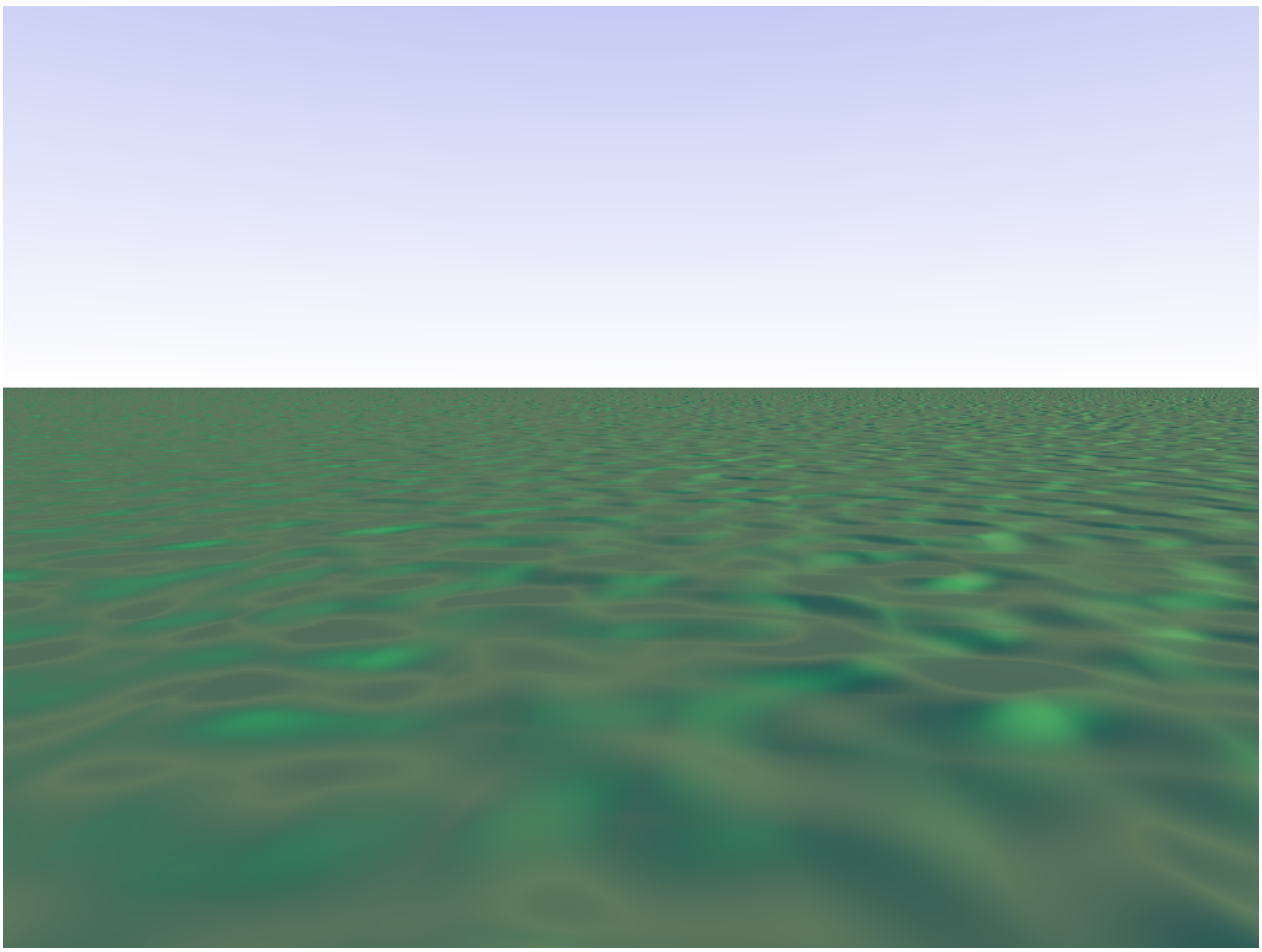
ASTR 324
Longitude, Time, and Navigation
Lecture Notes



Tom Bensky
Physics Department
Cal Poly, San Luis Obispo
San Luis Obispo, CA 93407
tbensky@calpoly.edu
Spring 2008 edition

Cover Figure

If you were a ship's captain in the 1700s, the scene on the cover would be a big problem for you. Just the sky and the sea, with no sight of land. In fact, a scene like this would be your biggest problem, and is the topic of this book: *WHERE ARE YOU?*

Acknowledgements

Special thanks to M. Moelter for his careful and professional proofreading of an early draft of this work (I love that green pen!). Thanks also to D. Doty, J. Keller, and R. Echols for helpful discussions about astronomy.

Contents

1 Longitude, Time, and Navigation **6**
1.1 Introduction 6
1.2 The world way back then 7
1.3 Latitude and Longitude 8
1.4 Perils of the ignorance of longitude as sea 12
1.4.1 The Galleon: a death trap 13
1.4.2 Sir Clowdisley Shovell 15
1.4.3 George Anson 15
1.4.4 La Salle 16
1.5 But still they set sail 17
1.5.1 Best guesses at one's location 17
1.5.2 Deduced Reckoning or DR 18
1.5.3 Milk your latitude 19
1.6 So how does one find longitude at sea? 19
1.6.1 Time and Longitude 19
1.6.2 Idea #1: The Lunar Method 21
1.6.3 Idea #2: Keep accurate time as sea 23
1.6.4 The moons of Jupiter 24
1.7 The Longitude Prize 25
1.8 The Remarkable John Harrison 26
1.9 On we go... 27

2 Excerpts from "The Quest for Longitude" **29**

3 The Earth, Moon, Stars, and Navigation **88**
3.1 Mapping the Earth 88
3.1.1 Locating yourself on the earth 89
3.1.2 Locating points with latitude and longitude 95
3.2 Angles in space 98
3.2.1 Use your hand as an "angular ruler" 98
3.2.2 Angular Measure 100
3.2.3 Angles and Distance 104
3.3 Celestial Navigation: Navigating using the Sun, Moon, and Stars 105

3.3.1 The Basic Premise of Celestial Navigation . . . 105
3.3.2 A basic measurement . . . 109
3.3.3 The Sky . . . 114
3.3.4 The Sun . . . 124
3.3.5 The Moon and Lunar Distances . . . 126
3.3.6 The Stars . . . 132
3.3.7 The Moons of Jupiter . . . 132
3.4 Celestial Navigation: a case study . . . 142
3.4.1 Let's find ourselves then . . . 143
3.4.2 What did we just do? . . . 158
3.5 How to find latitude . . . 160
3.5.1 Approximate Easy Latitude: A noon sight of the Sun . . . 166
3.5.2 Better Easy Latitude: A noon sight of the Sun . . . 169
3.5.3 Easy Latitude again: A sight of Polaris at night . . . 172
3.5.4 A Guess at History: Using Polaris + the Sun . . . 173

4 Time and Time Keeping 176
4.1 Introduction . . . 176
4.2 Measuring Time . . . 177
4.2.1 Short Times . . . 178
4.2.2 Longer times . . . 180
4.2.3 A summary of long and short times . . . 181
4.2.4 Heisenberg: The best we can do . . . 184
4.2.5 Time in Science . . . 185
4.3 Time Keeping with a Clock . . . 186
4.4 Nature's Clocks . . . 187
4.4.1 The Pendulum . . . 187
4.4.2 Mass on a Spring . . . 192
4.5 Core Elements of Mechanical Clocks . . . 203
4.5.1 Sustained Energy . . . 203
4.5.2 Escapements: Controlling the Energy Release . . . 207
4.6 More Bizarre Clocks . . . 216
4.6.1 Kepler's Laws . . . 216
4.6.2 Atoms . . . 219
4.6.3 Electronics . . . 231
4.6.4 Electrical circuits that have a period . . . 239
4.6.5 Crystals . . . 242

5 Modern Ideas on Time and Navigation 247
5.1 Introduction . . . 247
5.2 Albert Einstein . . . 249
5.2.1 Einstein's Ideas on Time . . . 249
5.2.2 Clocks on the move . . . 252

5.2.3 It moves so its life will slow . . . 258
5.2.4 Einstein and Harrison . . . 259
5.3 The Crown Jewel in Navigation: The Global Positioning System . . . 259
5.3.1 Introduction . . . 259
5.3.2 The GPS System . . . 261
5.3.3 Finding your location . . . 262
5.3.4 Why Atomic Clocks? . . . 266
5.3.5 Closing Remarks . . . 269

6 Could you do what Harrison did? 270
6.1 Introduction . . . 270
6.2 Problems . . . 270
6.2.1 Xprize.org . . . 271
6.2.2 The Copenhagen Consensus . . . 271
6.2.3 The Bill and Melinda Gates Foundation . . . 272
6.2.4 Obvious problems . . . 272
6.3 The Assignment . . . 272

7 Map Worksheets 274

Chapter 1

Longitude, Time, and Navigation

Dad: Longitude is so boring.
-Will Bensky (8 years old)

1.1 Introduction

Suppose you were asked you to list some of the biggest problems facing the world today. What would you say? Global warming? Terrorism? Something environmental? Hunger? Energy? Next, suppose $12 million was offered if you could solve one of these problems. Would you pick one and get started working on it? Or would you turn away shaking your head saying "How could I solve such a problem?" Would you be willing to dedicate your entire life to solving one of these problems?

In this class, we are going to take a whirlwind look at a huge problem facing the world in the 1700s. In fact until about 1780, this problem had been unsolved for more than 300 years, even though some of the greatest thinkers, such as Isaac Newton (inventor of calculus), Galileo Galilei (first to use a telescope), Edmund Halley (the comet), Huygens (calculus, optics, found Saturn), Euler (calculus, physics), and others worked on it. In the end, an uneducated, unknown, but extremely dedicated and brilliant carpenter, named John Harrison solved the problem after a lifetime of work. In fact, he only lived 3 more years (into his eighties) after winning the prize that would today be worth $12 million. He worked for stretches as long as 15 years straight on various attempts at a solution. Would you do this? Could you?

So what was this world-wide problem? **Finding longitude at sea.**

What you are thinking?

- Huh?
- What's that?
- What is longitude anyway?
- Why would this be an important problem?

- Why “finding longitude?”
- Why “at sea?”

1.2 The world way back then

The world in the 16th-18th Century (1500s-1700s) was obviously quite different than today. Societies were well developed in Europe, but the “new world,” consisting primarily of the North American continent was largely unclaimed and unexplored. Portugal kicked off what has come to be known as the “Age of Discovery,” in the mid-1400s. The westernmost country in Europe, Portugal was the first to significantly probe the Atlantic Ocean, colonizing the Azores and other nearby islands, then exploring the west coast of Africa. In 1488, Portuguese explorer Bartolomeu Dias was the first to sail around the southern tip of Africa, and in 1498 his countryman Vasco da Gama repeated the experiment, making it as far as India. Portugal would establish ports as far west as Brazil, as far east as Japan, and along the coasts of Africa, India and China[Smithsoninan, Sept. 2007].

For additional reference dates, think of Christopher Columbus setting sail from Spain in 1492. The 13 original English colonies were established between 1607 (Virginia) and 1732 (Georgia). The American Revolution culminated with a battlefield victory in 1781. The great westward exploration of Lewis and Clark took place in 1804. European societies (Portugal, France, England, etc.) knew of great riches in the new world, and the political gain that could be made by finding and staking claim therein. They just had to get there, and getting there was a problem.

At this time, the dominant ocean faring vessel was a “galleon,” which looked something like that shown in Figure 1.1. The look of these ships is probably somewhat familiar to you, as the first “long haul” sailing ship. They were large multi-decked sailing ships used primarily by the nations of Europe from the 16th to 18th centuries (1500s - 1700s).

Here was “the problem:” All of the great explorers you may recognize, who used these Galleons, from Columbus to de Gama to Magellan to Drake were essentially “sailing blind.” They had virtually no idea where they actually were as soon as they lost sight of land. Really? Yes, really. Any “great discovery” they made was at a great human cost and more-or-less by sheer luck.

Imagine your last trip to the beach. You look out over the water to the horizon, where the sky touches the sea. Now, imagine seeing this view in every direction; no matter where you look, for weeks or even months at a time. There are no landmarks in the ocean. No features and no reference points. Just waves, water, and more waves. Back in those years, knowing where you were in the vastness of the sea was anyone’s guess.

As an example, here is an excerpt from Rupert Gould’s book enetitled “The Marine Chronometer” about Christopher Columbus, the great explorer we all learn about in 2nd grade:

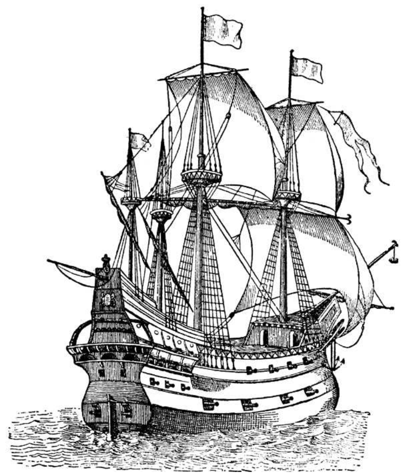


Figure 1.1: A typical 16th-18th century Galleon.

1.3 Latitude and Longitude

To locate yourself anywhere on the globe, you need to know two numbers, your latitude and longitude. Latitude is your distance north or south of the equator. Lines of latitude are shown in Figure 1.2.

The equator is the zero-reference line with a latitude of zero (0°). The north pole has a latitude of "90° North" and the south pole has a latitude of "90° South."' San Luis Obispo, California has a latitude of 35.33° North; the city is 35.33° North of the equator. Lines of latitude are circles drawn around the earth that run parallel to the equator. The circle that forms the equator is the largest; they get smaller and smaller as you travel north or south. Lines of latitude are sometimes called "parallels," since they all run parallel to the equator.

Longitude is your distance east or west on the globe. Lines of longitude are shown in Figure 1.3. The "problem of longitude" was with these lines, in that **during an ocean voyage, there was simply no practical method for finding the line of longitude on which you were located.** Why? Part of the problem was because unlike latitude, lines of longitude have no natural reference point (like the equator). The reference point of "zero" longitude can be any line of longitude one desires. It is a human set reference point, and has been arbitrarily set to pass from the north pole to the south pole, directly through Greenwich, England.

Every point on this particular line of longitude has a longitude of 0°. Lines of longitude are sometimes called "meridians" and are big circles, all with the same circumference, passing all the way around the earth, through the north and south poles. As mentioned, Greenwich, England is at 0° longitude, by international agreement. San Luis Obispo is at 120.74° West longitude. Moscow, Russia is at about 35° East longitude. London, England being so close to the Greenwich has a longitude of a mere 0.083°West

When put together, lines of latitude and longitude form an imaginary "grid" or "mesh" over the whole earth, giving definite points to reference location by, whether on land or sea

INTRODUCTION.

THE PROBLEM OF FINDING LONGITUDE AT SEA.

Will the reader be good enough to imagine, for a few minutes, that he is Christopher Columbus ?

It is the evening of Monday, February the 11th, 1493, and from the deck of the little " Nina " there is nothing to be seen but the " Pinta " labouring some distance astern, and the darkening rim of the sea horizon which surrounds them. Far behind them lies the new world of his discovery, and ahead is the old world, to which he is now returning. The ships are running bravely before a strong westerly breeze, and at intervals clouds of sea-birds sweep by, heralding the great gale that will burst upon him to-morrow.

As the sky fades and the stars come out, he turns his mind once more to the familiar and yet baffling task of determining the position of his little fleet. With his cross-staff he takes the altitude of the pole star, and, after some computation, obtains his latitude. Rough though this observation is, he can hope for no better, and it agrees tolerably well with his noon observation of the sun.

But in what longitude ? Ah, that is beyond even his skill to determine—beyond the skill of any living man. He can do no more than guess at it, as he has guessed ever since they left Madeira some six months earlier, steering into the unknown West. True, he has been able to inspire his men with the belief that he can keep an exact reckoning of his ship's course and the distance she traverses, and also determine the amount to which these are affected by currents—but this belief is due to a pious fraud, prepared, like the falsified reckonings of the outward voyage, for their encouragement, and in sober truth he is as helpless as any of them. He knows his ship's course, roughly, and he can guess at her speed—so can they. And accordingly, as to their longitude, there is much difference of opinion. The admiral thinks that they are south of the Azores. Vincenti Pinzon, his second in command, and others of experience, consider that they have already passed these, and are approaching Madeira—600 miles further eastward ! And meanwhile night has covered them, and the ships are running on blindly to meet their fate, not knowing whether the land is a hundred miles away or close at hand.

Next day comes the great storm, and all, even the Admiral himself, abandon hope. But his luck still holds, and after the pilgrimages and wax candles have been vowed, and the barrel with the tidings of the great discovery stealthily dropped overboard, the storm passes, and they find themselves in sight of—the Azores ! Whether through luck or intuition, his guess is the correct one, and the event emphasises the fact that discoveries generally go to the men best fitted to make them ; but it emphasises also the grave risks to which the navigators of his time were exposed by their utter inability to find their longitude when out of sight of land.

The case of Columbus is particularly striking, on account of the historical importance of his voyage, and the unusually long period, some six months, spent in total ignorance of his longitude*: but for many generations to come—in fact, until the close of the eighteenth century—navigators had practically no better method of finding their longitude than he had. All that they could do was to keep a reckoning, termed the "dead reckoning," of the courses they steered and the distances they ran, and to make such allowances as they thought fit for leeway, tide, current, variation of the compass, errors in estimating their speed, bad steering, and many other sources of error. When one considers the slow speed and comparatively enormous leeway of their clumsy vessels, the wonder is not that their dead reckoning was often absurdly wrong, but that it was ever anywhere near the truth. Practically their only advantage over Columbus was that they had a better means of estimating their speed—the hand log—than he had.

A few instances, taken almost at random during the period 1650-1750, will show the risks they ran.

In 1691 several warships were lost off Plymouth, having mistaken the Deadman for Berry Head.

Sir Cloudesley Shovel, returning from Gibraltar with his fleet in 1707, had cloudy weather during practically the whole passage, and, after some twelve days at sea, took the opinions of the navigators of all his ships as to his position. With one exception (which afterwards proved correct) their reckonings placed the fleet in a safe position some distance west of Ushant, and he accordingly stood on: but the same night, in fog, they ran on the Scillies. Four ships were lost, and nearly two thousand men, including the Admiral himself.†

Several transports were lost in 1711 near the entrance to the St. Lawrence River, having erred 45′ in their longitude in twenty-four hours.

Lord Belhaven was wrecked on the Lizard the same day that he sailed from Plymouth, November 17th, 1721.

The famous voyage of Commodore Anson, who took the Acapulco galleon in 1743 and came home round the world with over half a million in prize money, provides two particularly striking examples of the total inability of the navigators of his day to find their longitude when out of sight of land. In 1741 he spent over a month endeavouring to round Cape Horn to the westward, and having, by his reckoning, made good sufficient westing to place him 10° clear of the most western point of Tierra del Fuego, stood to the northward, only to sight land right ahead, and to find that owing to an unsuspected easterly current he was still on the eastern side of the Cape.

* Although the sighting of the Azores shows the *relative* accuracy of the reckonings kept on the outward and homeward voyages, yet Columbus' determinations of the *actual* longitude of his discoveries were marvellously erroneous. He firmly believed that Cuba was part of the mainland of Asia—nay, *he enacted that this was the case*, and compelled every member of his expedition, under heavy penalties, to make affidavit thereunto! Having thus abolished the Pacific Ocean by legislation, it is not surprising that he should remark to Queen Isabella "the earth is not so large as vulgar opinion makes it."

† A story was current, long afterwards, that a seaman of the flagship had kept his own reckoning, which showed that they were in a dangerous situation, and that on his making this known to his superiors he was hanged for mutiny, there and then. *Credat Judaeus Apella.*

Again, after rounding the Horn and parting company with his squadron, scurvy broke out aboard his flagship, the "Centurion," and Anson, with his men dying like flies, ran to the northward, hoping to make the island of Juan Fernandez, where he could land his sick. In the ordinary way he would have steered to get into the latitude of the island a long way east or west of it, and then have run along that parallel until he sighted it—a plan still practiced by many Pacific traders. To save time and lives, and urged by the terrible fact that a few more days of the present death rate would leave the ship too short-handed to go about, he sailed straight northward for the island, with the result that he reached its latitude without sighting it, and was uncertain whether it lay to the eastward or the westward. He ran westward until (unknown to him) he was within a few hours' sail of it: then, concluding he was wrong, he sailed eastward until he made the coast of Chile, and had to turn and run back westward over the same track until he finally sighted the island. This uncertainty as to his longitude cost him the lives of some seventy or eighty of his men, who would probably have recovered if they could have been got ashore.*

Enough has been said to show that the problem of finding longitude at sea was no academic exercise, but a matter of the most urgent and vital importance, and one which no nation which used the ocean highways could afford to ignore. It overshadowed the life of every man afloat, and the safety of every ship and cargo. Yet, for nearly three centuries after the great voyages of Columbus, Cabot, and the Portuguese navigators had focussed attention upon it, it defied all attempts at solution—not only those of seamen, but of astronomers, mathematicians, geographers; in short, it baffled the best brains of the civilised world.

It was as a solution—and, until the introduction of W/T time-signals, the best solution—of this problem that the marine chronometer came into being. And before we can properly estimate its value to the navigator, it is necessary to form some idea of the difficulties which it overcame, and of the various other methods which are, theoretically, available for the finding of longitude at sea.

So long as a ship remains in sight of land whose position is accurately shown on her charts, and which she can identify, her own position can be readily obtained by direct observation. But when once she is out of sight of land, her position must be obtained by observations either (1) of some terrestrial phenomenon, or (2) of the heavenly bodies.

* To add to the pathos of this story, the "Centurion" was the ship which, some years earlier, carried Harrison's first timekeeper to Lisbon, and she was thus the first vessel to be provided, even temporarily, with a practical means of finding her longitude accurately. The machine itself was going, in Harrison's house in London, during the whole period of her voyage, and much later—from 1736 until 1766.

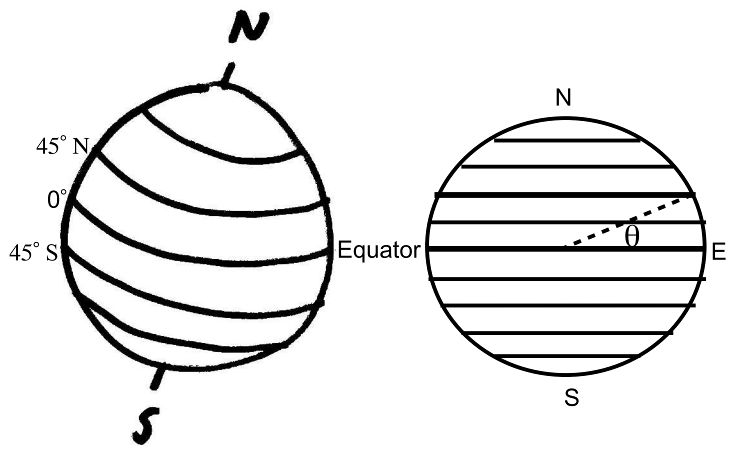


Figure 1.2: Lines of latitude. These lines all run parallel to the equator. The equator is the line of 0° latitude.

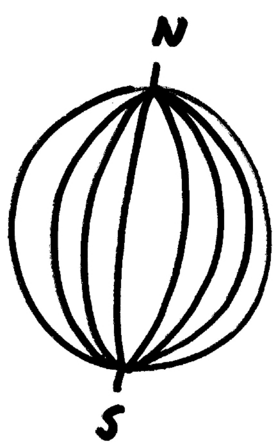


Figure 1.3: Lines of longitude. These lines all run around the earth, passing through the North and South poles. There is no natural 0° line of longitude. Determining which line you were on during an ocean voyage was at the heart of the longitude problem.

Figure 1.4: Lines of latitude and longitude put together on a globe.

(or in your living room). This is shown in Figure 1.4.

Keep in mind that lines of latitude and longitude as reference points are not restricted to the earth, but can be used to pin-point objects on any spherical surface. Planetary scientists, for example, also use latitude and longitude to map the surfaces of planets, such as Mars, Venus, and even our very own moon. For example, Apollo 11 landed on the moon in July 1969 at a latitude of 0.67 N and Longitude 23.49 E, right at Mare Tranquillitatis.

There was some hope however. In the vast, featureless ocean, about the only thing to look at is the sky; the sky still has unique features, even at sea. As we will learn in a later chapter, it is possible to use the sun and stars to find one's latitude. This is because as you move North or South on the globe, the height of the sun and stars, as you see them change with your north/south position. Their height at your location can be used to find your latitude, and this was well known even way back in the time of Columbus.

The stars and sun, however, do not easily reveal your longitude. As you move east and west, the sun and stars appear to rise and set at earlier and later times, depending on your east-west position and there was no other way of determining longitude. That is all, and this is not enough to find your longitude. So at best, early explorers only knew half of where they were. They could find their distance north or south of the equator, but not their position east or west on the globe.

So this was the 300-year-old problem: coming up with a method of accurately finding one's longitude at sea.

1.4 Perils of the ignorance of longitude as sea

Next, a few historical accounts on how exploration proceeded without the knowledge of accurate longitude. One might think that not knowing longitude is no big deal. If you know your latitude, you "sort of" know where you are. Isn't this good enough? Not really. These accounts will demonstrate the severity of the longitude problem and the desperate need for

a solution.

To start with, traveling on a galleon was simply dreadful as you will now see. Combine that with a captain who doesn't really know where he is, or where he's going, and you're in for some ocean voyages with pretty horrible outcomes.

1.4.1 The Galleon: a death trap

Anyone who has taken a long driving trip (particularly with kids) knows that, for many reasons, it's best to get to the destination with as few delays as possible. Traffic jams, bad weather, getting lost and dead iPod batteries all lead to hunger, boredom, and numerous bathroom breaks for those in the car. For a long driving trip, its best to have good weather, and precise directions, all to ensure a speedy arrival. However, a long road trip, even where the kids are screaming, is very comfortable compared to the Galleons from 300 years ago.

The word "galleon" has described many types of ships in the history of mankind, but its most recent meaning has settled on the type of ship shown in Figure 1.1: essentially a ship with a high superstructure and poor sailing qualities [Kirsch, p. 1]. These galleons were widespread in Europe in the 16th century, and as clumsy as they look now were actually a step forward in maneuverability and seaworthiness [Kirsch, p. 3]. They were also well equipped to carry the heavy cannons on the upper decks. A galleon was typically 100 feet long, with a 32 foot width and a hull depth of 12 ft[Ibid, p. 21]. The high "fortress-like" structure was undoubtedly to protect the ship from unwanted entry; getting on board from sea-level was like climbing a wall.

There is one thing we know for certain about the galleons: life aboard them was dreadful. The ships wore out the men who traveled on them. A hard cabin, cold and salt meats, broken sleeps, moldy bread, dead beer, wet clothes and a desire for a warm fire, which was strictly prohibited[Kirsch, p. 74]. Most of the men only had one set of clothing for an entire voyage. Darkness, dampness, and foul air below the desk presented lots of problems.

One might think severe weather made sea-travel hard. Sure it did but calm weather for a sailing ship is no fun either. At times, particularly near the equator (7-8 degrees North latitude), these great sailing ships got caught in 30-60 days of extremely calm, windless days; the "dreaded doldrums."[Kirsch, p.82] The air was still and the heat on board the ships would be unbearable. The heat would cause cracks to appear in the ship walls, and the hot, tarred wood would have to be sprayed with water. In one case:

> *The heat is strong and oppressive. It destroys most of the provisions, the water if foul and full of worms. Even very carefully salted meats and fish go rotten. Butter melts to oil, as do the candles. Tar melts as well. It is impossible to stay below deck, as it is like an oven down there.*[Kirsch, p. 83]

Elementary hygiene was lacking on the ships too. By the time ships leaving Europe made it to the equator, the ships resembled floating plague-houses. The inhabitants were dying like flies. Pots and jugs were used to "relieve oneself," and the weak and sick were often left lying for days in their own excrement.[Kirsch, p. 83].

Yes, virtually every voyage resulted in losses due to serious illness, accidents, and desertion. Therefore most voyages started out heavily manned. Also, since there were almost no mechanical aids on board, the ships required a lot of muscle power[Kirsch, p. 74]. Medicine was almost unknown on board, despite the high frequency of longer and longer voyages, in tropical waters. Dysentery (bloody, painful stool), scurvy (see below), typhoid, and beri-beri (mental disorder, painful limbs, irregular heart-beat due to lack of thiamine) were widespread.

The galleons from way back were typically loaded with food and fresh water, but it wasn't possible to keep these items fresh for very long. Food for each man consisted of one pound of biscuits, a gallon of beer, and two pounds of salted beef[Kirsch, p. 75]. Beer was difficult to keep and quickly went sour. As described below, not knowing one's longitude would cause many transoceanic journeys to increase significantly in duration. Fresh water was quickly depleted, and fruits and vegetables wouldn't keep for long (they weren't really part of the popular diet back then anyway). Scurvy would often set in, which is a gruesome way to die.

Lack of vitamin-C leads to weakening of the blood vessels and connective tissue in the body. Internal breaks would lead to blood leaking out of veins and arteries inside the body. Skin would appear to simply melt off of a body[Dash, p. 73]. A person suffering from scurvy would then have bruises all over their body, and any cut or injury would not heal. Teeth would become loose and fall out, and the tongue would swell; it would become hard to talk. Death would often follow after blood vessels in the brain burst. This disease could easily kill off most of a ship's crew (at a rate of 6 - 10 men day[Sobel p 18]). All of this while the captain was zig-zagging the ocean turning this way and that following a "hunch" on how to get "there." Never mind getting tossed around in a good storm, only to be totally lost when the calm returned. Indeed a delayed car trip results in passengers who haven't eaten in "a while" (a few hours), or in need of a bathroom break (if the bathroom at the rest-stop is clean). Dead iPod batteries can be charged using the cigarette lighter. A delayed transoceanic trip in the 18th century led to death.

Further, the command of a galleon was the responsibility of only a few people, considering there were about 300 people on board. The captain was clearly in charge; disagreeing with him was punishable by immediate death-by-hanging. When there was nothing but the sea for months at a time, desperation would set in; the crew would become antsy. Did the captain know where he was going? Are we going to survive this trip? How would two or three in charge maintain order among 100 other seamen on the ship? Each day at noon, the captain and his inner circle would emerge onto a deck of a ship, in something of a show, to a prominent place where they could be seen by all. They would put on an elaborate act, flashing books, charts, and navigation instruments, as if applying some great skill to determine the position of the ship [see image Andrews, p. 73]. The others could only stop and wonder, in relief and awe, that the captain is able to figure their position en route to a safe and timely arrival. A mutiny aboard the ship would certainly be a bad idea, for who would be able to perform these daily navigational duties?

After the captain, there was the pilot who actually steered the ship as he watched a compass needle. The masters, quarter masters, and mates commanded the stern (rear), middle, and forecastle (forward) parts of the ship. A "merinho" (petty officer) carried out

the captain's orders as they pertained to justice aboard the ship. Rigging masters would mend the sails. There was also the secretary, who was an important man on board, since he was appointed by the King and represented his power. Finally, there was a chaplain, master gunner, and barber (medical matters).

In closing, doesn't it sound best for a galleon to get to where it is going, *as soon as possible?* At sea, this can only be accomplished with correct, confident navigation.

1.4.2 Sir Clowdisley Shovell

Sir Clowdisley was headed due North, returning to England during a 12-day stretch of fog. Latitude measurements or any guides from the sky were impossible [Sobel, p. 11]. So together with his navigators, Sir Clowdisley guessed at their longitude and determined that they were safely west of Ile d'Ouessant (Ouessant Island), a small group islands off the west coast of France (near Brest). They were almost home. As they continued North, surprise and horror struck. They were headed straight for the Scilly Islands, a group of islands that dots the ocean near the southeast tip of England (about 20 miles from the coast). Large sailing ships are not easy to stop or turn. Collisions with the rocky islands commenced, and on October 22, 1707, four of Sir Clowdisley's five ships sank within minutes. Over two thousand lives were lost. England was in shock. Careful study of a map show this to be a clear problem with determining longitude. There is a safe northward passage west of the Ouessant Islands, and east of the Scilly Islands, but Sir Clowdisley's error in east-west position (longitude) caused him to be too far to the west, headed straight for the islands.

