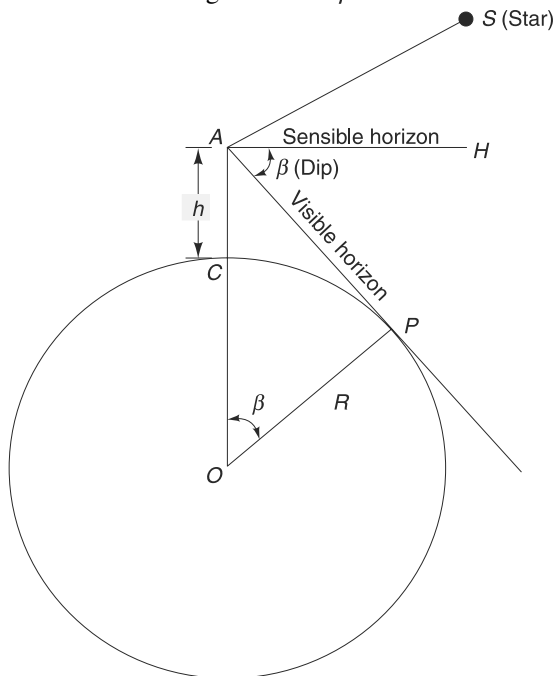


Fig. 10.35 Correction for dip

Angle of dip, $\beta = \angle HAP = \angle AOP$

From the right angle triangle AOP

$$\begin{aligned}
 \tan \beta &= \frac{AP}{OP} \\
 &= \frac{\sqrt{(R+h)^2 - R^2}}{R} \\
 &= \frac{\sqrt{R^2 + h^2 + 2Rh - R^2}}{R} = \frac{\sqrt{h(2R+h)}}{R} \\
 &= \sqrt{\frac{h(2R+h)}{R^2}} = \sqrt{\frac{2h}{R} + \frac{h^2}{R^2}} \\
 &= \sqrt{\frac{2h}{R}} \quad (\text{since } h^2/R^2 \text{ is very small})
 \end{aligned}$$

The correction for dip is always subtractive.

Example 10.28 An observation was made on the sun's upper limb and the observed altitude was $32^\circ 15' 18''$. The semi-diameter of the sun at the time of observation was $16' 11''$. Find the true altitude of the sun.

Solution

Correction for parallax

$$C_p = + 8.8'' \times \cos 32^\circ 15' 18'' = + 7.44''$$

Correction for refraction

$$C_R = - 58'' \times \cot 32^\circ 15' 18'' = - 1' 31.90''$$

Correction for semi-diameter

$$C_\gamma = -\frac{\gamma}{2} = - 16' 1''$$

It is negative since observation is made for upper limb of the sun.

Hence, total correction = $+ 7.44'' - 1' 31.9'' - 16' 1'' = - 17' 25.46''$

Corrected altitude = $32^\circ 15' 18'' - 17' 25.46'' = 31^\circ 57' 52.54''$

10.13 DETERMINATION OF TIME

The local time may be determined by extra-meridian observation of a star or sun, meridian transit of a star or sun, and by equal altitudes of a star or sun.

10.13.1 Time by Extra-meridian Observation (Altitude) of a Star

In this method the altitude of a star is observed with a theodolite. Observation is taken with both the faces and the time at the instant of observation is recorded. Knowing the altitude and the declination of the star and the latitude of the place of observation, the H.A. of the star may be computed by the relation

$$\sin \alpha = \sin \theta \sin \delta + \cos \theta \cos \delta \cos H$$

When the star is at prime vertical

$$\cos H = \frac{\tan \delta}{\tan \theta}$$

Then,

$$\text{L.S.T.} = \text{R.A.} \pm \text{H.A.}$$

Use plus sign if the star is to the west of the meridian and minus sign if it is to the east. Knowing the longitude of the place and the sidereal time of G.M.M./G.M.N., the L.M.T. of observation can be computed.

In the case of the sun, the process is the same as above. The altitude of the sun is measured twice, i.e., both for its lower and upper limbs, with telescope normal and inverted. The time is noted at the instant of each observation. Necessary corrections are applied to the observed altitudes. The best time of observation is when the sun is on the prime vertical.

10.13.2 Time by Meridian Transit of Star or Sun

In this method, a star is observed by a theodolite when it crosses the meridian. The chronometer time at this instant is noted. Since H.A. of star at transit is zero, L.S.T. = R.A. of star. Then, L.S.T. is converted into L.M.T.

In the case of sun, the time at which its east and west limb crosses the meridians are noted. The mean of these two times gives the time of transit of the sun's centre. Then, knowing the equation of time, L.M.T. at L.A.N. can be determined.

10.13.3 Time by Equal Altitude of a Star or Sun

This is the most simple and accurate method and is used when meridian is not known. The computed altitude of the star is set on the vertical circle and the star is followed with the vertical cross-hair. When it crosses the horizontal cross-hair, the time is recorded. When the star attains the same altitude on the opposite side of the meridian, the above procedure is repeated and time is recorded again. The average of the two observations give the time of transit of the star. This is compared with the L.M.T. computed from R.A. and the error of the watch is determined.

In the case of sun, the procedure is same except that the same limb of the sun is observed to avoid semi-diameter correction. The time of transit so found needs a correction, since the declination of the sun will change.

Example 10.29 After applying the correction for refraction, parallax and level, the mean observed altitude of the sun was $35^{\circ}10'15''$ at a place having latitude $35^{\circ}20'25''\text{N}$ and longitude $55^{\circ}25'10''\text{E}$. At the time of observation, the mean watch time was $15^{\text{h}} 39^{\text{m}} 14^{\text{s}}$ and the watch is known to be 4^{m} fast on L.M.T. If the declination of the sun at the instant of observation is $16^{\circ}35'44''$ and G.M.T. of Greenwich apparent noon is $11^{\text{h}} 55^{\text{m}} 30^{\text{s}}$, then find out the watch error.

Solution Here, $\alpha = 35^{\circ}10'15''$, $\delta = 16^{\circ}35'44''$, $\theta = 35^{\circ}20'25''\text{N}$
The hour angle H can be calculated from:

$$\begin{aligned}\sin \alpha &= \sin \theta \sin \delta + \cos \theta \cos \delta \cos H \\ \sin 35^{\circ}10'15'' &= \sin 35^{\circ}20'25'' \times \sin 16^{\circ}35'44'' \\ &\quad + \cos 35^{\circ}20'25'' \times \cos 16^{\circ}35'44'' \times \cos H\end{aligned}$$

$$\text{or} \quad \cos H = \frac{0.576016 - 0.165208}{0.78175} = 0.52549$$

$$\begin{aligned}\text{or} \quad H &= 58^{\circ}17'53'' = 3^{\text{h}} 53^{\text{m}} 11.5^{\text{s}} \\ \text{L.A.T.} &= 15^{\text{h}} 53^{\text{m}} 11.5^{\text{s}}\end{aligned}$$

Now,
$$\begin{aligned} \text{G.A.T.} &= \text{L.A.T.} - \text{longitude} \\ &= 15^h 53^m 11.5^s - 3^h 41^m 40.6^s \\ &= 12^h 11^m 30^s \end{aligned}$$

(Longitude = $55^\circ 25' 10'' = 3^h 41^m 40.6^s$)

But,
$$\text{G.A.T.} = \text{G.M.T.} + \text{E.T.}$$

or
$$\text{E.T.} = \text{G.A.T.} - \text{G.M.T.}$$

Now,
$$\text{G.M.T. of G.A.N.} = 11^h 55^m 30^s$$

or
$$\text{G.M.T. of } 12^h \text{ apparent time} = 11^h 55^m 30^s$$

$\therefore$
$$\text{E.T.} = 12^h - 11^h 55^m 30^s = 4^m 30^s$$

Now,
$$\begin{aligned} \text{G.M.T.} &= \text{G.A.T.} + \text{E.T.} \\ &= 12^h 11^m 30^s + 4^m 30^s \\ &= 12^h 16^m 0^s \end{aligned}$$

$$\begin{aligned} \text{L.M.T.} &= \text{G.M.T.} + \text{longitude} \\ &= 12^h 16^m 0^s + 3^h 41^m 40.6^s \\ &= 15^h 57^m 40.6^s \end{aligned}$$

Error in the watch = $15^h 39^m 14^s - 15^h 57^m 40.6^s$

$$= - 18^m 26.6^s$$

Hence, the watch is $18^m 26.6^s$ slow.

10.14 DETERMINATION OF TRUE MERIDIAN

The most important operation in the field of engineering astronomy is the determination of the true north. While determining the true meridian it is necessary to know the latitude of the place. It can be either taken from the map or may be found by observing the altitude of Polaris which is very nearly equal to the latitude of the place of observations. The most common methods of determining true meridian are by observation on Polaris at culmination or at elongation, the latter being the most accurate method.

Note *Since the celestial pole is not visible, it is necessary to make use of celestial bodies for establishing meridians.*

10.14.1 Observation on Polaris at Culmination

When Polaris is at its upper or lower culmination, a line of sight directed to it from any point on the earth's surface will be in a meridian plane through the point. This occurs twice every 24 hours. A day is selected when Polaris is at the upper culmination during the early part of the evening or during night. The exact time of its upper culmination is obtained from tables. The time of its lower culmination is calculated by the relation:

$$\text{Time of lower culmination} = \text{Time of upper culmination} \pm 11^h 58^m.$$

Use plus sign when the time of upper culmination is less than $11^h 58^m$ and minus if it is more than $11^h 58^m$.

It may also be noted that at the instant of upper culmination, L.S.T. is equal to R.A. of Polaris and at the instant of lower culmination, L.S.T. = R.A. + 12^h .

Procedure

1. About 15 min. before the culmination, a theodolite is set up and levelled over the place of observation. The vertical circle vernier is set to the latitude of the place to locate Polaris.
2. About 5 min. before the time of culmination, the telescope is directed to Polaris. With the help of the tangent screw, Polaris is followed continuously with the vertical cross-hair until the exact time.
3. The telescope is depressed and a stake is set on the line of sight at about 150 m from the theodolite. A tack is fixed to it in line with the vertical cross-hair. The lines joining the instrument station and the tack provides the meridian.

Since the meridian is fixed by only one observation, there is no check on the work. Also, if the transit is out of adjustment, the errors will be appreciable.

10.14.2 Observation on Polaris at Elongation

When Polaris is directly above the pole, it is in the plane of the meridian (bearing true north) and is said to be in its upper culmination (Fig. 10.36). About $11^h 58^m$ later, it will be directly below the pole (bearing true north) at its lower culmination. About halfway between these two positions, Polaris reaches its greatest east or west bearing and is said to be at its greatest elongation. When Polaris is at elongation, its bearing does not change since it moves vertically at that instant. Therefore, the instant of elongation is the most favourable time for accurate observations of meridian. However, Polaris moves very slowly during culmination and, therefore, its bearing can be obtained with fair accuracy. The following method will usually suffice to determine a meridian within a minute or two, of its true direction. Greater precision is seldom required.

The time of elongation of Polaris may either be obtained from the tables or may be calculated. It is at eastern elongation about $5^h 55^m$ before upper culmination and at western elongation about $5^h 55^m$ after upper culmination.

Procedure

1. About half an hour before the computed time of elongation, a theodolite is set up and levelled at the place of observation. The vertical circle vernier is set to the latitude of the place to locate Polaris.
2. The horizontal plates are set to 0° . The star is continually followed with the vertical hair by means of the upper tangent screw. Polaris may be found moving left or right at the eastern or western elongations, respectively. Just about the time of elongation, it stops moving horizontally, but appears to move vertically upwards for eastern elongation and downwards for western elongation. Polaris is followed until the time of elongation.
3. The telescope is depressed and a stake is set as explained in the previous method.
4. The telescope is transited and the Polaris is bisected again. Then the telescope is depressed and a second mark is set on the stake beside the first mark. The point exactly midway between these two marks gives the exact position of Polaris at elongation. The line joining the instrument station to the point so established gives the direction of Polaris at elongation.

5. The azimuth of Polaris is given by

$$\sin \text{azimuth} = \frac{\cos \text{declination}}{\cos \text{latitude}} \quad \text{or} \quad \sin A = \frac{\cos \delta}{\cos \theta}$$

This azimuth is laid off. The line of sight is thus directed along the true meridian.

6. A stake is driven at about 150 m on the line of sight, in line with the vertical cross-hair, The line joining the two stakes is the direction of the true meridian.

10.14.3 Comparison of the Two Methods

The culmination method is the most simple as it does not involve calculation of the azimuth of the star. But since the star appears to move fast at culmination, this necessitates the determination of time of culmination accurately and the use of an accurate timepiece by the observer. The other drawback of this method is that of a single observation and no checks over the work. The elongation method is independent of the timepiece of observer. Though calculation of azimuth is required, but the calculation of elongation is simple. The only disadvantage of this method is that elongation may occur at inconvenient times, very late at night or in the daytime.

10.15 DETERMINATION OF AZIMUTH

An azimuth is the horizontal angle a celestial body makes with the pole. The observation for azimuth determination of a survey line consists in measuring the horizontal angle between a reference mark, a triangulation station appropriately illuminated, and the celestial body. The azimuth of the reference mark is the calculated azimuth of the celestial body. The azimuth of the survey line is then obtained by adding the horizontal angle which the line makes with the reference mark to the azimuth of the latter.

There are several methods prevalent for the determination of the azimuth, but the one permitting both the left and the right face observations should be preferred. It is advisable to choose a close circumpolar star, since it changes very slowly with time in azimuth. In the northern hemisphere such a star is *Polaris* (Dhruva) which is the brightest star of Ursa Minor and is also known as *pole star* or *north star*. It can be seen by the naked eye and describes a circle of about 1° radius around the pole. It can be readily identified by means of the constellation called *Great Bear* or *Ursa Major* (Saptarishi) in summers (Fig. 10.36). This consists of seven bright stars, two of which point directly towards north and are called *pointers*. The line joining these two stars passes very nearly through Polaris. In winters, Polaris can be identified by the constellation *Cassiopea* (Kashyapi). It lies at the bisector of the large angle of W-shaped Cassiopea, the two constellations being on the opposite sides of Polaris. In case of Southern hemisphere, azimuth can be determined with the help of a star Alpha Crux of crux constellation also called *Jewel Box*.

In case some other star is chosen for determination of azimuth, its position may be located with the help of a star chart. In the absence of a star chart, the H.A.

of the star may be calculated and knowing the declination of the star and latitude of the place, the star position may be located by solving the ZPS triangle.

At upper culmination

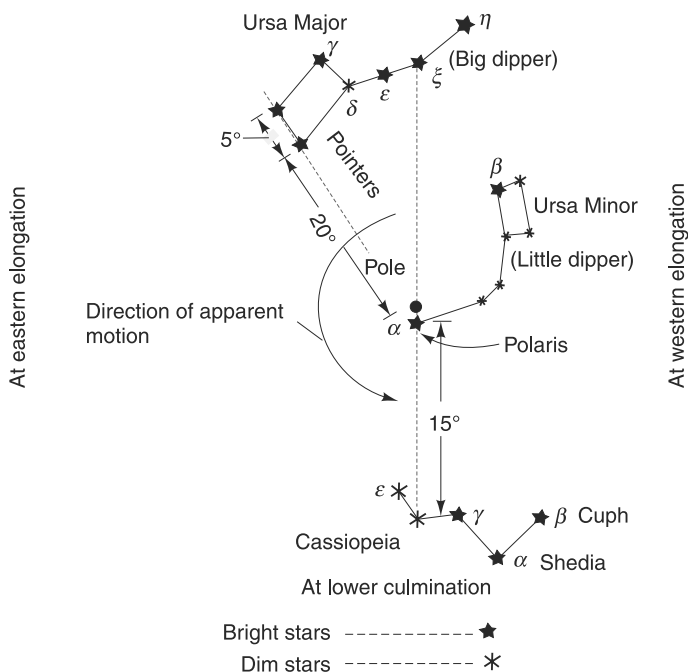


Fig. 10.36

The azimuth of a survey line can be determined by extra-meridian observation of the sun or of a circumpolar star, or of a star near prime vertical; circumpolar star at elongation; and equal altitudes of the sun or of a circumpolar star.

10.15.1 Azimuth by Extra-meridian Observation of Sun

The results obtained by observations to the sun are not as accurate as that computed from star observations, but since it is convenient to do the field work at daytime and since only one observation is involved, it is the most convenient method of determining the azimuth of a survey line.

The most common field method is by measuring altitude of the sun and computing the sun's azimuth by sine or tangent formula.

Procedure

1. A theodolite is set up at one end of the line where azimuth is required and the plate vernier is set to read 0° . Bisect the other end, the reference mark of the line with the vertical cross-wire.
2. The coloured glass shade is placed over the eyepiece and the telescope is pointed towards the sun by loosening the upper clamp.

3. The sun is first observed in the positions 1 and 3 (Fig. 10.37) and then in positions 2 and 4 with changed face. In all the positions, the horizontal angle, altitude and the time of observation are noted.

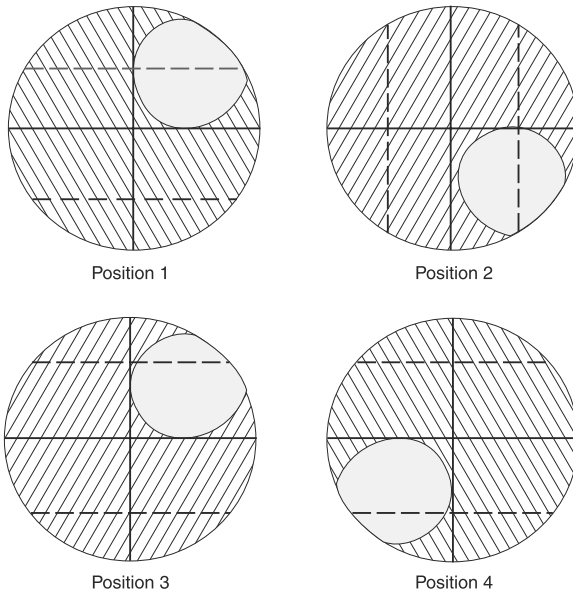


Fig. 10.37 Observations on sun for azimuth

4. During the above four observations, the sun changes its position considerably and accurate results cannot be obtained by averaging the measured altitude and time. However, the altitudes and time measured with sun in positions 1 and 3 can be averaged and similarly those measured for positions 2 and 4 can be averaged.
5. These two averaged altitudes are corrected for index error to give the apparent altitudes. The azimuth for these two values of the altitudes is calculated using the tangent formula. Here, reference may be made to Fig. 10.38.

$$\tan \frac{A}{2} = \sqrt{\frac{\sin(s-z)\sin(s-c)}{\sin s \sin(s-p)}}$$

$$\text{where } s = \frac{1}{2}(p + c + z)$$

$$p = \text{polar distance} = 90^\circ - \delta$$

$$c = \text{co-latitude} = 90^\circ - \theta$$

$$z = \text{zenith distance} = 90^\circ - \alpha$$

The two values of the azimuth obtained are averaged.