Later, it was learned that Sir Shovell's navigation was mostly correct; he was where he thought he was. It was *the map* that was in error[Dash, p. 8], which is an entirely new dimension to the longitude problem: bad maps. Cartographers didn't know longitude either and didn't really know where to draw the land masses!

1.4.3 George Anson

Commodore Anson left Portsmouth, England in September of 1740, on a trip to battle the Spanish in the Caribbean. The first leg of his trip was to round South America, to cross from the Atlantic to the Pacific. He crossed the Atlantic in decent shape, making it to Patagonia (the southern portion of South America) in May of 1741. This is where the trouble started. He passed through the Straights of Le Marie (south-eastern tip of South America, a 20 mile-wide channel, considered the beginning of a trip around the Horn[Dash, p. 69]) and rounded Cape Horn. A terrible storm then hit and batted him for about 58 days. After it calmed, he found himself at 60° south latitude (the tip of Cape Horn is at about 56° south latitude). He decided to run the parallel and head west for about 200 miles to round Cape Horn, then head north, where he hoped to find Juan Fernandez Island (and fresh food and water). As the sun rose on his northward path, he spotted land right away. Guess what? It was Cape Noir, a matter of miles from Cape Horn (all in the region of Tierra del Fuego). The storm, and his guess to head 200 miles "west" left him essentially where he started before the 58-day storm! The storm clearly blew him east of Cape Horn, well back into the

Atlantic. But without a way of determining his longitude, how could he have known this? But this was not the worst of it.

Scurvy was setting in, but the expedition persisted. Anson guided his ship northward to the 35th parallel, the latitude of the Juan Ferdandez Island. Now, the big question: was he east or west of the island? Does he turn left or right? Anson headed west for four days. After no land was sighted, he decided that he must be west of the island, so he turned the ship around and headed east. In two days land was sighted! But guess what? It was the Spanish-ruled, mountainous east coast of South America. Anson quickly turned his ship around again, and headed west. Later he acknowledged that he was probably within hours of Juan Fernandez Island when he made his first U-turn [Sobel, p. 19]. He found the island on June 9, 1971. His zig-zagging cost him another 80 lives. He was down to less than half of his original 500 crewmen.

1.4.4 La Salle

Rene Robert Cavelier, Sieur de La Salle was a Frenchman living in Ontairo, Canada[Andrews, p. 16]. He wanted to explore and build enough forts down the riverbanks of the central North American continent to keep the English contained to the region east of the Appalachian Mountains. His problem with longitude had nothing to do with fleets of great sailing ships, scurvy, or open ocean. He "simply" tried to navigate from Canada to the Gulf of Mexico, via the Mississippi River. There was land all around him!

Traversing the Mississippi River was his great dream. It was a mighty river and he wanted to conquer it. Despite many false starts and hardships, in canoes, he eventually paddled down 2,000 miles of the river and came to its mouth at the Gulf of Mexico on April 9, 1692. Here he planted a great Cross of Christ and the flag of France. He claimed the region on behalf of France and called it Louisiana; his new dream was to claim the whole middle and south-central portion of North America as belonging to France. La Salle was euphoric, but this would be his first, last, and only successful trip to the mouth of the Mississippi. La Salle returned to France, via Canada to secure King Louis's blessing for another expedition, again down the great Mississippi, and as the King said "to proceed and control the continent." La Salle wanted to establish a French colony at the mouth of the Mississippi. Now, here's where the trouble starts.

La Salle was ignorant of his longitude. He again descended the 2,000 miles south from Canada but missed a few key passes through to the Mississippi. He finally landed on the coast of Texas, about 400 miles west of the river's mouth (where he planted the cross and French flag two years earlier). There he and his men began to wander, looking for the mouth of the great river. Native Americans always pointed them west. La Salle begged his men, "just another day," "one more try"[Andrews, p. 17]. One day, 20 of his men agreed to one final push west, across a stretch of rolling prairie. Over that prairie, they found, another prairie. La Salle's men murdered him on the spot and left his naked body for the buzzards[Andrews, p. 18]. This all occurred near what is now Novasota, Texas. One of his ships, the La Belle, was (recently) discovered in 1995 in the muck of Matagorda Bay (on the coast of Texas), and has been the site of an archeological dig.

La Salle veered too far to the west during his expedition. This was an error in determining his longitude.

Activity

Get out a map of the world and chart the paths of Shovell, Anson, and LaSalle. Convince yourself their missions failed due to a lack of knowing longitude.

1.5 But still they set sail

It is perhaps not entirely fair to say that these early oceanographic voyages were run entirely blind. Experienced sailors were very smart, and their voyages were not entirely random. Ocean voyages had to have some hope of success from the start. As mentioned, it was possible to measure one's latitude by observing the height on the sun at local noon, or by observing the height of Polaris (the "north star") for those in the northern hemisphere. Also, there is certainly something to be said of seamanship, experience, and lessons passed from one sailor to another. There are also a few clues out there as to where you are and experienced sailors knew how to exploit them (and keep them secret from competitors)[Andrews p. 22].

1.5.1 Best guesses at one's location

The depth and composition of the ocean could be found by lowering and dragging a heavy object, and could indicate impending landfall. The ocean floor topography has three distinct transitions as the sea floor meets a continent: the rise, slope, and shelf. The ocean floor near the continental shelves is often covered with terrigenous sediments (rock and sand from land), or sediments mostly pushed from the land into the sea by glaciers during the last ice age. Ocean floor sediments also become increasingly fine with distance from a coast. Sand, like on the beach, is often limited to wave-agitated waters. Sunlight in shallower waters also leads to more near-coastal life forms. Floating vegetation or debris could also indicate impending landfall[Wikipeida]. Finally, sea water becomes less salty near land where it can mix with freshwater (run-off, rivers, glacier melt, etc.).

Ocean currents and prevailing winds in well traveled waters were somewhat categorized. For instance, when trying to round the Cape of Good Hope (the southern tip of Africa) from the Atlantic side, it was not a good idea to sail straight south, in sight of the west coast. It was best to swing way west, almost to the South American coast where a strong easterly current would swing you south-east to the Cape.

Study a map of the world carefully; it's oceans are littered with little islands (Azores, Canaries, Madeiras, Cape Verde). Short distance "island hopping" could accomplish a lot. Also, birds often fly near land and not too far into open ocean. Captains also released ravens [Andrews p. 22] and waited for their return. Cumulus clouds (puffy popcorn-like clouds) are formed in part by warm air that can rise from warmer land, which can be seen from far distances.

1.5.2 Deduced Reckoning or DR

There always existed a simple method for tracking both your latitude and longitude as you sailed. It was called Deduced Reckoning and here's how it worked. As a pilot, set your direction with a compass, say West. As you sail, throw a chunk of wood overboard. Suppose your 50 foot ship took 10 seconds to pass the wooden chunk. (Time this passage with an hourglass.) You would be going 300 feet per minute, or about 3.5 miles per hour, West. Suppose then you held this course for 2 hours. You'd be 7 miles west of your original position. The number of miles made in this segment were then added to those recorded in logs from a previous segment. This works for any given direction of the ship. You can imagine then a map with a bunch of little line segments, connected end-to-end, representing the progress of your journey. Deduced reckoning is the dead-simple practice of keeping an account of the coursed steered and distances run from a known departure and not breaking the "thread" until another known point is established[Schlereth, p. vii]. Sometimes deduced reckoning is called "dead reckoning." Most often it is referred to simply as "DR."

Another method worked as followsSobel and Andrews, p. 18]. A triangular log was attached to the end of a knotted line. Beginning at 60 feet from the log, knots were tied at regular intervals of 51 feet. When the log was cast over the side of the ship, the number of knots counted in a period of 30 seconds, as measured with a sand glass, would indicate the speed of the ship (this is how the term "knot" was adapted as the nautical measure of speed). The job required three people, one to hold the reel of rope, one to turn the sand glass, and one to count the knots.

With either technique, you then factor in ocean currents, fickle winds, and errors in judgment and you have one big erroneous estimate of your position. Hundreds of miles in error were typical. Even so, and even in modern day navigation, no captain ever gives up on his dead reckoning. Even if erroneous, it is nonetheless an approximate measure of your position at sea, which can then be compared to some other "fancy" technique you have, be it moon observations, accurate time, or even modern-day GPS. But technology can fail. GPS batteries can go dead or have selective availability, and accidents do happen. Whatever you do, never give up the DR[Schlereth, p. viii].

If you are familiar with high-school physics, than DR is just a huge vector addition problem. A vector, recall, is an arrow that points in your direction of travel with a length proportional to the distance covered. Your first position vector starts when you leave port. As you sail, you would construct new position vectors by multiplying your speed by the time spent moving at that speed, so long as your direction remained fixed. The tail of the new vector would be placed at the tip of the last vector. At any time, your current DR position is at the tip of the last vector you found.

Activity

Use DR log data to re-create a journey. Data consists of speed, direction, and duration data.

1.5.3 Milk your latitude

Finally, there is latitude (distance north or south of the equator), which one could make a good guess at, if not precisely determine, by the height of the Sun, stars, or length of the day. As we will see later, the height of the north star (Polaris) gives latitude directly; at least in the northern hemisphere. Similar star sightings also exist for the southern hemisphere. Early tables of the sun's motion would give latitude by observing the angle of the highest point of the sun (at local noon), then adding a single number (the sun's declination) read from the tables for that particular day of the year (see Chapter 2). So as long as skies were clear, latitude could be found and a ship could be guided to stay at a particular latitude for the duration of a voyage. This was called "running the parallel," and is how Columbus would have discovered a direct passage to the Indies, if the North American continent did not get in his way [Sobel p. 4].

Conceivably, an entire voyage could be planned by knowing only the latitude of your destination. As captain, your first task in open water would be to steer your ship north or south, onto the needed latitude. When done, you know that you are either directly east or directly west of your destination. All that's left then is to "run the parallel" to finish your trip. This sounds easy, and in some cases, maybe it was; in other cases it was not. Lacking knowledge of one's longitude, deciding which direction to sail, east or west, can be difficult (see the account above of George Anson's attempt to find the Juan Fernandez Islands).

Keep in mind though that well traveled parallels were known about, and often pirate or enemy ships would wait in ambush (see the 1592 story of the Spanish trip the *Madre de Deus* [Sobel, p. 15]). For example, the Maderia Islands, just South East of Portugal were at a latitude well known for a strong East-to-West current. So it was well trafficked for both noble and less-noble purposes (in other words: piracy was a thriving back then). On the contrary though, relying on running a parallel came with no guarantee of favorable winds, or ocean currents, and storms could make a "hugging" a parallel nearly impossible. All of these problems are, of course, related to only knowing half of your exact position.

1.6 So how does one find longitude at sea?

Hopefully the discussion above has convinced you of the importance of knowing longitude at sea. The all important quest to leave land and explore the earth by sea just couldn't continue in the manner presented. If it wasn't for the dreadful time aboard the galleons, then there was property loss, loss of life (10,000 died of scurvy during Queen Elizabeth I's reign)[Kirsch, p. 74], theft, and loss of national pride and confidence. So how did the problem of finding longitude eventually get solved?

1.6.1 Time and Longitude

Longitude is an arbitrarily referenced position on earth. The zero degree reference point (0°) was set by international agreement to be at Greenwich, England. As long as this reference stays fixed, maps and navigation throughout the whole world can be confidently stated.

Greenwich is at 0°. Moscow is at 37°42′ (East of Greenwich). San Luis Obispo, CA, USA is at 120.65° (west of Greenwich). Your ship's position can also be stated relative to these, or any other landmarks, such as your destination, or dangerous underwater rocks.

The key to determining longitude is that it's really the same as the passage of time. Why? The earth makes one complete rotation every 24 hours. This corresponds to 360° every 24 hours, 15° per hour, or about 1° every 4 minutes. So, suppose it's noon at Greenwich. The sun would be directly overhead; it would be directly over the line of longitude, or the meridian that sets 0° longitude for the entire world. Anyone on this meridian, anywhere between the north and south poles would see the sun directly overhead. Their time would be 12 noon, and they would have a longitude of 0°. One hour later, the sun would be 15° to the west. An hour earlier, the sun was at 15° to the east.

So, if you were located 15° west of Greenwich. You would see the sun directly overhead (your noon) one hour after the sun was directly overhead at Greenwich. If you were located at 45° west of Greenwich, you would see the sun directly overhead (your noon) 3 hours after noon happened in Greenwich. This is why the time is not the same all over the world, and is why we have time-zones. If you were east of Greenwich, you'd see the sun at its highest position in the sky at a time before it was seen at a highest point in Greenwich.

But the presentation of these numbers can also be reversed. Say you are in the middle of the ocean, with no land in sight. It is a clear day, and you observed the sun at its highest point in the sky (not hard to do). This once again means that it local noon where you are. The sun is directly over your meridian, or line of longitude. Everyone on your meridian, between the north and south poles, would observe the sun at its highest point in the sky. Suppose that at the same instant you see the sun at its highest point, you *magically* know that it is 4:30 pm in Greenwich, England. Your longitude must be (12 noon - $4:30$ pm)$\times 15° = 67.5°$ West. Presto! You just found your longitude at sea! It is noon for you now on your ship, and was noon in Greenwich in 4.5 hours ago. You must be 67.5° West of Greenwich.

Armed with your longitude, you can plot your location on a map. But, what is a map? A map is a picture of oceans and land, with their boundaries precisely drawn to their best known latitudes and *longitudes*. You may now compare your position with the map itself, and suddenly you will be aware of your location of your destination, or dangerous ship eating rocks. Sounds easy, huh? Do we really need a whole class dedicated to this? Yes! Why? Because what is the "magic" you needed to determine the time in Greenwich when you saw the sun at its highest point out on your ship? Think about it. How would you know or keep track of the time in Greenwich in the 1700s from a ship? In the 1700s, there was the need to sail across the great oceans, but there were no clocks, radios, cell phones, or satellites to communicate back to Greenwich.

This is the principle behind how keeping track of time is used to find longitude. For every hour your time is different to that at Greenwich, your longitude is an additional 15° farther from Greenwich. If your time is earlier than Greenwich, your longitude must be east of Greenwich. If your time is later than at Greenwich, your longitude must be wast of it.

People knew that **time difference would give longitude** well before the 1700s, and

had scientifically sound ideas for how to determine the time difference. The trouble was that no one knew how to implement any time-keeping methods aboard a ship at sea. It was the **implementation** of any of these ideas that was the problem. Of all of the ideas for finding longitude at sea, only two were worth pursuing: build a reliable sea clock (that could keep track of the time in Greenwich on a ship) or use the motion of the moon, the so called "lunar method," to figure out what time it must be in Greenwich. The best minds of the time all agreed one of these two methods would eventually solve this 300-year old longitude problem, if only one could implement one of these.

Activity

On a map, make a correspondence between time of day and longitude around the globe. Example: time here versus that back home.

1.6.2 Idea #1: The Lunar Method

One idea for finding longitude at sea was by observing how the moon moved across the backdrop of the stars at night. Against the backdrop of the stars, the moon moves approximately its diameter every hour. In other words, the moon moves rapidly enough that you can think of the moon and stars as a big clock in the sky. The stars are the "numbers" on the clock, and the moon is the "hand" of the clock. Using the moon in this way was called the "lunar method" for finding longitude. It worked as follows.

When the moon is visible, you'd find the angular distance between the moon and a star (or planet, or the sun, etc.). Suppose a given lunar distance X, say between the sun and the moon, is measured in Greenwich at 3 pm on a given day. If it's 3 pm in Greenwich, then a distance of X would be observed between the sun and the moon. Alternatively, if you are observing distance X in Greenwich, it must be 3 pm. But you measured the distance X thousands of miles away, in the middle of the ocean. It cannot be 3 pm since you are not in Greenwich. But, if you get out your tables, and look up your distance X, they will tell you at what time this distance is to be observed at Greenwich that day. Even though error-prone, such tables appeared in 1772, giving distances between the moon and various stars every 3 hours.

Now all you have to do is figure out your local time, and there are a variety of way to do this. For example, observing the sun's highest position to give 12 noon where you are; tables also existed allowing you to calculate your local time based on the height of the moon above the horizon. Suppose then you determine that you observed a distance X at 11 pm local time. You look up X in your tables and see that it occurred in Greenwich on that same day at 3pm. You can conclude that you must be $(11\text{pm} - 3\text{pm}) \times 15^\circ = 120^\circ$ East of Greenwich.

It was a popular idea because astronomy was *the* science back then, and this method had far reaching support. In other words, if someone were to dedicate time and resources to solving the longitude problem, astronomy was probably the way to do it. Observatories were built in Greenwich, Paris, and other countries for the sole purpose of tabulating the moon's motion relative to the stars, to be used to find longitude.

Problems with the moon

But the lunar method had its problems, primarily because nature itself seemed to be against using the moon to determine longitude (not just because of potentially cloudy skies, either). In short, the lunar method could be implemented only with tables outlining the exact position of the moon, on any given day of the year. This means some theory had to be able to predict the path of the moon on any day of the year, at any time of day. If this could be done, a book could be generated and printed, giving sailors something to take aboard their ships. It turns out though, that predicting the moon's orbit was so difficult that it was beyond the understanding of even Isaac Newton. Here's why.

The earth moves in an orbit around the sun due to gravity: the sun's gravity constantly pulls on the earth, to keep it in this perennial orbit. The earth's motion around the sun is not exactly a circle; it's actually an ellipse (which is a stretched out circle). The earth has a closest position to the sun (91 million miles in January) and farthest position(94 million miles in July). The physical theory of this so called "two body" problem (two bodies: the earth and the sun), is fairly simple. (It is often first taught in modern-day high school physics courses.) One part of the theory also states that the earth moves faster when it's closest to the sun, and slower when its farther from the sun.

Next, throw in the position of the moon, however, and now you have a three body problem: there is a gravitational pull between the sun-earth, sun-moon, and earth-moon. The moon moves in an elliptical orbit around the earth with unequal extreme distances and speeds (just like the earth). All of these motions put together make for a moon that moves across the sky in a somewhat non-uniform way. The non-uniform motion is barely perceptible, but it was enough to make predicting its motion nearly impossible, even for the great Isaac Newton, and made it even worse for finding longitude on earth; here's why.

Remember a theory is needed to predict the moon's path, so that extensive tables could be generated and given to sailors. The best theory Newton could muster differed from the best observations made by Edward Halley (of Halley's Comet) by as much as $0.06°$. In other words, if tables were generated for sailors, the numbers in it would be in error by as much as $0.06°$[Kollerstrom, p. 23 and 184]. The moon itself has a diameter of $0.5°$, so the tables would be in error by approximately 1/8th of the moon's diameter. Is this good or bad? Well, for finding longitude it turned out to be *horrible*.

The moon takes about one month (27.3 days) to orbit the earth, while the earth takes 1 day to make a complete revolution on its axis. Thus, observing the moon move in the sky from earth has 27.3 times more to do with the rotation of the earth than the orbital motion of the moon; and remember, it's the rotation of the earth which is linked to time of day, and longitude ($15°$ for every hour). Thus, if the moon's position is in error by $0.06°$, then a time (or longitude) measurement will be off by 27.3 times that amount, or $27.3 \times 0.06°$ or $1.62°$. For a ship at the equator, this would translate into an error of longitude of over 100 miles! No good! And this was the best theory available!

Those pursuing the lunar theory in the 1700s had this idea: let's start observing the moon for the next 18 years. The moon has a funny 18-year cycle to it, in that the sun, earth, and moon will return to almost an identical arrangement every 18 years. (Go look at the moon

tonight. If you look at it again, from the same spot 18 years from now, it will be in exactly the same position.) This is called the "Saros Cycle." It was thought that the observations could be used to correct Newton's theory, so his predictions could become better and better. And, if it could be done for 18 years, the corrections will hold over the next 18, and the next, and the next, since everything starts to repeat itself on this time scale. This is why the Greenwich observatory was built: to carefully observe the moon and correct Newton's theory. This is what Halley, Flamsteed, and Maskelyne did: carefully observe the moon and record its position.

So the moon had the potential to be the ideal universal clock, one that the entire world could read "time" from, but it just didn't happen; the moon's position just couldn't be predicted to a high enough precision. Instead it behaves like a clock with hands that always appear blurry; a clock that reads a slightly different time for everyone who looks at it. In fact, Newton wrote[Kollerstrom, p. 3]:

> *"The irregularity of the moon's motion hath been all along just a complaint of the astronomers; and indeed I have always look'd upon it as a great misfortune that a planet so near us as the moon is, and which might be so wonderfully useful to us by her motion, as well as her light and attraction (by which our tides are chiefly occasioned) should have her orbit so unaccountably various, that it is in a manner vain to depend on any calculation of an eclipse, transit, or an appulse of her, tho never so accurately made. Whereas could her place be but truly calculated, the longitudes of places would be found every where at land with great facility, and might be nearly guess'd at Sea without the help of a telescope, which cannot be used.*
>
> *The irregularity of the Moon's motion depends on the attraction of the sun, which perturbs the motion of the moon and makes her move sometimes after and sometimes slower in her orbit; and makes consequently an alternation in the figure of that orbit."*

It other words, the moon, as familiar to us as it is, has no simple use as a clock and no no simple theory to predict its motion (computers do it nowadays). Nevertheless, back in the 1700s, the moon remained a leading idea to solve the longitude problem. It was a scientifically sound idea to solve the longitude problem, and if you couple this with the deep-rooted beliefs in astronomy, you would get motivation enough to continue pursuing it as a solution. The method was eventually shown to work, and was even used on many ocean voyages of the famous Captain Cook. It was a difficult technique to use though, and was never widely adopted.

1.6.3 Idea #2: Keep accurate time as sea

The second idea for solving the longitude problem was the most concise. If you could keep accurate time at sea, you could determine your longitude. Period. It was sometimes called the "two clock method."[Kollerstrom, p. 22]. If one had two clocks, one on local time and

one on Greenwich time, you could simply observe the clocks, take the difference in time between them and from that your longitude could be determined.

Despite the simplicity of this idea, a clock (in the 1700s) that could keep accurate time in a harsh marine environment simply did not exist. The idea was sound, but it needed an implementation. Never mind the rocking motion of the ship. There's humidity, changes in the earth's gravity, pressure changes, and temperature changes; all of which wreak havoc with a clock's internal time keeping mechanisms.

There were philosophical problems with a clock too. You must think for a moment what a clock is. It's a box that you are to read and use numbers from without having to understand *how* it works. (Do you know how your wristwatch works?) People back then just didn't believe in such "magical" devices." People were suspect of them. Trust a ship and the longitude calculation to a magical machine? No way! In the 1700s, people had a hard time believing in man-made machines. The sun, stars, and planets, however, they were God's creation; you can trust them.

So anyone who decided to pursue the longitude problem in terms of clock making would, in addition to the construction of a reliable clock, have to battle the prevailing attitude that a machine would not and could not solve the longitude problem. Even the great and influential Isaac Newton was one of the harshest critics of the clock method. Keep in mind though, that a clock did eventually solved the longitude problem. It eventually won out over the lunar method, but only after a lifetime of work by a man named John Harrison. Even to this day, time (not the planets) plays an integral role in modern-day navigation by GPS.

Activity

How you might keep time at sea in the 1700s?

1.6.4 The moons of Jupiter

Some discussion is warranted at this point on using Jupiter, and its moons, as a method for finding Longitude.

To some degree, Jupiter stands alone as a planet. It is the largest in our solar system, and at least in the northern hemisphere, very easy to spot in the night sky many months of the year. Even the smallest, cheapest, telescope will reveal the apparent visible disc of Jupiter, color bands on its surface, and many small moons orbiting around it. Further, Jupiter's moons are all in the ecliptic plane, which is the plane that contains both the Earth and Jupiter. The many moons around Jupiter are constantly appearing in front of the planet, casting shadows on its surface, then disappearing behind it, before appearing again on the other side. And they do this with highly reliable regularity; the perfect clock! For example, here is what the moon Io did on Tuesday, August 28, 2007[skyandtelescope.com]:

```
00:50 UT, Io's shadow begins to cross Jupiter.
01:44 UT, Io ends transit of Jupiter.
03:04 UT, Io's shadow leaves Jupiter's disk.
20:40 UT, Io enters occultation behind Jupiter.
```

UT in this sample data stands for "Universal Time," which is the same as the time at Greenwich. From the data, Io alone had a pretty busy day on very short time scales (on the order of hours). Also, to within a high degree of precision, due to Jupiter's large distance from the earth, these times are universal, as observed everywhere on earth. This means if Io begins to cross Jupiter at 1:44 UT, as observed near Hawaii, it will very nearly also be seen doing so across the world, in the Indian Ocean. Thus anyone familiar with Jupiter's moons could spot it using a small telescope, and tell (from tables on the moons' motions) what time it must be in Greenwich. As with the other methods, after determining one's local time, the difference in the two times would give longitude.

As great as all of this sounds, the required observations were simply impossible from the deck of a rocking ship. This knocked this otherwise simple solution out of the realm of practicality at sea (again with the harsh sea conditions). It was flatly rejected as a solution to the longitude-at-sea problem for this reason (a problem with implementation). People tried building gimbal mounted viewing chairs, etc. but nothing worked. Even on solid ground, the earth's rotation will cause Jupiter to move out of the view of the telescope in a matter of minutes. Jupiter is mentioned here because it was the most successful way of determining longitude *on land*, for the sake of map making. It was also used to verify the success of John Harrison's clocks, once landfall was made at a given destination. (Side note: it is thought the moon's of Jupiter guided the 1804-1806 expedition of Lewis and Clark[AJP 55(2) Feb 1987, p. 103].

1.7 The Longitude Prize

Sir Clowdisley Shovell's accident (see above), right on England's doorstep, more than any other disaster, strengthened England's resolve to solve the problem of longitude once and for all. The Longitude Act was established 7 years later, by the English parliament, offering $12 million (in today's value) for anyone who could find a way of determining longitude at sea.

A committee on longitude was assembled. It put aside favoritism over particular solutions, or even British over foreign solutions[Sobel, p. 53]. It simply urged the English Parliament to welcome potential solutions from anyone, and to reward their success handsomely. With this urging, Queen Anne issued the "Longitude Act of 1714" as follows:

- $\pounds 20,000$ for a method that would determine longitude to 1/2 of a degree.
- $\pounds 15,000$ for a method accurate to within 2/3 of a degree.
- $\pounds 10,000$ for a method accurate to within 1 degree.

One degree of longitude on a great circle is 60 nautical miles (68 geographical miles). Note that 1/2 of a degree (which would win the most money) is still 30-ish miles. This can still easily crash a ship. Such relatively "relaxed" constraints, though, clearly demonstrate the pathetic and utterly desperate state of navigation at this time in history[Sobel, p. 54]. Compare this to modern day GPS which will tell you where you are to 30 cm or so. The 1/2

degree probably would NOT have helped Shovell (above), but most likely would have saved Anson and La Salle. So that was it then. Three-hundred years of efforts by scientists could not solve the longitude problem, so England decided to motivate the public at large

It is interesting that 300 years ago, a government simply "had enough" of this "nonsense" with longitude, and would offer a so called "inducement prize" with such a huge award. Inducement prizes are not for recognition or for "further study." They are not research grants. They are meant to motivate finding nothing other than an actual **solution** to an engineering-type problem, that usually extends what humans are able to do. Such "prizes" have been offered before. The Manhatten Project put the most brilliant scientists together to build an atomic bob. The X-prize foundation (xprize.org) had such an award for putting a spacecraft into outer space twice in 14 days. They are now offering a \$10 million for a 100 mpg automobile. Charles Branson (Virgin Airlines) is offering \$25 million for someone to get CO_2 out of the atmosphere. John F. Kennedy's famous speech put humans on the moon. There is talk for forming a "Manhatten Project" to come up with some "non fossil fuel" energy source. What global problem are you passionate about? Do you think a large inducement prize would solve it?

1.8 The Remarkable John Harrison

We close this chapter with one of the most remarkable aspects of the "story of longitude," which is also a testament to its central figure, John Harrison. The Longitude Act stipulated that to win the £20,000, a method of determining longitude-at-sea would enable one to do so to within 1/2 of a degree. This translates to an east-west distance of about 34 miles ($0.5° \times \pi/180° \times 3963$ miles=34 miles) at the equator. No person, for hundreds of years was able to solve the problem of finding longitude-at-sea, let alone to 34 miles. A solution eluded everyone in the history of the world, even those we consider as preeminent scientists today, such as Newton, Euler, Halley, Huygens, Galileo and Leibniz. None of them could solve the longitude problem, and neither could anyone else. Newton even stated that he didn't believe a clock would ever solve the longitude problem. So what business was it of Harrison's, a uneducated carpenter from the English countryside, to even try?

Harrison's belief in his ability to build a clock that would solve the longitude problem is astonishing. To find longitude to within half of a degree meant his clock would need to be true to with 119 seconds. (Why? The earth rotates at 15°/hour or 0.0042°/second. Thus 0.5° translates into 0.5°/0.0042°/second = 119 seconds. More on these numbers later.) The Act stipulated that the official test of any proposed solution would be done on a voyage from England to the West Indies (essentially the island of Jamaica), which one-way, was typically a two-month voyage (60 days). This meant Harrison thought he could build a clock that, at sea, would lose no more than 119 seconds in 60 days or about 2 seconds per day! The clock he built, that eventually won the longitude prize, after the full round trip lasting 147 days, lost 1 minute and 54.5 seconds[Andrewes, p. 243], or about 0.78 seconds per day. In other words after a lifetime of work, Harrison's produced a clock with twice the performance required of the Longitude Act!

The documented history of Harrison's work clearly demonstrates his passion, determination, and ability to solve problems as they arose. If he needed something for this clock that didn't exist, he would *invent it*. And he did this over and over again. All of his clocks contained mechanisms either previously unseen, or implemented in quality previously unseen for that time period, and Harrison's work had a *huge* impact on navigation. Within years of Harrison's successful marine clock, the marine timekeeper was standards on all sea-going vessels. Note that keeping time at sea to find longitude was not Harrison's idea; the idea was known for hundreds of years. Harrison was the first to *implement* the idea, in terms of a functioning marine clock.

If one looks back at the history of navigation, it is clear what a "quantum leap" the realization of a reliable marine chronometer truly was. From 1762 (when Harrison's winning clock was tested) until the early 1900s, there was no fundamentally new development in the science of navigation. Sure marine clocks (and sextants, compasses, navigation tables, etc.) got better, cheaper, and more widespread, but the clock still ruled the sea. It wasn't until the 1920s and into the 1940s that time signals were broadcast world-wide using radio waves. The GPS system wasn't turned on until the early 1990s. So the world has seen merely two new developments (radio and GPS) in maritime navigation since the time that Harrison gave the world a reliable marine chronometer.

Lastly, if you are looking for a modern-day problem like longitude in the 1700s, then look no further than "fusion," which would allow us to derive cheap and clean electrical power from ocean water. We know that fusion will work, but we do not know how to implement it. No one has been able to do it for about 50 years now. There are striking societal and political similarities between the two problems.

1.9 On we go...

This chapter was meant as a gentle introduction to the "problem of longitude" and some of the baggage attached to it. In the following chapters, we'll dig deeper into many aspects of the longitude problem, including:

Longitude and latitude. What are latitude and longitude, and how do they work? Why are these two numbers used to locate a position on earth? We'll peek a bit into the structure of the galaxy and solar system and see how the sun, stars, and planets can be used to determine latitude and longitude (this is called "celestial navigation").

John Harrison. This is the man who heard about the "Longitude Prize" when he was about 14 years old, and decided to dedicate his entire life to finding a solution. He was passionate about using the "clock method" for finding longitude, and that is what he pursued. For Harrison, the materials and technology to implement a sea clock just didn't exist; but one by one he invented what he needed to make it all work. He eventually won the prize by building a clock that could keep accurate time at sea, within a very tight precision, as dictated by the Longitude Act. Historically, he is credited with solving "the problem of longitude at sea." We'll examine the technical

aspects of Harrison's work and how he pioneered several "clock mechanisms" that nature provides to ultimately solve the longitude problem.

Time. Time itself is a fascinating subject that always stands on its own, in the context of science. We'll take a close look at time, including some of the ideas about time discovered by Albert Einstein.

Modern Navigation and Timekeeping How is navigation in our modern world done? How is time kept today? We'll take a look at universal time broadcasts, Global Positioning Satellites (GPS), and modern methods for keeping time.

Chapter 2

Excerpts from “The Quest for Longitude”

The following two chapters are reprinted (with permission) from a volume titled “A Quest for Longitude,” by. W. H. J. Andrewes (http://www.amazon.com/Quest-Longitude-Proceedings-University-Massachusetts/dp/0964432900). The first chapter, “Even Newton Could be Wrong” details Harrison’s lifelong work in solving the longitude problem. The chapter has a dubious title, but in the longitude problem, indeed the great Isaac Newton was wrong (he did not think a clock would be able to solve the problem). The chapter also discusses the sea trials and the politics Harrison faced. The second chapter, “The Timekeeper that won the prize” discusses Harrison’s fourth clock, than eventually won the longitude prize, and ushered in a whole new era of precise navigation.

The author wishes to thank the Department of the History of Science, Collection of Historical Scientific Instruments, at Harvard University, for permission to reprint the following two chapters.

Chapter 3

The Earth, Moon, Stars, and Navigation

This chapter will examine the meaning of latitude and longitude and also address why finding latitude can be done by observing the stars, but finding longitude cannot. In doing so, we'll look at the spherical structure of the earth, and how it sits in the solar system, in relation to the sun and the moon. We'll also look at how and why the stars seem to move as they do, as viewed from earth. In the last part of the chapter, we'll see how the moon, sun, and stars can be used to tell your local time of day and even your position (this is called celestial navigation). You'll find a lot of introductory astronomy in this chapter.

3.1 Mapping the Earth

The earth is a sphere, or nearly so. It bulges some at the equator because it is spinning, but that won't concern us here. It orbits in a near circle around the sun at a distance of approximately $93,000,000$ miles. It has a radius of

$$6,350,000 \text{ km} = 6.35 \times 10^6 \text{ km} = 3949.9 \text{ miles}, \tag{3.1}$$

and makes one complete revolution on its axis in 24 hours, or once per day. The fact that it takes 24 hours to complete one revolution is critical to the finding your longitude. The number 1 revolution (rev) in 24 hours can be written as

$$\frac{1 \text{ rev}}{24 \text{ hours}} \tag{3.2}$$

or 0.0417 rev/hour. Since there are 360° in one revolution the earth also rotates through 360° in 24 hours or

$$\frac{360°}{24 \text{ hours}} = \frac{15°}{\text{hour}}. \tag{3.3}$$

So the earth rotates through 15° in one hour. Try to remember this "15° per hour" rule. It is a fact about the earth that appears everywhere in the context of finding longitude, and

here you see where it comes from. This rule alone was Harrison's guiding principle for his belief in a clock as the solution to the longitude problem. Why? Because if you knew the time difference between your ship's location and Greenwich, England, you'd multiple this time difference (in hours) by 15°/hour to compute your longitude. This also translates to about

$$\frac{15^\circ}{\text{hour}} \times \frac{1\text{ hour}}{60\text{ minutes}} = \frac{1^\circ}{4\text{ minutes}} \tag{3.4}$$

or 1° every 4 minutes.

The earth is in a nearly circular orbit around the sun, and it takes about 365 days for it to make one complete orbit around the sun. So the earth spins around its axis 365 times faster than it moves around the sun. This means if you are tracking the sun's position from earth to make a longitude measurement, the earth's rotation by far dominates the apparent motion of the sun. We say apparent because the sun itself isn't moving in any way that is noticeable to us here on earth; it just appears to move because the earth is rotating. The same goes for the stars, and to some extent, the moon (see below). Thus the earth's rotation at 15° per hour seems to make the sun, stars, and moon also move across the sky at 15° per hour. Can we prove this? Yes! Try to observe the sun relative to some fixed point, perhaps the top of a tree, or edge of your window. Next, look at it one hour later. It will have appeared to move 15° to the west. You can try the same for stars at night. Remember that 15° is the angle subtended between your pinky and index finger with your arm stretched out.

Activity: One hour after last activity

Re-measure the position of the sun relative to the same fixed object. How far has the sun moved in one hour?

3.1.1 Locating yourself on the earth

Take a look at the globe (or sphere, or ball), meant to be a model of the earth shown in Figure 3.1. It has a north (top) and south (bottom) pole, just like the earth. Next, imagine that you are located at the spot where the black dot is. How would you uniquely describe your position? By "uniquely" we mean in such a way that there is *no ambiguity* as to your position; there's no way you or anyone would get confused about where you are, and you're completely sure that you are not about the hit a bunch of rocks with your ship.

The globe is conspicuously bare. There are no landmarks or reference points, but such is the lot of the earth. Oceans are particularly barren, as are deserts to some extent. Forests can look the same in all directions (as can densely populated cities). What we need is a good model for the earth; a good way of mapping the earth. We need something that will allow us to accurately plot our position relative to everything else on earth, so we can know where we are, in relation to everything else.

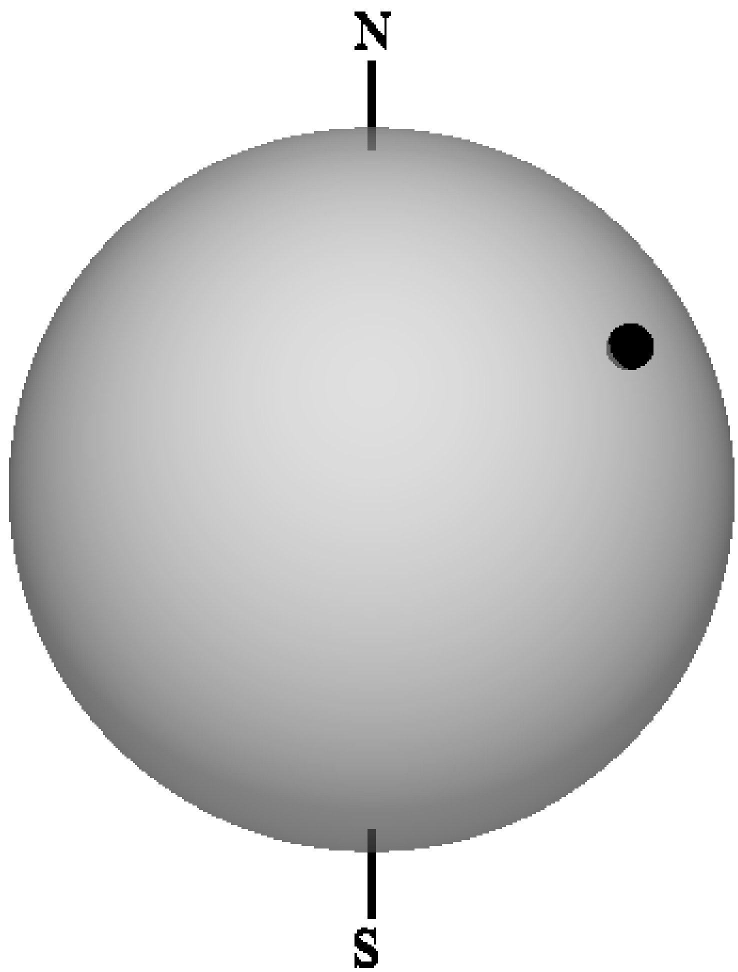


Figure 3.1: Model of the earth, including north and south poles. The dark spot is a location on earth that we'd like to describe.

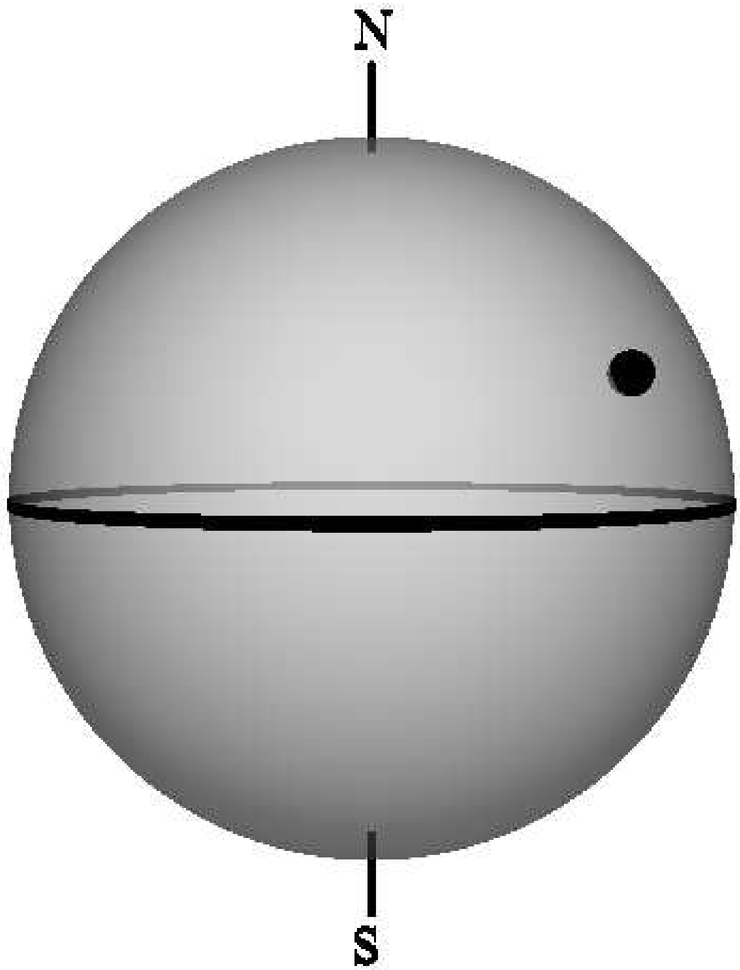


Figure 3.2: Model of the earth, including north and south poles, and the equator.

The Equator and Latitude

To start, let's draw a line on the bare globe from Figure 3.1 that might seem natural; a line all the way around the globe, that evenly splits the globe into two equal halves, as shown in Figure 3.2.

You probably know that we think of our earth as having an imaginary line around it called the equator. Of course the line isn't real, but we associate it with the middle of the earth. Living in the U.S., we are always north of the equator, and like to travel toward the equator for nice warm vacations (Cancun and Acapulco, Mexico, Costa Rica, etc.). Physically, the equator isn't just a made-made map-making feature. It actually serves as a logical reference point for how the earth sits in the solar system.

Imagine a plane, like a huge sheet of paper, containing both the earth and the sun. The earth orbits in a circle around the sun. The equator of the earth is tilted out of this big sheet

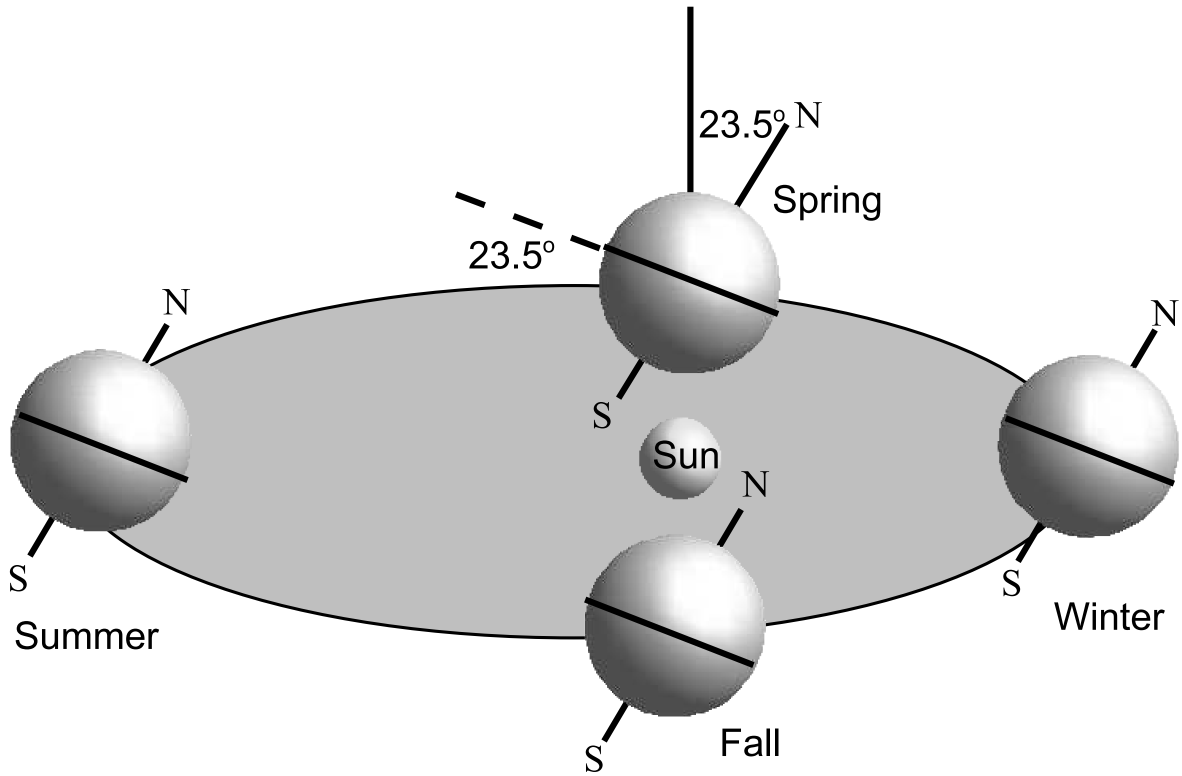


Figure 3.3: The earth's 23.5° tilt.

of paper by about 23.5°, as shown in Figure 3.3. In this figure, imagine the earth orbiting counterclockwise.

The plane (or huge sheet of paper) would be the black, solid oval (or elliptical) path the earth takes around the sun, filled in with the light gray color for emphasis. The black path is called the "ecliptic" and will come up again later. As a consequence of this tilt, the sun's path through the sky, as observed from earth, is quite bizarre over the course of a year (it's called an "analemma"–look it up.). The sun does not, for example, always shine directly on the equator, the way it would if the earth was not tilted. Also, the tilt, as you may know, is the reason for our seasons.

For example, in the northern hemisphere, summer occurs when the earth is farthest from the sun, but pointed toward the sun, and winter occurs when it is closest, but pointed away from the sun. The earth is the same distance from the sun during fall and spring (which is why they are called equinoxes), but on opposite sides of the sun, heading into or away from its "tilted toward" or "tilted away" position. As spring turns to summer, the sun's highest position in the sky (at noon) appears to head north over the earth, until it reaches the Tropic of Cancer, in the middle of summer. As summer turns into fall, the sun's noon position heads south again, passing the equator in the fall, eventually making it as far south as the Tropic of Capricorn in the middle of Winter in the northern hemisphere. The cycle

then repeats itself, as the sun's noon position crosses the equator as it moves north again, in the spring.

So the equator is a natural, and primary earth-bound reference point that links us to the sun's important influence on the earth. Even if humans didn't invent the word "equator" and draw it on globes, the earth would still have this funny, imaginary line that the sun "danced around" over the course of a year. When places above the equator are tilted toward the sun, they would still be warmer than the places below the equator that are tilted away from the sun at the same time, etc. Also, the distance of a location due north or due south of the equator is called "latitude." This *was not* the quantity that caused 300 years of trouble for sailors (*longitude* caused all of the trouble). As you might guess, due to the sun's involvement in all of this, one's latitude on earth can be found by observing the sun's highest position in the sky at one's local noon, and sailors well knew this.

As shown in Figure 3.2, an equator drawn on a barren globe, allows us to begin to develop a reference point for the location of the black spot in Figure 3.1. For instance, in Figure 3.2, we can at least say that our location is in the northern part of the globe. In fact, the equator does a bit more than this. It serves as the world-wide reference point for north-south navigation; any point on the equator is said to have "Latitude=0°."

Longitude and the Prime Meridian of the World

Let us now allow our figure of the barren globe to evolve even further, adding another reference line as shown in Figure 3.4.

Note carefully the properties of the new, vertical-ish line shown in Figure 3.4. It starts at the north pole and goes straight south, ending at the south pole. A line drawn on the earth that passes directly from one pole to another is called a "line of longitude," or a "meridian."

Unlike the equator, a line of longitude has no physical significance at all. It isn't related to the sun or the moon, or any other natural phenomena; a line of longitude is purely a human-made map-making tool. The lack of any such significance is perhaps why longitude ended up being the hard one to find; it can't be linked to anything going on, on earth or in the sky! So why not just ignore it? Why use it? It is only a figment of man's imagination! Two reasons: 1) there isn't a better way, and 2) longitude is critically important in navigation, since it is used to precisely locate a point's east-west location (recall from the last section that latitude handled the north-south location).

Finally, as long as longitude had to be completely invented by humans, they decided to set the primary line of longitude to run from the north pole, through Greenwich, England, then on to the south pole. This line is called the "prime meridian of the world" and arbitrarily sets the reference to which all world-wide longitude is measured. Any point on the prime meridian has its "Longitude = 0°."

So our barren globe now contains two key reference lines. The equator, which sets Latitude = 0° and the prime meridian of the world which sets Longitude = 0°. With these two reference lines set, we are now ready to do some navigation. Just where is that pesky black spot?

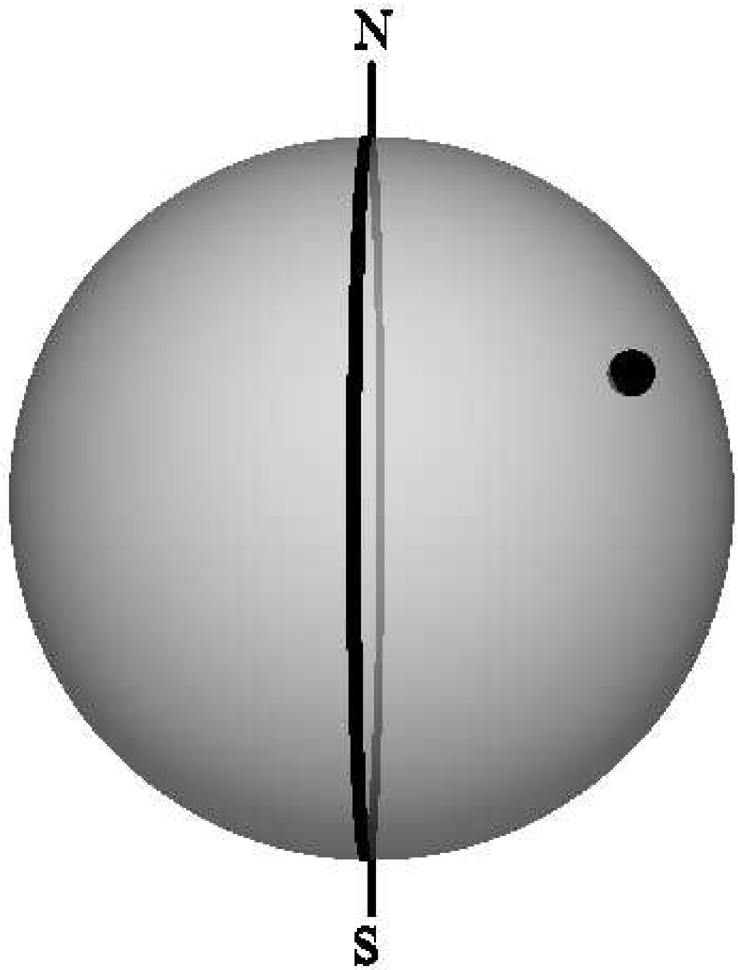


Figure 3.4: Model of the earth, including north and south poles, the equator, and a line of longitude.

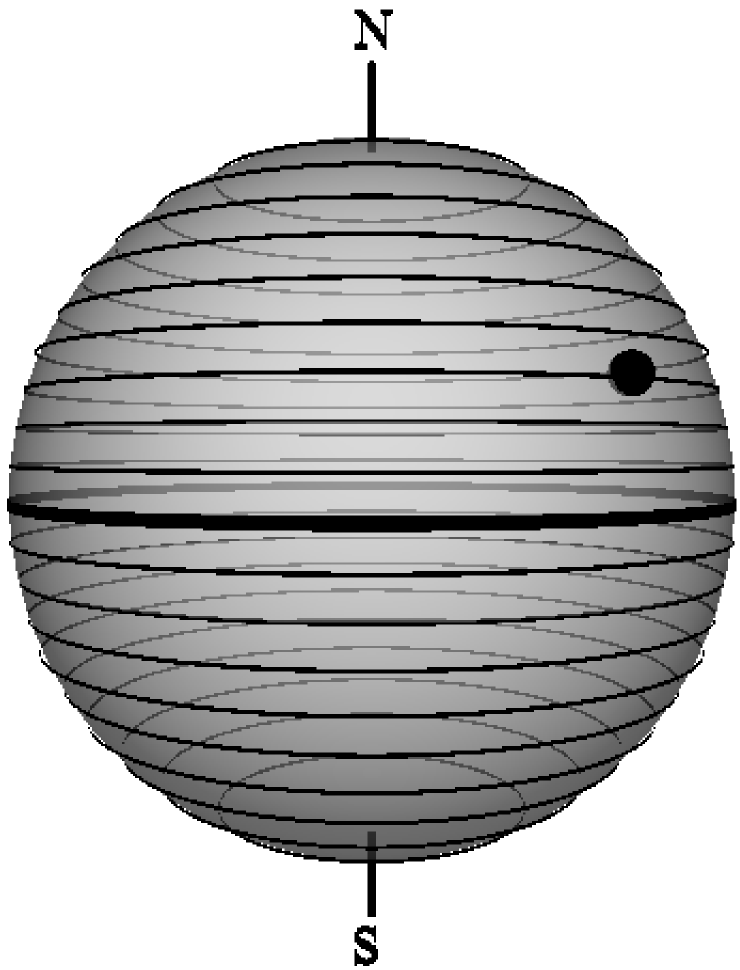


Figure 3.5: Model of the earth, with many lines of latitude drawn both north and south of the equator.

3.1.2 Locating points with latitude and longitude

Let's see now how lines of latitude and longitude can be used to pin-point any object on earth.

Lines of Latitude

Again, we let our globe evolve. Now we'll put lines of latitude on it, every 10° north and south of the equator, as shown in Figure 3.5.

Notice how a line of latitude just happens to pass through our seemingly "un-locatable" black spot! (This was done on purpose for our illustration here, but you can imagine that

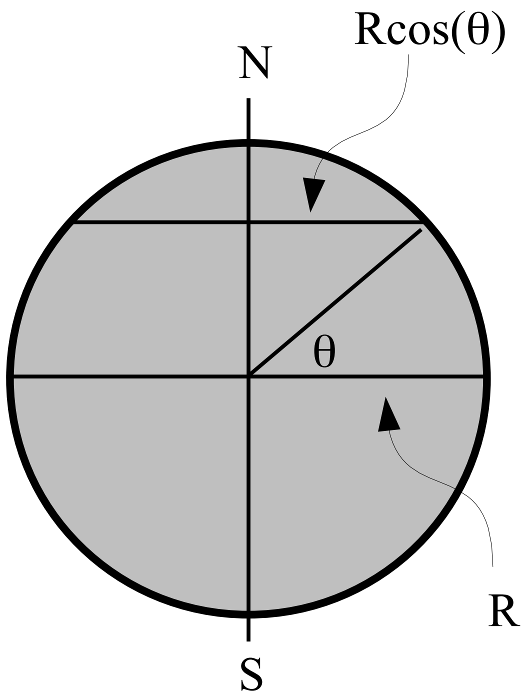


Figure 3.6: The radius of the earth at a latitude of θ is $R\cos\theta$.

there's *always* a line of latitude to pass through any point on a globe.) As mentioned, the lines of latitude are placed every 10° beginning at 0° on the equator. It looks like the 30° line of latitude passes through the black spot. There now, we've specified half of the location of the black spot; it is 30° north of the equator.

Had the spot been south of the equator, we still would have started counting with 0° at the equator and would have concluded that the spot was "so many degrees" *south* of the equator.

Latitude values then always start at zero on the equator, and increase and come to a maximum value of 90° at the north or south pole. Also notice something about circles of latitude. They are circles that are all parallel to the circle of the equator, but they get progressively smaller in radius as the latitude approaches the poles, which have a latitude of 90°. That is, circles of differing latitude do not have the same radius. In fact, the radius of a circle of latitude on the earth changes with the cosine of the latitude. So if the radius of the earth is R, the radius at at latitude θ will be

$$R(\theta) = R\cos\theta, \tag{3.5}$$

as shown in Figure 3.6.

This means if the circle of the equator has a radius of 3949.9 miles, the circle at a latitude (north or south) of 45° will have a radius of 2,793 miles. And remember, the nautical mile from Section 3.2.3? One travels 60 miles by traveling 1° on the equator. This distance drops

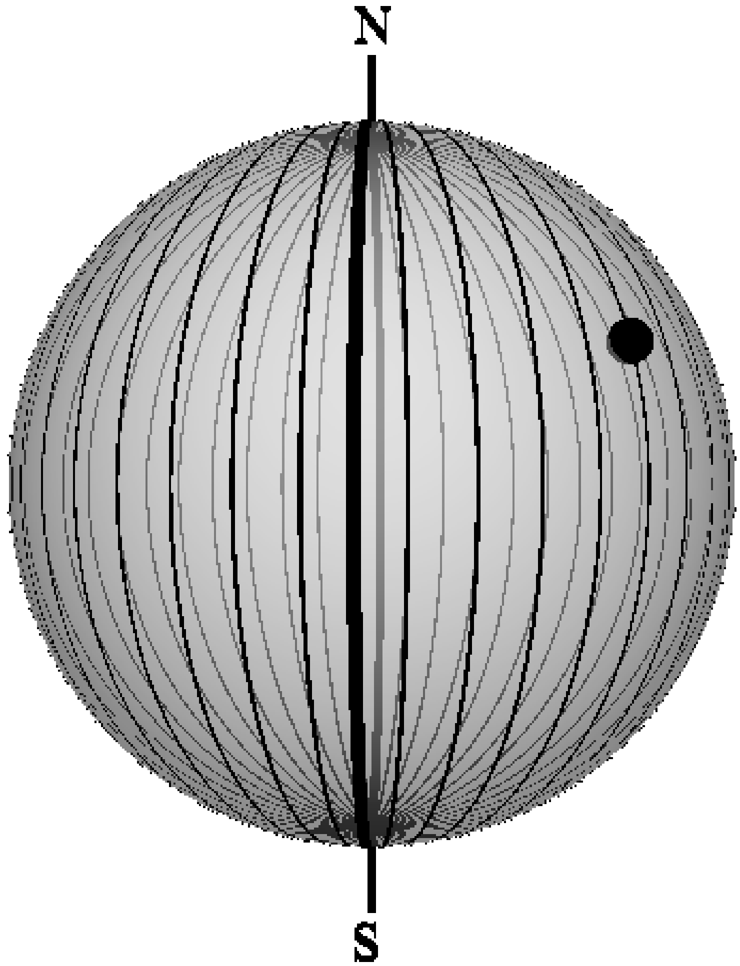


Figure 3.7: Model of the earth, with many lines of longitude drawn. The thick line of longitude is the prime meridian.

to 43 miles at a latitude of 45°, and 10 miles at a latitude of 80°.

Lines of Longitude

Now, we'll remove the lines of latitude and draw lines of longitude onto the globe, as shown in Figure 3.7.

We've drawn lines of longitude, starting at the prime meridian (the thickest line of longitude), every 10°. It looks like our black spot is along the line of longitude 50° east of the prime meridian. Next, let's combine the lines of latitude and longitude and see how things look (the astute reader will notice how we just precisely located the black spot!).

Put it all together now

See Figure 3.8. The earth is covered in a "checkerboard" made of lines of latitude and longitude. The thicker lines forming the equator and prime meridian are clearly visible as the world-wide reference of $0°, 0°$.

To finally locate our black spot, we simply see what lines of latitude and longitude pass through it. From the figure we see that

$$\text{Latitude} = 30° \text{ North} \tag{3.6}$$

and

$$\text{Longitude} = 50° \text{ East.} \tag{3.7}$$

So, that's it. Our black spot has been located, and the same latitude/longitude grid can be used to plot or locate *any* spot on earth. As you might guess then, the trick to safely navigating the seas is to locate your own latitude and longitude, and then plot it on a map, relative to other known features on earth, such as land, your destination, or known pirate hang-outs. Up to the late 1700s, sailors could confidently plot their latitude, but their longitude was only a guess. In other words, they were trying to locate their position on earth while only knowing half of the information they needed.

Activity

Use Google map application to find the latitude and longitude of 25 places on earth. Convert all numbers to degree, minute, and second format.

3.2 Angles in space

We're going to ease into the process of using the stars at night to find your position on earth. To do so means you have to assign a little meaning to what you see up there, other than "look at all of the little points of light." Perhaps the most basic measurement is to try and state how far apart two celestial objects appear to be. For this type of "observational astronomy," the *actual* distances (in miles, kilometers, etc.) are of no use (and are nearly impossible to know anyway). Instead astronomers use angles to classify celestial distances. The meaning of these angles is the same as from grade-school math. There are 360° in a circle, and a nice straight wall in a house would sit at 90° relative to the floor.

3.2.1 Use your hand as an "angular ruler"

Now, take a look at Figure 3.9[Harrington, p. 11]. It turns out that most people's hands and arm lengths grow in a particular proportion, independent of gender. As the figure shows

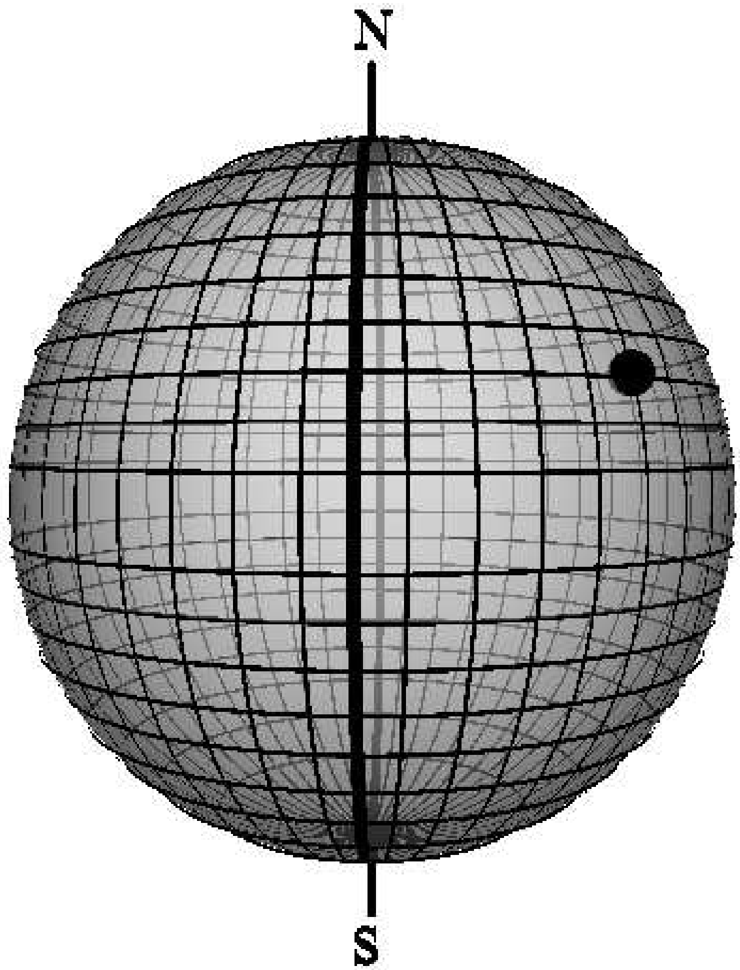


Figure 3.8: Model of the earth, with many lines of latitude and longitude drawn. The thick lines are the prime meridian and equator. The grid-like pattern makes it obvious now how to located a spot on the earth.

that your little finger held at arm's length subtends an angle of about 1°. Hold up your little finger and extend your arm. Close one eye and look directly at your little finger with the other eye. What object around you appears to be the same size as your little finger at arm's length? Whatever object you find can be said to have a "size" of 1°. Given the distance between you and the object, the object subtends an angle of 1°.

Groups of fingers, or even different hand positions, all at arm's length, as shown in Figure 3.9 represent different angles. An example of using your hand in observational astronomy is shown for the Big Dipper, which in the northern hemisphere, you can find to the north. The opening of the Big Dipper's "cup" is about 10°. From the tip of the "handle" to the bottom edge of the "bowl" is about 25°, and so on. Remember all of these angles are with your hand at arm's length, and you can probably imagine how your hand can be used to state the angular distance between any two objects in the sky. Verify these: The full moon and sun both subtend an angle of about 0.5°. And, if you are able to find a suitable reference point (a tree, building, etc.) you can verify that the sun and stars all move from east to west at about 15° per hour. For a trickier measurement, see if you can verify that the moon moves about 0.5° relative to the stars in the background every hour.