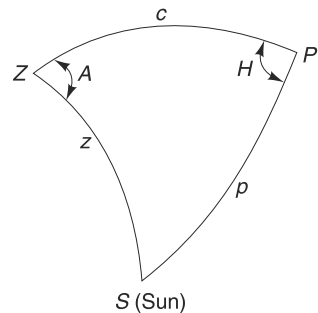


Fig. 10.38

Notes

- i. The formula gives the angle as measured from the north point in a clockwise or counter clockwise direction according to whether the sun is rising or

setting. Expressed as a clockwise whole circle bearing from the north, the azimuth in the former case is $\angle PZS$ (A), and in the latter case $360^\circ - \angle PZS$.

- ii. By observing the opposite limbs of the sun and averaging the observations, the instrumental errors and error due to semi-diameter are eliminated.
- iii. The best period for observations is between 8 to 10 a.m. and 2 to 4 p.m.

10.15.2 Azimuth by Extra-meridian Observation of a Circumpolar Star or of a Star near the Prime Vertical

The process consists in knowing the latitude of the place, the declination of the star, and solving the astronomical triangle for the azimuth.

A star should be observed when it is changing rapidly in altitude and slowly in azimuth. A favourable position is when the star is on the prime vertical. In such a position, the influence of errors in observed altitude will be small. In Fig. 10.39, S is the star for which α is the observed altitude. Azimuth (A) is given by

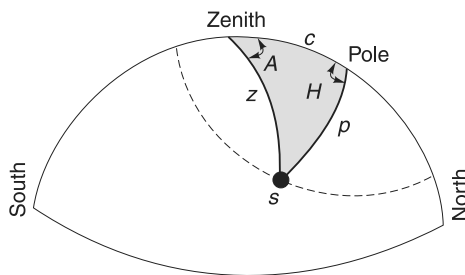


Fig. 10.39

$$\tan \frac{A}{2} = \sqrt{\frac{\sin(s-z)\sin(s-c)}{\sin s \sin(s-p)}}$$

Procedure

1. A theodolite is set up and levelled at one end of the line whose azimuth is to be obtained.
2. Both the plates are clamped to 0° and the reference mark, the other end of the line, is bisected with the vertical hair with face left.
3. The upper clamp is unclamped and the star is bisected. The horizontal and vertical angles are noted.
4. With changed face the star is again bisected, and the horizontal and vertical angles are once again noted.
5. The entire process is repeated to obtain one more set of observations.
6. The average values are used to calculate azimuth.

10.15.3 Azimuth by Observations on Circumpolar Star at Elongation

A circumpolar star is always above the horizon and therefore does not set. When the star is at elongation, it appears to move vertically and the vertical circle through the zenith is tangent to the circular path of the star (Fig. 10.40). The angle ZPS , i.e., the angle between the plane of the declination circle and the plane of the vertical circle, is a right angle.

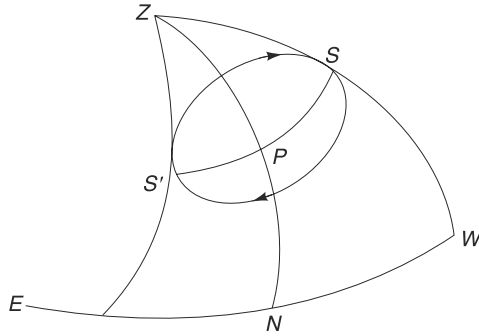


Fig. 10.40 *Circumpolar star at elongation*

Prior to making field observations for azimuth, it is necessary to calculate the time at which the star will elongate. The process has already been explained in Sec. 10.13.

10.15.4 Azimuth by Equal Altitude of a Circumpolar Star

In this method, knowledge of the latitude or local time is not necessary, but the duration of field work is of 4–6 hours at night which is most inconvenient. The results obtained are not very accurate because of considerable variation of refraction effects during the period of observations.

Procedure

1. A theodolite is set up and levelled at one end of the line whose azimuth is to be found. At the other end of the line, the reference mark is bisected by the vertical hair with both the plates clamped to 0° .

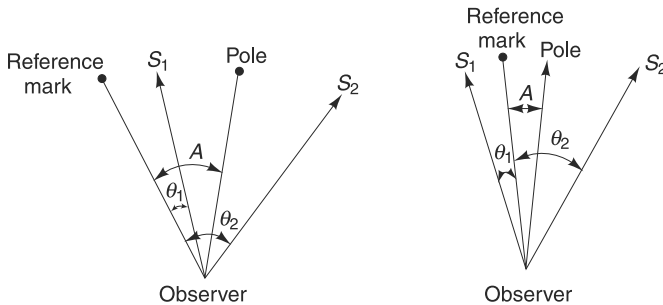


Fig. 10.41 *Observations on circumpolar star*

2. The upper clamp is loosened and the telescope is revolved clockwise to bisect the star at position S_1 (Fig. 10.41). The horizontal angle θ_1 and the altitude α of the star are noted.
3. The star in its due course reaches the other side of the meridian. The upper clamp is loosened and the star is bisected again when it attains the same altitude. In this observation the telescope is turned in azimuth without disturbing the reading on the vertical circle.

4. Since the direction of the meridian is midway between the two positions of the star, the azimuth of the line may be determined according to whether both the positions of star are to the same side of the reference mark or on its different sides. In the former case, azimuth

$$A = \theta_1 + \frac{\theta_2 - \theta_1}{2} = \frac{\theta_1 + \theta_2}{2}$$

In the latter case, azimuth

$$A = \frac{\theta_1 + \theta_2}{2} - \theta_1 = \frac{\theta_2 - \theta_1}{2}$$

10.15.5 Azimuth by Observation on Sun at Equal Altitudes

For observation on sun at equal altitudes, the process is the same as that for a star at equal altitude with the following differences:

1. The observations are made at daytime.
2. The observations are made on the right and left limbs of the sun with the telescope normal and inverted, both in the forenoon and the afternoon.
3. The time for each observation is noted.
4. A correction is applied to the azimuth so obtained because of the change in sun's declination during the interval between the forenoon and afternoon observations. This correction is given by

$$C = \frac{1}{2} (\delta_W - \delta_E) \sec \theta \operatorname{cosec} t$$

where C is the correction, t is half the interval between the times of equal altitude, θ is the latitude of the observer's place, δ_E is the sun's average declination of the forenoon observations, and δ_W is the sun's average declination of the afternoon observations.

Example 10.30 At a place of latitude $30^\circ 40' 48''$ N an observation was made on a star at eastern elongation. The horizontal circle reading observed for the star was $327^\circ 00' 03''$ and for the reference object was $6^\circ 50' 10''$. The declination of the star was $83^\circ 05' 14''$. Find the azimuth of the object.

Solution Refer to Fig. 10.42.

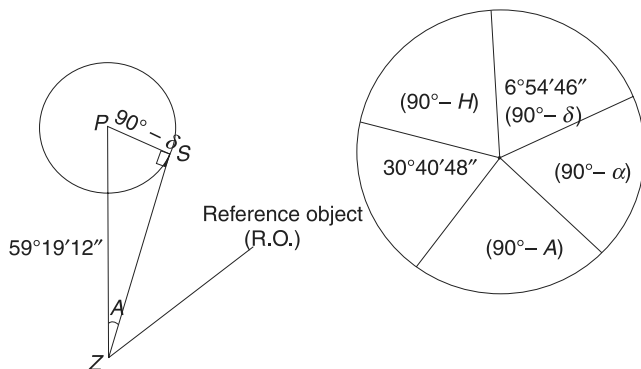


Fig. 10.42

The horizontal angle between the reference object and the star at elongation is

$$6^{\circ}50'10'' + (360^{\circ} - 327^{\circ}00'03'') = 39^{\circ}50'07''$$

From Napier's rules

$$\sin (90^{\circ} - \delta) = \cos \theta \cos (90^{\circ} - A)$$

$$\text{or } \sin (90^{\circ} - 83^{\circ}05'14'') = \cos 30^{\circ}40'48'' \times \cos (90^{\circ} - A)$$

$$\text{or } A = 8^{\circ}02'41''$$

Hence, azimuth of reference object = $39^{\circ}50'07'' + 8^{\circ}02'41'' = 47^{\circ}52'48''$.

Example 10.31 To determine the azimuth of a reference object and the latitude of a place, the following observations were made upon two stars at elongation. Obtain the desired azimuth and latitude.

Object	Elongation	Declination	Horizontal circle reading
Star S_1	Eastern	$79^{\circ}52'08''\text{S}$	$10^{\circ}20'15''$
Star S_2	Western	$84^{\circ}27'15''\text{S}$	$57^{\circ}29'40''$
Reference object	—		$98^{\circ}20'35''$

Solution Refer to Fig. 10.43.

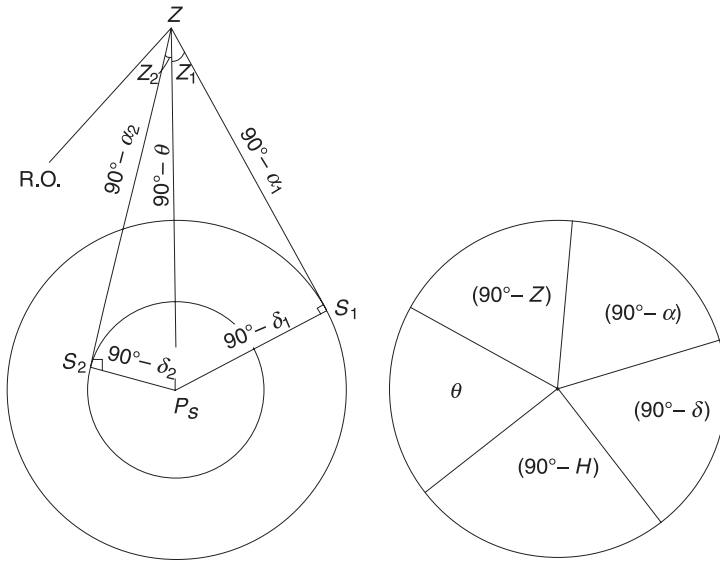


Fig. 10.43

Let the horizontal circle reading of stars S_1 and S_2 be H_1 and H_2 , respectively. Since the declination is south, the latitude is south. From the figure,

$$Z_1 + Z_2 = H_2 - H_1 = 57^{\circ}29'40'' - 10^{\circ}20'15'' = 47^{\circ}09'25''$$

$$\text{We know, } \cos \delta_1 = \sin Z_1 \cos A$$

$$\text{and } \cos \delta_2 = \sin Z_2 \cos A$$

$$\text{Hence, } \frac{\sin Z_1}{\sin Z_2} = \frac{\cos \delta_1}{\cos \delta_2}$$

$$\frac{\sin (47^{\circ}09'25'' - Z_2)}{\sin Z_2} = \frac{\cos 79^{\circ}52'08''}{\cos 84^{\circ}27'15''}$$

$$\text{or } \frac{\sin 47^{\circ}09'25'' \times \cos Z_2 - \cos 47^{\circ}09'25'' \times \sin Z_2}{\sin Z_2} = 1.8201332$$

$$\Rightarrow Z_2 = 16^{\circ}20'42''$$

$$\text{Hence, } Z_1 = 47^{\circ}09'25'' - 16^{\circ}20'42'' = 30^{\circ}48'43''$$

Azimuth to reference object

$$\begin{aligned} &= 180^{\circ} + 16^{\circ}20'42'' + (98^{\circ}20'35'' - 57^{\circ}29'40'') \\ &= 237^{\circ}11'37'' \end{aligned}$$

$$\cos \theta = \frac{\cos \delta_1}{\sin Z_1} = \frac{\cos 79^{\circ}52'08''}{\sin 30^{\circ}48'43''}$$

$$\text{or } \theta = 69^{\circ}54'55.21''\text{S}$$

Example 10.32 For the following data, determine the altitude of the star, azimuth of the survey line AB and the L.M.T. of observation.

Latitude of station $A = 48^{\circ}45'\text{N}$

Longitude of station $A = 52^{\circ}10'\text{W}$

Horizontal angle of star to the right of reference mark $B = 95^{\circ}20'45''$

Declination of the star = $61^{\circ}14'50'\text{N}$

R.A. of the star = $10^h 59^m 59^s$

G.S.T. of G.M.N. = $6^h 12^m 20^s$

Solution Since the star was observed at elongation, the angle ZPS in the astronomical triangle will be a right angle.

Altitude of the star

$$\sin \alpha = \frac{\sin \theta}{\sin \delta} = \frac{\sin 48^{\circ}45'}{\sin 61^{\circ}14'50''}$$

$$\text{or } \alpha = 59^{\circ}02'44''$$

Azimuth of the line AB

Azimuth of the line $AB = \text{Azimuth of star} + \text{Horizontal angle between the line and star}$

$$\sin A = \frac{\cos \delta}{\cos \theta} = \frac{\cos 61^{\circ}14'50''}{\cos 48^{\circ}45'}$$

$$\text{or } A = 47^{\circ}10'26''$$

$$\text{Hence, azimuth of } AB = 47^{\circ}10'26'' + 95^{\circ}20'45''$$

$$= 142^{\circ}31'11''\text{W}$$

$$= 360^{\circ} - 142^{\circ}31'11''$$

$$= 217^{\circ}28'49'' \text{ clockwise from north}$$

L.M.T. of observation

$$\text{G.S.T. of G.M.N.} = 6^h 12^m 20^s$$

$$\text{Longitude west} = 52^{\circ} 10' = 3^h 28^m 40^s$$

Acceleration of 9.856^s per hour of longitude = 34.28^s

$$\text{L.S.T. of L.M.N.} = 6^h 12^m 20^s + 34.28^s$$

$$= 6^h 12^m 54.28^s$$

$$\text{L.S.T.} = \text{R.A. of star} + \text{H.A. of star}$$

$$\cos H = \frac{\tan \theta}{\tan \delta} = \frac{\tan 48^\circ 45'}{\tan 61^\circ 14' 50''}$$

or

$$H = 51^\circ 16' 11'' = 3^h 25^m 4.7^s$$

$\therefore$

$$\text{L.S.T.} = 10^h 59^m 59^s + 3^h 25^m 4.7^s = 14^h 25^m 3.7^s$$

$$\text{S.I. from L.M.N.} = \text{L.S.T.} - \text{L.S.T. of L.M.N.}$$

$$= 14^h 25^m 3.7^s - 6^h 12^m 54.2^s = 8^h 12^m 9.5^s$$

Retardation at 9.8296 per hour

$$= 9.8296^s \times 8^h 12^m 9.5^s = 1^m 20.63^s$$

Mean time interval from L.M.N. $8^h 12^m 9.5^s - 1^m 20.63^s$

$$= 8^h 10^m 48.87^s$$

Hence, L.M.T. of observation = $8^h 10^m 48.87^s$ p.m.

Example 10.33 A star of declination $42^\circ 20' 15''\text{S}$ is to be observed at lower and upper transit at a place of latitude $78^\circ 30'\text{S}$. Find the approximate apparent altitudes at which the star should be sought.

Solution Refer to Fig. 10.44.

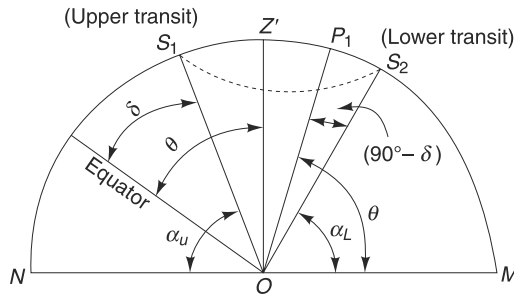


Fig. 10.44

$$\delta = 42^\circ 20' 15''\text{S}, \theta = 78^\circ 30'\text{S}, \alpha = \text{corrected altitude}$$

At lower transit

$$\alpha_L = \theta - (90^\circ - \delta)$$

$$= 78^\circ 30' - 90^\circ + 42^\circ 20' 15'' = 30^\circ 50' 15''$$

$$\text{Hence, apparent altitude} = 30^\circ 50' 15'' + 58'' \cot 30^\circ 50' 15''$$

$$= 30^\circ 51' 52''\text{S}$$

At upper transit

$$\alpha_u = 90^\circ - \theta + \delta$$

$$= 90^\circ - 78^\circ 30' + 42^\circ 20' 15'' = 53^\circ 50' 15''$$

$$\text{Apparent altitude} = 53^\circ 50' 15'' + 58'' \cot 53^\circ 50' 15''$$

$$= 53^\circ 50' 57''\text{N}$$

Example 10.34 Extra-meridian observations from a station in latitude $51^{\circ}38'24''\text{N}$ on Altair, when rising, yielded the following data:

Mean observed altitude of star	$= 32^{\circ}24'08''$
Mean observed horizontal angle on star	$= 119^{\circ}35'15''$
Mean observed horizontal angle on reference object (<i>R.O.</i>)	$= 0^{\circ}00'00''$
Refraction correction	$= 1'31''$ (negative)
Declination of star	$= 8^{\circ}43'51''\text{N}$
Find the azimuth of the <i>R.O.</i>	

Solution

Mean observed altitude of star $= 32^{\circ}24'08''$

The refraction correction to be applied is $1'31''$ (negative).

$\therefore$ corrected altitude of star $= 32^{\circ}24'08'' - 1'31'' = 32^{\circ}22'37''$

$\therefore$ co-altitude $= 90^{\circ} - 32^{\circ}22'37'' = 57^{\circ}37'23'' = ZS$

Latitude $= 51^{\circ}38'24''$

$\therefore$ co-latitude $= 38^{\circ}21'36'' = ZP$

Declination $= 8^{\circ}43'51''$

$\therefore$ co-declination $= 90^{\circ} - 8^{\circ}43'51'' = 81^{\circ}16'09'' = PS$

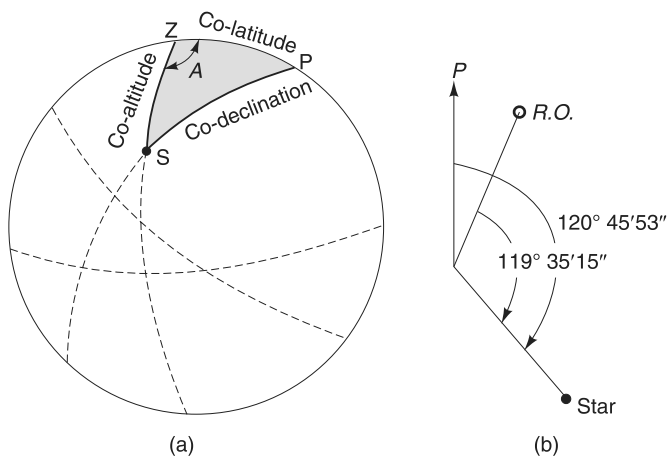


Fig. 10.45

The three sides of the astronomical triangle being available (Fig. 10.45(a)) the solution of the angle A is

$$\tan \frac{A}{2} = \sqrt{\frac{\sin(s - ZP) \sin(s - ZS)}{\sin s \sin(s - PS)}}$$

where $s = \frac{1}{2} (ZP + ZS + PS)$

$$= \frac{1}{2} \times 177^{\circ}15'08''$$

$$\begin{aligned}
 \therefore s &= 88^{\circ}37'34'' \\
 \text{Now, } (s - ZP) &= 50^{\circ}15'58'' \\
 (s - ZS) &= 31^{\circ}00'11'' \\
 (s - PS) &= 7^{\circ}21'25'' \\
 \tan \frac{A}{2} &= \sqrt{\frac{\sin 50^{\circ}15'58'' \times \sin 31^{\circ}00'11''}{\sin 88^{\circ}37'34'' \times \sin 7^{\circ}21'25''}} \\
 \text{or } A &= 120^{\circ}45'53''
 \end{aligned}$$

From Fig. 10.45(b)

$$\begin{aligned}
 \text{Azimuth of } R.O. &= 120^{\circ}45'53'' - 119^{\circ}35'15'' \\
 &= 1^{\circ}10'38''
 \end{aligned}$$

10.16 DETERMINATION OF LATITUDE

Latitude of a place may be determined by any of the following methods.

10.16.1 Latitude by the Altitudes of a Circumpolar Star

We know that the latitude of a place is equal to the altitude of the pole. The altitude of the pole is the mean of the altitude of a star at its upper and lower transit. If the two altitudes of a circumpolar star at its transits are known, then the latitude θ of the place of observation will be

$$\theta = \frac{1}{2}(\alpha_1 + \alpha_2)$$

where α_1 and α_2 are the observed altitudes, corrected for refraction, of the circumpolar star. The most suitable star for this method is Polaris, since it is very near to the pole and is bright enough for identification.

This method requires the calculation of time of culmination of the star. The disadvantage of the method is that the time interval between two observations is 12 sidereal hours necessitating one of the star observations to be made in the daylight requiring special instruments.

10.16.2 Latitude by Meridian Altitude of a Star

In this method the altitude of a star is measured when it crosses the meridian. If the declination, and hence the polar distance of the star is known the latitude can be easily computed. The observed altitude must be corrected for refraction. For accurate results, both face observations are taken and the value is averaged and this, therefore, requires taking observation on close circumpolar stars. The direction of the meridian of the place is established before making the observations. Four cases (Fig. 10.46) arise according to the position of the star.

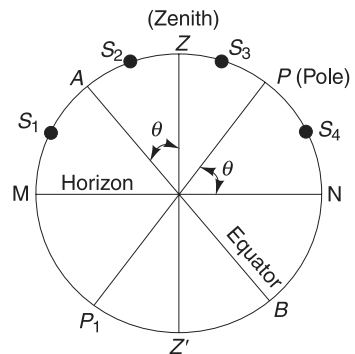


Fig. 10.46

Case I: Star (S_1) between the horizon and the equator

$$AZ = \text{latitude} = \theta$$

MS_1 = altitude of the star = α_1

$ZS_1 = z_1$ = zenith distance of the star = $90^\circ - \alpha_1$

AS_1 = declination of star = δ_1 (south)

Now,

$AZ = ZS_1 - AS_1$

or

$\theta = (90^\circ - \alpha_1) - \delta_1$

or

$\theta = z_1 - \delta_1$

Case II: Star (S_2) between the equator and the zenith

AZ = latitude = θ

MS_2 = altitude of the star = α_2

$ZS_2 = z_2$ = zenith distance of star = $90^\circ - \alpha_2$

AS_2 = declination of the star = δ_2

Now,

$AZ = ZS_2 + AS_2$

or

$\theta = (90^\circ - \alpha_2) + \delta_2$

or

$\theta = z_2 + \delta_2$

Case III: Star (S_3) between the zenith and the pole

AZ = latitude = θ

NS_3 = altitude of the star = α_3

$ZS_3 = z_3$ = zenith distance of the star = $90^\circ - \alpha_3$

AS_3 = declination of the star = δ_3

Now,

$AZ = AS_3 - ZS_3$

or

$\theta = \delta_3 - (90^\circ - \alpha_3)$

or

$\theta = \delta_3 - z_3$

Case IV: Star (S_4) between the pole and the horizon

NP = latitude = θ

NS_4 = altitude of the star = α_4

$ZS_4 = z_4$ = zenith distance of the star = $90^\circ - \alpha_4$

BS_4 = declination of the star = δ_4

Now,

$NP = NS_4 + PS_4$

or

$\theta = \alpha_4 + (90^\circ - \delta_4)$

or

$\theta = (90^\circ - z_4) + (90^\circ - \delta_4)$

or

$\theta = 180^\circ - (z_4 + \delta_4)$

10.16.3 Latitude by Altitude of the Sun

Observations are made on the sun when it is on the meridian. The declination of the sun at the instant of observation can be calculated if the longitude of the place is known. The observed altitude of the sun should be corrected for parallax, refraction and semi-diameter. Refer to Fig. 10.47.

AZ = latitude = θ

MS_1 = altitude of the sun = α

AS_1 = declination of the sun = δ

$ZS_1 = z$ = zenith distance of the sun Z
 $= 90^\circ - \alpha$

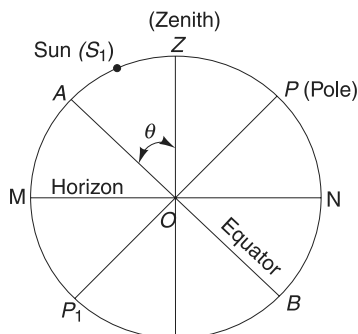


Fig. 10.47

Now, $AZ = ZS_1 + AS_1$
 or $\theta = 90^\circ - \alpha + \delta$
 or $\theta = z + \delta$

Note Declination is positive or negative according to whether the sun is north or south of the equator.

10.16.4 Latitude by Extra-meridian Altitude of Sun or Star

In this method the altitude of the sun or star is determined at any convenient time when it is not on the meridian.

Procedure

1. The altitude of the star is observed and due corrections are applied to it.
2. From the known declination, the polar distance ($90^\circ - \delta$) is calculated.
3. From the local mean time (L.M.T.) of observation, the hour angle of the sun or star is determined.
4. If observations are made of the sun, the interval since L.M.N. should be converted to L.A.N. using the equation of time. L.A.N. converted to arc gives the H.A. of the sun. If observations are made of a star, the L.M.T. of the observations is converted to S.T. The hour angle of the star may be obtained by the relation:

$$L.S.T. = R.A. + H.A.$$

5. By solving the astronomical triangle, the azimuth is computed:

$$\sin A = \frac{\sin(90^\circ - \delta)}{\sin(90^\circ - \alpha)} \sin H$$

6. The co-latitude is calculated by the relation:

$$\tan \frac{\text{co-latitude}}{2} = \frac{\sin \frac{A+H}{2}}{\sin \frac{A-H}{2}} \tan \frac{(90^\circ - \delta) - (90^\circ - \alpha)}{2}$$

7. Knowing the co-latitude ($90^\circ - \theta$), the value of latitude can be found

$$\theta = 90^\circ - \text{co-latitude}$$

10.16.5 Latitude by Extra-meridian Observation of Polaris

In this method the observations of Polaris are made at any convenient time when it is not on the meridian. Both, the face left and the face right observations are made. Accurate L.M.T. of observations are noted. The latitude θ is calculated from the relation:

$$\theta = \alpha - p \cos H + \frac{1}{2} \sin 1'' (p \sin H)^2 \tan \alpha$$

where α is the corrected altitude of Polaris, p is the polar distance ($90^\circ - \delta$), and H the hour angle of Polaris in arc (obtained from L.M.T. of observation of Polaris).

Example 10.35 The apparent altitude of Alpha Crux at its upper transit was observed as $22^\circ 20' 20''$ S. If its declination was $50^\circ 31' 25''$ S, what is the observer's latitude?

Solution Let S be the star at south of the equator OA , as shown in Fig. 10.48.

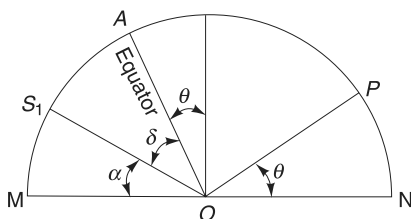


Fig. 10.48

$$\begin{aligned}\text{True altitude, } \alpha &= \text{Apparent altitude} - 58'' \times \cot 22^\circ 20' 20'' \\ &= 22^\circ 20' 20'' - 58'' \times \cot 22^\circ 20' 20'' \\ &= 22^\circ 20' 20'' - 2' 21.14'' \\ &= 22^\circ 17' 58.85''\end{aligned}$$

$$\begin{aligned}\text{Latitude, } \theta &= (90^\circ - \alpha) - \delta \\ &= 90^\circ - 22^\circ 17' 58.85'' - 50^\circ 31' 25'' \\ &= 17^\circ 10' 36.15''\text{N}\end{aligned}$$

Example 10.36 A star of declination $34^\circ 12' 28''\text{S}$ is observed at its upper transit. The altitude of the star was observed as $56^\circ 50' 53''\text{N}$. Find the latitude of the place.

Solution Refer to Fig. 10.49.

$$\delta = 34^\circ 12' 28''\text{S}, \alpha_1 = 56^\circ 50' 53''\text{N}$$

At upper transit

Corrected altitude,

$$\begin{aligned}\alpha &= 56^\circ 50' 53'' - 58'' \times \cot 56^\circ 50' 53'' \\ &= 56^\circ 50' 15''\text{N}\end{aligned}$$

$$\begin{aligned}\text{Latitude, } \theta &= (90^\circ - \alpha) + \delta \\ &= 90^\circ - 56^\circ 50' 15'' + 34^\circ 12' 28'' = 67^\circ 22' 13''\end{aligned}$$

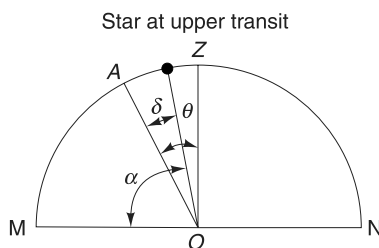


Fig. 10.49

Example 10.37 Two stars were observed while crossing the meridian and the following observations were recorded.

Star	Declination (δ)	Observed altitude at transit (α)
S_1	$22^\circ 10'\text{N}$	$53^\circ 20' 20''\text{N}$
S_2	$82^\circ 11'\text{S}$	$52^\circ 50' 10''\text{S}$

Determine the latitude.

Solution Refer to Fig. 10.50.

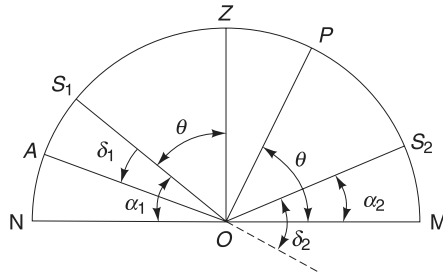


Fig. 10.50

Star S_1

$$\delta_1 = 22^\circ 10' \text{N}, \alpha'_1 = 53^\circ 20' 20'' \text{N}$$

$$\text{Corrected altitude, } \alpha_1 = 53^\circ 20' 20'' - 58'' \times \cot 53^\circ 20' 20'' = 53^\circ 19' 37''$$

$$\begin{aligned} \text{Latitude, } \theta &= 90^\circ - \alpha_1 + \delta_1 \\ &= 90^\circ - 53^\circ 19' 37'' + 22^\circ 10' = 58^\circ 50' 23'' \end{aligned}$$

Star S_2

$$\delta_2 = 82^\circ 11' \text{N}, \alpha'_2 = 52^\circ 50' 10'' \text{S}$$

$$\text{Corrected altitude, } \alpha_2 = 52^\circ 50' 10'' - 58'' \times \cot 52^\circ 50' 10'' = 52^\circ 49' 26''$$

$$\begin{aligned} \text{Latitude, } \theta &= 90^\circ - \delta_2 + \alpha_2 \\ &= 90^\circ - 82^\circ 11' + 52^\circ 49' 26'' = 60^\circ 38' 26'' \end{aligned}$$

$$\text{Hence, latitude} = \frac{58^\circ 50' 23'' + 60^\circ 38' 26''}{2} = 59^\circ 44' 24.5''$$

10.17 DETERMINATION OF LONGITUDE

The longitude of a place on the earth's surface may be defined as the angle at the pole between the meridian of Greenwich and the meridian of observation. It is thus the arc of the equator intercepted between these two meridians. Since the earth rotates uniformly about its axis with respect to the mean sun or star, this angle is proportional to, and is measured by, the time intervening between the respective transits of these bodies. Hence, terrestrial longitude may be defined as the difference of the L.A.T./L.M.T. of the meridian concerned and the time at the same instant at the Greenwich meridian.

The following methods are used for the determination of longitude:

1. By triangulation
2. By transportation of chronometers
3. By wireless signals
4. By observation of moon and stars which culminate at same time
5. By lunar distance
6. By celestial signals

Of these six methods, only the first three methods are in use and rest are absolute and outdated.

10.17.1 Longitude by Triangulation

This method of determining longitude is very costly, exceedingly intricate, and involves knowledge of the earth's figure, but is the most accurate method.

10.17.2 Longitude by Transportation of Chronometers

This method of determining longitude is the most common method and consists of expressing the difference between G.M.T. (by chronometer) and the local mean time (determined by the observation of star or sun). Knowing this difference, the longitude of the place can be found.

The difficulty in this method arises from the fact that the chronometer's rate, which is defined as the amount by which it gains or loses in 24 hours, is not constant for stationary and heavenly positions. Hence, the travelling rate of the chronometer has to be determined for precise calculations.

10.17.3 Longitude by Wireless Signals

The use of wireless signals has done away with the necessity of transporting the chronometer between various stations. Time signals are emitted from various stations at known intervals and the surveyor stationed anywhere in the world can check his chronometer with the aid of these signals. Knowing the difference between the chronometer readings at two stations serves as a direct measure of the longitude of the place.

10.18 NAUTICAL SEXTANT

It is an instrument for measuring angle between two distant objects from the observer's position. It can be used to measure horizontal as well as vertical angles. It is especially useful because it requires no fixed support, is light, and easy to handle. A sextant used for measuring the angles of elevation of stars is known as a *nautical sextant* and that used for locating the soundings in hydrographic surveys is known as a *sounding sextant*.

10.18.1 Construction

A sextant essentially consists of the following:

1. **A graduated arc ABC (Fig. 10.51) about which the vernier moves** It is of a radius of about 15 cm with centre at O . The arc is graduated to $10'$ and the vernier reads to $10''$.
2. **An index arm OB pivoted at O and carrying the index mirror M perpendicular to the plane of arc** It can be moved by means of a milled headed screw.
3. **A fixed horizon glass H in the line of sight of the telescope T** It is fixed at right angles to the plane of the arc and is parallel to the index mirror when the index arm is set to read zero. The lower half of the glass is silvered while the upper half is left plain.
4. **A set of coloured glasses** These serve to reduce the brightness of the sun if sighted.

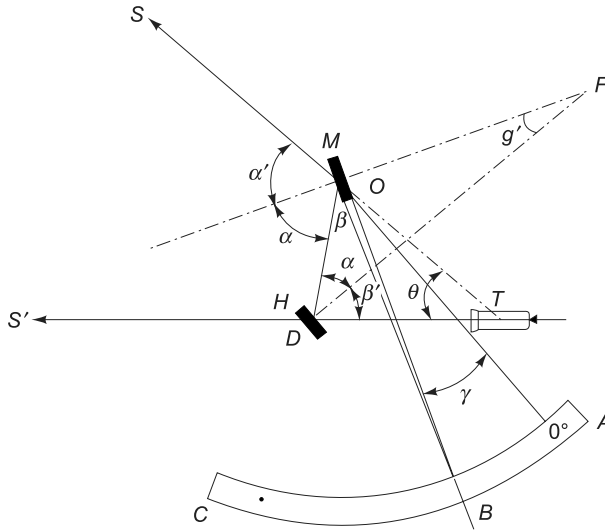


Fig. 10.51 Principle of nautical sextant

10.18.2 Principle

The principle of a sextant is ‘when a ray of light is reflected successively from two plane mirrors, the angle between the first incident ray and the last reflected ray is twice the angle between the plane of the two mirrors’.

To illustrate the principle of a sextant, let us measure the angle STS' (Fig. 10.51) between two objects or two stars S and S' . Point the telescope at the star S' . The ray $S'T$ comes through the clear part of the horizon glass H and forms the image of the star S' at the focal plane of the telescope. Rotate the sextant about the axis of the telescope and bring the star S into the plane of the instrument. Swing the index arm until the ray $SODT$ from S is brought into the field of the telescope forming an image in the focal plane. Clamp the index arm and bring the images of the two stars into coincidence by the tangent screw. The reading on the arc gives the angle STS' . The fact that the two images will not coincide unless S and S' are in the plane of the instrument, permits accurate measurement with the sextant even without a firm fixed support.