How do these angles relate to a circle, and the 360° is contains? Imagine that you are the center of a circle and that a line goes from you, through your extended arm, to one of the two objects that you are finding the angular distance between. This is shown in Figure 3.10.

Let the circle be drawn now from this first object, through the second one, then all the way around you and back to the first one again from the other side of it. This is the 360° circle that you are familiar with. The smaller angular distance you've measured is always a fraction of this big circle. If you measure a 25° angle between two objects, then the objects take up about 6.9% of that big circle. The radius of the circle is always how far away the objects appear to be. For observing stars, we say they are all on "the sky" which can be thought of as a big sphere surrounding the earth, containing all of the stars that we see.

Note also that these angles, and your hand-technique for measuring them, can be used to measure the "height" of objects above your horizon. We'll discuss later how the height of the sun at noon, or the height of Polaris (the north star, seen at night) will tell you things about your local time or even your latitude. To measure such a height, do the same thing as before, but put one side of your hand (or fingers) in touch with the horizon. This doesn't yield as accurate an angle as we'd like, but it's a start and we'll make a small instrument to help us get better results for the height measurements later on.

And that's it for measuring "distances" between objects in space! The distances are really angles, but who cares? The larger the angle, the larger the distance between the objects, which you can use in your discussion or analysis. Also, notice that you are stating [correct] distances without every knowing an actual distance. You just don't need to!

3.2.2 Angular Measure

As mentioned, the size and scale of astronomical objects are specified by angles. Here are some facts about these angles[Chaisson, p. 9]:

- A full circle contains 360 arc degrees (360°). Therefore, a half-circle stretches from the

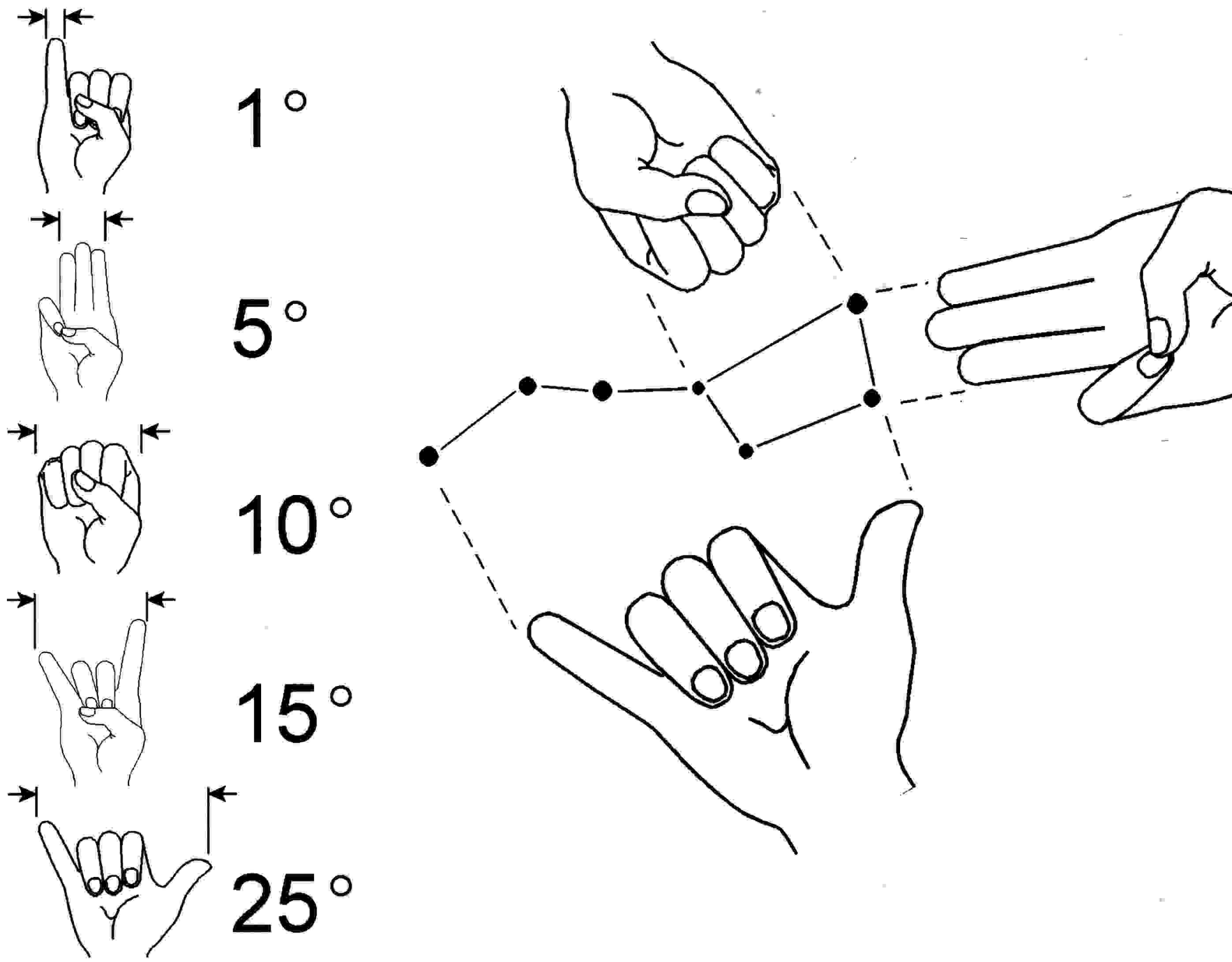


Figure 3.9: How your hand, held at arm's length, can be used to estimate angles in space.

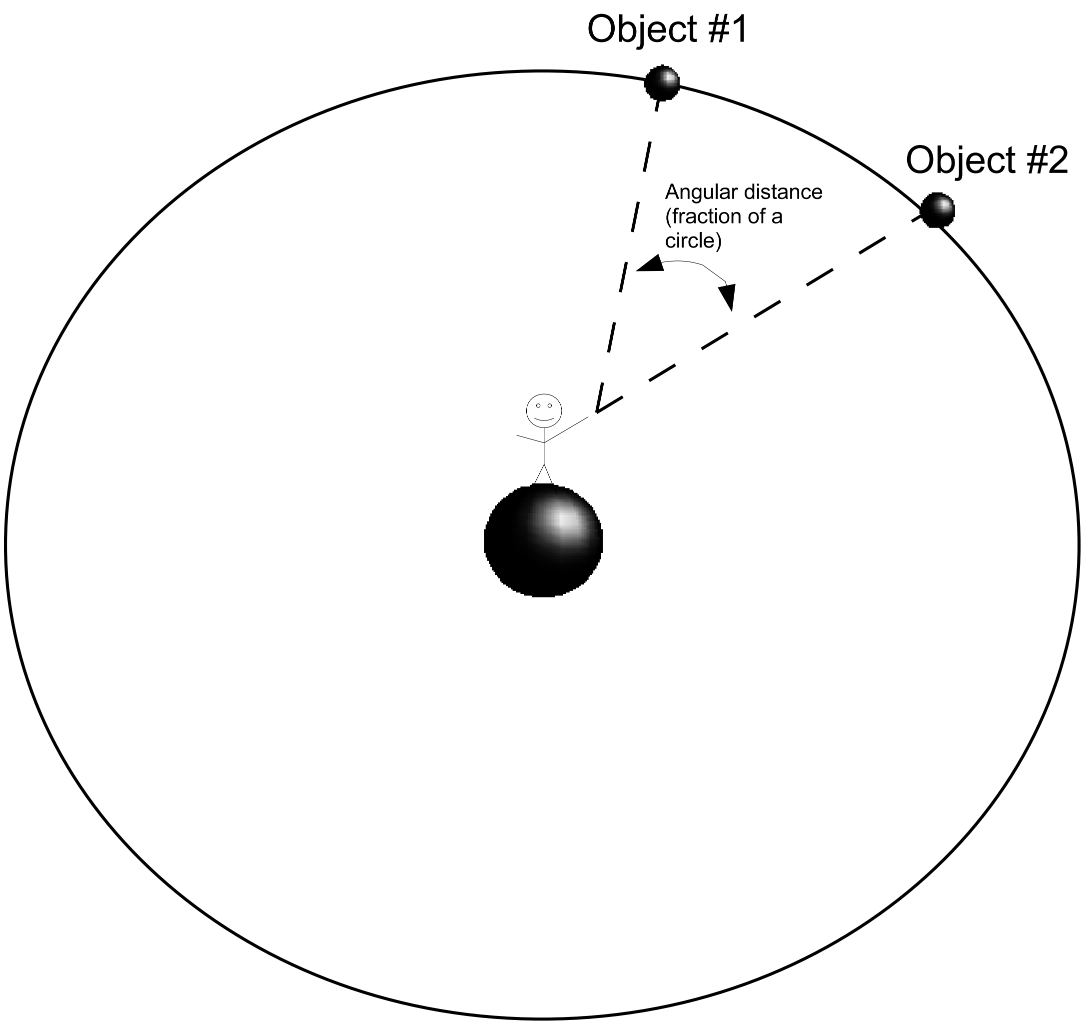


Figure 3.10: What angular distances mean. They can be thought of as fractions of a circle.

horizon in front of you, to the horizon in back of you spanning the entire portion of the sky visible to you at one time, or 180°.

- Each 1° increment can be further subdivided into fractions of an arc degree, called arc minutes; there are 60 arc minutes (60') in one arc degree.

  As an example, the full moon or sun are about 0.5° in size. This will be half of your little finger at arm's length. This half of a degree can also be written at $30'$ or 30 arc minutes. With $60'$ in one degree,

$$30' \times \frac{1^\circ}{60'} = 0.5^\circ \tag{3.8}$$

  as we said.

- Next, an arc minute can be divided into 60 arc seconds ($60''$). In other words, 1 arc minute is 60 arc seconds, or $1' = 60''$. As you might guess, an arc second is an extremely small unit of angular measure; it's the angular size of a U.S. dime at about a mile away.

Relation to time?

Keep in mind that the minutes and seconds mentioned above about angles **have nothing to do with time!** Originally they probably had something to do with round clock faces and seconds and minutes on a clock, but they have no relation to time whatsoever.

Angles of Latitude and Longitude

The minutes and seconds are also used in latitude and longitude measurements, or anywhere angles are used. So, a location might have a longitude of 120.8°, but this would be the same as

$$120.0^\circ + 0.8^\circ \times \frac{60'}{1^\circ} = 120^\circ 48'. \tag{3.9}$$

The result here is read "one hundred and twenty degrees and forty-eight minutes." As another example, what about $32^\circ\ 12'\ 34''$, read "thirty two degrees, twelve minutes, and thirty-four seconds?" The 32 stays as it is; it's already in degrees. The $12'$ gets converted to degrees like this:

$$12' \times \frac{1^\circ}{60'} = 0.2^\circ. \tag{3.10}$$

Lastly, the $34''$ gets converted to degrees like this:

$$34'' \times \frac{1'}{60''} \times \frac{1^\circ}{60'} = 0.0094^\circ. \tag{3.11}$$

Note that you always divide seconds ($''$) by 3600 (=60×60) to get degrees. So the final result here is that $32^\circ\ 12'\ 34'' = 32^\circ + 0.2^\circ + 0.0094^\circ = 32.2094^\circ$.

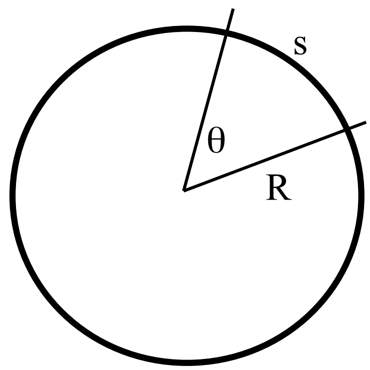


Figure 3.11: Relationship between distance and angular distance.

3.2.3 Angles and Distance

Angles (in degrees) are not the same as a distance, say that you would measure with a ruler. So if you measure the angular distance of the Big Dipper to be 25°, this is *not* the same as the number of miles between the two stars. However, when distances are very large (they always are in this class: the radius of the earth, the distance to the moon, sun, or stars), there is a simple relationship between angular distance (in degrees) and distance, as shown in Figure 3.11.

This figure shows a circle with a "pie-slice" cut into it. The angular distance around the circle is, as always, 360°. Notice the three variables in the figure too, R, θ, and S. R is the large distance, like the radius of the earth, or distance to the sun, etc. θ is the angular distance, that you would measure with your fingers (Figure 3.9). S is the actual, real distance, along the circle, that you'd measure with a ruler (in inches or miles). (It's the distance between the two etch marks that extend a little bit past the circle itself). For the illustration shown, the relationship between the variables is

$$S = \frac{\pi R \theta}{180}. \tag{3.12}$$

In other words, multiplying the angle θ (in degrees) by the distance to the objects (R), would give you the distance between the objects (remember $\pi = 3.1415..$). To use this equation, always find the center of your circle. R will always be the distance from the center of your circle to the object you are measuring. Here's an example.

Suppose you are on the equator, and sail $1'$ (one minute) due east. By $1'$, we mean you've traveled $1' = 0.0166°$ of the 360° that make up the big circle of the equator that goes all the way around the earth. How far have you traveled? To figure this out, you can use the formula above. The center of the circle of the equator is the center of the earth, so the distance between the center of the circle and the object you are measuring (your ship on the surface of the earth) is just the radius of the earth or $R = 20,855,472$ feet. Plug this in to the above formula to get

$$S = \frac{\pi(20,925,637 \text{ feet})(0.0166°)}{180} \approx 6,062 \text{ feet}. \tag{3.13}$$

Thus, traveling $1'$ on the equator is the same as traveling $6,062$ feet. Notice that this is a bit bigger than one mile at $5,280$ feet. The figure $6,062$ feet (should be $6,076$ feet) is the definition of one nautical mile. In other words, you travel one nautical mile by traveling $1'$ on a great circle of the earth (a great circle is a circle whose center is also the center of the earth itself)[Kelch, p. 14-15]. This is why a nautical mile is larger and different than a geographical mile. A nautical mile always refers to traveling $1'$ on a great circle of the earth. As a last example to try with the above formula, if you travel $1°$ on a great circle, you'll travel about 60 miles.

Activity

Go out and measure the angular distance of a few objects. Report back to discuss. Measure the position of the sun relative to a fixed object.

3.3 Celestial Navigation: Navigating using the Sun, Moon, and Stars

So you've just had an introduction to latitude and longitude, and finding your specific location on earth. Did you know that it's possible to find your location by looking up in the sky and observing the sun, moon, and stars? This is called "celestial navigation."

Celestial navigation is important to the longitude story because "back then" (the 1700s and earlier), *it's all navigators had.* There was no GPS, cell phone, radio, radar, satellites, or the like. If you wanted to know where you were, you had to "read" your position from the sky.

People have always been fascinated with the thought that by observing the positions of the sun, stars, or other celestial bodies, their exact position on earth can be found. But, an air of mystery has always surrounded this process. How is it done? Does someone need a supernatural ability? If not, they must at least be a mathematical genius![Kelch, p. 7]. In this section, we say neither! Just a basic understanding of the geometry of the sun, earth, and stars are all one needs to find your "fix," or location on earth, by observing the celestial bodies above.

3.3.1 The Basic Premise of Celestial Navigation

Suppose you are in a new city, surrounded by high-rise buildings. You are generally lost, but notice that there is one very tall building in the center of the city. It seems that no matter where you are, you can always look up and see the top of this building. This is illustrated in Figure 3.12(a), where you have to elevate your chin by an angle θ above the flat ground to spot the top of the building.

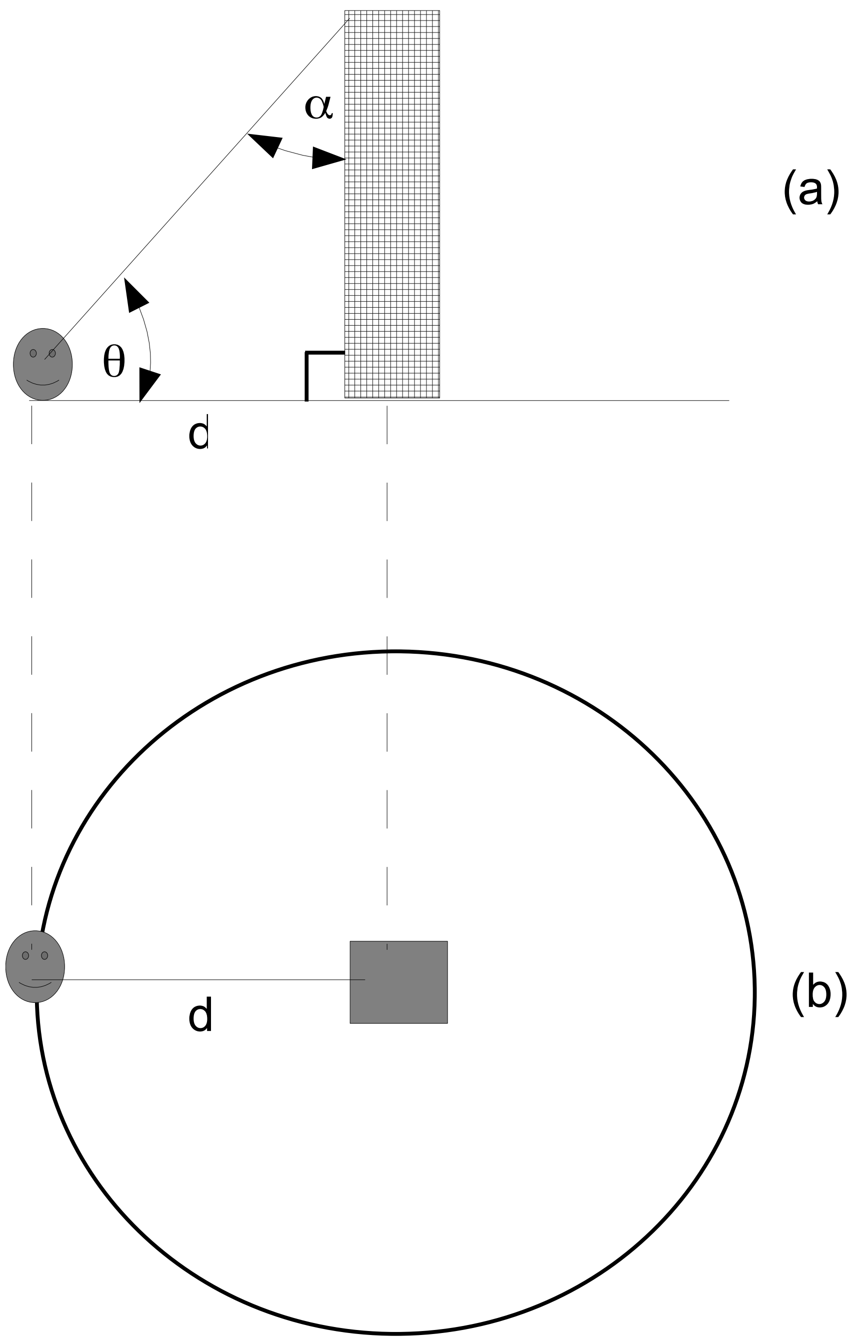


Figure 3.12: Finding your circle of position by shooting the top of a tall building.

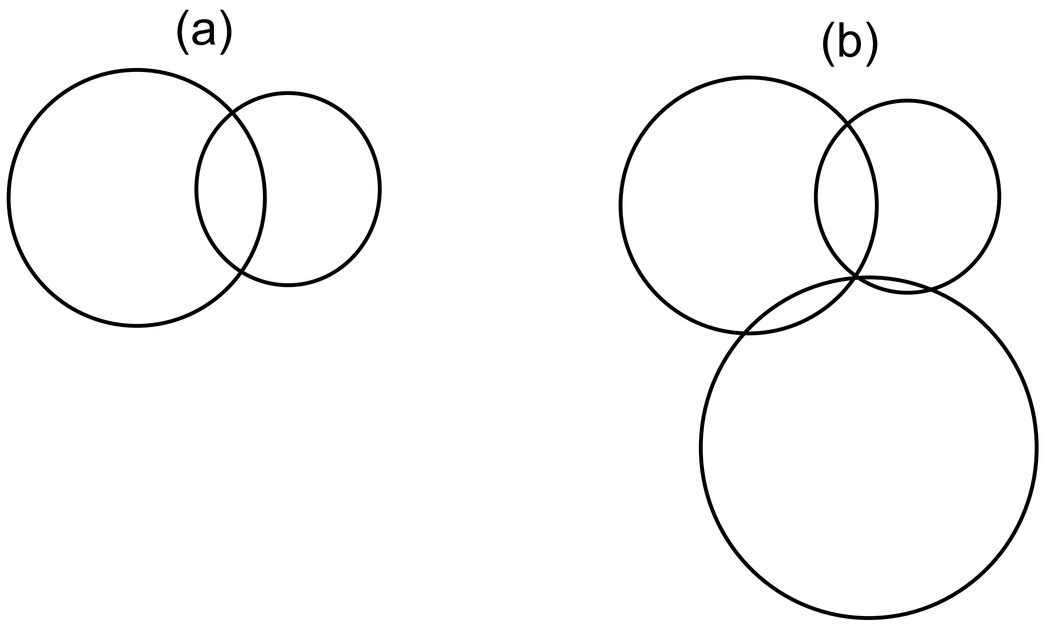


Figure 3.13: Shooting the top of two buildings narrows down your position to two points in (a). Three shots tells you exactly where you are in (b), where the three circles intersect.

It turns out, that with such a landmark, you are not really totally lost, for look at the "bird's-eye" view of you in the city, as shown in Figure 3.12(b). Why are you not lost? Let's choose a value for θ, say 30°. You are not lost because as long as you can hold your chin at 30°, and see the top of the building, then you must be somewhere on the circle shown. Stated another way, no matter where you are, if you look up and see the top of the building, and your chin is at an angle of 30°, then you must be somewhere on the circle shown. Naturally, different angles will have different so called, "circles of position." The radius of the circle is related to the "zenith angle" α, which is just $90° - \theta$. The larger α is (the smaller θ), the larger the radius of the circle will be. If you had a map of the city, you could draw your circle on the map, very neatly and to scale.

Being sure you are somewhere on a circle still leaves you somewhat lost, but this can be narrowed down with a simple compass. For instance, when you look up at 30° and see the top of the building, in which direction are you looking? If, for instance, you are looking to the North, then you can eliminate just about an entire circle, save for a small southern arc of it.

In order to obtain a more precise "fix" on your location, you'd need another landmark. Suppose there were two tall buildings in the town. You first established yourself to be on the first circle by looking up at 30° to see the top of the first building. Next, you turned and looked up at 67° to spot the top of a second building. The bird's-eye view of your situation is shown Figure 3.13(a).

You've established yourself to be on two independent circles of position, one for each

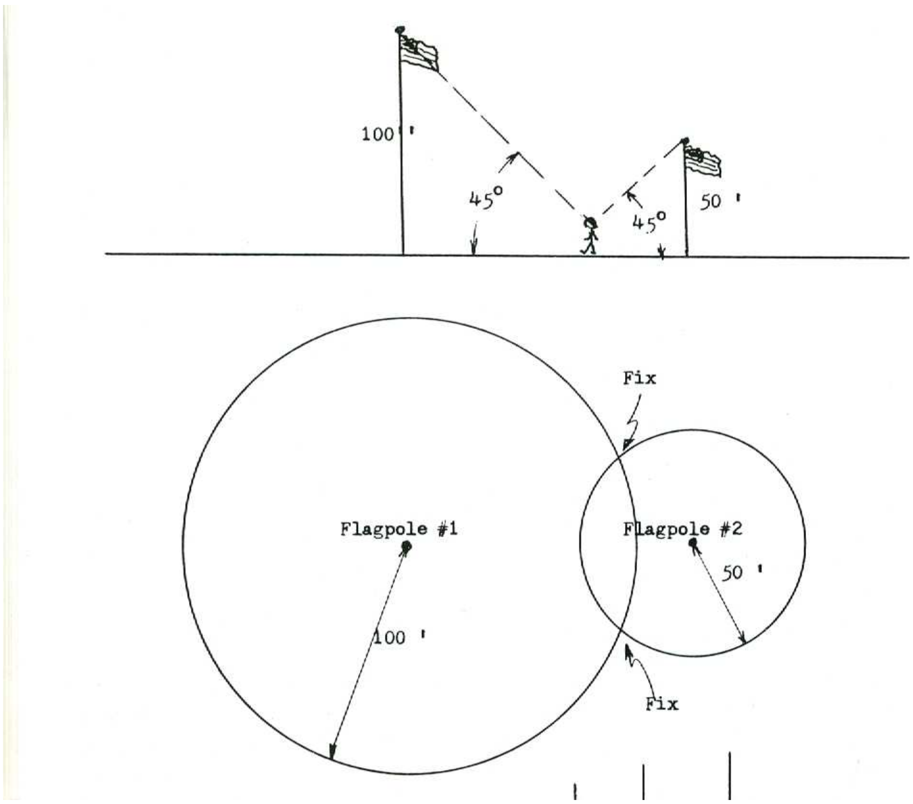


Figure 3.14: Shooting the top of two flag poles. You are at one of the two points where the circles of position intersect. From Kelch.

building (the center of each circle is where one or the other building exists). You must be at either point where the two circles intersect. You could easily reject one point based on a compass reading or simple common sense. Another example of this is shown in Figure 3.14, where you might have two flag poles nearby, whose tips you sighted.

If there was a third building, you could establish a third circle of position and where all three circles intersect would be your location, as shown in Figure 3.13(b).

This is the basic working premise of celestial navigation. Only out at sea, there are no buildings. Instead you sight the angles of celestial objects, like the Sun, Moon, planets, or stars. Also, the distances involves are much larger. A typical "circle of position" could easily have a radius of $3,000$ miles or more. Figure 3.15 shown two circles of positions obtained by sighting the angular height of two celestial bodies.

Imagine then, you are lost at sea. You sight the angular positions of two celestial bodies (Mars and Sirius, perhaps). You draw the circles of position, just like you did for the flag poles in Figure 3.14 on the playground at your kid's school. Out of the whole, vast Earth, you now know that you are at one of the two points where the circles intersect. Usually some common sense, or even a compass would help you eliminate one point or the other (i.e. was Mars to your west, etc.)? This is the basic premise of celestial navigation. But, as usual, it

isn't this simple in practice, because there is no practical map or method of precisely drawing these circles, particularly to pin-point your location to within a few miles or so. Also, over such large distances, the flat circles of position would have to be projected onto the curved earth's surface. Other methods have been developed to deal with these problems as we'll see. Read on, but keep in mind, that these circles of position are all there is to celestial navigation; the rest is a practical implementation of this idea.

Activity

Make the straw/protractor/string/plumb-bob device. Go out and measure some buildings. Find distance from observation point to building. Pairs of groups communicate and verify finding the same angle at different positions puts you both on the same circle. Report findings back in class.

3.3.2 A basic measurement

To start with, let's review the same measurement you did for tall buildings in the previous section, but instead illustrate it as will apply to celestial navigation. Figure 3.16 shows this measurement. As shown, you are somewhere on the earth and want to determine the your position using, not the top of a building, but a bright star you see in the sky. Just like you did for the tall building, you need to be able to accurately determine the angular distance of the celestial object (the star) above the horizon.

This is normally done with a device called a sextant. You use a sextant by leveling it with the ground, then looking through a little telescope with a cross-hair in it. While maintaining the level of the device, you rotate the telescope up until you see the celestial object in the cross-hair. At this point, you'd read the angle from a dial on the sextant and you'd have your angle, which would just be a plain number, like 45° for example. Or more realistically, one might say the star in Figure 3.16 is $36°15'32''$ above the horizon.

Next, let's look a bit more closely at just what you are measuring in Figure 3.17, supposing that the object whose angle you are measuring in the sun, although it could be the moon, a star, or a planet. The point labeled "You" is your current position on the earth. Notice that directly below the celestial object is a point on the earth labeled "GP" This stands for "geographical position." The GP is the point on the earth directly below the celestial object you are "shooting." Keep this in mind.

With your sextant, you measure angle θ as shown. Since the triangle formed between you, the sun, and the sun's GP is a right triangle, you can compute angle α to be $\alpha = 90° - \theta$. Knowing α, here's an important fact about something else you just measured: α is also the angular distance (on the earth) between what your position and the GP of the celestial object. Take a close look at Figure 3.17 to see this. The angle α is the angle in the triangle that is opposite to the side of the triangle formed by the line between "You" and the GP. In other words, finding θ, which gives you α, has also given your distance from the GP!

Further, since all points on this arc are on the surface of the earth, you can compute the distance between you and the GP knowing that every degree of separation between you and the GP equals another 60 nautical miles. So by measuring a simple angle, you can always

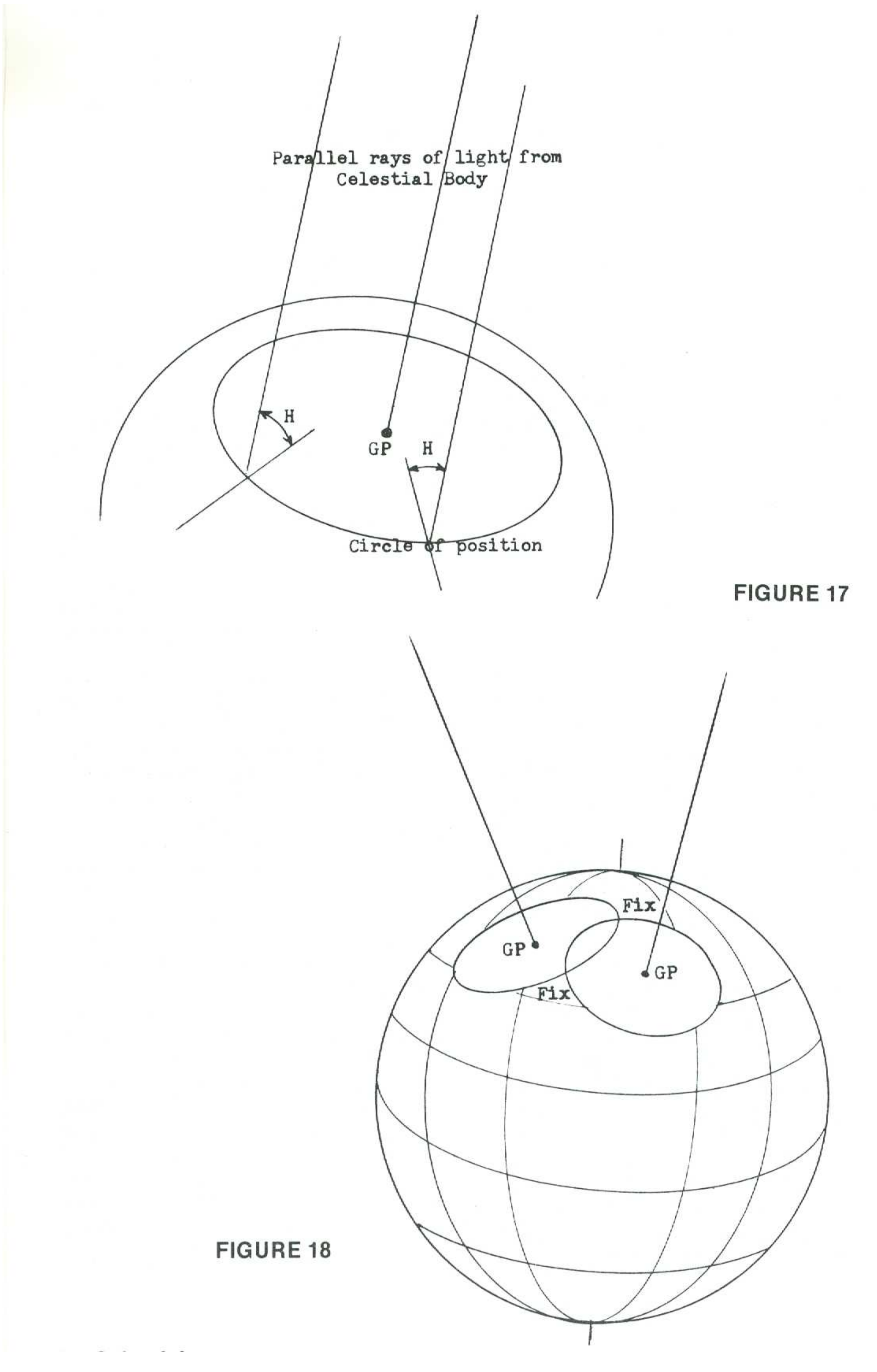


Figure 3.15: Shooting two celestial objects to find two circles of positions. On the entire earth you are at either one or the other "fix" points. From Kelch.

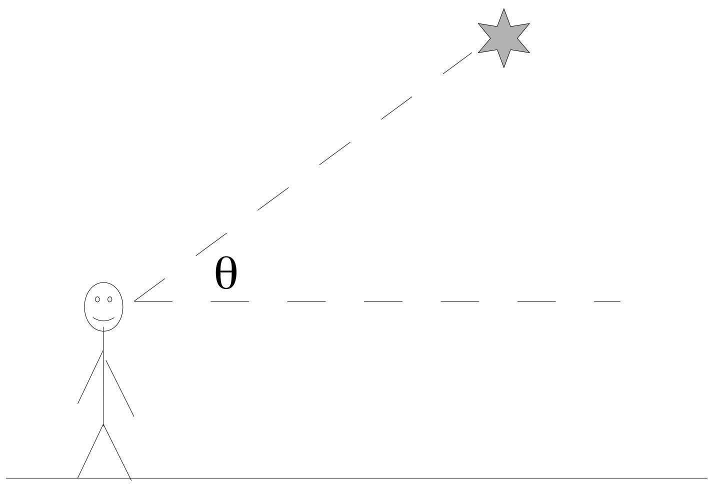


Figure 3.16: The most basic measurement for celestial navigation: measuring the angle of a celestial body above the horizon.

find where you are relative to the GP of some heavenly object. Congratulations! You are a bit “less lost.” Figure 3.18 shows some arcs of “great circles” between you and a GP for a few places you might be on earth.

Here is an example[Schlereth, p. 12]. Suppose you measured the altitude of the sun to be $\theta = 59°30'$. This makes $\alpha = 30°30'$ making the distance between you and the GP of the sun $1,830$ miles. So you can draw a (scaled) circle of radius $1,830$ nautical miles with the sun's GP at the center. You are no longer lost; you are somewhere on the circle shown in Figure 3.19.

Now, believe it or not, this is all there is to celestial navigation. We all know we're not done though. For where is the GP? You don't know, but it is something you can look up.

The most boring book ever written

A book called the *Nautical Almanac* is published every year. All the book contains are pages and pages of tabulated numbers. There is almost nothing to *read* in the book at all. The purpose of the book is the list the GP for every hour of the entire year for the sun, moon, major planets, and stars. Two sample pages, illustrating the information it provides for each day of the year, are shown in Figures 3.20 and 3.21. Here is a quick tour of these pages.

The leftmost column, labeled “UT” stands for “universal time” which is the same as the time in Greenwich, at the prime meridian. Note that data is given for all 24 hours in a day, and three days worth of information is given on a single page. Don't worry about the column labeled “Aries;” it will be discussed shortly. Notice columns for common planets, like Venus, Mars, Jupiter, and Saturn. These are listed because these planets, when visible, are easy to spot and very bright. Figure 3.21 gives information on the Sun and Moon. All of the little numbers in the columns marked “GHA” and “Dec” are essentially the longitude and latitude of these objects' GP So on Tuesday, December 9, 2003, at 05:00 UT, the sun

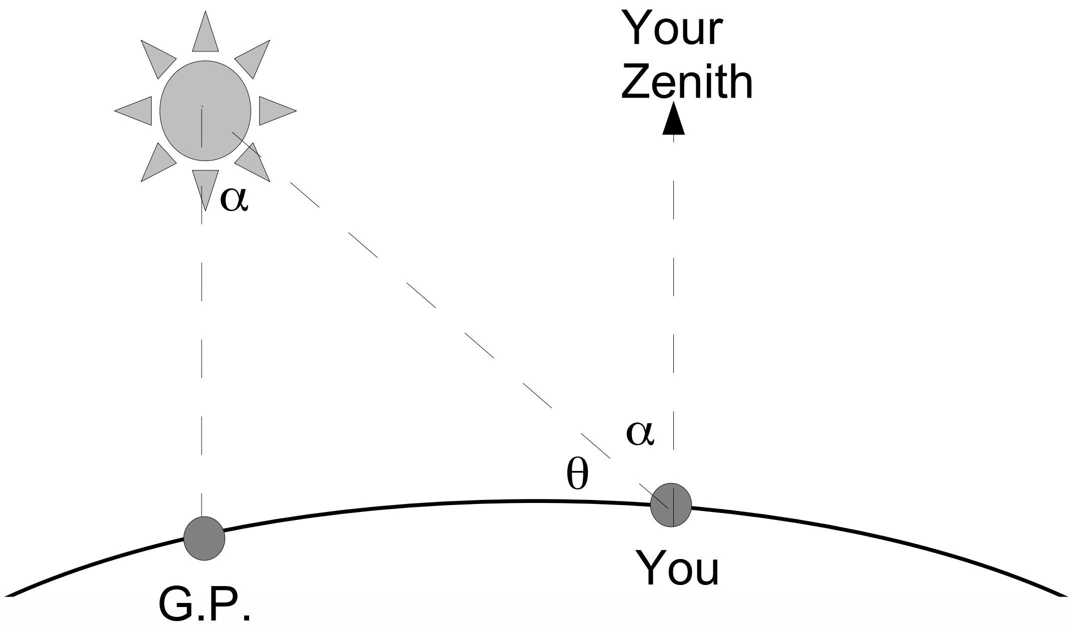


Figure 3.17: Celestial Navigation's triangle that allows you to determine how far you are from a celestial object's GP (α), by simply observing the height of an object (θ) above the horizon, then computing $\alpha = 90 - \theta$. α is opposite to the side of the triangle connecting "You" and the GP.

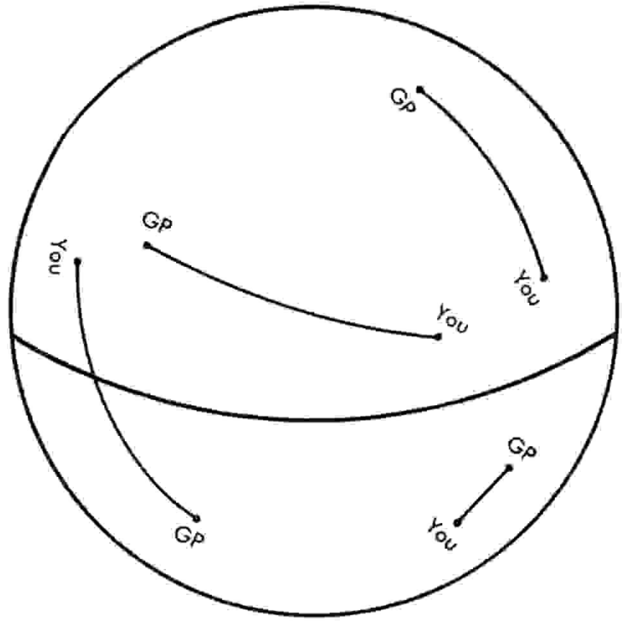


Figure 3.18: Examples of arcs, with the radius of the earth, between you and some hypothetical GPs for positions all over the globe[From: Schlereth, p. 5].

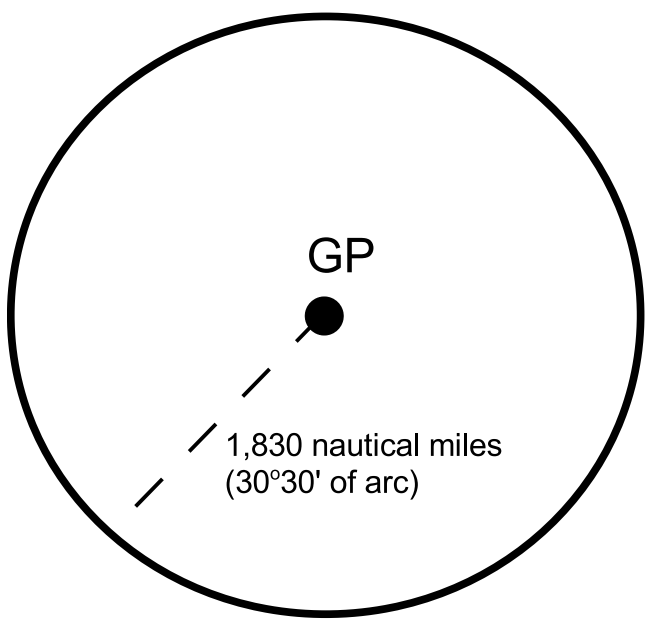


Figure 3.19: The most basic measurement for celestial navigation: measuring the angle of a celestial body above the horizon.

was directly over the point longitude=256°59.3′ and latitude=22° S 46.5′.

So, if you knew what time it was in Greenwich when you took your original altitude reading of the sun, you could look up the sun's GP and then know where the center of your circle was. You are getting "less lost" all the time.

In the example worked above, you've seen that these circles can be thousands of miles in radius. Can't you do better than this? Yes, because as a dedicated sailor you were always keeping track of your dead-reckoning or DR. So, you would draw the object's GP and your distance circle on the same map that you are using to plot your DR. The closet approach between your DR and distance circle is most likely where you are.

So you are getting less and less lost, but we can still do better. The best scenario is that some kind of celestial navigation would give you your latitude and longitude directly. As you'll see, longitude will always remains the elusive one, but in the sections that follow, we'll study enough celestial navigation to achieve a reasonable "fix."

We'll pause now and take a break from more specifics on celestial navigation. The topic will be picked up again in Section 3.4 because a a few words are needed about the sky, that contains the sun, moon, stars and oddly enough, the moons of the planet Jupiter.

Also, a final note about the geographical position. In Figure 3.17 the GP was drawn right "under" the sun, but the radius of the earth was grossly off scale in this figure, for the sake of illustration. Technically, the GP is defined this way: draw an imaginary line from the center of the earth, that pierces the surface of the earth, on its way out to the celestial body in question. The point where the line pierces the earth is the GP of the body. This is shown in Figure 3.22; the correct illustration for where the GP comes from.

Activity

Go out with protractor and find the GP of the sun. Estimate the distance between you and the GP of the sun. Need watch and current version of the Nautical Almanac.

3.3.3 The Sky

Go out and look at the "stars" some night. You'll see many "little points of light" way up there. Of course what are you seeing are not all stars. The points of light are planets, stars, other galaxies, huge groups or stars, or many other things we know about. The first question you might ask is: "how far away is a given star?" Of course each star is a different distance from earth (so there's your answer). For celestial navigation, however, the only answer you really need is "all of the stars are very far away." They are all so far from us, that for celestial navigation, their precise distances don't matter. What is important about them are three things. First, they don't appear to be moving relative to one another. Second, they appear to move across the sky at 15° per hour. Third, they are so far away that light they emit arrives on earth as parallel rays.

Of course the stars are moving but they are so far away that their individual motions aren't noticeable. For example, the big dipper has looked pretty much the same for the past 25,000 years, and will continue to do look as it does for another 25,000 years. Also, the

238 **2003 DECEMBER 9, 10, 11 (TUES., WED., THURS.)**

UT	ARIES	VENUS −3.9		MARS −0.2		JUPITER −2.1		SATURN −0.3		STARS		
	GHA	GHA	Dec	GHA	Dec	GHA	Dec	GHA	Dec	Name	SHA	Dec
d h	° ′	° ′	° ′	° ′	° ′	° ′	° ′	° ′	° ′		° ′	° ′
9 00	77 19.5	151 02.4	S24 13.0	81 02.4	S 2 15.2	268 04.3	N 5 49.3	334 49.6	N22 14.2	Acamar	315 23.6	S40 17.4
01	92 22.0	166 01.5	12.7	96 03.6	14.6	283 06.6	49.2	349 52.2	14.2	Achernar	335 31.7	S57 13.3
02	107 24.4	181 00.6	12.4	111 04.8	14.0	298 08.8	49.2	4 54.9	14.2	Acrux	173 18.3	S63 06.9
03	122 26.9	195 59.7	. . 12.1	126 06.0	. . 13.3	313 11.1	. . 49.1	19 57.5	. . 14.2	Adhara	255 18.1	S28 58.5
04	137 29.4	210 58.8	11.8	141 07.3	12.7	328 13.4	49.0	35 00.2	14.2	Aldebaran	290 57.7	N16 31.1
05	152 31.8	225 57.8	11.6	156 08.5	12.1	343 15.7	49.0	50 02.8	14.3			
06	167 34.3	240 56.9	S24 11.3	171 09.7	S 2 11.4	358 17.9	N 5 48.9	65 05.5	N22 14.3	Alioth	166 27.3	N55 56.1
07	182 36.8	255 56.0	11.0	186 10.9	10.8	13 20.2	48.8	80 08.1	14.3	Alkaid	153 05.0	N49 17.5
T 08	197 39.2	270 55.1	10.7	201 12.1	10.2	28 22.5	48.8	95 10.8	14.3	Al Na'ir	27 53.1	S46 56.8
U 09	212 41.7	285 54.2	. . 10.4	216 13.3	. . 09.5	43 24.8	. . 48.7	110 13.5	. . 14.3	Alnilam	275 53.7	S 1 11.9
E 10	227 44.2	300 53.3	10.1	231 14.5	08.9	58 27.1	48.6	125 16.1	14.4	Alphard	218 03.3	S 8 40.4
S 11	242 46.6	315 52.4	09.8	246 15.7	08.2	73 29.3	48.6	140 18.8	14.4			
D 12	257 49.1	330 51.5	S24 09.5	261 16.9	S 2 07.6	88 31.6	N 5 48.5	155 21.4	N22 14.4	Alphecca	126 17.7	N26 42.0
A 13	272 51.5	345 50.5	09.2	276 18.1	07.0	103 33.9	48.4	170 24.1	14.4	Alpheratz	357 51.3	N29 06.8
Y 14	287 54.0	0 49.6	08.9	291 19.4	06.3	118 36.2	48.4	185 26.7	14.4	Altair	62 15.8	N 8 52.7
15	302 56.5	15 48.7	. . 08.5	306 20.6	. . 05.7	133 38.4	. . 48.3	200 29.4	. . 14.4	Ankaa	353 22.8	S42 17.4
16	317 58.9	30 47.8	08.2	321 21.8	05.1	148 40.7	48.2	215 32.0	14.5	Antares	112 35.9	S26 26.4
17	333 01.4	45 46.9	07.9	336 23.0	04.4	163 43.0	48.2	230 34.7	14.5			
18	348 03.9	60 46.0	S24 07.6	351 24.2	S 2 03.8	178 45.3	N 5 48.1	245 37.3	N22 14.5	Arcturus	146 02.8	N19 09.7
19	3 06.3	75 45.1	07.3	6 25.4	03.2	193 47.6	48.0	260 40.0	14.5	Atria	107 45.2	S69 02.1
20	18 08.8	90 44.2	07.0	21 26.6	02.5	208 49.9	48.0	275 42.7	14.5	Avior	234 20.8	S59 31.0
21	33 11.3	105 43.3	. . 06.7	36 27.8	. . 01.9	223 52.1	. . 47.9	290 45.3	. . 14.6	Bellatrix	278 39.7	N 6 21.3
22	48 13.7	120 42.4	06.4	51 29.0	01.3	238 54.4	47.8	305 48.0	14.6	Betelgeuse	271 09.1	N 7 24.6
23	63 16.2	135 41.4	06.1	66 30.2	00.6	253 56.7	47.8	320 50.6	14.6			
10 00	78 18.7	150 40.5	S24 05.7	81 31.4	S 2 00.0	268 59.0	N 5 47.7	335 53.3	N22 14.6	Canopus	263 59.0	S52 41.7
01	93 21.1	165 39.6	05.4	96 32.6	1 59.4	284 01.3	47.6	350 55.9	14.6	Capella	280 45.1	N46 00.2
02	108 23.6	180 38.7	05.1	111 33.8	58.7	299 03.5	47.6	5 58.6	14.6	Deneb	49 37.0	N45 17.8
03	123 26.0	195 37.8	. . 04.8	126 35.0	. . 58.1	314 05.8	. . 47.5	21 01.2	. . 14.7	Denebola	182 41.3	N14 33.0
04	138 28.5	210 36.9	04.5	141 36.2	57.5	329 08.1	47.5	36 03.9	14.7	Diphda	349 03.2	S17 58.0
05	153 31.0	225 36.0	04.1	156 37.5	56.8	344 10.4	47.4	51 06.6	14.7			
06	168 33.4	240 35.1	S24 03.8	171 38.7	S 1 56.2	359 12.7	N 5 47.3	66 09.2	N22 14.7	Dubhe	194 00.5	N61 43.6
W 07	183 35.9	255 34.2	03.5	186 39.9	55.5	14 15.0	47.3	81 11.9	14.7	Elnath	278 21.7	N28 36.7
E 08	198 38.4	270 33.3	03.1	201 41.1	54.9	29 17.2	47.2	96 14.5	14.8	Eltanin	90 50.2	N51 29.3
D 09	213 40.8	285 32.4	. . 02.8	216 42.3	. . 54.3	44 19.5	. . 47.1	111 17.2	. . 14.8	Enif	33 54.7	N 9 53.5
N 10	228 43.3	300 31.5	02.5	231 43.5	53.6	59 21.8	47.1	126 19.8	14.8	Fomalhaut	15 32.2	S29 36.3
E 11	243 45.8	315 30.6	02.2	246 44.7	53.0	74 24.1	47.0	141 22.5	14.8			
S 12	258 48.2	330 29.7	S24 01.8	261 45.9	S 1 52.4	89 26.4	N 5 46.9	156 25.2	N22 14.8	Gacrux	172 09.8	S57 07.7
D 13	273 50.7	345 28.8	01.5	276 47.1	51.7	104 28.7	46.9	171 27.8	14.8	Gienah	176 00.2	S17 33.6
A 14	288 53.2	0 27.9	01.2	291 48.3	51.1	119 31.0	46.8	186 30.5	14.9	Hadar	148 59.3	S60 23.2
Y 15	303 55.6	15 27.0	. . 00.8	306 49.5	. . 50.5	134 33.2	. . 46.8	201 33.1	. . 14.9	Hamal	328 09.1	N23 29.0
16	318 58.1	30 26.1	00.5	321 50.7	49.8	149 35.5	46.7	216 35.8	14.9	Kaus Aust.	83 54.2	S34 23.1
17	334 00.5	45 25.1	24 00.1	336 51.9	49.2	164 37.8	46.6	231 38.4	14.9			
18	349 03.0	60 24.2	S23 59.8	351 53.1	S 1 48.6	179 40.1	N 5 46.6	246 41.1	N22 14.9	Kochab	137 20.3	N74 08.2
19	4 05.5	75 23.3	59.5	6 54.3	47.9	194 42.4	46.5	261 43.8	15.0	Markab	13 45.9	N15 13.6
20	19 07.9	90 22.4	59.1	21 55.5	47.3	209 44.7	46.4	276 46.4	15.0	Menkar	314 22.6	N 4 06.3
21	34 10.4	105 21.5	. . 58.8	36 56.7	. . 46.6	224 47.0	. . 46.4	291 49.1	. . 15.0	Menkent	148 16.9	S36 23.2
22	49 12.9	120 20.6	58.4	51 57.9	46.0	239 49.3	46.3	306 51.7	15.0	Miaplacidus	221 41.1	S69 43.6
23	64 15.3	135 19.7	58.1	66 59.1	45.4	254 51.5	46.2	321 54.4	15.0			
11 00	79 17.8	150 18.8	S23 57.7	82 00.3	S 1 44.7	269 53.8	N 5 46.2	336 57.0	N22 15.0	Mirfak	308 50.8	N49 52.7
01	94 20.3	165 17.9	57.4	97 01.5	44.1	284 56.1	46.1	351 59.7	15.1	Nunki	76 07.9	S26 17.6
02	109 22.7	180 17.0	57.0	112 02.7	43.5	299 58.4	46.1	7 02.4	15.1	Peacock	53 31.3	S56 43.7
03	124 25.2	195 16.1	. . 56.7	127 03.9	. . 42.8	315 00.7	. . 46.0	22 05.0	. . 15.1	Pollux	243 36.6	N28 01.0
04	139 27.7	210 15.2	56.3	142 05.1	42.2	330 03.0	45.9	37 07.7	15.1	Procyon	245 07.3	N 5 13.0
05	154 30.1	225 14.3	56.0	157 06.3	41.5	345 05.3	45.9	52 10.3	15.1			
06	169 32.6	240 13.4	S23 55.6	172 07.5	S 1 40.9	0 07.6	N 5 45.8	67 13.0	N22 15.2	Rasalhague	96 13.8	N12 33.4
07	184 35.0	255 12.5	55.3	187 08.7	40.3	15 09.9	45.8	82 15.7	15.2	Regulus	207 51.4	N11 56.9
T 08	199 37.5	270 11.6	54.9	202 09.9	39.6	30 12.2	45.7	97 18.3	15.2	Rigel	281 18.9	S 8 11.7
H 09	214 40.0	285 10.8	. . 54.5	217 11.1	. . 39.0	45 14.4	. . 45.6	112 21.0	. . 15.2	Rigil Kent.	140 02.9	S60 50.8
U 10	229 42.4	300 09.9	54.2	232 12.3	38.4	60 16.7	45.6	127 23.6	15.2	Sabik	102 21.5	S15 43.8
R 11	244 44.9	315 09.0	53.8	247 13.5	37.7	75 19.0	45.5	142 26.3	15.2			
S 12	259 47.4	330 08.1	S23 53.5	262 14.7	S 1 37.1	90 21.3	N 5 45.4	157 29.0	N22 15.3	Schedar	349 49.1	N56 33.7
D 13	274 49.8	345 07.2	53.1	277 15.9	36.4	105 23.6	45.4	172 31.6	15.3	Shaula	96 32.6	S37 06.5
A 14	289 52.3	0 06.3	52.7	292 17.1	35.8	120 25.9	45.3	187 34.3	15.3	Sirius	258 40.1	S16 43.1
Y 15	304 54.8	15 05.4	. . 52.4	307 18.3	. . 35.2	135 28.2	. . 45.3	202 36.9	. . 15.3	Spica	158 39.4	S11 10.8
16	319 57.2	30 04.5	52.0	322 19.5	34.5	150 30.5	45.2	217 39.6	15.3	Suhail	222 57.8	S43 26.6
17	334 59.7	45 03.6	51.6	337 20.6	33.9	165 32.8	45.1	232 42.3	15.4			
18	350 02.2	60 02.7	S23 51.3	352 21.8	S 1 33.3	180 35.1	N 5 45.1	247 44.9	N22 15.4	Vega	80 44.5	N38 47.2
19	5 04.6	75 01.8	50.9	7 23.0	32.6	195 37.4	45.0	262 47.6	15.4	Zuben'ubi	137 14.1	S16 03.4
20	20 07.1	90 00.9	50.5	22 24.2	32.0	210 39.7	45.0	277 50.2	15.4		SHA	Mer. Pass.
21	35 09.5	105 00.0	. . 50.1	37 25.4	. . 31.3	225 42.0	. . 44.9	292 52.9	. . 15.4		° ′	h m
22	50 12.0	119 59.1	49.8	52 26.6	30.7	240 44.3	44.8	307 55.6	15.4	Venus	72 21.9	13 58
23	65 14.5	134 58.2	49.4	67 27.8	30.1	255 46.6	44.8	322 58.2	15.5	Mars	3 12.8	18 32
	h m									Jupiter	190 40.3	6 03
Mer. Pass.	18 43.7	*v* −0.9	*d* 0.3	*v* 1.2	*d* 0.6	*v* 2.3	*d* 0.1	*v* 2.7	*d* 0.0	Saturn	257 34.6	1 36

Figure 3.20: Sample page from the Nautical Almanac, p. 238, 2003 edition.

2003 DECEMBER 9, 10, 11 (TUES., WED., THURS.) 239

	UT	SUN GHA	SUN Dec	MOON GHA	MOON v	MOON Dec	MOON d	MOON HP
	d h	° ′	° ′	° ′	′	° ′	′	′
	9 00	182 00.7	S22 45.2	0 36.1	11.1	N25 19.2	5.1	54.1
	01	197 00.4	45.4	15 06.2	11.0	25 24.3	4.9	54.1
	02	212 00.2	45.7	29 36.2	10.9	25 29.2	4.8	54.1
	03	226 59.9	. . 46.0	44 06.1	10.9	25 34.0	4.7	54.1
	04	241 59.6	46.2	58 36.0	10.8	25 38.7	4.5	54.1
	05	256 59.3	46.5	73 05.8	10.8	25 43.2	4.5	54.1
	06	271 59.1	S22 46.7	87 35.6	10.8	N25 47.7	4.3	54.1
	07	286 58.8	47.0	102 05.4	10.7	25 52.0	4.2	54.1
TUESDAY	08	301 58.5	47.2	116 35.1	10.6	25 56.2	4.1	54.1
	09	316 58.2	. . 47.5	131 04.7	10.6	26 00.3	4.0	54.1
	10	331 58.0	47.7	145 34.3	10.6	26 04.3	3.8	54.1
	11	346 57.7	48.0	160 03.9	10.5	26 08.1	3.7	54.1
	12	1 57.4	S22 48.2	174 33.4	10.5	N26 11.8	3.6	54.1
	13	16 57.1	48.5	189 02.9	10.4	26 15.4	3.4	54.1
	14	31 56.8	48.7	203 32.3	10.4	26 18.8	3.4	54.2
	15	46 56.6	. . 48.9	218 01.7	10.4	26 22.2	3.2	54.2
	16	61 56.3	49.2	232 31.1	10.3	26 25.4	3.1	54.2
	17	76 56.0	49.4	247 00.4	10.3	26 28.5	2.9	54.2
	18	91 55.7	S22 49.7	261 29.7	10.3	N26 31.4	2.9	54.2
	19	106 55.5	49.9	275 59.0	10.2	26 34.3	2.7	54.2
	20	121 55.2	50.2	290 28.2	10.2	26 37.0	2.5	54.2
	21	136 54.9	. . 50.4	304 57.4	10.1	26 39.5	2.5	54.2
	22	151 54.6	50.6	319 26.5	10.1	26 42.0	2.3	54.2
	23	166 54.3	50.9	333 55.6	10.1	26 44.3	2.2	54.2
	10 00	181 54.1	S22 51.1	348 24.7	10.0	N26 46.5	2.0	54.2
	01	196 53.8	51.4	2 53.7	10.1	26 48.5	1.9	54.2
	02	211 53.5	51.6	17 22.8	10.0	26 50.4	1.8	54.3
	03	226 53.2	. . 51.8	31 51.8	9.9	26 52.2	1.7	54.3
	04	241 52.9	52.1	46 20.7	10.0	26 53.9	1.5	54.3
	05	256 52.7	52.3	60 49.7	9.9	26 55.4	1.4	54.3
	06	271 52.4	S22 52.5	75 18.6	9.9	N26 56.8	1.3	54.3
WEDNESDAY	07	286 52.1	52.8	89 47.5	9.9	26 58.1	1.1	54.3
	08	301 51.8	53.0	104 16.4	9.8	26 59.2	1.0	54.3
	09	316 51.5	. . 53.2	118 45.2	9.8	27 00.2	0.9	54.3
	10	331 51.3	53.4	133 14.0	9.9	27 01.1	0.7	54.3
	11	346 51.0	53.7	147 42.9	9.8	27 01.8	0.6	54.3
	12	1 50.7	S22 53.9	162 11.7	9.7	N27 02.4	0.5	54.3
	13	16 50.4	54.1	176 40.4	9.8	27 02.9	0.3	54.4
	14	31 50.1	54.4	191 09.2	9.7	27 03.2	0.2	54.4
	15	46 49.9	. . 54.6	205 37.9	9.8	27 03.4	0.1	54.4
	16	61 49.6	54.8	220 06.7	9.7	27 03.5	0.1	54.4
	17	76 49.3	55.0	234 35.4	9.7	27 03.4	0.2	54.4
	18	91 49.0	S22 55.3	249 04.1	9.7	N27 03.2	0.3	54.4
	19	106 48.7	55.5	263 32.8	9.7	27 02.9	0.5	54.4
	20	121 48.4	55.7	278 01.5	9.7	27 02.4	0.6	54.4
	21	136 48.2	. . 55.9	292 30.2	9.7	27 01.8	0.7	54.4
	22	151 47.9	56.1	306 58.9	9.7	27 01.1	0.9	54.5
	23	166 47.6	56.4	321 27.6	9.7	27 00.2	1.0	54.5
	11 00	181 47.3	S22 56.6	335 56.3	9.7	N26 59.2	1.2	54.5
	01	196 47.0	56.8	350 25.0	9.6	26 58.0	1.3	54.5
	02	211 46.7	57.0	4 53.6	9.7	26 56.7	1.4	54.5
	03	226 46.5	. . 57.2	19 22.3	9.7	26 55.3	1.5	54.5
	04	241 46.2	57.5	33 51.0	9.7	26 53.8	1.7	54.5
	05	256 45.9	57.7	48 19.7	9.7	26 52.1	1.8	54.5
	06	271 45.6	S22 57.9	62 48.4	9.7	N26 50.3	2.0	54.6
	07	286 45.3	58.1	77 17.1	9.7	26 48.3	2.1	54.6
THURSDAY	08	301 45.0	58.3	91 45.8	9.7	26 46.2	2.2	54.6
	09	316 44.7	. . 58.5	106 14.5	9.7	26 44.0	2.4	54.6
	10	331 44.5	58.7	120 43.2	9.7	26 41.6	2.5	54.6
	11	346 44.2	58.9	135 11.9	9.7	26 39.1	2.6	54.6
	12	1 43.9	S22 59.1	149 40.6	9.8	N26 36.5	2.8	54.6
	13	16 43.6	59.4	164 09.4	9.8	26 33.7	2.9	54.6
	14	31 43.3	59.6	178 38.2	9.7	26 30.8	3.0	54.7
	15	46 43.0	22 59.8	193 06.9	9.8	26 27.8	3.1	54.7
	16	61 42.7	23 00.0	207 35.7	9.8	26 24.7	3.3	54.7
	17	76 42.5	00.2	222 04.5	9.9	26 21.4	3.5	54.7
	18	91 42.2	S23 00.4	236 33.4	9.8	N26 17.9	3.5	54.7
	19	106 41.9	00.6	251 02.2	9.9	26 14.4	3.7	54.7
	20	121 41.6	00.8	265 31.1	9.8	26 10.7	3.8	54.8
	21	136 41.3	. . 01.0	279 59.9	9.9	26 06.9	4.0	54.8
	22	151 41.0	01.2	294 28.8	10.0	26 02.9	4.1	54.8
	23	166 40.7	01.4	308 57.8	9.9	N25 58.8	4.2	54.8
		SD 16.3	d 0.2	SD 14.8		14.8		14.9

Lat.	Twilight Naut.	Twilight Civil	Sunrise	Moonrise 9	Moonrise 10	Moonrise 11	Moonrise 12
°	h m	h m	h m	h m	h m	h m	h m
N 72	08 11	10 28	■	□	□	□	□
N 70	07 52	09 36	■	□	□	□	□
68	07 37	09 04	■	□	□	□	□
66	07 25	08 41	10 14	□	□	□	□
64	07 15	08 22	09 37	□	□	□	16 28
62	07 05	08 07	09 11	13 54	14 24	15 34	17 10
60	06 57	07 54	08 50	14 32	15 11	16 15	17 38
N 58	06 50	07 42	08 33	14 59	15 41	16 42	17 59
56	06 44	07 33	08 19	15 20	16 04	17 04	18 17
54	06 38	07 24	08 07	15 38	16 23	17 22	18 32
52	06 32	07 16	07 56	15 53	16 38	17 37	18 45
50	06 27	07 09	07 47	16 06	16 52	17 50	18 57
45	06 16	06 53	07 27	16 32	17 20	18 17	19 20
N 40	06 06	06 40	07 10	16 54	17 42	18 38	19 39
35	05 57	06 29	06 57	17 11	18 00	18 55	19 55
30	05 49	06 18	06 45	17 26	18 16	19 11	20 08
20	05 33	06 00	06 24	17 52	18 43	19 36	20 31
N 10	05 17	05 43	06 06	18 15	19 06	19 58	20 51
0	05 00	05 26	05 49	18 36	19 27	20 19	21 10
S 10	04 42	05 09	05 32	18 57	19 49	20 40	21 29
20	04 20	04 49	05 13	19 19	20 12	21 02	21 48
30	03 51	04 24	04 52	19 45	20 38	21 27	22 11
35	03 33	04 10	04 39	20 01	20 54	21 42	22 25
40	03 11	03 52	04 24	20 19	21 12	22 00	22 40
45	02 41	03 30	04 07	20 41	21 35	22 20	22 58
S 50	01 58	03 01	03 45	21 08	22 03	22 47	23 21
52	01 33	02 47	03 35	21 22	22 16	22 59	23 32
54	00 54	02 30	03 23	21 37	22 32	23 14	23 44
56	////	02 09	03 09	21 56	22 51	23 31	23 58
58	////	01 42	02 53	22 19	23 15	23 51	24 14
S 60	////	01 00	02 34	22 48	23 45	24 17	00 17

Lat.	Sunset	Twilight Civil	Twilight Naut.	Moonset 9	Moonset 10	Moonset 11	Moonset 12
°	h m	h m	h m	h m	h m	h m	h m
N 72	■	13 17	15 34	□	□	□	□
N 70	■	14 09	15 52	□	□	□	□
68	■	14 41	16 07	□	□	□	□
66	13 31	15 04	16 20	□	□	□	□
64	14 08	15 23	16 30	□	□	□	13 35
62	14 34	15 38	16 39	10 46	12 03	12 41	12 52
60	14 55	15 51	16 47	10 08	11 16	12 00	12 24
N 58	15 12	16 03	16 55	09 42	10 46	11 32	12 02
56	15 26	16 12	17 01	09 21	10 23	11 11	11 43
54	15 38	16 21	17 07	09 04	10 05	10 53	11 28
52	15 49	16 29	17 13	08 49	09 49	10 38	11 15
50	15 58	16 36	17 18	08 36	09 35	10 24	11 03
45	16 18	16 52	17 29	08 10	09 07	09 57	10 39
N 40	16 35	17 05	17 39	07 49	08 45	09 36	10 19
35	16 48	17 16	17 48	07 32	08 27	09 18	10 03
30	17 00	17 27	17 57	07 17	08 11	09 02	09 49
20	17 21	17 45	18 13	06 52	07 45	08 36	09 24
N 10	17 39	18 02	18 28	06 30	07 22	08 13	09 03
0	17 56	18 19	18 45	06 10	07 00	07 52	08 44
S 10	18 14	18 37	19 03	05 49	06 39	07 31	08 24
20	18 32	18 57	19 26	05 28	06 16	07 08	08 03
30	18 54	19 21	19 54	05 03	05 50	06 42	07 38
35	19 06	19 36	20 12	04 48	05 34	06 26	07 24
40	19 21	19 54	20 35	04 31	05 16	06 08	07 07
45	19 39	20 16	21 05	04 10	04 54	05 46	06 46
S 50	20 00	20 44	21 48	03 45	04 26	05 18	06 21
52	20 11	20 59	22 14	03 32	04 12	05 04	06 08
54	20 23	21 16	22 53	03 18	03 56	04 48	05 54
56	20 37	21 37	////	03 02	03 38	04 30	05 37
58	20 53	22 05	////	02 42	03 15	04 06	05 17
S 60	21 12	22 48	////	02 18	02 45	03 36	04 51

Day	SUN Eqn. of Time 00^h	SUN Eqn. of Time 12^h	SUN Mer. Pass.	MOON Mer. Pass. Upper	MOON Mer. Pass. Lower	MOON Age	Phase
d	m s	m s	h m	h m	h m	d %	
9	08 03	07 50	11 52	24 48	12 23	16 100	
10	07 37	07 23	11 53	00 48	13 14	17 97	○
11	07 10	06 56	11 53	01 40	14 06	18 94	

Figure 3.21: Sample page from the Nautical Almanac, p. 239, 2003 edition.