If OF and FD are the respective normals to the mirrors M and H then $\alpha = \alpha'$ and $\beta = \beta'$. From triangles FDO and ODT

$$\gamma' + \beta = \alpha$$

$$\text{and} \quad 2\beta + \theta = 2\alpha$$

Eliminating α from these equations gives

$$\gamma' = \frac{\theta}{2}$$

If γ represents the angle of rotation of the index mirror, then $\gamma = \gamma'$ and

$$\gamma = \frac{\theta}{2}.$$

Thus, to read the angle between the two objects directly on the arc, the graduations of half degrees are numbered as whole degrees.

10.18.3 Index Correction

The index error is usually determined before each series of observations. The index of vernier is said to be 'on the arc' when it is to the left of the zero mark of the graduated arc and 'off the arc' when it is to the right.

If the index is on the arc, the index correction $i = 0^\circ - R$ and i is negative; when the index is off the arc, $i = 360^\circ - R$ and i is positive (Fig. 10.52), where R is the reading on the arc.

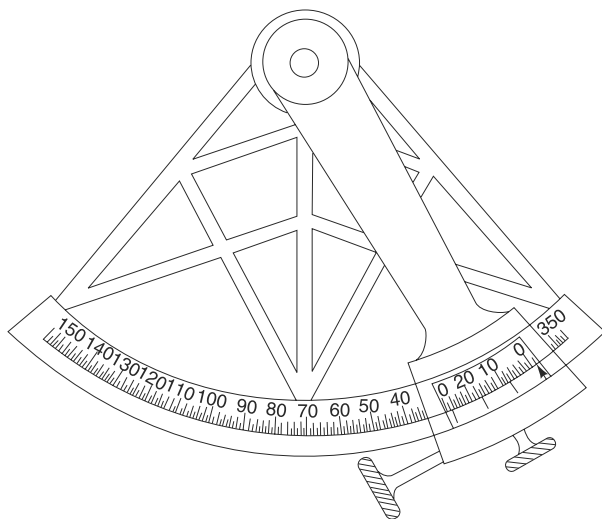


Fig. 10.52 Vernier index of the arc (approximate reading of angle shown is $359^\circ 15'$)

10.18.4 Solar Observations

Point the telescope to the sun and make the two images externally tangent. First put the direct image below the reflected image, in which case the vernier index will be on the arc (Fig. 10.53(a)). The angular diameter of the sun D will equal vernier reading R_1 plus i . Then put the direct image above the reflected image (Fig. 10.53(b)), the vernier index will be off the arc. Then, the angular diameter of sun D will equal 360° minus the vernier reading R_2 minus the index error i . Thus we get,

$$D = R_1 + i$$

and

$$D = 360^\circ - R_2 - i$$

Eliminating D and solving for i , we have

$$i = 180^\circ - \frac{1}{2}(R_1 + R_2)$$

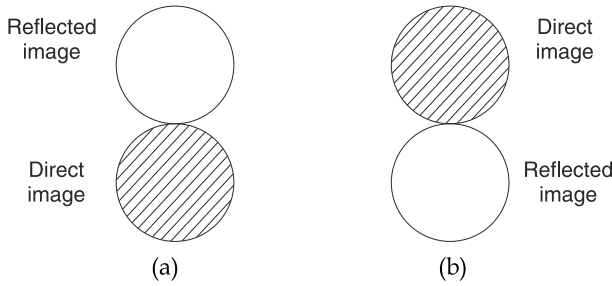


Fig. 10.53 *Position of two images of sun*

10.18.5 Measurement of Altitude

When a sextant is used at sea to obtain the altitude of sun, colour shades are placed in front of the index mirror to diminish the image brightness and, if necessary, in front of the horizon glass to diminish the brightness of the field. The sextant is held in the right hand with the telescope pointed at the horizon directly below the sun. The index arm is moved so that the reflected image of the sun is brought to the horizon near the middle of the field (Fig. 10.54(a)). To bring the instrument into a vertical plane, rotate it slightly about the axis of the telescope by a small movement of the wrist. This will cause the sun's image to describe an arc. By means of the slow motion screw, make the arc described by the lower limb of the sun tangent to the horizon (Fig. 10.54(c)). When the altitude is changing rapidly as in the case of the sun away from the meridian, proceed thus. For the sun east of the meridian, with the slow motion screw, slightly increase the altitude (Fig. 10.54(b)) and wait until the arc described by the sun's lower limb is tangent to the horizon (Fig. 10.54(c)). For the sun west of the meridian, diminish the altitude, and wait for the tangency as before. The most favourable time to observe the altitude of a planet or a star at sea is during twilight when the horizon is clearly visible. At this time, also, the beginner is not confused by too many stars in the sky. The star or planet appears as a point of light, and the arc described by it is made tangent to the horizon in the same manner as has been explained with reference to the sun.

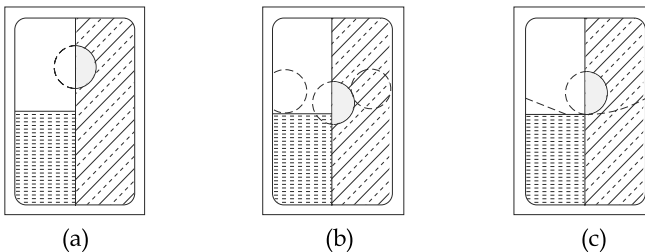


Fig. 10.54 *The sun and the horizon observed through a direct vision telescope*

When the altitude α of a heavenly body is measured on land, the artificial horizon is used. This is a rectangular shallow basin of mercury, protected from wind by a sloping roof of glass. The observer so places himself (Fig. 10.55) that he can see the image of the body, whose altitude is to be measured, reflected in the mercury and points the telescope towards the artificial horizon so that for 'direct image' he uses a second image of the heavenly body, in place of the horizon. By moving the index arm, the two images are made to coincide. The reading of the circle gives twice the altitude; this is clear from Fig. 10.56. The index correction must be applied to the double angle.

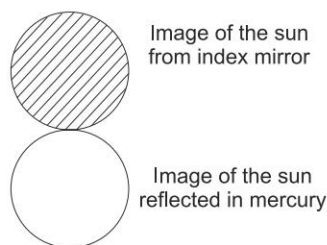
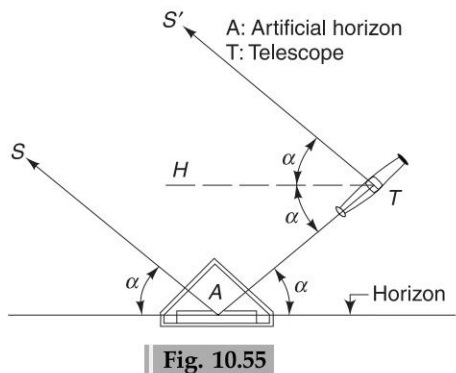


Fig. 10.56 Altitude of the upper limb of the sun observed with the inverting telescope

When the sun's altitude is measured, the two images are made tangent externally or allowed to move into external tangency. This gives the double altitude of the upper or lower limb. Suppose that an inverting telescope is used and that the image reflected from the mercury is made red by means of the red colour shade, and the other image is made blue. Then, if the red image of the sun is below (Fig. 10.57), the upper limb of the sun has been observed; if the red image is above, the lower limb, has been observed.

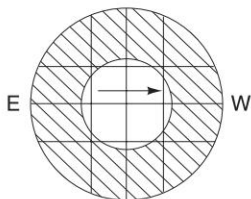


Fig. 10.57 System of wires of solar telescope showing the image of the sun so placed that its centre is in the line of sight of the telescope

Adjustments The following adjustments should be made in the order given:

1. **Make the index mirror perpendicular to the plane of the sextant:** Remove the telescope and place the instrument on a table with the arc away from the body. Set the index at about the middle of the arc and look into the index mirror in such a way as to see, at the same time, part of the arc by reflection and part directly. If the two images form a continuous arc,

no adjustment is necessary; if not, tip the mirror by whatever means has been provided on that particular instrument.

2. *Make the line of sight of the telescope parallel to the plane of the sextant:* Place the instrument on a table about 6 m from a wall and sight along the arc. Mark on the wall a line in the plane of the arc. Measure the distance from the centre of the telescope to the plane of the arc and mark a second line on the wall as much above the first as the distance just measured. Usually, the telescope of the sextant is provided with four cross-wires to mark the approximate centre of the field. Rotate the eyepiece until two of the parallel wires are horizontal. Now see if the image of the second mark bisects the space between the parallel wires. If so, the telescope is adjusted; if not, adjust by means of the screws in the ring that carries the telescope.
3. *Make the horizon glass perpendicular to the plane of the sextant:* Point the telescope to a well-defined object—a star, the sun, or a distant terrestrial mark. Move the index arm slowly back and forth past the zero mark. If the reflected image of the object does not coincide with the direct one in passing it, the horizon glass is not perpendicular to the plane of the sextant. Adjust it by means of the screw provided for the purpose at the back of the horizon glass.
4. *Make the two images equally distinct:* If the direct image of the object is brighter than the index image, the telescope receives more light through the plane glass and therefore the distance between the telescope and the frame of the instrument should be decreased. The adjustment is made by means of a screw in the reverse side of the frame.
5. *Make the vernier read zero when the index mirror and horizon glass are parallel:* The error arising from this is usually called the index error. Set the index at zero and observe a very distant, well-defined object. If the two images coincide, no adjustment is necessary. If not, turn the horizon glass about the axis perpendicular to the frame by means of the proper screw, until the two images coincide.

10.19 THEODOLITE ATTACHMENTS

Most of the observations in field astronomy are made with a theodolite provided with axis illumination. The two other instruments used to facilitate observations are solar attachment and prismatic astrolabe.

10.19.1 Solar Attachment

It is a device, usually mounted on the transit, for determining a true meridian by an observation on the sun and thus to determine the azimuth of the line. Time may also be determined with this attachment. Its use is based on the solution of the astronomical triangle mechanically (rather the trigonometrically), so that at the end of the observation, the line of sight through the transit telescope lies in the true meridian.

In the use of a solar attachment, it is necessary to know only the latitude. (or co-latitude) of the place of observation and the apparent declination of the sun at the hour of observation corrected for refraction. These are set off on the attachment and the astronomical triangle is solved solely by the manipulation of the instrument. It was invented by Burt in 1836. The other forms are Smith solar attachment and Saegmuller solar attachment. Although these differ in appearance and in the method by which these are mounted on the transit, they are alike in the principle. In addition, the Smith and the Burt attachments are provided with a circle for setting off the hour angle.

It is very useful when only a few observations are to be made and is principally used when accuracy is subordinate to speed. It is capable of attaining an accuracy of about one minute of arc. The two major parts of a solar attachment are the hour circle and the declination arc. The hour circle, the axis of which is called *polar axis*, is provided with a clamp and a reading index. The polar axis is perpendicular to the plane formed by the line of sight of the telescope of the transit and the horizontal axis of that telescope. The declination arc, which also contains the polar axis, is divided to read minutes in its own plane. The latitude arc is formed by the vertical circle of the theodolite, to which the attachment is attached, by its polar axis to the centre of the telescope. The solar attachment can be rotated about the polar axis.

Principle Suppose that the telescope of the transit is pointed at the intersection of the meridian with the plane of the equator with all the motions clamped except the one allowing the solar telescope to turn about the polar axis. The line of sight of the solar telescope will then describe a diurnal circle. If the angle between the two telescopes, when both are in meridian, equals the declination of the sun at the instant (neglecting refraction), then the line of sight for the solar telescope will follow the sun for a time until the sun's declination changes appreciably. Now suppose that the exact direction of the meridian is not known, that the horizontal motion of the transit is unclamped, and that the sun is at least an hour from the meridian; then the only time when the sun may be centred in the solar telescope is when the transit telescope is in the plane of meridian. This is the fundamental principle of the use of a solar telescope.

1. *Smith solar attachment*: It was designed by Benjamin H. Smith, a United States surveyor, in 1880. It consists of a solar telescope of low magnifying power mounted on a horizontal axis parallel to, but independent of, the supporting axle of the transit telescope. The solar telescope can be rotated in its supporting collars to permit the hour angle of observation to be set off by means of the graduated drum near the eyepiece at the upper end of the telescope. The index of the hour drum is attached to the upper collar. The eyepiece is provided with a coloured glass to cut down the intensity of sunlight. The declination arc is attached to the shell of the solar telescope and rotates with it when the hour drum is turned. When the declination is set off on its arc and the transit is turned horizontally, the sun's image can be reflected into the telescope. The image is centred accurately by means of cross-hairs which are spaced equal to the sun's apparent diameter. The latitude arc is mounted close to the solar telescope in such a manner as to rotate freely

about the outer end of the latitude axis. The collars supporting the solar telescope are so attached to the latitude arc that when the latitude is set off, the axis of the solar telescope is made parallel to the earth's polar axis in only one position, namely, when the transit telescope is turned, so as to lie in the meridian. At the same instant, the sun's image will be bisected in the solar telescope, provided the declination is set off correctly and the hour drum is rotated to show local apparent time. Conversely, if the latitude of the station is set off and the declination and the hour angle are turned off correctly for a given instant, the main telescope will lie in the meridian.

Procedure

1. Note the declination for the day and hour of the observation.
2. Correct the declination for refraction in the declination.
3. Set up the transit over a convenient station, say *A*, where the sun can be viewed without obstruction. Level it approximately.
4. Set off the final value of declination (step 2) on the declination arc. If the declination is negative, the sun will be below the plane of the celestial equator and the declination arm will have to be raised to tilt the reflecting mirror downwards towards the sun.
5. Set off the latitude of the station on the latitude arc.
6. The solar telescope is then rotated, thus tilting the reflecting mirror towards the sun, until the hour angle corresponding to the time of observation is set off on the drum.
7. Level the transit carefully with the telescope level.
8. Vernier *A* of horizontal circle is set to zero. The lower clamp screw is left unclamped so that both the plates, clamped together, can turn freely.
9. Rotate the instrument until the image of the sun is resected by the cross-hairs of the solar telescope. The final centring is completed by the lower clamp and the tangent screw. The line of sight in such a position will be in the true meridian.
10. The upper plate clamping screw is unclamped and the other end *B* of the line *AB* is sighted; the azimuth of the line *AB* is the angle turned off, i.e., the reading on the horizontal circle.

2. *Saegmuller solar attachment*: It is very simple in construction, has no graduated arcs, uses the vertical circle of the transit to set off both the latitude and declination angles. It also has no hour circle. The attachment is mounted on a polar axis fastened to a small base on the top of the transit telescope.

It consists of solar telescope supported by short standards in such a manner so as to allow the solar telescope to rotate about the polar axis. To control the rotation, a polar clamp and tangent screw are provided. A spirit level is attached to the solar telescope to aid in setting off the declination. The solar telescope is provided with cross-hairs spaced to include the sun's image. The sun is viewed directly through the solar telescope. A diagonal prism at the eyepiece is provided to allow sights to be taken without interfering with the main telescope.

Procedure

1. The final declination for the sun is found at the time of observation (see steps 2 and 3 for Smith solar attachment).
2. Both the solar and transit telescopes are brought into the same vertical plane. To do this, the instrument is levelled carefully over a station *A* and the index error, if any, of the vertical circle is noted. The solar attachment is also levelled (similar to the transit). Transit telescope is backsighted on a station *B*. The solar telescope is also backsighted on *B*, thus bringing both the telescopes into the same vertical plane.
3. The transit telescope is then inclined equal to the value of the sun's declination found in step 1. If the declination is north, the objective end of the telescope is depressed; and if south, it is elevated. Now, without disturbing the transit telescope, the solar telescope is brought to a horizontal position by means of its level. The angle between the two telescopes will be thus equal to the sun's declination.
4. Without disturbing the relative positions of the two telescopes, the vertical circle is unclamped and is set to the co-latitude of the station.
5. The upper horizontal plate is unclamped and is turned around the vertical axis. At the same instant, the solar telescope is turned around the polar axis until the sun's image is brought into the field of the solar telescope. The image is accurately centred. In such a position, the transit telescope must be in the meridian. The azimuth of *AB* is then known from the reading on the horizontal circle.

3. *Burt's solar attachment*: It was invented by William J. Burt, in 1835, as an attachment with an ordinary compass for setting out boundaries along meridians in land surveys and has since then been adopted to the engineers transit. It consists of (i) a polar axis mounted on an adjustable base attached to the top of the telescope directly above the centre of its supporting axle, so that the entire attachment, except the hour drum near the base, is free to rotate about the polar axis, (ii) a declination arc, and (iii) a wide movable solar arm about 5" in length having a small lens and small silver plate at each end. Each plate is ruled with fine lines defining small squares for the purpose of centring the sun's image formed by the solar lens. The arm carrying the lenses and the declination arc is reversible about the polar axis. When the declination is south, i.e., the sun is below the equator, the arm is turned end for end from the normal position. This is the reason for two lenses and two-ruled plates.

Procedure

1. Steps 1–4 are same as for the Smith solar attachment, except that after setting off the declination, due care is taken to turn the proper lens towards the sun.
2. The instrument is carefully levelled using the telescope level and the index error, if any, on the vertical circle is noted.
3. The local apparent time of the observation is set off on the hour drum.
4. Other steps required for backsighting, turning the transit so as to bring the sun's image within the ruled square, and reading the azimuth of

the line to the backsight are similar to those explained for the other attachments.

Determination of time When the transit telescope is in the meridian and the sun's centre is in the line of sight of the solar telescope, the angle between the two telescopes measured in a plane perpendicular to the polar axis is equal to the meridian angle of the sun. To measure it, the transit telescope is pointed to a distant mark, nearly on the same level as the instrument, and the horizontal circle is read. The transit telescope is pointed to the same mark. The horizontal circle is read again. The difference between the two readings gives the meridian angle of the sun. Hence, by comparing the clock time of observation with the time computed from the meridian angle of the sun, the correction for the clock may be obtained.

10.19.2 Prismatic Astrolabe

It can only be used to determine latitude, longitude, and time with reference to stars; it cannot be used for observation of azimuth. It is a French invention by M.M. Claude and Driencourt, who used a 60° prism to control the fixed altitude at which the stars are observed.

Its main feature is an equilateral prism P fitted in front of the objective O of a theodolite (Fig. 10.58) so that the face nearest to the telescope is vertical where an artificial horizon A is shown in position. Two images of a star are seen in the field of view: (i) directly from a face of the prism and (ii) indirectly from the artificial horizon and the other face of the prism. Since the first image is reflected once and the second twice, these images will move in opposite directions, according to whether the star is rising or setting. When the altitude is exactly 60° , the images coincide. Since the rate at which the images move is twice that at which a star would appear to move in the field of an ordinary telescope, the instant at which coincidence occurs can be determined more accurately with it.