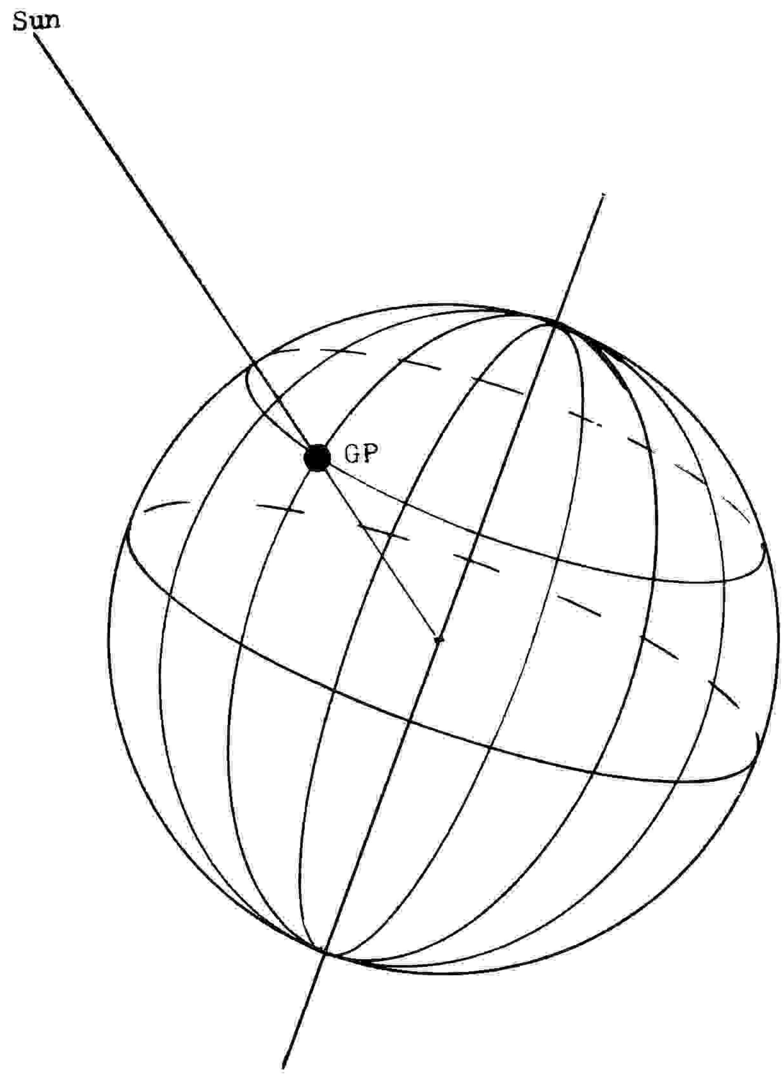


Figure 3.22: The nature of a body's geographical position, or GP.

reason they appear to be moving across the sky is because of the earth's rotation. The bit about the parallel rays means that light coming from a given star arrives at the earth with the same angle from all parts of the star, so your observations of it won't be distorted. This isn't necessarily true for the planets and is certainly not true for the moon.

Let's focus now on the idea that for celestial navigation, all the stars are "far away." This allows us to think of the sky as follows. Take a look at Figure 3.23. It shows the earth and a bunch of stars that you would be able to see from the surface of the earth. Remember that looking at the stars from earth, they all look like little points, that don't seem to move relative to one another, and are very far away. In this view, you can "wrap" a large sphere around the earth, so that all of the stars appear to fall on the inside "skin" of the large sphere, as shown in Figure 3.24

The large sphere with the stars on its inner skin is called the "celestial sphere." You can also just refer to it as the "sky." Remember, that physically of course all of the stars do not fit on the inner skin of a single sphere, because they are all at different distances from the earth. But as viewed from earth, they all appear to be the "same, large distance" away.

The celestial sphere has a lot in common with the earth. First, if you extend the north and south poles of the earth way out to the celestial sphere, you'll get the "celestial north and south poles." If you let the equator of the earth extend out until it touches the celestial sphere, you'll get the celestial equator. This is illustrated in Figure 3.25.

What is the advantage of having the celestial sphere? To answer this, notice that because the celestial sphere is "a sphere" that we can map the stars on, then we can we can also use latitude and longitude to specify the location of the stars, just as we do for point on the earth. There is a slight name change here, however. Instead of latitude, the north/south position of a celestial body is called the "declination." Also, instead of longitude, the east/west position of a celestial body is called the "right ascension" or RA for short. Unfortunately RA is not measured with respect to the prime meridian being projected onto the celestial sphere. Instead, another arbitrary reference is used that is a little "less arbitrary." To see where this zero-reference is for right ascension, take a look at Figure 3.26.

See where the Υ symbol is? This is the point in the Spring (about March 20th), where the sun's orbit intersects the celestial equator (and the equatorial plane of the earth). The line of longitude running through this point is where RA is zero. This point is sometimes called Aries, or the "first point or Aries." Stated another way, it is the point where the sun is on the celestial equator as it moves from south of the celestial equator to north of it, on March 20th, as spring heads into summer in the northern hemisphere (of the earth). There is no actual star at Aries. It is just an empty point on the celestial sphere. The point itself is in the constellation of Pisces, and the nearest star is Alpheratz[Kelch, p. 27]. Alpheratz has a declination of about $25°40'$ and a right ascension of $0.13°$ (a small number ≈ 0).

As far as the Nautical Almanac goes, the columns in Figures 3.20 and 3.21 labeled "GHA" and "Dec" stand for the "Greenwich hour angle" and "declination" of a given celestial object. GHA is the same a longitude and declination is the same as latitude. The column labeled "Aries," for instance, gives you the location on the celestial sphere of Aries, relative to the projection of the prime meridian onto the celestial sphere. This is also the GP of Aries, or the point on earth that the zero for right ascension is directly above on the celestial sphere.

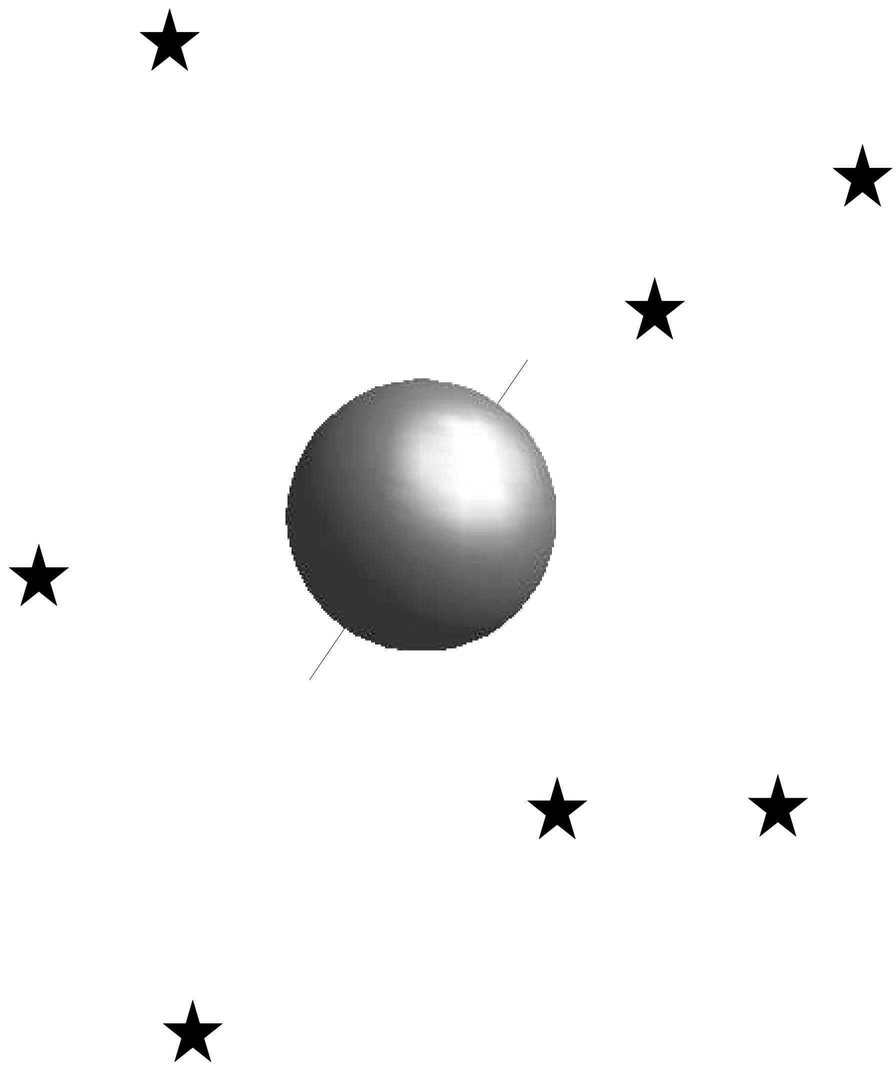


Figure 3.23: The earth and a few stars you would be able to see from the surface of the earth.

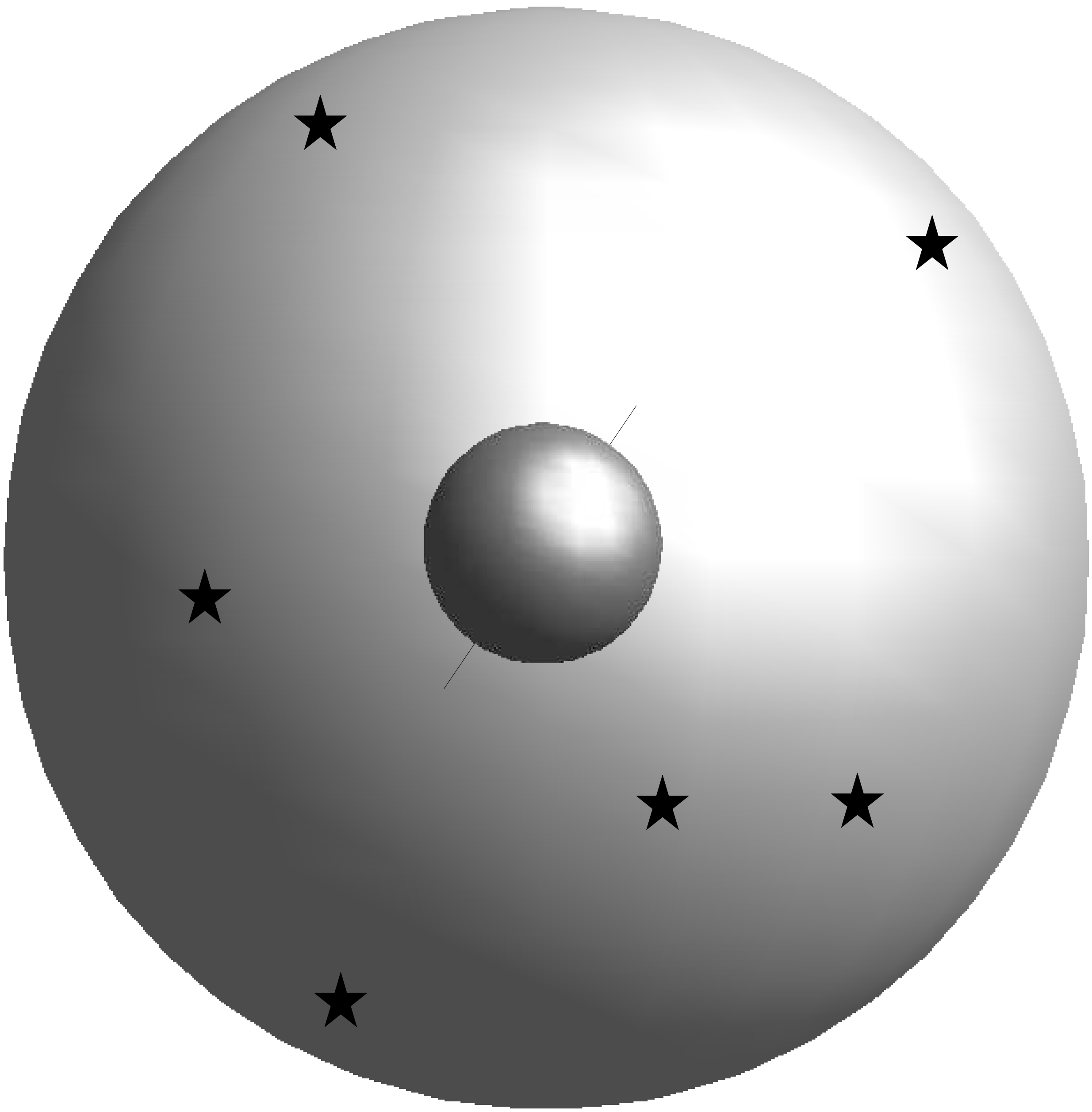


Figure 3.24: The earth, and few stars mapped onto the celestial sphere, which we'll call the sky.

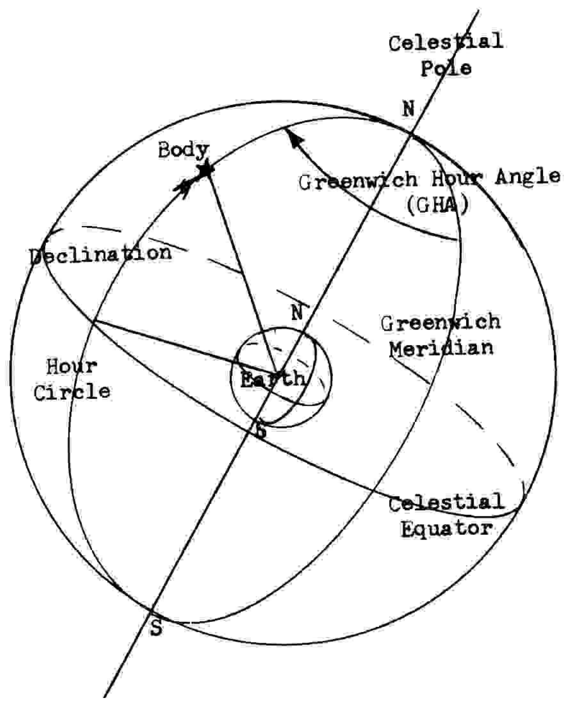


Figure 3.25: Features of the celestial sphere.

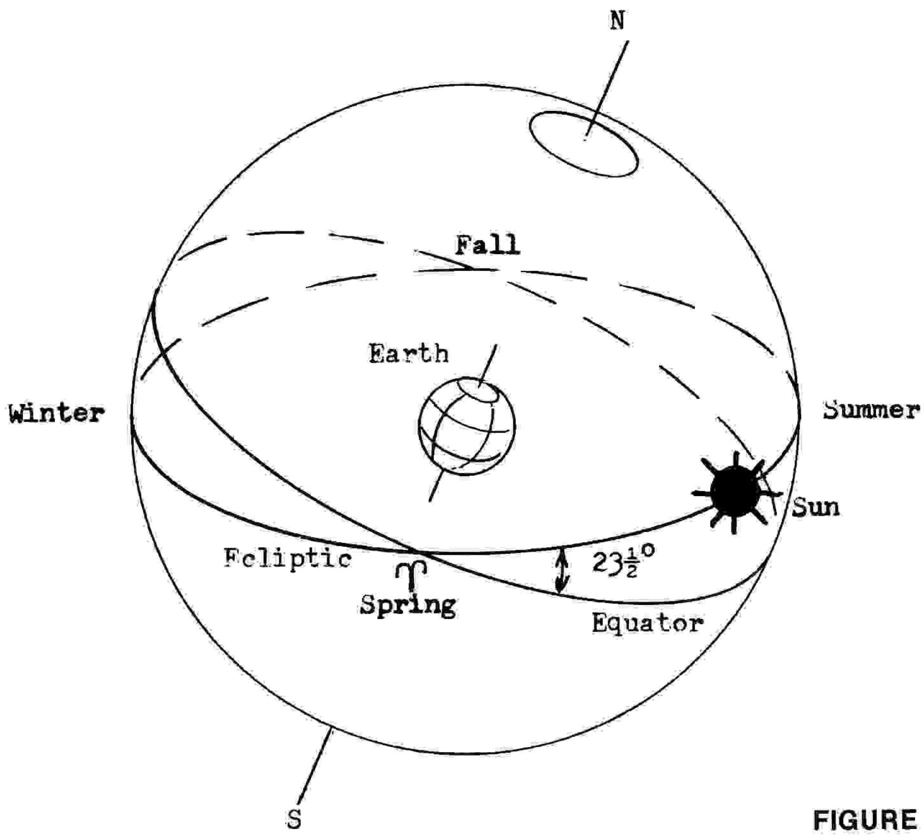


FIGURE 9

Figure 3.26: The point where the sun is at vernal equinox on March 20th. This is the point where the sun's orbit intersects the celestial equator. The line of longitude running through this point is where right ascension is zero.

There is one final note about RA: it is not measured in our usual units of $^\circ$,$'$, and $''$. It is measured in actual hours, minutes, and seconds that you use to tell time. This is because the celestial sphere seems to rotate due to the earth's rotation of 15°/hour. In other words, 360° of RA happen in 24 hours. So, an hour of right ascension is equal to $1/24$ of this, or 15° of arc, a single minute of right ascension equal to $15'$ of arc, and a second of right ascension equal to $15''$ of arc.[Wikipedia and Chaisson, p. 9].

Finally, a few closing words about the sky and navigation, summarized from [Gould, "The Marine Chronometer", p. 5-6].

> *As you read above, any point on the earth's surface can be defined by the intersection of its lines of latitude and longitude. The displacement from this point, upon any change in position can always be expressed as the resultant of the change in latitude and the change in longitude. In other words, if a ship starts from a known position and sails any given distance in any given direction, its total excursion can be stated as how far north or south it sailed, and how far east or west it sailed. The question, then, in using the sky for navigation is, what observable change, if any, is produced by either of these motions with respect to the heavenly bodies visible from the ship?*
>
> *Consider a ship in a north latitude, sailing southward. As it advances, all heavenly bodies southward of its original zenith (point directly overhead) will gradually become more and more elevated above the horizon. Stars which were at first below the horizon will come into view, and all the southern stars–those whose declinations is southerly–will rise earlier, set later, and cross her meridian at an increased altitude. The northern stars, on the other hand, will rise later and set earlier, and those to the northward of her original zenith (like the pole star) will gradually sink lower and lower. Thus, when the ship reaches the equator, the pole star will no longer be visible.*
>
> *Change in latitude, then, causes a corresponding change in the apparent altitude of the heavenly bodies, quite independent of their diurnal motion, and since the sextant enables altitudes to be observed at sea with considerable accuracy, a ship's latitude can be easily found with the help of such observations, in combination with tables, such as the "Nautical Almanac," of the celestial positions, at the time of observation, of the bodies observed. This was the method used, in a rudimentary fashion, by the early navigators.*
>
> *But when a ship sails eastward and westward, no change is produced by such motion in the apparent altitude of the heavenly bodies. The rotation of the earth will bring them across the ship's meridian, for example, at exactly the same altitude as before, and the only alteration produced by the ship's change of position will be that such transits will occur earlier if the ship has traveled eastward, and later if it has traveled westward. Thus if the ship alters its longitude by 90° to the east, a star which previously crossed the meridian at 11 pm will now do so at 5 pm, local time.*

> *We have seen then that a celestial observation taken by a ship in two different places will exhibit a relative change in altitude if it has gone northward or southward, and a relative change in local time if it has gone eastward or westward. Also by such observations, the ship can obtain its latitude and local time.*
>
> *But in order to know her change in longitude, the ship must know its local time and also the local time of the place it sailed from : or to know her actual longitude, which is more convenient, the ship must know the local time of some standard meridian. How is a knowledge of this time, which we may term "standard time," to be obtained?*

So, as you can see, the sky can be used to find one's latitude, and local time, (even easily in some cases). But you must still know your local time and the time at the prime meridian, to compute your longitude. This was, of course, the essence of the longitude problem: keeping the local time at the prime meridian, while at sea.

Activity

Use Google sky to find the RA and Dec. of common objects. Get a feeling for N-S/E-W movements.

3.3.4 The Sun

The sun was (and is) the "crown jewel" for celestial navigation. It's big and bright and unmistakable in the sky. Even if you were totally lost at sea, and can't see it because of fog, clouds, etc., you would still be able to at least judge the passage of days (via day/night transitions).

You already saw in Section 3.3.2 how the GP of a celestial body can be found. The most straightforward measurements relating to celestial navigation are always done with the sun. On a nice, sunny day you can:

1. Measure the sun's altitude (θ) above the horizon,
2. find α, the zenith angle of the sun, which is $\alpha = 90^\circ - \theta$,
3. look up the sun's GP in the Nautical Almanac at your time of observation,
4. find how many nautical miles you are from the sun's GP by simply computing $60 \times \alpha$, since there are 60 nautical miles in 1° (on a great circle of the earth).

Now you know how far you are from the sun's GP Later, in Section 3.4 we will proceed further with this calculation to actually find your location from this basic measurement of the sun.

Another useful measurement that can come from the sun is to determine when it is noon at your location. This is illustrated in Figure 3.27. Imagine that you are out at sea at some unknown line of longitude, as shown by the dark vertical line in Figure 3.27. You notice that

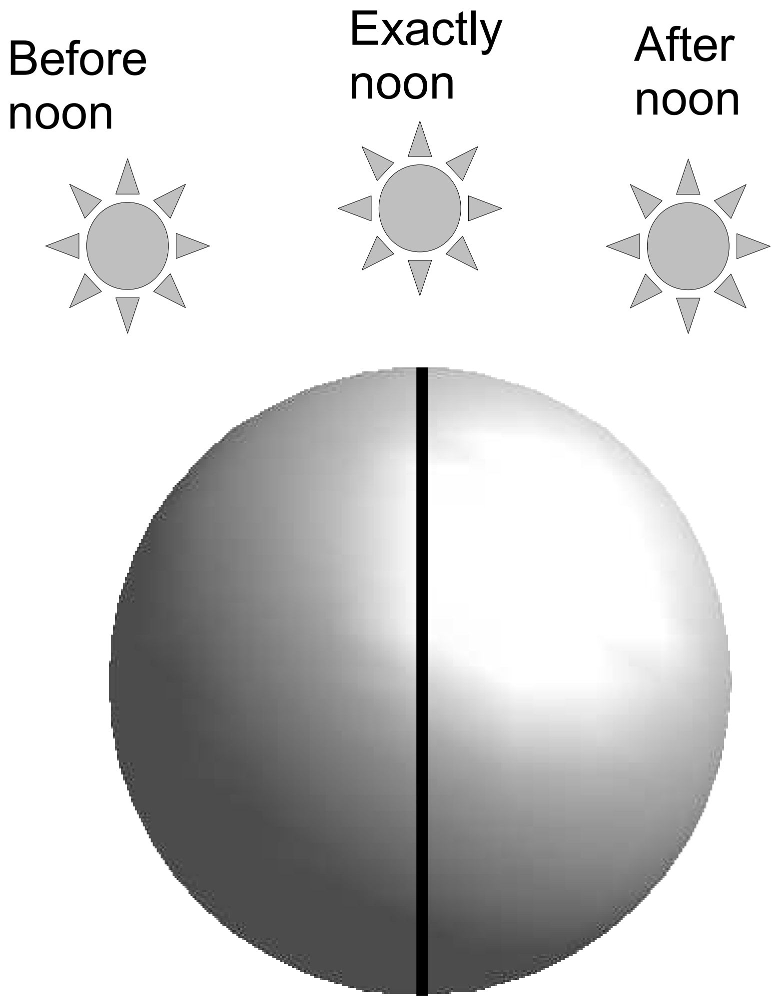


Figure 3.27: How local noon is found using the sun.

the sun is rising in the sky, and getting quite high. Using your sextant, you begin taking its altitude. You take larger and larger altitude readings, then they start getting smaller again. The largest altitude signifies the sun's highest point in the sky, when it was your local noon.

On land (and perhaps even at sea on a calm day), local noon can be found by planting a cross in the ground, with the horizontal portion of the cross oriented directly north/south (as found with a compass). When the sun is at its highest point, the shadow of the horizontal part of the cross will not be visible. (Without a cross, all objects will show their shortest shadow at noon. You could track the extent of shadows as the sun rises. They will get shorter and shorter, then begin growing again. The shortest you measure is when your local noon occurred.)

The importance of this measurement is that at the instant of noon, the sun is directly over

the meridian (line of longitude) on which you are located. It is noon for you and everyone else on this meridian. So you just found the time with the sun; it is exactly 12 noon for you, at your location. This is what is meant by your "local time." Suppose that you also had a way of knowing the time in Greenwich, England. You could subtract your local noon from Greenwich time, multiply by 15°/hour, and you'd have your longitude.

Hopefully this discussion gives you the clearest picture yet on the connectedness between longitude and time. Suppose that you are 60° east of Greenwich and the sun is over your meridian. It will take 4 more hours of earth rotation (at 15°/hour) for the sun to move off of your meridian and over the meridian of Greenwich. It will be 12 noon when the sun is over your meridian, but only 8 am in Greenwich. It will be 4 pm for you, when the sun is over the meridian at Greenwich, signifying 12 noon over there.

Naturally, any object on the celestial sphere has a meridian passage, or a time when it is directly over the line of longitude of your position. Take a look at the chart in Figure 3.28 from the Nautical Almanac. You can see that for all the days of the year, the sun will be over your local meridian at about 12 noon (makes sense). Look also, for example, at Mars on June 20th (another object on the celestial sphere). It will pass over your meridian at about 4:30 in the morning. Planets, however, cannot be relied upon for meridian timing because not all planets are visible all of the time, depending on your latitude (which is not taken into account in the chart). Also, planets simply might not be visible because of the bright sun (the planet might make a meridian passage during the day). But in principle, one could track the altitude of Mars, determine when it is highest in the [night] sky, and then determine the local time from a chart, compare with Greenwich, and compute longitude.

3.3.5 The Moon and Lunar Distances

The moon orbits the earth once every 27.3 days, which we call a "lunar month." Like the sun and stars, the moon rises in the east and sets in the west, because of the earth's west to east rotation. So like the sun and stars, you might expect the moon to move through the sky at about 15° per hour. But the moon also revolves around the earth from west to east, so its motion in the sky is a little less than the 15° per hour; it wanders across the sky imperceptibly slower than the sun and stars. This so called "retardation" makes the moon's schedule a bit erratic[Rey, p. 136]. It rises about 50 minutes later each day. The full moon has an angular size of approximately 0.5°.

For our needs in this class, there is nothing too remarkable about the moon's motion relative to the earth. It is mentioned here because by the early 1700's, using the moon and "the lunar distance" method (see Chapter 1) was one of two best ideas for solving the longitude problem. Although this method did not win the longitude prize and hasn't been used to navigate since about 1800, it plays a very important historical role in the longitude story. (Note: Joshua Slocum discussed this in *Sailing Around the World Alone*, as late at 1890.)

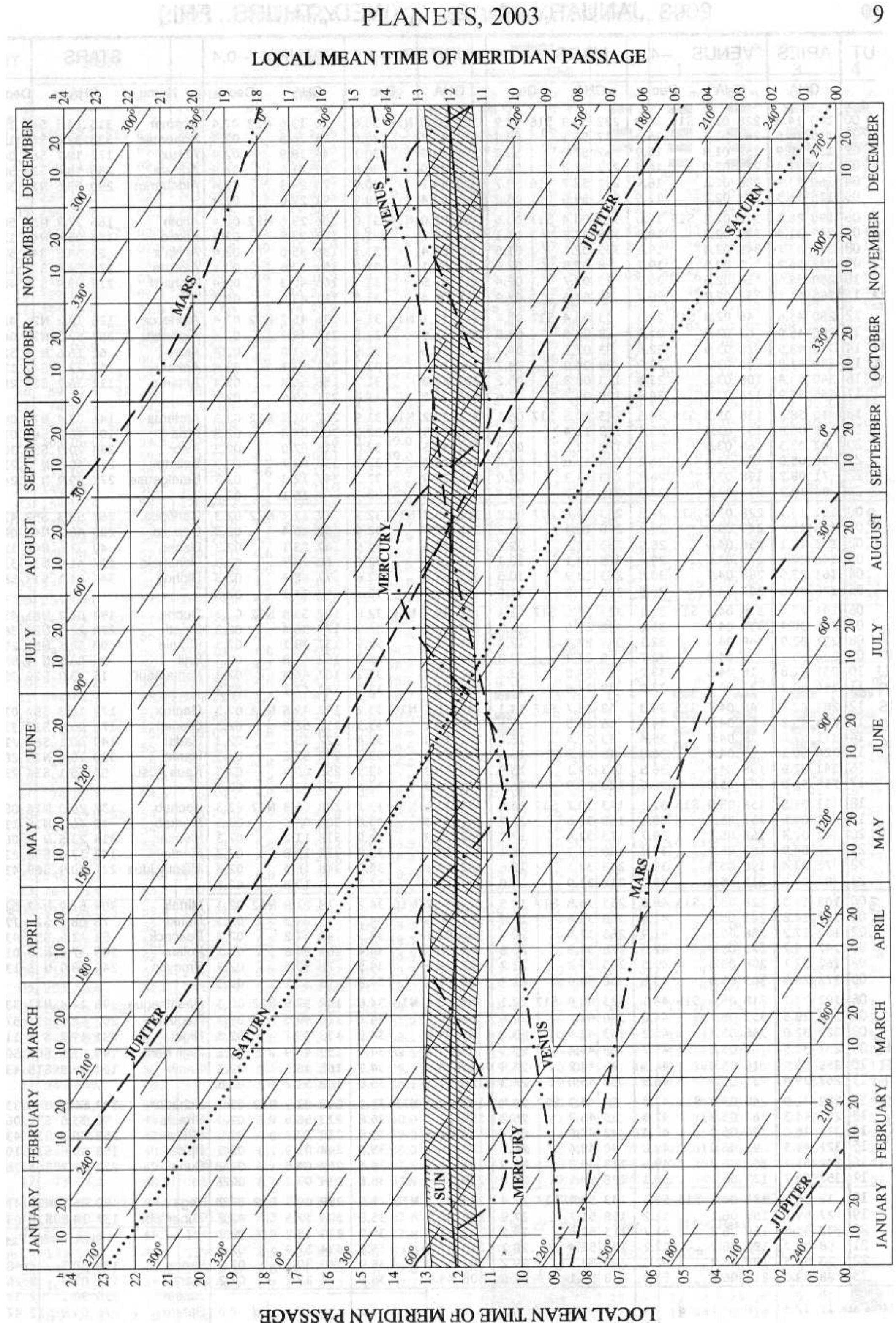


Figure 3.28: Meridian passage schedule of the sun and planets.

Find Greenwich time out on your ship

The lunar distance method to find longitude starts with cataloged positions of the moon, as shown in Figure 3.29. This figure shows an excerpt from Edmund Halley's tables derived from his efforts to observe the moon over an 18 year "saros" cycle.

In the table, columns one and two (from the left) are the date and time in Greenwich, England (where Halley worked, and made his measurements). The third column is the moon's declination (see below), which is half of what one must know to pin-point its position in the sky. The fourth column is the moon's longitude, as observed by Halley. The fifth is the moon's longitude as computed by Isaac Newton's theory. The sixth is the error between Halley's observations and Newton's theory.

The idea of the tables is that a given longitude of the moon only happens once a day[Schlereth, VK555.S34, p. 134], sort of like how the sun hits a high-point in the sky only once per day (at noon). Looking at Halley's table, you can see that the moon's longitude was $20°$, $50'$ and $35''$ on June 22, 1732, at $8:21:22$ in the morning. These lunar positions could then be combined with positions of other heavenly bodies to compute distances from the moon to these bodies every three hours of every day, as shown in Figure 3.30. Such tables were compiled by Maskelyne in the 1700s (the one in the Figure was compiled by S. Wepster, at http://www.math.uu.nl/people/wepster/ldtab.html). The names are all of easy-to-find stars, with the exception of the planet Saturn.

Note how the leftmost column is the time it must be in Greenwich for the given distance to be occurring. This column is labeled "UT" which stands for "Universal Time," the modern day name for "the time in Greenwich." This table then, was used to compute what time it must be in Greenwich based on the position of the moon. It went as follows[Navigation and Nautical Astronomy, H.W. Jeans, 1853, p. 86 Rule XV.]:

1. Find the distance between the moon and a celestial object in the tables. Call this distance d.

2. Look in the tables, and find what two distances d is between. Call the lower distance D_1, and the higher distance D_2. You already have the time in Greenwich to within 3 hours; but you have to be more precise to get a good longitude estimate. Computing a result, based on known low and high results is called "interpolation."

3. To interpolate between the two distances below to "hone in" on the more exact Greenwich time, you can set up a simple proportionality between what you know and what you'd like to know. For this lunar result, it will look like this

$$\frac{10,800}{t-t_1} = \frac{D_2 - D_1}{d - D_1}, \tag{3.14}$$

   where $10,800$ is the number of seconds in 3 hours (the resolution of the tables), t is the number of seconds past t_1, which is Greenwich time at which you made your measurement, d is your distance measurement, and D_2 and D_1 are the low and high

LUNÆ MERIDIANÆ LONGITUDINES
GRENOVICI OBSERVATÆ
CUM COMPUTO NOSTRO COLLATÆ.

Anno JULIANO MDCCXXXII. Currente.

Transitûs Limbi Lunæ T. æq.					Argument. Annuum.			Distantia ☾ à ☉			Longitudo Centri Lunæ Observata.			Longitudo Centri Lunæ Comput.			Error Comp.	
M.	D.	H.	′	″	S.	°	′	S.	°	′	°	′	″	°	′	″	′	″
Junii.	21	7	32	44	9	11	58	3	27	28	♏ 8	3	36	♏ 8	5	8	+1	32
	22	8	21	22	9	12	50	4	9	26	♏20	50	35	♏20	52	6	+1	31
	23	9	13	15	9	13	43	4	21	32	♐ 3	58	20	♐ 3	59	18	+0	58
	24	10	7	57	9	14	36	5	4	0	♐17	29	31	♐17	29	41	+0	10
	25	11	4	22	9	15	29	5	16	47	♑ 1	24	13	♑ 1	24	10	−0	3
Junii.	29	14	46	18	9	19	0	7	10	49	♓ 0	11	19	♓ 0	9	33	−1	46
Julii.	2	17	19	38	9	21	38	8	22	51	♈14	57	41	♈14	54	1	−3	40
	6	20	58	32	9	25	9	10	16	58	♊12	18	24	♊12	13	55	−4	29
	14	2	37	31	10	1	18	1	10	37	♍13	50	26	♍13	50	31	+0	5
: :	15	3	18	32	10	2	10	1	21	53	♍26	7	28	♍26	8	20	+0	52
: :	17	4	41	54	10	3	55	2	14	32	♎20	32	35	♎20	35	3	+2	28
	20	7	1	59	10	6	33	3	19	47	♏28	4	48	♏28	6	53	+2	5
	23	9	46	0	10	9	12	4	27	44	♑ 8	49	1	♑ 8	47	36	−1	25
	24	10	42	43	10	10	5	5	11	9	♑23	20	20	♑23	19	23	−0	57
	25	11	38	39	10	10	58	5	24	52	♒ 8	17	45	♒ 8	14	44	−3	1
	28	14	21	35	10	13	38	7	7	18	♓24	27	48	♓24	26	26	−1	22
	30	16	7	46	10	15	24	8	5	44	♈24	52	16	♈24	50	1	−2	15
Julii.	31	17	2	20	10	16	17	8	19	42	♉ 9	37	46	♉ 9	34	50	−2	56
Aug.	1	17	58	4	10	17	10	9	3	24	♉24	2	38	♉23	58	55	−3	43
	2	18	54	18	10	18	4	9	16	45	♊ 8	6	56	♊ 8	2	53	−4	3
	3	19	50	5	10	18	57	9	29	44	♊21	52	10	♊21	48	14	−3	56
	5	21	35	57	10	20	43	10	24	34	♋18	31	36	♋18	30	4	−1	32
	18	6	37	1	11	1	20	3	12	27	♐18	41	24	♐18	43	2	+1	38
	19	7	31	27	11	2	14	3	25	9	♑ 2	10	33	♑ 2	10	22	−0	11
	20	8	26	46	11	3	8	4	8	15	♑16	9	7	♑16	7	23	−1	44
	22	10	17	11	11	4	55	5	5	37	♒15	41	46	♒15	37	49	−3	57
	25	13	2	6	11	7	36	6	18	41	♈ 2	43	17	♈ 2	39	59	−3	18
	26	13	57	7	11	8	30	7	3	11	♈18	25	45	♈18	22	57	−2	48
	27	14	53	17	11	9	24	7	17	32	♉ 3	50	55	♉ 3	48	37	−2	18
	28	15	50	34	11	10	18	8	1	40	♉18	52	41	♉18	49	55	−2	46
Aug.	29	16	48	15	11	11	12	8	15	27	♊ 3	27	0	♊ 3	23	37	−3	23
Sept.	1	19	33	4	11	13	53	9	24	22	♋14	32	17	♋14	29	25	−2	52
	2	20	22	35	11	14	47	10	6	30	♋27	31	54	♋27	30	13	−1	41
	4	21	53	21	11	16	34	10	29	45	♌22	50	7	♌22	51	23	+1	16

Figure 3.29: Reproduction of Halley's moon observations versus Newton's predictions[Kollerstrom, p. 184].

Precomputed Lunar Distances 1

FEBRUARY 2007

	° ′	P.L.	° ′	P.L.	° ′	P.L.	° ′	P.L.	° ′	P.L.
1 UT	+Aldebaran		+Elnath		−Saturn		−Regulus		−Spica	
0	48 10.0	2840	34 25.4	2706	25 21.1	2682	32 53.9	2709	86 54.2	2707
3	49 43.6	2843	36 1.9	2715	23 44.0	2692	31 17.4	2718	85 17.7	2716
6	51 17.1	2846	37 38.2	2724	22 7.2	2701	29 41.1	2727	83 41.4	2725
9	52 50.6	2850	39 14.4	2733	20 30.5	2710	28 5.1	2736	82 5.3	2734
12	54 24.0	2854	40 50.3	2742	18 54.1	2720	26 29.2	2745	80 29.4	2742
15	55 57.2	2859	42 26.0	2751	17 17.9	2730	24 53.6	2755	78 53.7	2752
18	57 30.4	2864	44 1.5	2760	15 41.9	2739	23 18.1	2764	77 18.2	2761
21	59 3.5	2870	45 36.9	2769	14 6.1	2751	21 42.9	2775	75 42.9	2770
2 UT	+Aldebaran		+Capella		+Elnath		+Pollux		−Spica	
0	60 36.4	2877	50 27.7	3137	47 12.0	2778	16 57.9	2824	74 7.8	2779
3	62 9.2	2882	51 55.1	3131	48 46.9	2787	18 31.8	2826	72 32.9	2789
6	63 41.9	2890	53 22.7	3126	50 21.6	2797	20 5.7	2830	70 58.2	2798
9	65 14.4	2896	54 50.3	3123	51 56.2	2806	21 39.5	2835	69 23.7	2807
12	66 46.8	2903	56 18.0	3120	53 30.5	2815	23 13.2	2840	67 49.4	2816
15	68 19.1	2911	57 45.7	3119	55 4.6	2824	24 46.8	2847	66 15.3	2825
18	69 51.1	2918	59 13.5	3119	56 38.5	2833	26 20.2	2853	64 41.3	2834
21	71 23.1	2926	60 41.2	3119	58 12.3	2842	27 53.5	2861	63 7.6	2843
3 UT	+Capella		+Elnath		+Pollux		−Spica		−Jupiter	
0	62 9.0	3121	59 45.8	2852	29 26.6	2868	61 34.1	2853	111 46.6	2935
3	63 36.7	3122	61 19.2	2861	30 59.6	2876	60 0.8	2862	110 15.0	2944
6	65 4.4	3124	62 52.3	2869	32 32.4	2883	58 27.7	2871	108 43.6	2952
9	66 32.1	3128	64 25.3	2879	34 5.1	2892	56 54.7	2880	107 12.4	2961
12	67 59.7	3131	65 58.0	2887	35 37.6	2899	55 22.0	2889	105 41.4	2969
15	69 27.2	3135	67 30.6	2896	37 9.9	2907	53 49.5	2897	104 10.5	2978
18	70 54.6	3140	69 3.0	2905	38 42.0	2916	52 17.1	2907	102 39.9	2987
21	72 22.0	3144	70 35.2	2913	40 14.0	2923	50 45.0	2915	101 9.4	2995
4 UT	+Pollux		+Saturn		+Regulus		−Spica		−Jupiter	
0	41 45.8	2932	12 37.5	2903	4 55.0	2993	49 13.0	2924	99 39.1	3003
3	43 17.5	2939	14 9.7	2909	6 25.3	2974	47 41.2	2932	98 8.9	3011
6	44 48.9	2948	15 41.8	2917	7 56.0	2969	46 9.6	2941	96 39.0	3020
9	46 20.2	2955	17 13.8	2924	9 26.9	2969	44 38.1	2949	95 9.2	3028
12	47 51.4	2963	18 45.6	2931	10 57.7	2972	43 6.8	2957	93 39.6	3036
15	49 22.4	2970	20 17.2	2937	12 28.5	2976	41 35.7	2965	92 10.1	3043
18	50 53.2	2978	21 48.7	2946	13 59.2	2981	40 4.8	2973	90 40.8	3051
21	52 23.9	2985	23 20.1	2952	15 29.8	2987	38 34.1	2981	89 11.6	3059
5 UT	+Pollux		+Saturn		+Regulus		−Spica		−Jupiter	
0	53 54.4	2992	24 51.3	2959	17 0.3	2993	37 3.4	2988	87 42.7	3066
3	55 24.8	2999	26 22.3	2966	18 30.7	2999	35 33.0	2996	86 13.8	3074
6	56 55.0	3005	27 53.2	2973	20 0.9	3005	34 2.7	3002	84 45.1	3080
9	58 25.1	3012	29 24.0	2979	21 31.0	3011	32 32.5	3010	83 16.6	3088
12	59 55.0	3018	30 54.7	2986	23 0.9	3017	31 2.5	3016	81 48.2	3094
15	61 24.9	3024	32 25.2	2991	24 30.8	3022	29 32.7	3022	80 19.9	3100
18	62 54.6	3030	33 55.6	2997	26 0.5	3028	28 2.9	3029	78 51.8	3107
21	64 24.1	3036	35 25.8	3003	27 30.2	3034	26 33.3	3035	77 23.7	3113

Figure 3.30: Precomputed lunar distances, from S. Wepster, at http://www.math.uu.nl/people/wepster/ldtab.html

measurements that bound your observed d. (Note: you should convert all of your D's and d into pure degrees.) You can solve the above equation for t to get

$$t = t_1 + 10,800 \left[\frac{d - D_1}{D_2 - D_1} \right]. \tag{3.15}$$

So there! You have Greenwich time, from way out on your ship. Now, if you can just find your local time, you'd be able to find your longitude.

Find your local time

Continuing with the lunar method, you now need to find the local time aboard your ship, or the local time of your lunar distance observation. If you had this, you'd see how many hours off of Greenwich time you are, multiply by 15°/hour, and you'd have your longitude.

It is possible to find your local time by observing the height of the sun, moon, or stars in the sky. This is not an unrealistic claim. Take the sun (or any one of these bodies) as an example. It first rises from the horizon at sunrise, reaches its highest point in the sky at noon, then disappears below the horizon again at sunset. And, it spends equal time rising before noon as setting after noon. Here's how to find your local time by observing the sun[Navigation and Nautical Astronomy, H.W. Jeans, 1853, p. 210-211 Rule L.]:

> *To find ship mean time. Under the sun's declination put the latitude of the ship; take the sun if their names be unlike, the difference if the names be alike. Under the result put the sun's zenith distance; take the sum and difference of the last two lines put down. Add together the log secants of the two quantities in this form (omitting to put down the tens in the index) and half of the log. haversines of each of the last two quantities. The sum will be the log. haversine of the ship apparent time. To this apply the equation of time with its proper sign and the result will be the ship mean time.*

Huh? What? Shall we try to interpret this and run through and example? No! It's too difficult, and this is one of the reasons why the "lunar distance method" didn't win the longitude prize. It simply required too many mathematical computations, and this single part of determining your local time is an example. Many references estimate that it took up to 4 hours of computation to determine one's longitude after the moon observations were recorded. Let's just know that it is possible to find your local time by observing a celestial object and leave it at that.

Finally, for a few closing words about using the moon for navigation from [Gould, p. 8], see Figures 3.31, 3.32, 3.33. From [Schlereth, Yellow Book, p. 115] we note

> *The moon is close to earth—a mere couple of hundred thousand miles off—and the vagaries of its orbit are only too evident. The Nautical Almanac has a hard time keeping track of it. The basic difficulty is that the almanac calculates the instantaneous hourly positions of celestial bodies, which works fine for the stars,*

> *sun, and planets, because their GHA and declinations run along at steady rates or change quite slowly. The maximum rate of change of the sun's declination is, for example, $1'$ per hour, and its GHA varies so little from the adopted $15°$/hour that the difference is negligible.*
>
> *In the case of the moon, however, the basic rate of the GHA varies from hour to hour and is sometimes greater than $15'$. You need to interpolate, and you need another table. The declination behaves the same way, increasing or decreasing by $15'$ per hour or more. This means more work for the navigator.*
>
> *The moon's proximity to earth is the source of another complication for celestial navigators (and another table, naturally). This complicating factor is called parallax.*

So, in principle, you can can get your longitude by observing a distance between the moon and another celestial body but it's prohibitively difficult. Not very practical. G.W. Littlehales [Bulletin of the American Geophysical Society, Vol 41, No. 2 (1909), pp. 83-86] states "it is found hardly one navigator in two hundred has ever carried out the calculation of a longitude by the lunar-distance method in practice." Philip Van Horn Weems [Proceedings of the American Philosophical Society, Vol 98 No 4 (Aug 16 1954), pp. 270-272] states "I have never known a navigator who knew a person who made practical use of lunar distances."

3.3.6 The Stars

On any given evening, about $4,000$ stars are visible to the naked eye. You are free to use any of them for navigation purposes. They all move from east to west at about $15°$/hour. They rise earlier the further east you are and later the further west you are. Their height in the sky will vary with your north-south position. The Nautical Almanac helps you decide which of these $4,000$ stars to choose from by giving the GP of only 57 of them, as shown in the rightmost column of Figure 3.20. The most familiar ones (to us in the northern hemisphere) are found by "hopping" to them from more familiar features in the night sky, such as the Big Dipper or Orion. How to find some of these more common stars is shown in Figure 3.34 for the Orion constellation and in Figure 3.35 for the Big Dipper.

3.3.7 The Moons of Jupiter

Jupiter is a unique planet. It is the largest in our solar system and, when visible, it is very bright and easy to find in the night sky. It has long been a darling for navigators, even a possible solution to the longitude problem for one simple reason: it has many very bright moons orbiting around it, the first four being particularly big and bright. They are the "Galilean" moons and are named Io, Europa, Ganymede and Callisto.

Jupiter's Moons

The words "easy to find in the night sky" need a bit of context. Jupiter looks like a big bright "star" when visible. If Jupiter is visible at your location, and you know where to look for it,

making use of it.

The one most often proposed* is that of lunar transits—and theoretically it is quite sound. Unlike a star, which crosses every meridian at the same local time, the local time at which the moon crosses each meridian is affected by her motion round the earth, and if the time of her transit on any particular day be calculated, say, for Greenwich (as it can be), it will be found to get later and later for places further westward, and *vice versa*, by a regular amount, which is called her "retardation." Now if we could observe at sea the exact local time when the moon crossed the meridian, and compare it with the tabulated time of her crossing that of Greenwich, the difference between the two would give the retardation. And since this always bears a fixed relation to the difference of longitude between the standard meridian and the observer, the latter's longitude could readily be obtained.†

In practice, however, this method fails utterly, for the simple reason that there is no known means of determining, with anything like the accuracy required, when the moon, or any other heavenly body is on the meridian, whether by a direct transit observation or by the mean of sights taken on either side of the meridian. If it were not so, indeed, finding longitude would be a very simple matter.

But although this method of using the moon fails, there is another which is quite practicable, although no longer used—the method of "lunar distances." If the moon's motion be known with sufficient accuracy, tables can be drawn up forecasting her position in the heavens for a long time in advance, and also her angular distance, as observed on some standard meridian, from suitable fixed stars. These distances can also be observed, by means of the sextant, on board ship, and, by interpolation, the Greenwich time corresponding with that distance can be taken out of the tables. The local time of observing such "lunar distances" can be obtained by the ordinary observations, and the difference, of course, gives the longitude of the ship.

Figure 3.31: From Gould, *The Marine Chronometer, p. 8*

It was a later Astronomer-Royal, Maskelyne, who brought the lunar method into general use. He published in 1763 the " British Mariners' Guide," which gave a general outline of its principles, and sufficient tabular information for their application (though by a very laborious method†), while four years later he instituted the " Nautical Almanac," in which he gave, for the first time in the history of navigation, lunar distances of the sun and seven selected stars computed for every threc hours at Greenwich. The " Nautical Almanac " continued to publish such distances uninterruptedly, and several years in advance, until 1907, when they were discontinued, as no longer worth the trouble of computing.

But at no time during that period had they been an entirely satisfactory solution of the problem. In the hands of a good observer and computer they were excellent, but for general use they were unreliable. The reason was twofold. In the first place, the observations had to be extremely accurate—an error, such as the best observer could hardly make certain of avoiding, of only 1′ of arc produced, owing to the moon's comparatively slow motion, one of 30′ or so of longitude in the result. Secondly, the calculations were long and intricate, and although many efforts were made to simplify them, there remained many pitfalls and chances of committing some slight error which might easily pass unnoticed, and yet convert the result from a safeguard to a source of fearful danger. The combination of a good observer and a good computer was not very usual,‡ while even those expert in both branches could not guarantee, however favourable the observing conditions, that the mean results of several sets of distances would not exhibit considerable discrepancies.§

The method of lunar distances, as we have seen, was early suggested, but remained inapplicable till 1764 for lack of fundamental data,

Figure 3.32: From Gould, *The Marine Chronometer, p. 9*, continued below.

and no other method of obtaining a standard of time by celestial observations appeared feasible. It was natural, therefore, that enquiry should be directed towards finding some other means of obtaining such a standard.

The obvious method, of course, is to carry some clock or other timekeeper on board, which will give the standard time. The longitude can then be found very simply by comparing it with the local time found by observation.

Figure 3.33: From Gould, *The Marine Chronometer, p. 10*

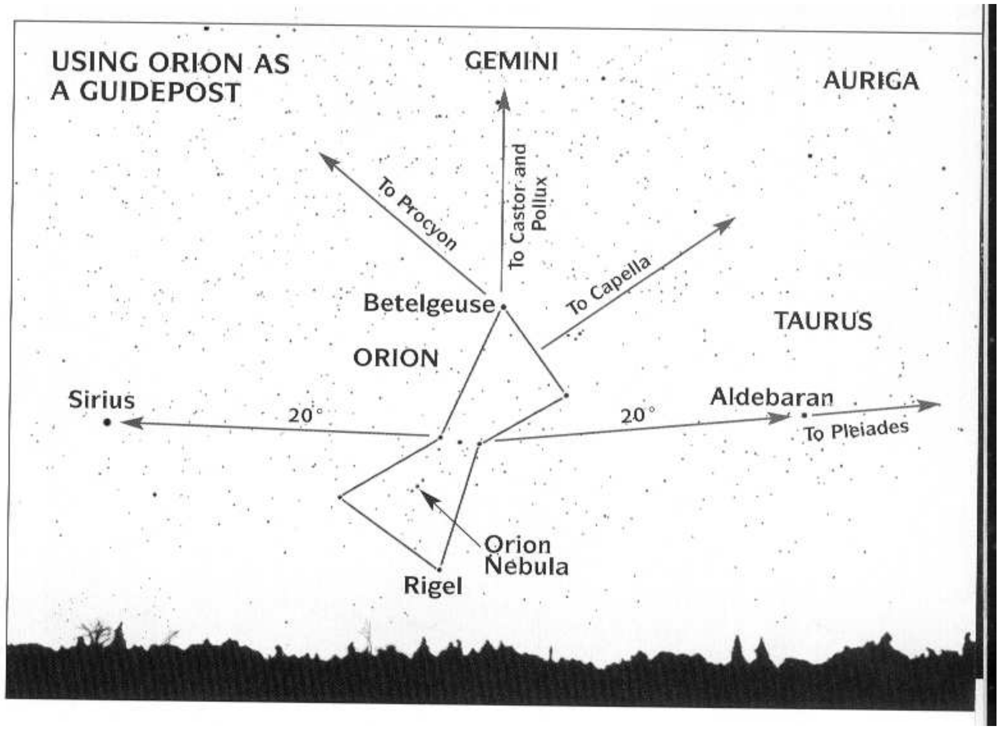


Figure 3.34: Finding stars using Orion (from T. Dickinson *Nightwatch*)

you can easily spot it with the naked eye. To see the moons, one needs nothing more than a cheap, plastic, $50 telescope from Walmart, like a 60 mm refracting type (such a telescope is probably of worse quality than the one Galileo used around 1605). When viewed through such a telescope, Jupiter will transform from a bright "dot" in the sky, to a readily apparent circular disc with four (or more) bright little dots distributed around it, which are its moons (the first four are the largest). You might even be able to see color bands on the surface of the planet. Also, the planes of the moons' orbits lie in the ecliptic (the plane containing the Earth and Sun), so they are viewed always head-on.

Jupiter is attractive for navigation because its moons (its first four in particular) orbit around the planet in a very regular easy-to-see motion. Being a bit more careful, one can observe the moons disappearing behind the planet (called an occultation), suddenly reappearing, or even eclipsed by Jupiter's shadow relative to the sun. Times of the events between Jupiter and its moons are summarized in Figure 3.36. Galileo first discovered the moons of Jupiter in 1610, and almost immediately saw Jupiter's moons as a solution to the longitude problem[Andrews, p. 89]. The moons could act as a "celestial timepiece."

All told, by watching and timing the action of Jupiter's moons, the planet *essentially becomes a clock in the sky, visible to the world at large.* We say a "clock" because the orbits of Jupiter's moons are very regular and can be predicted with relative ease. This means that tables can be constructed for every day of the year, stating when (in Greenwich time) a particular moon can be observed doing "something" around the planet (disappearing,

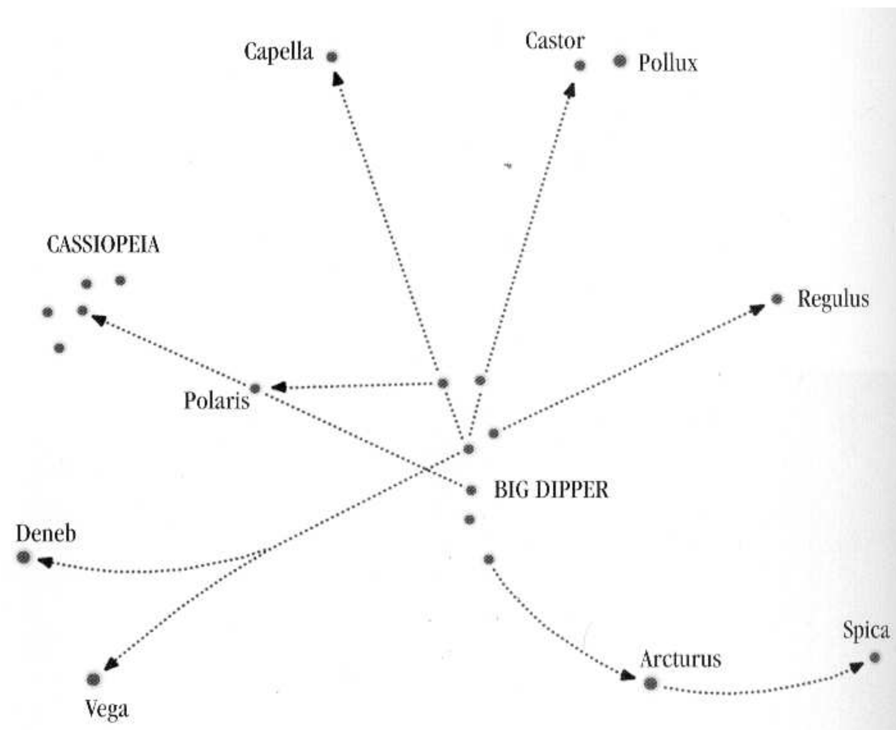
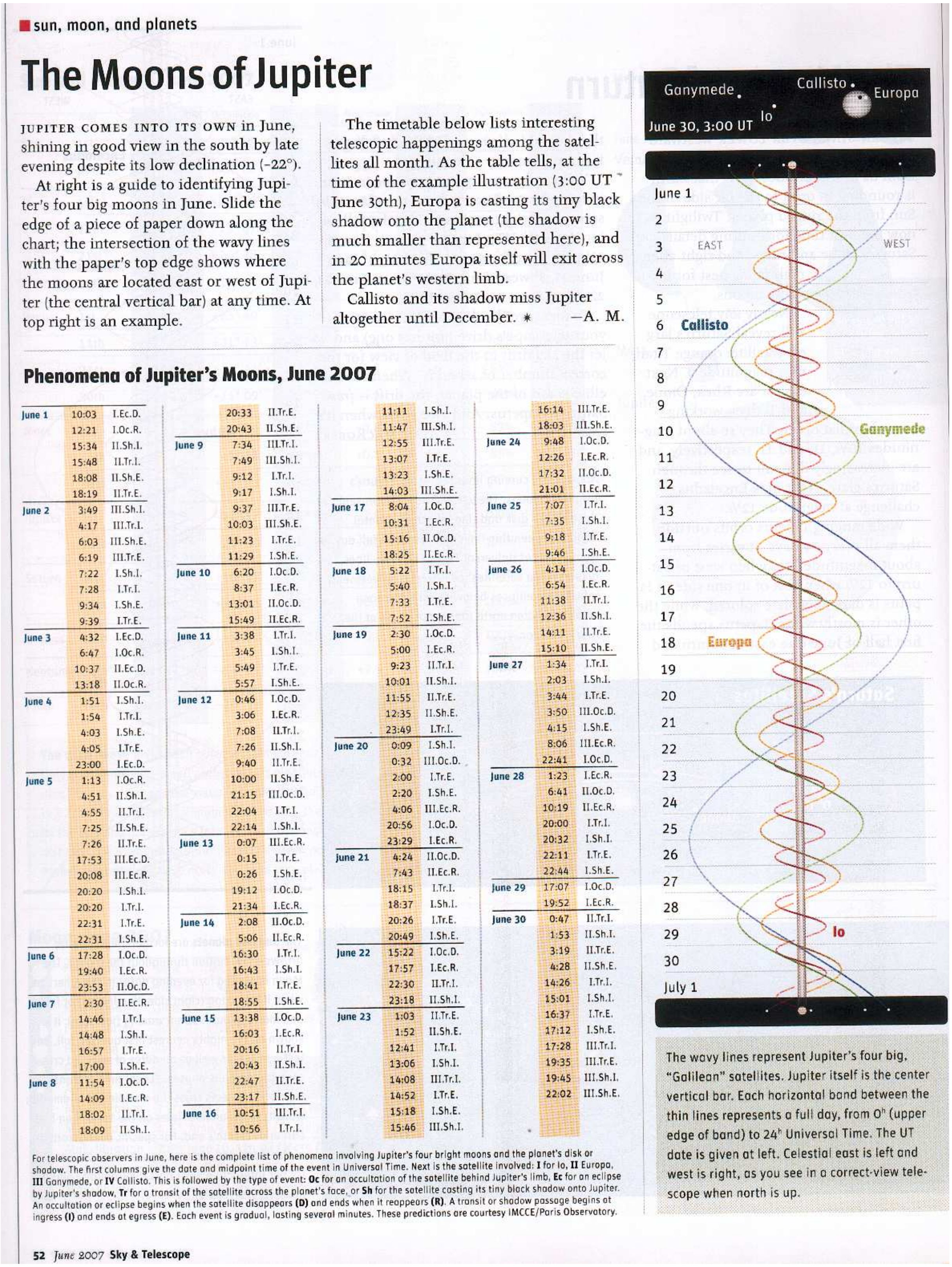


Figure 3.35: Finding stars using the Big Dipper (from T. Dickinson *Nightwatch*)



sun, moon, and planets

The Moons of Jupiter

JUPITER COMES INTO ITS OWN in June, shining in good view in the south by late evening despite its low declination (–22°).