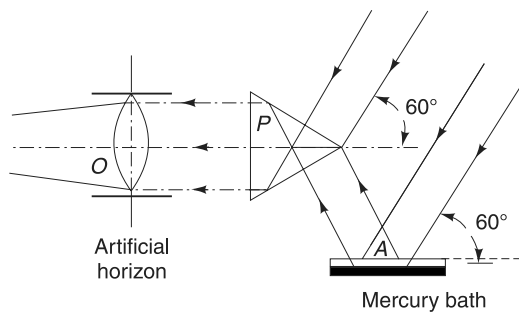


Fig. 10.58 Principle of Prismatic Astrolabe

The 60° Prismatic Astrolabe was improved by Captain T.Y. Baker, R.N., of England. In the improved model, the observations are made at a fixed altitude of 45° and is called 45° Prismatic Astrolabe (Fig. 10.59). One advantage of observing at an altitude of 45° instead of 60° is that there is a greater choice of stars for the lower altitude. Another is that the computations are somewhat simpler.

Stars near the meridian are preferred for latitude observations, whereas for time observations, those near prime vertical are suitable.

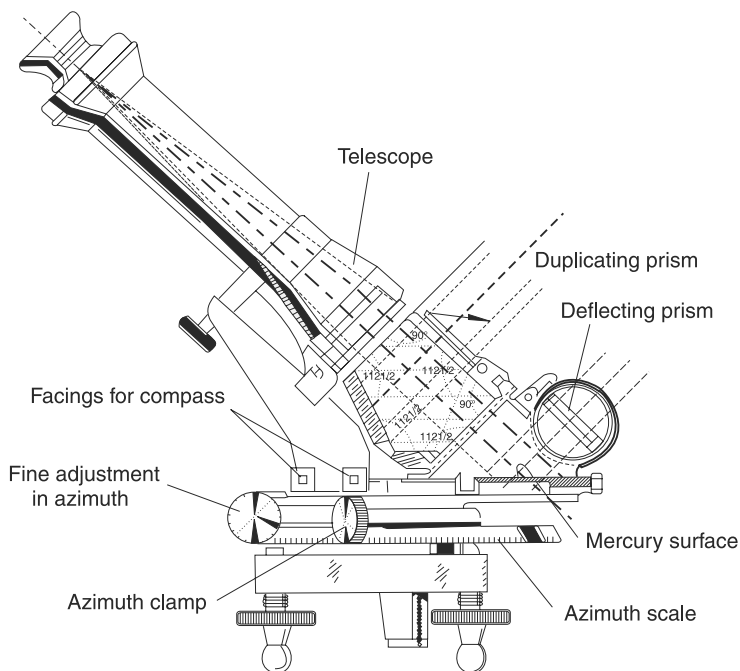


Fig. 10.59(a) *45° Prismatic Astrolabe (Side elevation)*

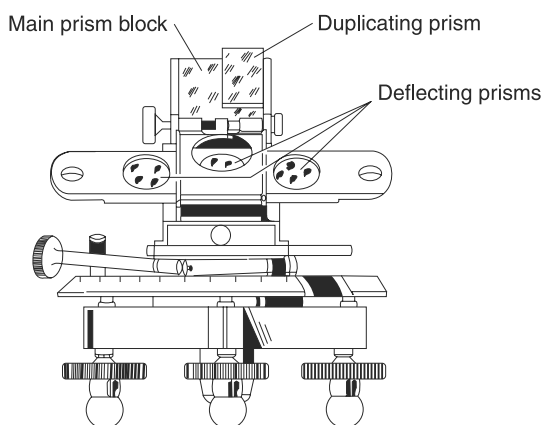


Fig. 10.59(b) *45° Prismatic Astrolabe (Front elevation Telescope removed)*

10.20 ZENITH TELESCOPE

A zenith telescope (Fig. 10.60) is an instrument designed to measure zenith distances. By the use of this, the latitudes can be worked out conveniently.

altitude levels are attached to a graduated circle, which is fixed to the telescope. The levels are used to measure accurately the small deviations in zenith distance and the circle reads zenith distance. The horizontal axis coincides with the east and west line and, thus, when the telescope is rotated about it, the line of sight describes the meridian. By reversing the entire instrument about the vertical axis, the meridian may be described on the other side of the vertical axis. Thus, two stars on south and north of the zenith of approximately same zenith distance, may be observed conveniently without resetting the telescope. The instrument is provided with a horizontal circle at the foot of the vertical axis and a striding level to rest on the horizontal axis. The eye end of the zenith telescope is fitted with a micrometer which is designed to measure difference in zenith distance of the stars. The two stars forming a pair, for the observation, are so chosen that their zenith distance is less than 30° .

Determining Latitude The following equation may be used for calculating the latitude, the observations being made with a zenith telescope:

$$\phi = \frac{1}{2}(\delta' + \delta'') + \frac{1}{2}(m' - m'') R + \frac{1}{2}(b' + b'') + \frac{1}{2}(r' - r'') \quad (10.30)$$

where δ' and δ'' are the declinations of the two stars,
 m' and m'' are the micrometer readings for the two stars,
 b' and b'' are the corrections for levels, positive when north reading is high, and
 r' and r'' are the refraction corrections.

Notes

- i. The “primes” refers to stars south and “seconds” refer to stars north.
- ii. The formula for reduction is based on the assumption that the micrometer head is inclined upwards, in which case, if the micrometer reading is greater, the zenith distance is also greater. If the reading were made with the micrometer head down, the sign of the term must be changed.

The level correction term $\left[\frac{1}{2}(b' + b'') \right]$ may be found as below.

Let n' and s' be the north and south readings of the level for the south star, n'' and s'' be the readings for the north star, d be the value of one division of the level, and x be the error of the level.

Case I If the graduations of the level are numbered in both directions from the middle, then

$$b' = \frac{1}{2}(n' - s') d + x$$

and
$$b'' = \frac{1}{2}(n'' - s'') d - x$$

Hence,
$$\frac{1}{2}(b' + b'') = \frac{1}{4}[(n' - s') - (n'' - s'')] d$$

Case II If the graduations of the level are numbered continuously from one end to the other, with the numbers increasing towards the eyepiece (when the telescope is inclined), then

$$\frac{1}{2}(b' + b'') = \frac{1}{4}[(n' + s') - (n'' + s'')] d$$

Case III If the graduations are numbered from one end, but with the numbers increasing towards the objective (when the telescope is inclined), then

$$\frac{1}{2}(b' + b'') = \frac{1}{4}[(n'' + s'') - (n' + s')] d$$

The refraction correction $[1/2 (r' - r'')]$ may be obtained as below.

$$\begin{aligned} r' - r'' &= 60.6'' (\tan z' - \tan z'') \\ &= 60.6'' \frac{\sin(z' - z'')}{\cos z' \cos z''} \\ &= 60.6'' \frac{\sin(z' - z'')}{\cos^2 z_0} \quad (z' \approx z'' \approx z_0) \\ \frac{1}{2}(r' - r'') &= 30.3'' \frac{\sin(z' - z'')}{\cos^2 z_0} \end{aligned}$$

Advantages The method of determining latitude by the zenith telescope yields the best results for the following reasons:

1. It measures differences of zenith distances instead of absolute zenith distances.
2. All errors of graduation of the circles are eliminated.
3. Since the pair of stars used have nearly the same meridian zenith distances and only the difference of these zenith distances is used in the computation, the refraction effect is very small.

Summary

Survey methods generally aim at determining the relative positions of points on the earth's surface with respect to each other. When required, the absolute positions of the earth surface can be obtained by astronomical observations. Astronomy from surveying perspective deals with determination of time, azimuth, latitude and longitude. The celestial bodies with which the surveyor is concerned are the so-called fixed stars (Polaris) and the sun.

Since a surveyor has to deal with only the angular positions of the stars, these are considered to be studded upon the surface of an imaginary sphere, called the celestial sphere; observer is stationed at the centre of this sphere. The astronomical determinations involve knowledge of spherical geometry and trigonometry; spherical triangle formed on the surface of celestial sphere, bounded by three arcs of the great circles forms the back-bone of the astronomical calculations. If three of the six parts of astronomical triangle are known, the methods of spherical trigonometry enable the triangle to be solved for the quantity under determination. The Napier's rule of circular parts is presented and rules are explained with the help of solved examples.

To designate the position of celestial body on the celestial sphere coordinate system is adopted similar to as a point on earth surface can be described by the coordinates—latitude

and longitude. Three systems of coordinates—altitude and azimuth (the reference plane is horizon), declination and right ascension, and declination and hour angle are in use. For the latter two coordinate systems the reference plane is equator.

The rotation of earth about its axis, and its complete revolution in its orbit results in a day and a year, respectively. The day may be defined as the interval between successive transits of a celestial body in the same direction across the meridian. Depending upon the celestial body the star, or the sun, or the moon a day is termed as sidereal, or solar, or lunar day respectively. The civil and astronomical time reckoning is described. Also, the conversion of times are presented and explained with the help of solved examples.

In the solution of most astronomical problems the true altitude of the centre of the celestial body is required. The observed altitude of the celestial body is not the true altitude because of the observational errors due to parallax, refraction, semi-diameter, and dip. The corrections required to be applied to the observed altitude are discussed and derived.

Determination of quantities such as time, azimuth, latitude and longitude is essential part of field astronomy. Each of these quantities can be determined by use of number of methods depending upon the instruments available and the degree of precision required. All the prevalent methods are described in detail. The chapter ends with the description of solar attachment, sextant, Prismatic astrolabe, zenith telescope.

Exercises

10.1 Define the following:

- (a) Altitude, (b) Latitude, (c) Right Ascension, (d) Declination, (e) Hour angle, (f) Azimuth

10.2 What do you understand by the following:

- (a) Aphelion, (b) Equinox, (c) Solstice, (d) Ecliptic, (e) Sidereal time

10.3 What are the various systems of coordinates employed to locate the position of a celestial body? Why is it necessary to have several systems instead of one?

10.4 What are the corrections to be applied to the observed altitude of sun?

10.5 What do you understand by sidereal day, apparent solar day, and mean solar day? State the relation between sidereal time, right ascension and hour angle.

10.6 Prove that the latitude of a place is equal to the altitude of the pole.

10.7 Write short notes on the following:

- (a) Circumpolar stars, (b) Star at culmination, (c) Star at elongation

10.8 Enumerate the principal methods for determining latitude in the field of astronomy and explain any one of them. State the advantages of the method preferred.

A sun observation was made for latitude at mid-day at a place in the northern hemisphere. Its corrected altitude was $61^{\circ}30'46''$ and the declination $21^{\circ}25'N$. Determine the latitude of the observer.

[Ans. $49^{\circ}52'14''$]

10.9 From the known fact that a solar year consists of 365.24 days, obtain relations for conversion of mean to sidereal and sidereal to mean time. Using these, deduce a rule to obtain L.S.T. of L.M.N. at a place, given G.S.T. of G.M.N.

10.10 What is equation of time? Show that it vanishes four times a year.

10.11 Write short notes on the following:

- (a) First point of aries (b) Equation of time
- (c) Apparent solar time (d) Solar system

10.12 Show with the aid of sketches, where necessary, the relationship between the following:

- (a) R.A. of star, H.A. of star at any instant, and sidereal interval at that instant.
- (b) L.M.T., L.A.T., and E.T.

- 10.13** Describe briefly the order of work in the field necessary for the correction of a clock. What other data is necessary for this correction and how is this obtained.
- 10.14** Deduce expressions for the following in terms of the latitude of the observation station and the declination of the star:
- The azimuth, altitude and hour angle of a star when at elongation.
 - The hour angle and altitude of a star when on the prime vertical.
 - The azimuth and altitude of a star, six sidereal hours after upper transit.
- For what reason would you make observations of elongation at prime vertical and in what way would you use the expressions you have deduced?
- 10.15** Calculate the azimuth of a star of declination $72^{\circ}17'21''\text{N}$ at eastern elongation in latitude $43^{\circ}53'53''\text{N}$.
[Ans. $24^{\circ}58'21''$]
- 10.16** In an extra-meridian observation of a star for determining azimuth of a reference mark, the following data was recorded:
Mean apparent altitude = $39^{\circ}41'26''$
Bearing from reference lineup (R.M.) = $10^{\circ}29'50''$ (right)
Declination of star = $30^{\circ}29'15''\text{N}$
Latitude = $65^{\circ}30'27''\text{N}$
Determine the azimuth.
[Ans. $246^{\circ}10'24''$]
- 10.17** Find the right ascension of a meridian at 4.30 p.m. given that the longitude of the place is $72^{\circ}48'46.8''\text{E}$ and G.S.T. of G.M.M. is $10^{\text{h}} 10^{\text{m}} 40.73^{\text{s}}$ on that day.
[Ans. $26^{\text{h}} 42^{\text{m}} 35.51^{\text{s}}$]
- 10.18** An observation of time was made on a star in altitude $52^{\circ}12'50''\text{N}$. The mean of the two observed altitudes being $28^{\circ}36'20''$. The average sidereal time was $0^{\text{h}} 15^{\text{m}} 28.4^{\text{s}}$ as given by a sidereal chronometer. Determine the chronometer error if the R.A. of star and declination were $4^{\text{h}} 32^{\text{m}} 31.1^{\text{s}}$ and $16^{\circ}23'30.5''$, respectively. The star was east of the meridian.
[Ans. slow by 12.52^{s}]
- 10.19** Observations were made on two stars simultaneously by a theodolite. The one having a declination of $84^{\circ}42'\text{N}$ at eastern elongation, had a clockwise horizontal angle from the survey line as $118^{\circ}20'$. The other star had a declination $72^{\circ}24'\text{N}$ at western elongation with a clockwise horizontal angle with the line as $94^{\circ}6'$. Determine the azimuth of the line.
[Ans. $12^{\circ}43'56''$]
- 10.20** A star was observed at western elongation at a station A in latitude $54^{\circ}30'\text{N}$ and longitude $52^{\circ}30'\text{W}$. The declination of the star was $62^{\circ}12'21''\text{N}$ and its right ascension $10^{\text{h}} 58^{\text{m}} 36^{\text{s}}$, the G.S.T. of G.M.N. being $4^{\text{h}} 38^{\text{m}} 32^{\text{s}}$. The mean observed horizontal angle between the referring object P and the star was $65^{\circ}18'42''$. Find (a) the altitude of the star at elongation, (b) the azimuth of the line AP, and (c) the local mean time of elongation.
[Ans. (a) $66^{\circ}58'7.13''$; (b) $241^{\circ}16'19.5''$; (c) $13^{\text{h}} 48^{\text{m}} 24.87^{\text{s}}$]
- 10.21** The meridian altitude of a star was observed to be $75^{\circ}18'25''$ on 15th October 1916, the observation being made with face left, the star lying between the zenith and the pole. The declination of the star on the given data was $58^{\circ}41'43''\text{N}$ and the index correction was $-5''$. Find the latitude of the place of observation.
[Ans. $44^{\circ}0'3''$]
- 10.22** An observation was made on a star lying west of the meridian at a place in latitude $40^{\circ}20'36''\text{N}$ to determine the azimuth of the survey line AB. The mean observed altitude was $42^{\circ}10'24''$ and the clockwise horizontal angle from AB to the star was

108°18'48". The declination of the star was 24°54'35"N. Find the azimuth of the survey line AB .

[Ans. 168°19'40.5"]

- 10.23** Find the L.S.T. of L.M.N. in the following cases:

G.S.T. of G.M.N.	10 ^h	0 ^h	0 ^h	0 ^h
Longitude	90°W	90°E	180°E	180°W

[Ans. 10^h 0^m 510.1^s, 23^h 0^m 59.139^s, 23^h 58^m 1.7^s, 1^m 58.3^s]

- 10.24** If G.S.T. of G.M.M. on a certain date is 1^h, what will be the L.S.T. of L.M.N. at a place in longitude 180°E, 180°W.

[Ans. 13^h, 13^h 3^m 56.6^s]

- 10.25** Compute the L.M.T. at which a star having a declination 58°25'20"N and right ascension 14^h 22^m 10^s will elongate in the evening of a day for which G.S.T. of G.M.N. is 4^h 05^m 25^s. The latitude of the place of observation is 51°35'30"N and the longitude is 15°30'45". Will the elongation be eastern or western?

[Ans. 7^h 38^m 8.77^s, eastern]

- 10.26** An observation was made on a star at a place in latitude 54°00'N and longitude 3°W, and a mean corrected altitude of 40°36' determined when the azimuth was 218°30'. If G.M.T. of the observation was 20^h 18^m 10^s, determine the right ascension of the star. G.S.T. of G.M.N. = 20^h 52^m 46^s.

[Ans. 15^h 07^m 27^s]

- 10.27** At a place of 39°N, the declination and hour angle of a star were 19° and 42°, respectively. Find the altitude and azimuth of the star.

[Ans. 48°40'23.2", 73°21'09"]

Objective-type Questions

- 10.1** The inclination of ecliptic with the equator is
 (a) 23°27' (b) 27°23' (c) 32°27' (d) 72°23'
- 10.2** The number of planets of the sun are
 (a) 6 (b) 8 (c) 9 (d) 11
- 10.3** The number of natural satellites the earth has is (are)
 (a) 1 (b) 2 (c) 4 (d) 6
- 10.4** In India, the standard meridian is at the following longitude from Greenwich:
 (a) 5^h 30^m E (b) 5^h 30^m W (c) 7^h 30^m E (d) 7^h 30^m W
- 10.5** One nautical mile is equal to
 (a) 1000 m (b) 1853.18 m (c) 2000 m (d) 2012.2 m
- 10.6** The distance equivalent to 1° of latitude is
 (a) 11 km (b) 111 km (c) 1111 km (d) 1 km
- 10.7** The first point of aries is on
 (a) 21st March (b) 21st June
 (c) 22nd September (d) 23rd December
- 10.8** If S is the sum of three angles of a spherical triangle, then the spherical excess equals
 (a) $S - 90^\circ$ (b) $S - 180^\circ$ (c) $S - 270^\circ$ (d) $S - 360^\circ$
- 10.9** The sun's declination as well as right ascension equal zero on
 (a) 21st March (b) 21st June
 (c) 22nd September (d) 22nd December
- 10.10** The Polaris remains below the horizon at
 (a) 10°N latitude (b) 50°N latitude
 (c) 5°S latitude (d) 10°S latitude

- 10.11** A tropical year is equal to
 (a) 365.2422 mean solar days (b) 366.2564 mean solar days
 (c) 366 mean solar days (d) 366 mean sidereal days
- 10.12** One solar day is less than a mean solar day by
 (a) $1^m 55.91^s$ (b) $2^m 55.91^s$ (c) $3^m 55.91^s$ (d) $4^m 55.91^s$
- 10.13** The L.M.T. at a place having a longitude 93°E when the standard time is 10^h and the standard meridian is 78°E is
 (a) 9^h (b) 10^h (c) 11^h (d) 25^h
- 10.14** The L.M.T. of a place at G.M.T. 21^h and with a longitude 60°W is
 (a) 5^h p.m. (b) 5^h a.m. (c) 1^h p.m. (d) 1^h a.m.
- 10.15** A sidereal day is divided into
 (a) 23^h (b) 24^h (c) 36^h (d) 12^h
- 10.16** How many circular parts are there in a right angled spherical triangle
 (a) 3 (b) 5 (c) 7 (d) 4
- 10.17** The sum of the three angles of a spherical triangle is
 (a) $= \pi$ (b) $< \pi$ (c) $> \pi$ (d) $> 3 \pi$
- 10.18** A star culminates in zenith when
 (a) $\delta < \theta$ (b) $\delta > \theta$ (c) $\delta \leq \theta$ (d) $\delta = \theta$
- 10.19** For a star to the west of meridian, the local sidereal time is given by
 (a) L.S.T. = R.A. + H.A. (b) L.S.T. = R.A. - H.A.
 (c) L.S.T. = R.A. $\pm$ H.A. (d) L.S.T. = R.A. + Declination
- 10.20** When the earth in its orbit is nearest to the sun, it is known as
 (a) prolate (b) oblate (c) perihelion (d) aphelion
- 10.21** Day and night on earth are of equal duration at the instant of
 (a) vernal equinox (b) autumnal equinox
 (c) solstices (d) Both (a) and (b)
- 10.22** The sun at rise and set has an H.A. of less than 90° at the instant known as
 (a) vernal equinox (b) autumnal equinox
 (c) summer solstice (d) winter solstice
- 10.23** A great circle passing through celestial poles and the vernal equinox is
 (a) hour circle (b) declination circle
 (c) equinoctial colour (d) parallel of latitude
- 10.24** A star having a polar distance less than the latitude of the place of observation is known as
 (a) star at elongation (b) star at culmination
 (c) star at prime vertical (d) circumpolar star
- 10.25** The declination and R.A. of stars
 (a) suffer from nutation (b) suffer from precession
 (c) Both (a) and (b) (d) are invariable
- 10.26** For field observations, the system of coordinates used is
 (a) azimuth and altitude
 (b) declination and hour angle
 (c) declination and right ascension
 (d) azimuth and hour angle
- 10.27** The star for which the zenith distance is 90° will be at
 (a) equator (b) horizon
 (c) elongation (d) prime vertical
- 10.28** For a star at its upper transit, the local sidereal time is equal to
 (a) H.A. of the star (b) declination of the star
 (c) R.A. of the star (d) azimuth of the star

- 10.29** For a circumpolar star, declination must be
(a) = co-latitude (b) > co-latitude
(c) < co-latitude (d) ≤ co-latitude
- 10.30** In the altitude and azimuth system of coordinates, the reference plane is
(a) equator (b) horizon (c) ecliptic (d) prime vertical
- 10.31** The angle between the observer's meridian and the declination circle of a heavenly body is known as
(a) azimuth (b) hour angle
(c) right ascension (d) declination
- 10.32** Right ascension of a heavenly body is its equatorial angular distance measured
(a) westwards from the first point of libra
(b) eastwards from the first point of aries
(c) westwards from the first point of aries
(d) eastwards from the first point of libra
- 10.33** A latitude of the pole is always equal to the
(a) latitude of the observer's position
(b) longitude of the observer's position
(c) hour angle of the heavenly body
(d) declination of the heavenly body
- 10.34** The latitude of the observer's position is equal to
(a) altitude of the celestial body at upper culmination
(b) altitude of the celestial body at lower culmination
(c) mean of the altitude of the celestial body at upper and lower transits
(d) none of the above.
- 10.35** A star is said to be at elongation when its distance east or west of the observer's meridian is
(a) least (b) greatest
(c) zero (d) same
- 10.36** The circle in which the plane tangent to the earth's surface at the point of observation intersects the celestial sphere, is known as
(a) visible horizon (b) sensible horizon
(c) true horizon (d) celestial horizon
- 10.37** The great circle that the sun appears to trace on the celestial sphere with the earth as its centre, during the year, is called
(a) equator (b) equinox
(c) ecliptic (d) prime vertical
- 10.38** Which of the following is correct?
(a) One degree of longitude has the greatest value at the equator
(b) One degree of longitude has the greatest value at the poles
(c) One degree of longitude has the same value at the equator and the poles
(d) None of the above
- 10.39** A nautical mile is equal to
(a) one minute arc of great circle joining two points
(b) one minute arc of the longitude
(c) 1853.18 m
(d) Both (b) and (c)
- 10.40** In a spherical triangle, which of the following statement is not true?
(a) Every angle is less than two right angles
(b) Sum of three angles is equal to two right angles

- (c) If the sum of any two sides is π , then the sum of the angles opposite to them is also π
(d) Sum of any two sides is greater than the third.
- 10.41** According to Napier's rule of circular parts for a right angled triangle, sine at the middle part equals the product of
(a) tangent of two adjacent parts (b) sine of two adjacent parts
(c) cosine of two opposite parts (d) both (a) and (c)
- 10.42** Longitude of a place is the angular distance between the meridian of the place and the
(a) standard meridian (b) Greenwich meridian
(c) Both (a) and (b) (d) None of these
- 10.43** The altitude of a star is zero, therefore the star is at
(a) culmination (b) horizon (c) meridian (d) ecliptic
- 10.44** The point at which the sun's declination changes from north to south is known as
(a) first point of aries (b) first point of libra
(c) vernal equinox (d) Both (b) and (c)
- 10.45** Summer solstice is the position of the sun where its
(a) north declination is maximum
(b) south declination is maximum
(c) declination is zero
(d) None of the above
- 10.46** Which one of the following represents a circumpolar star?
(a) Upper culmination above horizon, lower culmination below horizon
(b) Both upper and lower culminations below horizon
(c) Both upper and lower culminations above horizon
(d) Altitude at upper culmination is minimum
- 10.47** G.M.M. stands for
(a) Global midmorning (b) Greenwich mean midnight
(c) Greenwich mean midmorning (d) Global midnight
- 10.48** For a celestial body, in astronomy, which is a function of time?
(a) Azimuth (b) Altitude
(c) Hour angle (d) All of these
- 10.49** When the time measurement is based upon the diurnal motion of the first point of aries, it is known as
(a) apparent solar time (b) mean solar time
(c) sidereal time (d) equation of time
- 10.50** The time to which the watches and clocks are regulated is determined by
(a) Polaris (b) mean sun (c) sun dial (d) All of these
- 10.51** The watches and clocks maintain
(a) local sidereal time (b) mean solar time
(c) apparent solar time (d) standard time
- 10.52** At the first point at aries, the sun travels
(a) northwards (b) southwards
(c) south to north of the equator (d) north to south of the equator
- 10.53** International date line is located along
(a) Greenwich (b) equator
(c) 180° longitude (d) standard meridian
- 10.54** Equation of time, which is the difference between apparent and mean solar time at any instant, vanishes
(a) once in a year (b) twice in a year
(c) thrice in a year (d) four times in a year

- 10.55** Difference in longitude of two places expressed in terms of time is equal to the difference in their
 (a) sidereal time (b) apparent solar time
 (c) mean solar time (d) All of these
- 10.56** The most accurate method to locate the true meridian is by making observations of
 (a) Polaris at culmination (b) Polaris at elongation
 (c) sun (d) Polaris at prime vertical
- 10.57** Correction for parallax is not applied to which of the following celestial bodies?
 (a) star (b) moon (c) sun (d) earth
- 10.58** If the declination of a star is 30°N , its polar distance is
 (a) 30° (b) 60° (c) 120° (d) 150°
- 10.59** The longitude of two places at latitude 60°N are 93°E and 97°W , respectively. Their departure is
 (a) 5100 nautical miles (b) 3700 nautical miles
 (c) 1200 nautical miles (d) 1500 nautical miles
- 10.60** Match List-I (Term) with List-II (Definition) and select the correct answer using the code given in the lists:

List-I (Term)

- (i) Apparent solar day
- (ii) Sidereal day
- (iii) Tropical year
- (iv) Sidereal year

List-II (Definition)

1. The time interval between two successive upper transits of the first points of Aries over the same meridian
2. The interval between two successive lower transits of the centre of the Sun across the same
3. Time interval between two successive passages of the Sun over the meridian of any one of the fixed stars
4. Time interval between two successive vernal equinoxes

Codes:

(i)	(ii)	(iii)	(iv)
(a) 1.	2.	3.	4.
(b) 2.	1.	4.	3.
(c) 2.	1.	3.	4.
(d) 1.	2.	4.	3.

Answers to Objective-type Questions

10.1 (a)	10.2 (c)	10.3 (a)	10.4 (a)	10.5 (b)	10.6 (b)
10.7 (a)	10.8 (b)	10.9 (a)	10.10 (c)	10.11 (a)	10.12 (c)
10.13 (c)	10.14 (d)	10.15 (b)	10.16 (b)	10.17 (c)	10.18 (d)
10.19 (a)	10.20 (c)	10.21 (d)	10.22 (d)	10.23 (c)	10.24 (b)
10.25 (c)	10.26 (a)	10.27 (b)	10.28 (b)	10.29 (b)	10.30 (b)
10.31 (b)	10.32 (b)	10.33 (a)	10.34 (c)	10.35 (b)	10.36 (b)
10.37 (c)	10.38 (a)	10.39 (c)	10.40 (b)	10.41 (d)	10.42 (c)
10.43 (d)	10.44 (b)	10.45 (a)	10.46 (c)	10.47 (b)	10.48 (a)
10.49 (c)	10.50 (c)	10.51 (d)	10.52 (a)	10.53 (c)	10.54 (d)
10.55 (d)	10.56 (b)	10.57 (a)	10.58 (b)	10.59 (a)	10.60 (b)

CONVERSION TABLES

1 mile	= 1.6093 km
	= 8 furlongs
	= 1760 yards
	= 5280 ft
	= 63360 in.
1 furlong	= 0.2012 km
	= 220 yards
1 yard	= 0.9144 m
1 ft	= 0.3048 m
1 in.	= 25.4000 mm
1 km	= 0.6214 miles
	= 1093.6106 yards
1 m	= 1.0936 yards
	= 3.2808 ft
	= 39.36998 in.
1 mm	= 0.0394 in.
1 nautical mile	= 6080 ft
	= 1853.1886 m
1 cable	= 100 fathoms
1 fathom	= 6 ft
	= 1.8288 m
1 bar	= 10^5 N/m^2
1 knot	= 1 nautical mile
1° latitude	= 111.21 km
1° longitude	= 111.21 km (at equator)
	= 0 (at poles)
1 h	= 15°
1 min	= 15'
1 s	= 15'

DIGITAL TERRAIN MODEL

The Digital Terrain Model (DTM) also referred to as *digital elevation model* (DEM) or a *digital terrain elevation model* (DTEM) is a discrete representation of a topographic surface (a surface representing height data). A DTM is the simplest digital representation of a portion of the earth's surface. The surface of the earth is a continuous phenomenon. To model such a surface or its part accurately, storage of an infinite number of points (observation) would be necessary. Since this is impossible, a surface model approximates a continuous surface using a finite number of observations. Further, since the resolution of a DTM will depend upon the frequency of observations used, an appropriate number of observations to approximate the continuous surface will be required.

A DTM surface is created from a series of either regularly or irregularly spaced (x, y, z) data points, where x, y are the horizontal coordinates and z is the vertical or height coordinate. DTMs may be derived from number of sources such as topographic maps (contour and spot height information), stereoscopic aerial photography, satellite images and field surveys. These models can also be used to store other kind of continuous phenomenon such as isarithmic maps used for isobars (barometric pressures), isotherms (temperature), or isohyets (precipitation). Some of the other examples are the surfaces derived from point data for social, economic or demographic characteristics, like distribution of average income, education levels, etc.

DTM's can be constructed using either raster-based or vector-based approach; the former being more common. In raster-based approach, a DTM is simply grid of height values in which each cell contains a single value representative of the height of the terrain that is covered by the cell. The complexity of the terrain surface and the resolution of the grid dictate the accuracy of the model. The lines are drawn to connect the centre points of the ground array, which by virtue of their varying z values appear as undulating terrain (Fig. AII.1). A disadvantage of this method is the wastage of computer memory in storing highly redundant information, for example, uniformly sloping terrain. In the vector GIS, the terrain surface is represented by a regularly spaced set of spot heights (Fig. AII.2). Another advanced and better form of vector DTM is the triangulated irregular network (TIN) as shown in Fig. AII.3. The basic entity, the triangle can be stored as set of segments or as a set of vertices, the latter being more common. In this model,

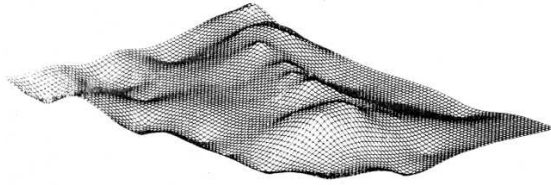


Fig. AII.1 *Three dimensional representation of DEM*

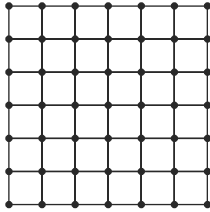


Fig. AII.2 *Grid model*

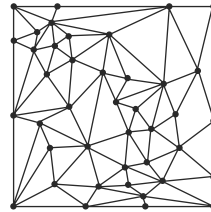


Fig. AII.3 *TIN model*

the height observations are joined together with straight lines to create a mosaic of irregular triangles. The lines are drawn between points in closest proximity to each other without any lines intersecting. This method is known as *Delaunay Triangulation*. In the TIN model of the surface, the vertices of the triangles stored represent terrain features such as peaks, depressions, while the edges represent ridges and valleys. The surfaces of individual triangles provide area, gradient (slope) and orientation (aspect). These values can be stored as attribute data. The TIN model is more efficient representation of the terrain in terms of data storage as spot elevation can be acquired at critical points.

The input data models and algorithms required by DTMs are quite different from those used in representing planimetric or two dimensional data. DTM encompasses generation sampling of original terrain data and formation of relations among the diverse observations; manipulation, modification and refinement of DTMs and derivation of intermediate model; interpretation, analysis and extraction of information derived from DTMs; and application development of appropriate application model for specific discipline. For details specific literature may be referred.

PHYSICAL LAWS GOVERNING EMR

Stefan–Boltzmann Law

The sun is the primary source of electromagnetic radiation for remote sensing. In fact, all the terrestrial objects above absolute zero (0 K) are also sources of electromagnetic radiations, but with considerably different magnitudes and spectral compositions than that of the sun. According to the Stefan–Boltzmann law, the amount of energy radiated by any object is a function of the surface temperature of the object and is given by

$$M = \sigma T^4 \quad (\text{AIII.1})$$

where

M = total radiant exitance from the surface of a material (W/m^2)

σ = Stefan–Boltzmann constant ($5.6697 \times 10^{-8} \text{W}/\text{m}^2/\text{K}^4$)

T = absolute temperature (K) of the emitting material

Thus, the total energy emitted from an object varies as T^4 . It should be noted that this law is expressed for an energy source that behaves as a *black body*; a hypothetical, ideal radiator that totally absorbs and radiates all energy incident upon it. Actual objects rarely approach this ideal body.

The spectral distribution of the emitted energy also varies with temperature. Figure 1 shows energy distribution curves for black bodies at temperatures ranging from 200 to 6000 K. The area under these curves equals the total radiant exitance, M . These curves graphically illustrate the Stefan–Boltzmann law. These also show that there is a shift toward shorter wavelengths in the peak of a black body radiation distribution as the temperature increases.

Wien's Displacement Law

The dominant wavelength, or the wavelength at which a black body radiation curve reaches a maximum, is related to its temperature by Wien's displacement law,

$$\lambda_m = \frac{A}{T} \quad (\text{AIII.2})$$

where

λ_m = wavelength of maximum spectral radiant exitance (μm)

$A = 2898 \mu\text{m K}$

T = temperature (K)

Thus, for a black body, the wavelength at which the maximum spectral radiant exitance occurs, varies inversely with the absolute temperature of the black body.

The temperature of the earth's surface material (such as soil, water, etc.) is about 300 K. From Wien's displacement law, this means that the maximum spectral radiant exitance from the earth features occurs at a wavelength of about $9.7 \mu\text{m}$. This radiant energy peaks in the thermal band of the infrared region. This energy can neither be seen nor photographed, but can be sensed with radiometers and scanners. The sun emits radiation in the same manner as a black body radiator whose temperature is 6000 K and has a much higher energy peak that occurs at about $0.5 \mu\text{m}$ (Fig. AIII.1). Our eyes and photographic films are sensitive to energy of this magnitude and wavelength. Thus in sunlight, the earth features can be observed by reflected solar energy.

Planck's Law This law enables us to assess the proportion of total radiant exitance within a selected wavelength. The total energy radiated in all directions by unit area in unit time in a spectral band is given by

$$M_{\lambda} = \frac{C_1}{\lambda^{5[\exp(C_2/\lambda T)]}} \quad (\text{AIII.3})$$

where

M_{λ} = spectral exitance per unit wavelength

C_1 = first radiation constant = $3.742 \times 10^{-16} \text{ W/m}^2$

C_2 = second radiation constant = $1.4388 \times 10^{-2} \text{ mK}$

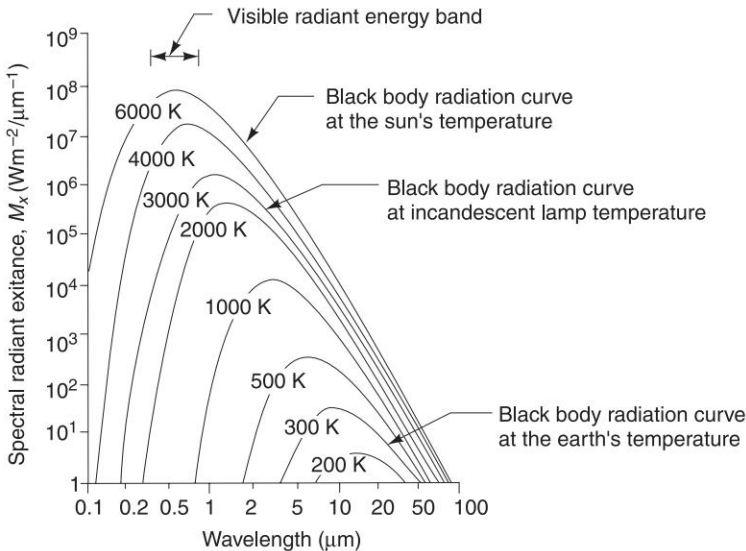


Fig. AIII.1

Spectral distribution of energy radiated from black bodies of various temperatures

IV

TOPOLOGY

The data in a database are conventionally connected by spatial relationships (usually coordinates). Further, as the attribute data is unique to each point in a database, it is also included. However, in addition to the spatial connection, the data can also be connected in a relational sense. Topology in GIS is used to describe the object's¹ geometric characteristics that do not change under transformations and are independent of any coordinate system and scale of measurement. It is used to describe as to how spatial objects are related to one another in relational sense. The topological relationships are normally listed in tables and stored within a database of GIS.

In GISs, the most important topological relationships are the connectivity, adjacency, containment and metric aspects. *Connectivity* identifies which chains are connected at which nodes. It is a geometric property used to describe the linkages between the features. An example is a network of roads wherein roads are linked together. *Adjacency* indicates which polygons have common boundaries. *Containment* describes simple spatial objects wholly contained within another spatial object such as an island within a lake. As another example, districts are contained in states and the states in a Nation. Adjacency and containment describe the geometric relationship which exists between area features. Containment is an expansion of the adjacency theme and describes area features. *Metric aspects*, for example are size, shape, distance and direction. Here, nodes, chains and polygons are additional simple spatial objects commonly used for specifying the topological relationships of information entered into the GIS database. *Nodes* are the beginnings and endings of *chains*. A chain is similar to lines and strings used to delineate specific boundaries; *polygons* are similar to areas defined to be a series of connected chains.

An understanding of the geometric relationship between spatial entities is important for analysis and integration in GIS. Without knowledge of these, it will not be possible to answer most of the queries.

¹ Entities are the things in the real world while objects are in the digital world.

PSEUDO-RANDOM CODES

The GPS system does not use numbers. However, both the satellites and the receivers actually generate a set of digital codes which are very complicated. The codes are made complicated purposely so that they can be compared easily and unambiguously, and for some other technical reasons also. The codes almost look like a long string of random pulses (Fig. AV.1). They are not really random though, they are carefully chosen pseudo-random sequences that actually repeat every millisecond. So they are often referred to as pseudo-random codes. The GPS system uses pseudo-random codes for the following reasons:

1. A pseudo-random code gives us a way of very clearly recognising a very faint signal. It means the GPS satellites do not have to be very powerful (so they cost less) and that receivers on the ground can get away with using very small antennas.
2. It provides a way to control access to the satellite system. In the time of war, the code can be changed and so the system can be prevented from being used by the enemy. Even in peace time, some exclusivity to the system can be retained.
3. Another benefit of the pseudo-random code scheme is that all the satellites in the system can share the same frequency without interfering with each other. Each satellite has its own distinct pseudo-random code, so differentiating between them is only a matter of using the right code during the comparison process at the receiver. Since all transmissions are of low power, no satellite overpowers any other.

The pseudo range and Delta-range measurements are made using the two L-band frequencies; L_1 at 1575.42 MHz and L_2 at 1227.6 MHz. The L_1 signal is modulated by both C/A and P codes and L_2 is generally modulated by P-code. The satellite also broadcasts navigation information modulated on L_1 signal. These signals are modulated by pseudo-random binary codes (PN).

Each satellite has a specific code, thus when the GPS receiver reads the code, it knows that the signal is from which satellite, though all the satellites transmit in the same frequency. The PN code is also fundamental to determine the time taken by the signal to reach the receiver, and hence the distance. There are two separate forms of pseudo-random binary code, one called the *C/A code* and the other called the *P-code*. The C/A code is the one that all the civilian receivers use. It has lower

frequency than the P-code, and so is thought by some users to be less precise. The P-code can be encrypted so only military users are guaranteed access to it. In addition, the P-code is almost impossible to jam. The current accuracy of the C/A code may even be degraded by using Selective Availability (S/A). Selective Availability is essentially a method for artificially creating a significant clock error in the satellites.

1. *Coarse Acquisition Code (C/A Code)*: This consists of 1023 bits, which repeats itself every millisecond. If the timing is converted into distance, the code length is ~ 300 km with each bit ~ 300 m. The positioning service using C/A code ranging data is referred to as the *Standard Positioning Service (SPS)*.

2. *Precise Code (P-Code)*: It is a long binary code, which repeats only once a week. The bit is equivalent to 30 m of length and the code length is 181.44×10^7 km. In order to guard against fake transmission of this code, it is encrypted and the resulting code is termed as the *Y-code* (the process of encryption is called Anti-spoofing (AS)). The positioning service based on P-code ranging data is known as *Precise Position Service (PPS)*.

GPS Error Contribution to the single range observation

<i>Source</i>	<i>P-Code</i>	<i>C/A Code</i>
<i>Satellite</i>		
Orbit	5 metres	5 metres
Clock	1 metre	1 metre
<i>Signal propagation</i>		
Ionosphere	cm to dm	
Troposphere	dm	
Multipath	1 metre	5 metres
<i>Receiver</i>		
Observation noise	10 cm to 1 metre	1 to 10 metres
Hardware delay	dm to m	m
Antenna phase centre	mm to cm	mm to cm



Fig. AV.1 *Pseudo-random code*

ORBITS OF
PLANETS

<i>Name</i>	<i>Semi-major axis ($\times 10^6$ km units)</i>	<i>Mean distance from the sun ($\times 10^6$ km)</i>	<i>Sidereal period in tropical years</i>	<i>Eccen- tricity</i>	<i>Inclination to the ecliptic</i>	<i>Synodic period in days</i>
Mercury	69.7	57.9	0.24085	0.20562	7°0'14"	115.88
Venus	109	108.2	0.61521	0.000680	3°23'39"	583.92
Earth	152.1	149.6	1.00004	0.001673	0°0'0"	...
Mars	249.1	227.9	1.88089	0.09336	1°51'0"	779.94
Jupiter	815.7	778.3	11.86223	0.04842	1°18'21"	398.88
Saturn	1507	1427.0	29.4577	0.05562	2°29'25"	378.09
Uranus	3004	2869.6	84.0139	0.04718	0°46'23"	369.66
Neptune	4537	4496.6	164.793	0.00857	1°46'28"	367.49
Pluto	7375	5900.0	247.7	0.24864	17°8'38"	366.73

Appendix VII

PHYSICAL CHARACTERISTICS OF THE SUN, MOON AND THE PLANETS

<i>Name</i>		<i>Diameter (Earth = 1)</i>	<i>Mass (Earth=1)</i>	<i>Sidereal Rotational period</i>	<i>Density (Water=1)</i>	<i>Sym- bol</i>
Sun		109.18	333 434	25 ^d 3 ^h 22 ^m	1.43	
Moon		0.272	0.0123	27 ^d 7 ^h 43 ^m 11.5 ^s	3.33	
Mercury		0.403	0.037	58.65 days	2.86	♿
Venus		0.990	0.826	243.01 days	4.86	♀
Earth	Equatorial	1.000	1.000	23 ^h 56 ^m 4.09 ^s	5.52	⊕
	Polar	0.997				
Mars	Equatorial	0.538	0.108	24 ^h 37 ^m 22.65 ^s	3.84	♂
	Polar	0.535				
Jupiter	Equatorial	11.27	318.4	9 ^h 50 ^m 30 ^s	1.30	♃
	Polar	10.58		9 ^h 55 ^m 41 ^s		
Saturn	Equatorial	9.47	95.2	10 ^h 2 ^m	0.69	♄
	Polar	8.55				
Uranus		4.19	14.6	10 ^h 48 ^m	1.10	♅
Neptune		3.90	17.3	15 ^h 48 ^m	1.62	♆
Pluto		0.50	1			♇

Appendix VIII

PHYSICAL CHARACTERISTICS OF THE SATELLITES OF THE PLANETS

<i>The satellites of the planets</i>	<i>Diameter (km)</i>	<i>Mass (kg)</i>	<i>Density</i>
Earth			
Moon	3476	7.366×10^{22}	3.33
Mars			
Phobos	$27 \times 21 \times 19$	9.6×10^{15}	≈ 2
Deimos	$15 \times 12 \times 11$	2.0×10^{15}	≈ 2
Jupiter			
Metis	(40)	?	?
Adrastea	(30)	?	?
Amalthea	$270 \times 170 \times 150$	?	
Thebe	(70)	?	?
Io	3632	8.916×10^{22}	3.55
Europa	3126	4.873×10^{22}	3.04
Ganymede	5276	1.490×10^{23}	1.93
Callisto	4820	1.064×10^{23}	1.79
Leda	(10)	?	?
Himalia	(170)	?	?
Lysithea	(20)	?	?
Elara	(80)	?	?
Ananke	(20)	?	?
Carme	(20)	?	?
Pasiphae	(30)	?	?
Sinope	(20)	?	?

(Contd.)

(Contd.)

Saturn			
Atlas	$40 \times 20 \times ?$	?	?
1980 S 27	$140 \times 100 \times 80$	?	?
1980 S 26	$110 \times 90 \times 70$	?	?
Epimetheus	$140 \times 120 \times 100$	?	?
Janus	$220 \times 200 \times 160$	?	?
Mimas	390	(3.76×10^{19})	1.2
Enceladus	510	(7.40×10^{19})	1.1
Tethys	1050	6.26×10^{20}	1.0
1980 S 25	$34 \times 22 \times 22$	?	?
1980 S 13	$34 \times 28 \times 26$	?	?
Dione	1120	1.05×10^{21}	1.4
1980 S 6	$36 \times 32 \times 30$	?	?
Rhea	1530	2.28×10^{21}	1.3
Titan	5150	1.359×10^{23}	1.9
Hyperion	$410 \times 260 \times 220$	(1.10×10^{20})	?
Lapetus	1460	(1.93×10^{21})	1.2
Phoebe	200	?	?
Uranus			
Miranda	(300)	(1.7×10^{19})	(1.3)
Ariel	(1300)	(1.6×10^{21})	(1.3)
Umbriel	(1100)	(9.3×10^{20})	(1.3)
Titania	(1600)	(2.7×10^{21})	(1.3)
Oberon	(1600)	(2.9×10^{21})	(1.3)
Neptune			
Triton	(3800)	5.7×10^{22}	(2.0)
Nereid	(940)	1.3×10^{15}	(2.6)
Pluto			
Charon	≥ 1200		

ASTRONOMICAL CONSTANTS

(For 1900, unless otherwise stated)

$$\begin{aligned}\text{Sidereal day} &= 1 - \frac{1}{366.25} \text{ mean solar day} \\ &= 23^h 36^m 4.1^s \text{ (mean solar units)}\end{aligned}$$

$$\begin{aligned}\text{Mean solar day} &= 1 + \frac{1}{365.25} \text{ sidereal day} \\ &= 24^h 3^m 56.5^s \text{ (sidereal units)}\end{aligned}$$

Year

$$\begin{aligned}\text{tropical} &= 365^d 5^h 48^m 45.98^s \text{ (mean solar units)} \\ \text{sidereal} &= 365^d 6^h 9^m 8.97^s \text{ (mean solar units)} \\ \text{anomalistic} &= 365^d 6^h 13^m 48.09^s \text{ (mean solar units)}\end{aligned}$$

Lunar month

$$\begin{aligned}\text{sidereal} &= 27.32166 \text{ days} = 27^d 7^h 43^m 11.5^s \\ \text{synodic} &= 29.53059 \text{ days} = 29^d 12^h 44^m 2.8^s\end{aligned}$$

period of rotation of moon's nodes

$$\begin{aligned}&= 18.60 \text{ years} \\ \text{saros} &= 6585.32 \text{ days} \\ &= 18.03 \text{ years} = 18 \text{ years } 10 \text{ or } 11 \text{ days}\end{aligned}$$

Equation of time

$$\text{maximum due to eccentricity} = 7.6 \text{ min}$$

$$\text{maximum due to obliquity} = 9.9 \text{ min}$$

Earth's

$$\text{equatorial radius} = 6378 \text{ km}$$

$$\text{polar radius} = 6357 \text{ km}$$

$$\text{compression} = \frac{1}{300}$$

$$\text{eccentricity} = 0.0820$$

$$\text{mass} = 5.98 \times 10^{27} \text{ g}$$

$$\text{density} = 5.52 \text{ g/cm}^3$$

$$\text{obliquity of ecliptic (1950)} = 23^\circ 26' 45''$$

$$\text{annual precession} = 50.26'$$

$$\text{period of precession} = 25695 \text{ years}$$

$$\text{constant of nutation} = 9.23'$$

constant of aberration	= 20.47'
velocity of light	= 300000 km/s
1 parsec	= 3.26 light years
1 light year	= 5.88×10^{12} miles

Sun's

mean parallax	= 8.790'
semi-diameter (angular)	= 16'1"
distance from earth	= 1.496×10^8 km
	= 1 AU $\approx$ 500 light second
diameter	= 14×10^5 km
density	= 1.4 g/cm ³ (taking water as 1)
mass	= 331100 times earth = 2×10^{30} kg
period of axial rotation	= $25^d 5^h 37^m$
age	= 4.5 billion years

Moon's

mean equatorial horizontal parallax	= 57'2.63"
semi-diameter (angular)	= 15'34"
distance	= 60.27 times earth's radius = 384000 km
diameter	= 3475 km
density	= 3.4 (taking water as 1)
mass	= $\frac{1}{81.5}$ times earth
eccentricity of orbit	= 1/18
inclination to ecliptic orbit	= 5°8'40"

Ecliptic limits

lunar	= 110°38' and 9°39'
solar	= 17°25' and 15°23'
1 radian	= $180/\pi$ degrees = 206265"

INDEX

A

- Absolute parallax 201
- Absolute positioning 366
- Acceleration 414
- Accidental error 82
- Active system–remote sensing 239
- Adaptive filter 289
- Adjustment in triangulation 105
- Aerial
 - photograph 169, 170
 - photography 164
 - surveying 164
- Aerial photogrammetry 164, 178
 - displacements in 178
 - errors in 178
- Aerial survey
 - procedure of 186
- Air base 190
- Air photo-interpretation keys 197
- Air station 170
- Altitude 386
 - measurement of 449
- Altitude and azimuth system 390, 391
- Angle adjustment 105
- Angle equations 112
- Antenna 350
- Aphelion 379
- Apses of the orbit 379
- Areal data 308
- Arundel's method 210
- Asterism 382
- Astronomical corrections 423
- Astronomical triangle 397
- Astronomy 378
- Atmospheric corrections 272
- Atmospheric scattering 232

- Atmospheric windows 230
- Atomic time scale 358
- Attendants 382
- Attribute data 307
- Autumnal equinox 382
- Averaging filter 288
- Axis signal correction 3
- Azimuth 386
 - determination of 431

B

- Band ratioing 290
- Base
 - broken 63
 - extension 62
 - interpolating 64
 - measurement problems 62
- Base line
 - measurement of 44
 - site selection 44
- Base measurement
 - by flexible apparatus 46
 - by rigid apparatus 44
 - problems 62
- Beacon lamps 35
- Bias 327
- Biases 325
- Buffering 320
- Buffer operation 319

C

- C/A code 349
- Cairn 33
- Captain G.T McCaw's solution 40
- Captured data 306
- Carrier phase measurement technique 358

- Cartographic
 - generalization 311, 328
 - symbolization 311
- Celestial sphere 380
- Cellular data structure 311
- Central system 30
- Check base 24
- Choice of triangulation figure 24
- Circumpolar star 389
- Classification of triangulation 31
- Co-altitude 386
- Co-declination 387
- Co-latitude 387
- Coefficient of refraction 13
- Colby apparatus 45
- Collimation marks 212
- Contrast enhancement 282
- Colour composites 245
 - compositing 281
- Composite analysis 323
- Conjugate principle points 200
- Constellation 382
- Contact apparatus 44
- Control stations 186
- Convolution filtering 287
- Corrections for
 - axis signal 3
 - change of gravity 50
 - cubic convolution method 280
 - curvature 2
 - inclination of end readings 51
 - index error 50, 448
 - parallax 423
 - pull 49
 - refraction 2, 424
 - semi-diameter 425
- Correlates, method of 103
- Cossiopea
- Crab 188
- Culmination 389
- Cunningham equation 150
- Current meter 151
- D**
- Danger circle 149
- Data
 - acquisition 242
 - areal 307
 - attribute 307
 - encoded 306
 - graphical 306
 - interpretation 244
 - line 308
 - models 309
 - point 307
 - reference 244
 - spatial 306
 - string 308
- Data format conversions 315
- Data structure 309
 - raster 311
 - vector 309
- DBMS 304
- Declination 387
- Declination and hour angle system 390, 392
- Declination and right ascension system 390, 391
- Declination circle 386
- Delta-range 349
- De-stripping 272
- Density slicing 286
- Detector response calibration 270
- Differential corrections 363
- Differential GPS 363
 - inverted 365
 - principle of 363
 - working with 364
- Differential parallax 200
- Digital
 - data 303
 - elevation model 215, 239
 - image 243
 - image processing 268, 289
 - terrain model 307
- Digitiser 305
- Digitising table 305
- Dilution of precision 373
- Direction theodolite 64
- Discrepancy 81
- Displacement
 - due to ground relief 178
 - due to tilt 183
- Distance angles 25
- Doppler effect 147, 345
- Doppler shift 147
- Doughnut buffer 320
- Drift 188
- Drift angle 188
- Dynamic survey technique 367

E

Earth 379
 Earth sun distance correction 273
 Ecliptic 381
 Eccentric stations 65
 Echo sounding 140
 Edge detection filter 289
 EDM 21
 Electromagnetic
 energy 226, 228
 radiation 227
 spectrum 229
 spectral regions 230
 waves 228
 Elongation 389
 Encoded data 306
 Energy-balance equation 235
 Ephemerides 351
 Ephemeris 349
 Ephemeris data 347
 Equation of time 412
 Equator 379
 Equatorial system 412
 Equinoctial colour 385
 Equipment segment 350
 Error
 accidental 82
 bit 271
 cumulative 82
 distribution 99
 mean square 83
 phase 36
 probable 85
 random 86
 residual 82
 standard 85
 systematic 82
 true 81
 Exposure
 interval 190
 station 170
 Eye and object correction 3

F

False colour composites 245
 False stations 66
 Fathom 135
 Fathometer 140
 Feature identifier 310
 Fiducial marks 212

Field astronomy 378
 Field checks in triangulation 70
 Figure adjustment 108
 Filtering 287, 320
 First point of aries 382
 First point of libra 382
 Flare triangulation 70
 Flight
 planning 186
 strips 165, 187
 Floating marks 206
 Floating triangulation 71
 Floats
 double 149
 rod 150
 submerged 149
 subsurface 149
 surface 149
 Fore and aft overlap 187
 Forward overlap 187
 Framing system 237
 Frequency domain filtering 289
 Fiducial marks 166

G

Galileo system 346
 Geocentric horizon 385
 Geocoding 309
 Geodetic quadrilateral adjustment 112
 Geodetic surveys 1
 Geographical information system 166
 Generalization 328
 Geometric corrections 274
 Geometric dilution of precision 367
 Georeferencing 278
 Geostationary satellite 239
 Geographic Information System 301, 372
 applications 332
 capabilities/functionalities 317
 data for 306
 data structure 309
 error sources in 328
 hardware 305
 software 334
 subsystems 303
 Global positioning system 343
 accuracy 368, 373
 applications 369
 error sources 359
 overview 347

- surveying techniques 366
- system time 358
- uses 369
- Glonass system 345
- Graphical elements 309
- Graphical entity 310
- Grey level thresholding 286
- Grid cells 308
- Grid iron system 30
- Ground principle point 173

H

- Half marks 206
- Height distortion 178
- Heliographs 35
- Heliotropes 35
- Helmert formula 50
- High pass filter 289
- Histogram equalization 285
- Homologous points 171
- Horizon 385
 - celestial 385
 - glass 141
 - system 391
- Horizontal control 21
- Hour angle 388
- Hour angle system 392
- Hour circles 386
- Hunter's short base method 47
- Hydrographic survey 134
- Hydrography 134

I

- Ideal remote sensing system 240
- IFOV 248
- Image
 - classification 291
 - element 243
 - enhancement 280
 - magnification 281
 - rectification 270
 - reduction 280
 - restoration 270
 - transformation 290
- Imaging 361
- Index correction 50, 448
- Index glass 141
- Index map 214
- Indian co-ordinate system 369
- Inertial surveying system 343
- Instrument and signal correction 3

- International atomic time 358
- Interpreted data 306
- Intervalometer 187
- Isocentre 171

J

- Jaderin's method 47
- Jewel box 431

K

- Kinematic positioning 366
- Kinematic surveying 368

L

- Latitude 386
 - determination 441
- Laws of
 - accidental errors 85
 - weight 83
- Lead line 139
- Lens stereoscope 200
- LIDAR 239
- Line data 308
- Line drop out 272
- Line-in-polygon 323
- Locating sounding 141
- Longitude determination 445
- Longitudinal overlap 187
- Low pass filter 288
- Luminous signals 35
- Lunar day 412

M

- Map 170
 - algebra 325
 - index 214
 - layers 312
 - overlay analysis 322
 - photographic 214
 - stereo-topographic 214
 - topographic 214
- Mapematics 325
- Marigram 138
- Mean anomaly 348
- Mean principal base 206
- Mean sea level 136
- Mean solar day 411
- Measurement of
 - angles 64
 - discharge 150
 - velocity 150
- Median filter 288

- Meridian 385
- Method of
 - correlates 103
 - differences 101
 - equal shifts 119
 - least squares 122
- Mie scatter 233
- Minimum-maximum stretch 283
- Minimum mapping unit 312
- Mirror stereoscope 200
- Mistake 81
- Mode filter 288
- Moon 384
- Mosaic 213
- Most probable value 82
- Multipath error 360
- Multiplex 211
- N**
- Nadir 385
- Nature of error 81
- Nadir point 171
- Napier's rule 399
- Nautical sextant 446
- Navy navigational satellite system 343
- Neap tide 316
- Neighbourhood functions 319
- Nodes 384
- Noise 271
- Noise correction 271
- Non-spatial data 307
- Non-symmetrical distortions 277
- Normal equations 100
- North star 431
- Number of photographs 189
- Nutation 391
- O**
- Oblique photograph 170
- Observation equation 96
- On-the-fly ambiguity resolution technique 368
- Opaque signals 34
- Operational
 - control segment 349
 - space segment 346, 347
- Organised data 306
- Orthophotos 214
- Overlap 187
- P**
- P-code 349
- Panaromic distortion 274
- Parallactic angle 398
- Parallax 201
 - bar 207
 - measurement 206
 - absolute 202
 - angles 198
 - equation 203
 - linear horizontal 167
 - stereoscopic 202
- Particle theory 231
- Pel 243
- Perihelion 379
- Phase error 36
- Photo-interpretation 196
 - equipment 198
- Photo-theodolite 166
- Photographic maps 214
- Photogrammetric
 - map 166
 - products 212
- Photogrammetry, aerial 164
- Photographic
 - base 206
 - maps 214
- Photons 231
- Photo plumb point 171
- Photo triangulation 210
- Picture control 209
- Picture element 243
- Piecewise stretch 285
- Pixels 242, 308
- Planets 382
- Platforms 239
- Plotting the soundings 147
- Point data 307
- Point-in-polygon 323
- Point positioning 366
- Pointers 431
- Polar
 - axis 379
 - distance 387
- Polaris 429
- Poles 379
- Pole star 429
- Polygon-on-polygon 323
- Position determination principle 351
- Precession 382
- Prime vertical 386
- Principal
 - base 206

- point 166, 171
- Principle of least squares 94
- Point location 211
- Prismatic astrolabe 455
- Probability curve 86
- Prolate spheroid 379
- Proximity analysis 320
- Pseudo-kinematic technique 367
- Pseudo-range 354
- Pseudo-static surveying 367
- Q**
- Quadrilateral 112
 - with central station 115
- Quantisation 243, 248
- Queries 318
- R**
- Radar
 - distance measurement by 60
 - principle 60
 - ranging 59
 - triangulation 59
- Radial line method 210
- Radiometric
 - corrections 270
 - enhancement 282
- Range 138
- Rapid static surveying 367
- Raster overlay capabilities 325
- Rationed prints 209
- Rayleigh scatter 232
- Real remote sensing system 240
- Receiver errors 360
- Reciprocal observation method 2, 11
- Reclassification 319
- Reconnaissance 37
- Rectification 211
- Reduction
 - of sounding 147
 - to centre 65
- Reference data 244
- Reflectance 235, 245
- Reflectors
 - lambertian 235
 - specular 235
- Relative positioning 366
- Relief displacement 178
- Remote sensing 225
 - applications 249
 - in India 258
- process 226
- Reoccupation technique 368
- Resampling 278
- Resolution
 - concepts 246
 - ground 247
 - radiometric 248
 - spatial 246
 - spectral 249
 - temporal 249
- Retardation 414
- Right ascension 387
- Right ascension and declination system 391
- Rigid bars 44
- S**
- Satellite
 - clock corrections 349
 - constellation 347
 - data products 264
 - doppler 146
 - geostationary 239
 - stations 66
 - sun-synchronous 240
- Satellite-related errors 360
- Satellite-to-user range 353
- Satellites 382
- Saturation stretch 284
- Scaffolds 37
- Scale
 - datum 174
 - factor 247
 - of vertical photograph 172
 - of tilted photograph 174
- Scan skew 274
- Scanner 237, 238
- Scanning system 237
- Scattering 232
- Selective availability 362
- Semi-diameter correction 425
- Sensible horizon 385
- Sensors
 - across track 238
 - active 236
 - along-track 238
 - framing system 237
 - hyperspectral 238
 - imaging 237
 - non-imaging 237
 - passive 236

- scanning system 237
 - super 241
 - Sextant
 - nautical 446
 - sounding 446
 - Side equation 113
 - Side overlap 188
 - Sidereal
 - day 409
 - year 409
 - Signals 32
 - characteristics 34
 - height 39
 - luminous 35
 - opaque 34
 - phase of 36
 - propagation errors 360
 - Signatures 244
 - Simple spatial objects 306
 - Single observation method 2, 9
 - Slivers 332
 - Slotted template method 211
 - Software
 - DIP 293
 - GIS 334
 - Solar attachment 451
 - Burt's 454
 - Saegmuller 453
 - Smith 452
 - Solar
 - day 411
 - observations 448
 - system 381
 - Solstice 382
 - Sounding 134, 138
 - boat 139
 - equipment 139
 - locating 135, 141
 - plotting 147
 - reduced 135
 - reduction 147
 - rod 139
 - sextant 141
 - Space mark 206
 - Space segment 347
 - Spatial data 306
 - Spectral
 - reflectance 235
 - rationing 290
 - Spectrometer 236
 - Spherical
 - excess 112, 384
 - triangle 384, 394
 - trigonometry 394
 - Spheroid 379
 - Spring tide 136
 - Standard meridian 413
 - Star
 - at horizon 398
 - at elongation 399
 - at prime vertical 401
 - Stars 382
 - Static positioning 366
 - Static Survey Technique 366
 - Station
 - adjustment 107
 - description of 33
 - marks 33
 - intervisibility 39
 - selection 39
 - Stereoautographs 168
 - Stereocomparator 168
 - Stero pair 166
 - Stereoplanigraphs 168
 - Steroplotter 211
 - Steroscope 199
 - Stereoscopic
 - fusion 199
 - pair 166, 187
 - parallax 201
 - plotting method 211
 - vision 198
 - Stream measurements 149
 - Strength of figure 25
 - String data 308
 - Summer solstice 382
 - Subsystems of GIS 303
 - Sun 381
 - Sun elevation correction 273
 - Sun-synchronous satellite 240
 - Surface phenomena 234
 - Surveys, geodetic 1
 - Systematic
 - distortions 274
 - errors 82, 360
- T**
- Tape corrections 47
 - modified 48

- Telemetry 328
 - Tellurometer 62
 - Terrestrial
 - photogrammetry 165
 - streophotography 166
 - Theodolite attachments 451
 - Theory of errors 81
 - Three-dimensional positioning 351
 - Tide gauges 137
 - Tides 135
 - Tilt 171
 - displacement due to 183
 - Tilted photograph 170
 - scale 174
 - Time 409
 - apparent solar 411
 - conversion 413
 - determination 427
 - equation of 412
 - greenwich mean 411
 - local sidereal 409
 - mean solar 411
 - measurement 409
 - sidereal 409
 - standard 413
 - Tip 171
 - Towers 32, 37
 - Traditional
 - kinematic technique 368
 - static surveying 366
 - Transforming printer 212
 - Transit 389
 - TRANSIT system 345
 - Transmittance 235
 - Traversing 21, 134
 - Triangular misclosure 23
 - Tringulation 21
 - adjustments 105
 - afloat 71
 - choice of figure 24
 - classification 31
 - field checks 70
 - figures 22
 - first order 31
 - flare 70
 - floating 71
 - misclosure 23
 - precision of network 32
 - second order 31
 - systems 22
 - third order 31
 - Trigonometric levelling 1
 - Trilateration 69
 - Tropical year 367
 - Tropic
 - of cancer 382
 - of capricorn 382
 - True colour composites 245
 - True meridian, determination of 429
- U**
- Ursa major 431
 - User equipment segment 350
 - User Position 355
- V**
- Vector data structure 309
 - Vector overlay capabilities 323
 - Vernal equinox 381
 - Vertical circles 386
 - Vertical photograph 170
 - View finder 187
 - Volume phenomena 234
- W**
- Water glass 140
 - Weight 82
 - laws 83
 - Wheeler's method 46
 - Whisk-broom scanner 238
 - Winter solstice 382
 - World geodetic system 369
- X**
- x-parallax 202
- Y**
- Yellow dwarf star 381
- Z**
- Zenith 385
 - Zenith distance 386
 - Zenith telescope 456
 - Zodiac, signs of 384

PLATE 1

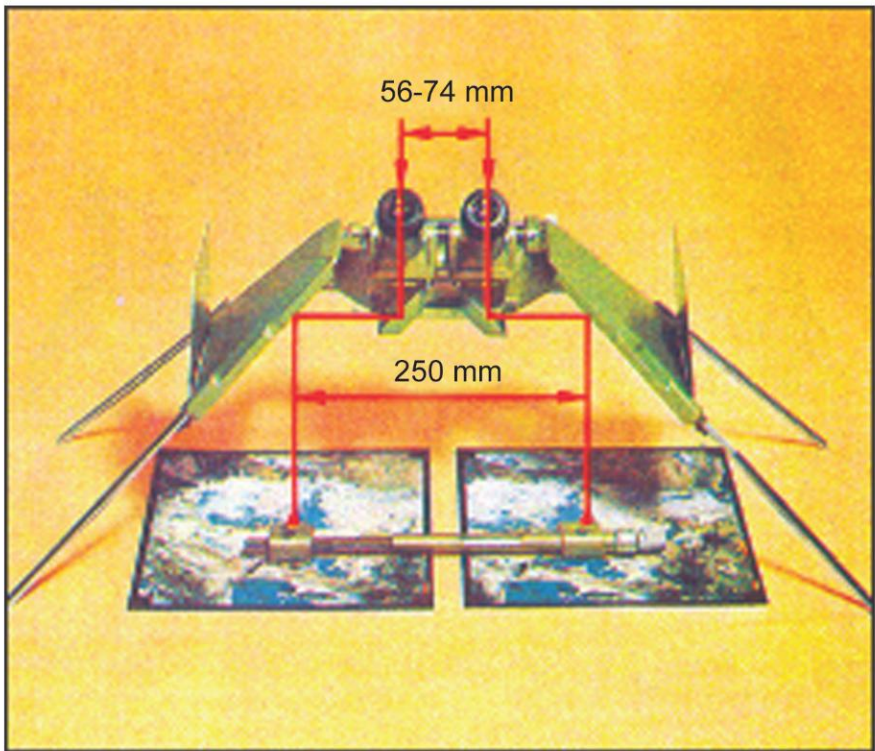


(a) Large size vessel



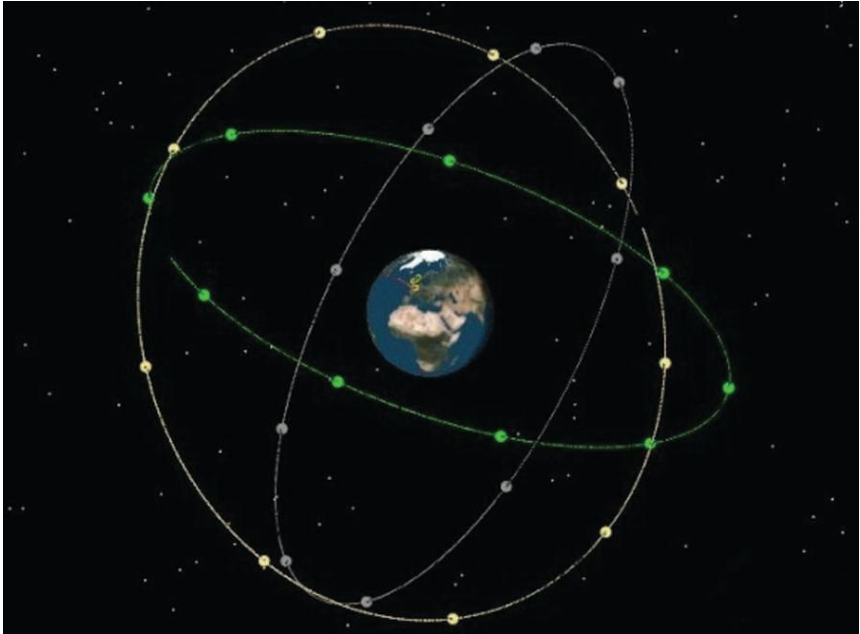
(b) Indian Navy ship, MAKAR

PLATE 2



Stereoscope: Wild ST4

PLATE 3

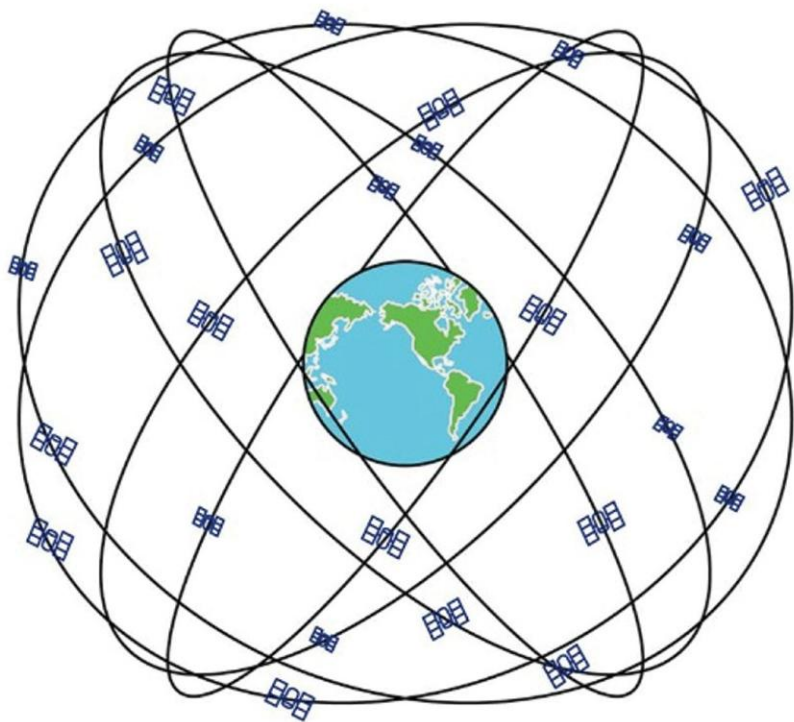


(a) Galileo satellite constellation



(b) GLONASS satellite constellation

PLATE 4



GPS satellite constellation

PLATE 5



GPS receiver and antenna