At right is a guide to identifying Jupiter's four big moons in June. Slide the edge of a piece of paper down along the chart; the intersection of the wavy lines with the paper's top edge shows where the moons are located east or west of Jupiter (the central vertical bar) at any time. At top right is an example.

The timetable below lists interesting telescopic happenings among the satellites all month. As the table tells, at the time of the example illustration (3:00 UT June 30th), Europa is casting its tiny black shadow onto the planet (the shadow is much smaller than represented here), and in 20 minutes Europa itself will exit across the planet's western limb.

Callisto and its shadow miss Jupiter altogether until December. ∗ —A. M.

Phenomena of Jupiter's Moons, June 2007

June 1	10:03	I.Ec.D.
	12:21	I.Oc.R.
	15:34	II.Sh.I.
	15:48	II.Tr.I.
	18:08	II.Sh.E.
	18:19	II.Tr.E.
June 2	3:49	III.Sh.I.
	4:17	III.Tr.I.
	6:03	III.Sh.E.
	6:19	III.Tr.E.
	7:22	I.Sh.I.
	7:28	I.Tr.I.
	9:34	I.Sh.E.
	9:39	I.Tr.E.
June 3	4:32	I.Ec.D.
	6:47	I.Oc.R.
	10:37	II.Ec.D.
	13:18	II.Oc.R.
June 4	1:51	I.Sh.I.
	1:54	I.Tr.I.
	4:03	I.Sh.E.
	4:05	I.Tr.E.
	23:00	I.Ec.D.
June 5	1:13	I.Oc.R.
	4:51	II.Sh.I.
	4:55	II.Tr.I.
	7:25	II.Sh.E.
	7:26	II.Tr.E.
	17:53	III.Ec.D.
	20:08	III.Ec.R.
	20:20	I.Sh.I.
	20:20	I.Tr.I.
	22:31	I.Tr.E.
	22:31	I.Sh.E.
June 6	17:28	I.Oc.D.
	19:40	I.Ec.R.
	23:53	II.Oc.D.
June 7	2:30	II.Ec.R.
	14:46	I.Tr.I.
	14:48	I.Sh.I.
	16:57	I.Tr.E.
	17:00	I.Sh.E.
June 8	11:54	I.Oc.D.
	14:09	I.Ec.R.
	18:02	II.Tr.I.
	18:09	II.Sh.I.
	20:33	II.Tr.E.
	20:43	II.Sh.E.
June 9	7:34	III.Tr.I.
	7:49	III.Sh.I.
	9:12	I.Tr.I.
	9:17	I.Sh.I.
	9:37	III.Tr.E.
	10:03	III.Sh.E.
	11:23	I.Tr.E.
	11:29	I.Sh.E.
June 10	6:20	I.Oc.D.
	8:37	I.Ec.R.
	13:01	II.Oc.D.
	15:49	II.Ec.R.
June 11	3:38	I.Tr.I.
	3:45	I.Sh.I.
	5:49	I.Tr.E.
	5:57	I.Sh.E.
June 12	0:46	I.Oc.D.
	3:06	I.Ec.R.
	7:08	II.Tr.I.
	7:26	II.Sh.I.
	9:40	II.Tr.E.
	10:00	II.Sh.E.
	21:15	III.Oc.D.
	22:04	I.Tr.I.
	22:14	I.Sh.I.
June 13	0:07	III.Ec.R.
	0:15	I.Tr.E.
	0:26	I.Sh.E.
	19:12	I.Oc.D.
	21:34	I.Ec.R.
June 14	2:08	II.Oc.D.
	5:06	II.Ec.R.
	16:30	I.Tr.I.
	16:43	I.Sh.I.
	18:41	I.Tr.E.
	18:55	I.Sh.E.
June 15	13:38	I.Oc.D.
	16:03	I.Ec.R.
	20:16	II.Tr.I.
	20:43	II.Sh.I.
	22:47	II.Tr.E.
	23:17	II.Sh.E.
June 16	10:51	III.Tr.I.
	10:56	I.Tr.I.
	11:11	I.Sh.I.
	11:47	III.Sh.I.
	12:55	III.Tr.E.
	13:07	I.Tr.E.
	13:23	I.Sh.E.
	14:03	III.Sh.E.
June 17	8:04	I.Oc.D.
	10:31	I.Ec.R.
	15:16	II.Oc.D.
	18:25	II.Ec.R.
June 18	5:22	I.Tr.I.
	5:40	I.Sh.I.
	7:33	I.Tr.E.
	7:52	I.Sh.E.
June 19	2:30	I.Oc.D.
	5:00	I.Ec.R.
	9:23	II.Tr.I.
	10:01	II.Sh.I.
	11:55	II.Tr.E.
	12:35	II.Sh.E.
	23:49	I.Tr.I.
June 20	0:09	I.Sh.I.
	0:32	III.Oc.D.
	2:00	I.Tr.E.
	2:20	I.Sh.E.
	4:06	III.Ec.R.
	20:56	I.Oc.D.
	23:29	I.Ec.R.
June 21	4:24	II.Oc.D.
	7:43	II.Ec.R.
	18:15	I.Tr.I.
	18:37	I.Sh.I.
	20:26	I.Tr.E.
	20:49	I.Sh.E.
June 22	15:22	I.Oc.D.
	17:57	I.Ec.R.
	22:30	II.Tr.I.
	23:18	II.Sh.I.
June 23	1:03	II.Tr.E.
	1:52	II.Sh.E.
	12:41	I.Tr.I.
	13:06	I.Sh.I.
	14:08	III.Tr.I.
	14:52	I.Tr.E.
	15:18	I.Sh.E.
	15:46	III.Sh.I.
	16:14	III.Tr.E.
	18:03	III.Sh.E.
June 24	9:48	I.Oc.D.
	12:26	I.Ec.R.
	17:32	II.Oc.D.
	21:01	II.Ec.R.
June 25	7:07	I.Tr.I.
	7:35	I.Sh.I.
	9:18	I.Tr.E.
	9:46	I.Sh.E.
June 26	4:14	I.Oc.D.
	6:54	I.Ec.R.
	11:38	II.Tr.I.
	12:36	II.Sh.I.
	14:11	II.Tr.E.
	15:10	II.Sh.E.
June 27	1:34	I.Tr.I.
	2:03	I.Sh.I.
	3:44	I.Tr.E.
	3:50	III.Oc.D.
	4:15	I.Sh.E.
	8:06	III.Ec.R.
	22:41	I.Oc.D.
June 28	1:23	I.Ec.R.
	6:41	II.Oc.D.
	10:19	II.Ec.R.
	20:00	I.Tr.I.
	20:32	I.Sh.I.
	22:11	I.Tr.E.
	22:44	I.Sh.E.
June 29	17:07	I.Oc.D.
	19:52	I.Ec.R.
June 30	0:47	II.Tr.I.
	1:53	II.Sh.I.
	3:19	II.Tr.E.
	4:28	II.Sh.E.
	14:26	I.Tr.I.
	15:01	I.Sh.I.
	16:37	I.Tr.E.
	17:12	I.Sh.E.
	17:28	III.Tr.I.
	19:35	III.Tr.E.
	19:45	III.Sh.I.
	22:02	III.Sh.E.

For telescopic observers in June, here is the complete list of phenomena involving Jupiter's four bright moons and the planet's disk or shadow. The first columns give the date and midpoint time of the event in Universal Time. Next is the satellite involved: **I** for Io, **II** Europa, **III** Ganymede, or **IV** Callisto. This is followed by the type of event: **Oc** for an occultation of the satellite behind Jupiter's limb, **Ec** for an eclipse by Jupiter's shadow, **Tr** for a transit of the satellite across the planet's face, or **Sh** for the satellite casting its tiny black shadow onto Jupiter. An occultation or eclipse begins when the satellite disappears **(D)** and ends when it reappears **(R)**. A transit or shadow passage begins at ingress **(I)** and ends at egress **(E)**. Each event is gradual, lasting several minutes. These predictions are courtesy IMCCE/Paris Observatory.

The wavy lines represent Jupiter's four big, "Galilean" satellites. Jupiter itself is the center vertical bar. Each horizontal band between the thin lines represents a full day, from 0^h (upper edge of band) to 24^h Universal Time. The UT date is given at left. Celestial east is left and west is right, as you see in a correct-view telescope when north is up.

52 June 2007 Sky & Telescope

Figure 3.36: Times of various events of Jupiter's 4 largest moons.

reappearing, transiting across, etc.) The "ticks" are quite useful too. In one example, Europa will appear at one edge of Jupiter's face. Over the next 2 hour and 35 minutes[S&T, Sept. 2007, p. 51] Europa will cross the planet's face, then disappear again.

One might think now, that if you are out at sea, you could observe Jupiter's moons and retrieve (from tables) at what time the particular event would be observed in Greenwich that day, and this would be Jupiter's role as a timepiece. But it doesn't work this way. Unlike all of the celestial objects we've discussed so far, the motion of Jupiter's moons are not bound by the earth's 15°/hour rotation; their motion is set by their gravitational interaction between themselves and Jupiter itself. Jupiter is about 800 million kilometers from the sun, almost 10 times the earth/sun distance. The view of Jupiter is the same everywhere on earth. That is, if everyone on earth looked up at the same instant, they would see Jupiter and its moons in exactly the same configuration well *almost.*

The view of Jupiter from points on Earth

Figure 3.37 shows a diagram of what is happening when two people, at positions x and y look at Jupiter from opposite sides of the earth[J. Keller and D. Doty, personal communication]. We chose opposite sides of the earth to illustrate the most extreme possible distances two observers on earth could possibly achieve, while both looking at Jupiter. A is the distance between the earth and Jupiter, B is the radius of the moon's orbit, and D is the diameter of the earth.

Suppose that we are trying to view the reappearance of the moon coming around the top of Jupiter from the right, as in the figure. To judge when the moon would reappear, we use the top edge of Jupiter as our reference point. The line from person x to the top edge of Jupiter is his or her line of sight, and the line from person y to the top edge is theirs.

Right off, notice something "interesting." Look at the distance labeled C. This is a small gap between the two viewer's line of sight along the moon's orbital path. Viewer x sees the portion of the moon's orbit where his or her line intersects the moon's orbital path and viewer y sees the orbit where their line crosses the orbit's path. Because of this gap, the two viewers cannot, and will not, observe the exact same view of Jupiter's moon if they look up at Jupiter at the same instant (from opposite sides of the earth). Thus Jupiter cannot precisely be used as a global celestial clock–two viewers will get different "readings" solely due to their positions on earth. Person x for example could be somewhere between Hawaii and Japan and y could be Greenwich, England.

The reason for this gap is called parallax and is a different view of an object due to the observer's position. (As an example, hold up your index finger with your arm outstretched and look at it with one eye closed and then the other–the view of your finger seems to shift due to the position of one of your eyes relative to the other.) Let's do some calculations using the figure and see what we can learn about the gap C, and find out just how bad this parallax problem really is. To analyze this parallax problem, we'll set up two similar triangles. Take a look at Figure 3.38 to see the triangles and keep Figure 3.37 nearby.

Triangle XYZ is similar to triangle efZ. This allows us to relate various angles and sides of the triangles. In particular the side length XY is to ef as A is to B. Mathematically, we

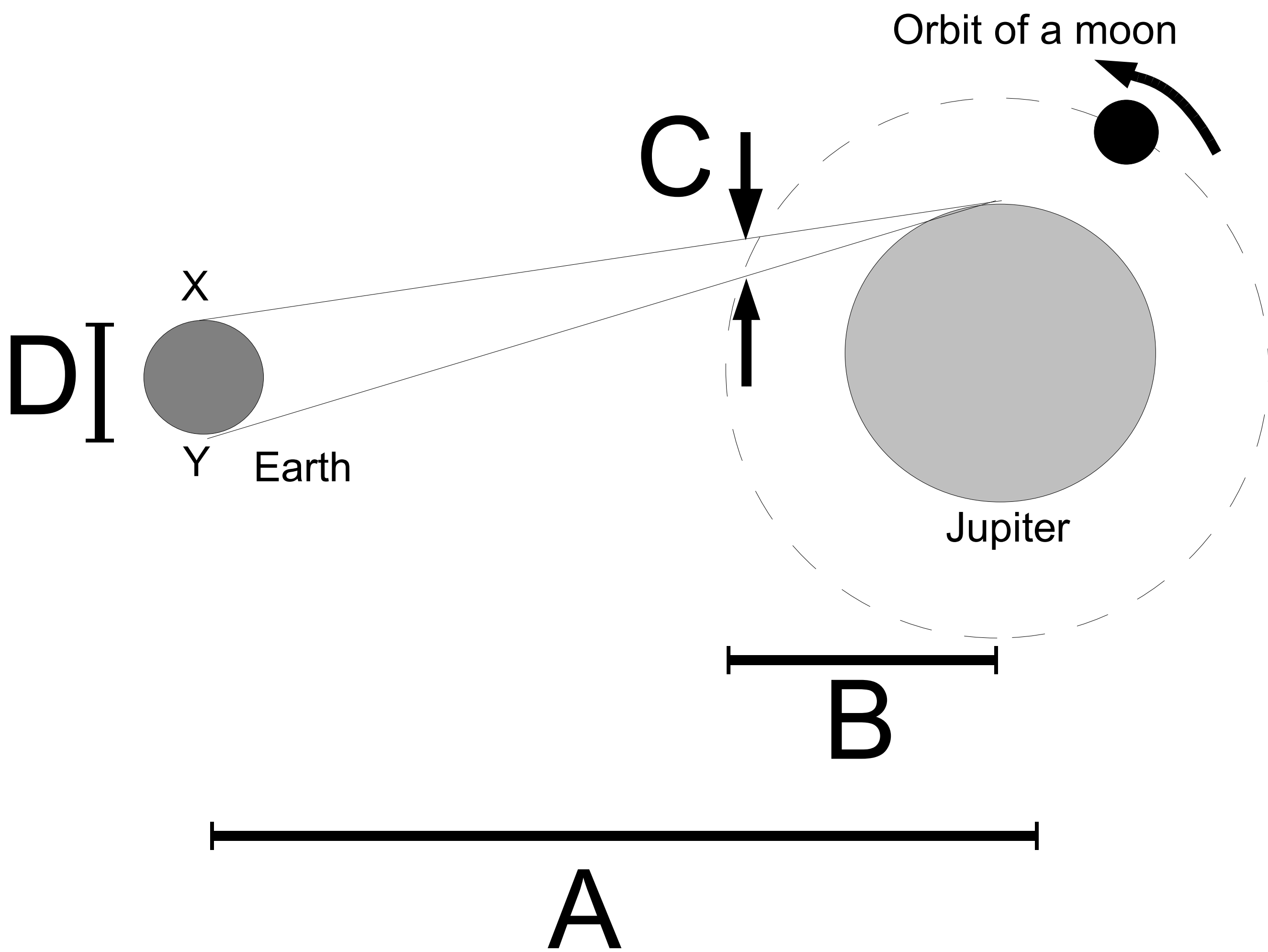


Figure 3.37: Viewing Jupiter from opposite sides of the earth.

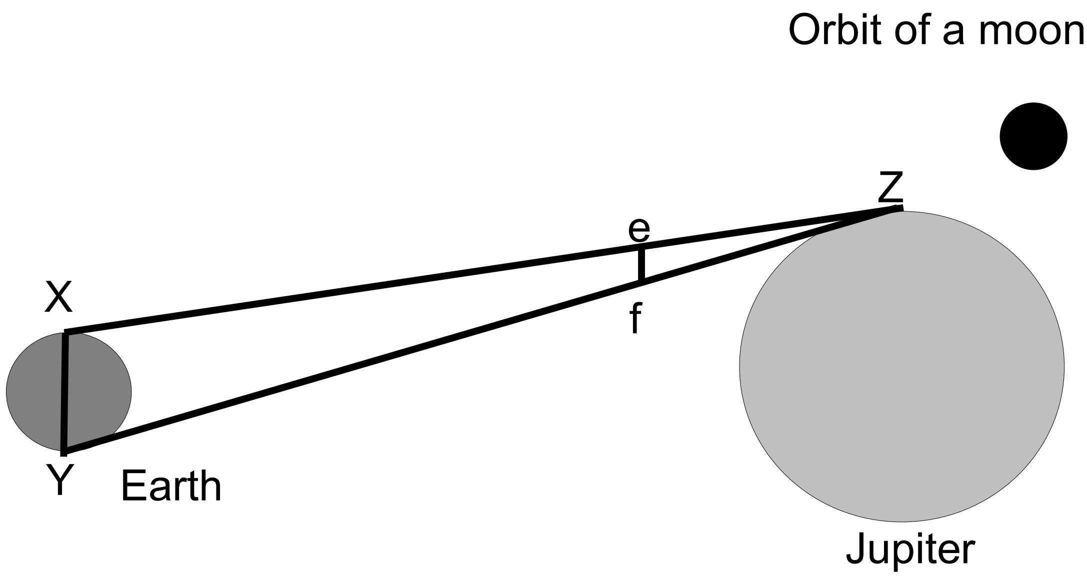


Figure 3.38: Triangles used to figure out the parallax problem when viewing Jupiter from the Earth.

can write this as

$$\frac{XY}{ef} = \frac{A}{B}. \tag{3.16}$$

But XY is D, the diameter of the earth, and ef is the gap distance, C, so this equation becomes

$$\frac{D}{C} = \frac{A}{B}, \tag{3.17}$$

or the gap distance is

$$C = \frac{DB}{A}. \tag{3.18}$$

So we have a measure of the gap, C, but let's remind ourselves what we are after. We want to know what time difference this gap will present to observers on the earth. This will involve how fast the moon is moving, or in other words, how quickly it passes through the gap. A very fast moving moon will zip through the gap, presenting the smallest time difference. A very slow moving moon will present a larger time gap, because it will take longer to go from viewer x's line of sight to viewer y's line of sight. Using the fact that speed is distance over time, we can write that

$$v = \frac{2\pi B}{T} \tag{3.19}$$

Moon	T (days)	Δt (Seconds)
Io	1.77	0.5
Eurpoa	3.55	1
Ganymede	7.16	2.01
Callisto	16.69	4.68

Figure 3.39: Time differences for two observers viewing Jupiter from opposite sides of the earth.

where v is the speed, $2\pi B$ is the circumference of the moon's orbit (total circular distance it travels), and T is the period of the moon, or the amount of time it takes the moon to go around Jupiter one time. But this speed could also be measured as the moon moves across the gap length, C. If it moves across the gap in a time Δt, then v would also be

$$v = \frac{C}{\Delta t}. \tag{3.20}$$

Pay attention now, because Δt is going to be our time gap. It will be the best timing possible using a given moon of Jupiter. Equating the two v's we get that

$$\frac{2\pi B}{T} = \frac{C}{\Delta t}. \tag{3.21}$$

Next, we know what C is from Equation 3.18, so we can eliminate it from this equation to get

$$\frac{2\pi B}{T} = \frac{DB}{A\Delta t}. \tag{3.22}$$

You can see that the B's will cancel across the top and we can solve for Δt to get

$$\Delta t = \frac{TD}{2\pi A}. \tag{3.23}$$

So we have it! We have Δt, the time gap that two viewers on earth must experience when observing Jupiter's moons because of parallax. In other words, the two earth-bound viewers will not be able to agree on the "time" they are reading from Jupiter and its moons to any better than Δt. For some numerical results, see Figure 3.39.

So as you can see observers on earth will see moon events on Jupiter happening in the range of seconds apart. As a clock then, this means a sailor at sea can't determine the time in Greenwich, for example, to any better than about 0.5 seconds. This is pretty good. Doing a bit more math with the earth's 15°/hour, we get

$$0.5 \text{ s} \times \frac{1 \text{ hr}}{3600 \text{ s}} \times \frac{15^\circ}{\text{hour}} = 0.002^\circ \tag{3.24}$$

an error in longitude of 0.002°, which would be 0.12 Nautical miles or about 768 feet (1° on a great circle of the earth is 60 nautical miles). So Jupiter would do quite well for keeping

time at sea, assuming your observations of the moons were precise enough so that Δt was your largest error and you could determine and keep your local time fairly well.

Lastly, notice that A, the distance from the earth to Jupiter is in *the bottom* of Equation 3.23. This means the closer the planet, the larger Δt will be, and this is the nature of parallax. The closer an object is, the worse parallax becomes. Since the moon, for example, is $1,500$ times closer to earth than Jupiter is, Δt would be $1,500$ times larger for some observation involving the moon, for the two observers on opposite sides of the earth. The "lunar method" for finding longitude, for example, involved finding the angular distance between the moon and an object on the celestial sphere. Hence while the two observers could observe a moon event for Jupiter, more or less simultaneously, they could not for the moon.

Failure as a Longitude Solution

So as fantastic as Jupiter could be for time keeping, the trouble was that is simply wasn't possible to observe it from the deck of a rolling ship. Such a beautifully simple solution to the longitude problem had to be abandoned for this logistical reason. Gimbal mounted "marine" chairs, and the like were all tried and failed, in attempts to stabilize the observation platform at sea. By the early 18th century (1700s), it was clear that the moons of Jupiter held little promise of finding longitude at sea[Andrews, p. 100]. The high magnifications needed of a telescope and the stability issues put an end to all hopes, more or less, by 1774.

Despite this, work continued on using Jupiter to find longitude at sea, up until the time John Harrison developed a suitable marine clock. Why? Why continue working on a longitude solution that would never work at sea? Because using Jupiter to determine longitude *on land* was so spectacularly successful[Andrews, p. 86]. As seen above, you can (in principle) keep time to 0.5 s if your observations of Jupiter's moons were careful enough. On land, you have the solid and motionless ground on which to place your telescope. After all, the problem of longitude at sea takes on meaning only when one considers it in the context of longitude on land. In other words, sailors needed good maps! For what good would it be for a sailor to know his position at sea, as he tried for California, but does not know say, the location of San Diego habor? So Jupiter did not present a method for finding longitude at sea; it just wasn't practical. But it did serve as the center for a revolution in map-making over the 17th, 18th, and 19th centuries[Andrews, p. 100].

3.4 Celestial Navigation: a case study

This section comes at a somewhat desperate moment. We are frustrated and exasperated. We've discussed Jupiter, the Sun, and the moon, and it seems like everything in between. We have examined all of nature's obvious heavenly sign-posts that could potentially be used to navigate at sea. For the last time: *is it possible to use celestial objects to find where we are at sea?* Is it? The answer is **YES**, but **you need to be able to keep accurate time at sea. With a precise, sea-worthy clock, you can figure out where you are. You cannot do so without one.**

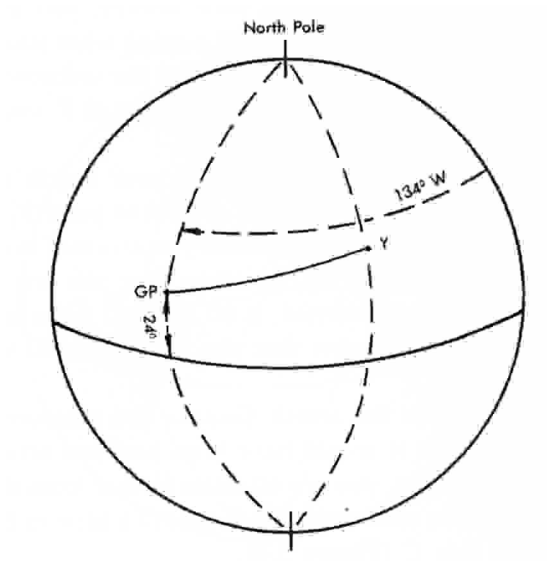


Figure 3.40: You ("Y") versus the GP of the sun.

3.4.1 Let's find ourselves then

So you are at sea, with nothing but water around, and you want to know where you are. You understand the basic premise of "circles of positions" from Section 3.3.1, but there are no tall buildings out at sea. The sun, however, is nice and bright to the West[Schlereth, chapter 3]. You decide to use the Sun as your "tall building." Here's what you would do.

First, get out your sextant and measure the angle of the sun in the sky. At the moment you take your measurement, yell "mark." On this cue, someone you are sailing with writes down the exact time in Greenwich, say 5 pm. You then read your sextant and see that at the time you yelled "mark," the sun was 31° above the horizon. This means its zenith distance was 59°. Refer again to Figure 3.17. Already you are not "so" lost anymore; at least you know that you are (using $1° = 60$ nautical miles) 59° or $59 \times 60 = 3540$ miles from the GP of the sun. OK, this doesn't help much, but we'll do better.

Next, you get out the Nautical Almanac, and look up the GP of the sun at 5 pm on the current date. You read that the sun is over a spot on the earth at 24° North latitude, and 134° west. This is the sun's GP at 5 pm, for your current date. Remember, you are desperately lost at sea. Why not get a sheet of paper and draw a circle on it, representing the earth? Mumble to yourself "Where am I?" You just looked up the Sun's GP, so put that on your sketch of the earth, perhaps like that shown in Figure 3.40. The leftmost dotted line (of longitude) contains the GP of the sun, as labeled "GP" in the Figure. Relative to the equator, the GP is up north 24°, and west 134°. So much for sun, what about you in your boat?

Well, at least you know you are on earth (you're not *that* lost), so you must be on a line of longitude *somewhere*. But think now, you know the sun is to your west, so draw your line of longitude to the right (or east) of the one you just drew for the sun. So we draw it in, like

the right dotted line of longitude shown in Figure 3.40. Now think a bit more. Are you at least in the same hemisphere as the sun's GP? Use common sense, or check your best guess using your DR records. Let's say you are, so draw a dot on your line of longitude and label it "Y" for "you." Great! Things our sketch is taking shape a little. We see the sun's GP and a guess of where we are.

Next, think about what your sextant reading means; the 59° zenith distance. You can draw a line between the "GP" and "Y" points and label it 59°. And as long as you are labeling lines on your sketch, if the sun's GP is 24° from the equator, it must be 66° down from the north pole. Label this too. Your sketch now might look like that in Figure 3.41. As you look at this sketch, note the beginnings of a triangle forming on the surface of the earth. The triangle is set by the north pole, the sun's GP, and your point "Y." This triangle is called the "eternal triangle" of celestial navigation, because its geometry is what allows you to ultimately find where you are.

At first glance of Figure 3.41, you might think your location is close at hand. You know how far you are from the GP of the sun, you've formed a triangle on the earth, and you have a couple of sides of the triangle labeled. Just like you know the latitude and longitude of the "GP" point, you'd like to know the latitude and longitude of point "Y" now, since this is your position. Does this come from the budding triangle? Yes it does, but not quite yet. You're out of information for now. Despite the great progress on your sketch, you simply don't have enough to go on. To prove this, imagine the triangle in Figure 3.41 is a good old triangle on the chalkboard of the last math class you took in high school. All you know are two sides of a triangle. Remember that Trigonometry class from high school? Whether you try to use the Pythagorean Theorem, Law of Sines, or Law of Cosines, it is mathematically impossible to compute any other sides or angles of your triangle at this point.

You keep thinking though. Can't the angle between the 66° line and 59° line be found using a compass? Isn't the angle of the 59° line, as it leads from "Y" to "GP" the direction of which you observe the sun? Yes, it is, but you would never be able to specify an observation direction with the needed precision. Also, a compass works by the little needle being pushed and pulled by the earth's magnetic field, which is known to vary and deviate from "true north" all over the world. A compass gives you your general direction; nothing precise or absolute enough for navigational computations.

Well, what next then? You need more numbers on your triangle. Where will they come from? The only thing you can do at this point is, believe it or not, *to guess* your current position. This is the only way of introducing more information onto the eternal triangle, so let's do it.

Guessing your position might not be as bad as you think. For one, we've already stated if we're north or south of the equator. That eliminates 50% of the globe right off. What about the prime meridian? You should be able to tell what ocean you're in and if you are east or west of it. Even with anything resembling a map of the world and some reasonably careful DR, you should be able to make at least an educated guess as to our location. Granted DR is never known to be exact, but we are not looking for exactness right now. We are looking for something with which to label the spot "Y." Suppose then we check our DR records and it puts us at about 67° West longitude and 37° North latitude. Our eternal triangle now

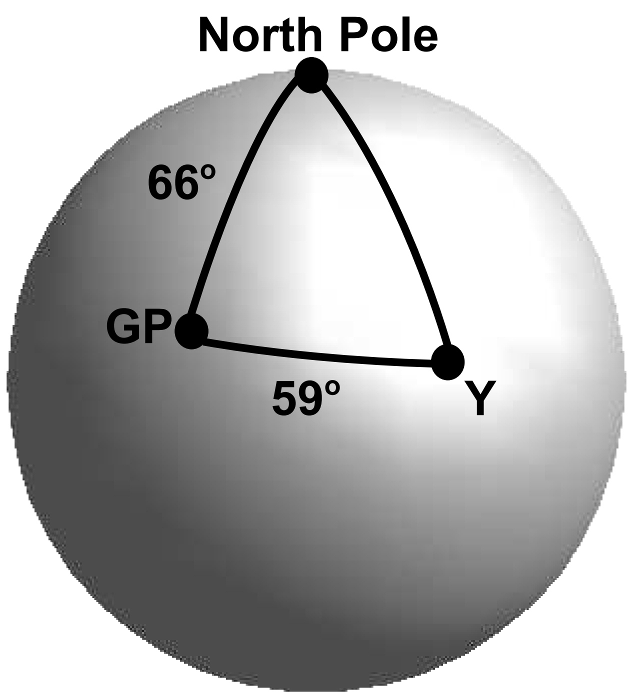


Figure 3.41: What you know about your location and the GP of the sun thus far.

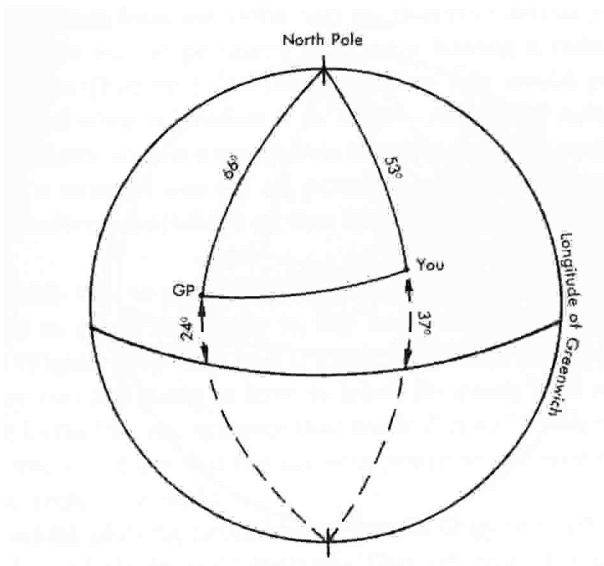


Figure 3.42: The eternal triangle, with the sun's GP and our guess position.

looks like that shown in Figure 3.42.

In practice, your exact DR position isn't even typically the guess you would put on your sketch at this point. Since you are guessing anyway, you'd "refine" your guess and choose numbers that would make your calculations simpler later on. No one, after all, wants to deal with messy decimals out at sea. So you'd choose, for example, 30° instead of 30.55°. All told, your guessed position on the map is referred to as your "assumed position" or AP for short. In the end, we'll see when everything is done is that your DR or AP position won't really matter. The goal of celestial navigation is to *correct your guess*, which will reveal your actual position, and that's how we'll find where you are; by correcting your initial guess.

So with our AP in hand, our eternal triangle now looks like that shown in Figure 3.43. It shows how the apex angle of the triangle is now thought to be 67°. (Again, no one is claiming our AP is correct, but the triangle is at least becoming something we can work on.)

With the apex angle and two sides of our triangle known, we can do a little bit of mathematics to compute the distance between ourselves and the GP of the sun based on our guess. It works like the triangle shown in Figure 3.44

The triangle we are dealing with is not strictly flat like the ones your teacher drew on the flat chalkboard in high school. They are triangles that exist on the surface of a sphere (the earth). This means that the lengths of the sides of the triangle can be kept as angles, measured in degrees. The three key relationships that relate the sides and angles of such "spherical triangles" are the following.

$$\cos(a) = \cos(b)\cos(c) + \sin(b)\sin(c)\cos(A) \tag{3.25}$$

$$\cos(b) = \cos(a)\cos(c) + \sin(a)\sin(c)\cos(B) \tag{3.26}$$

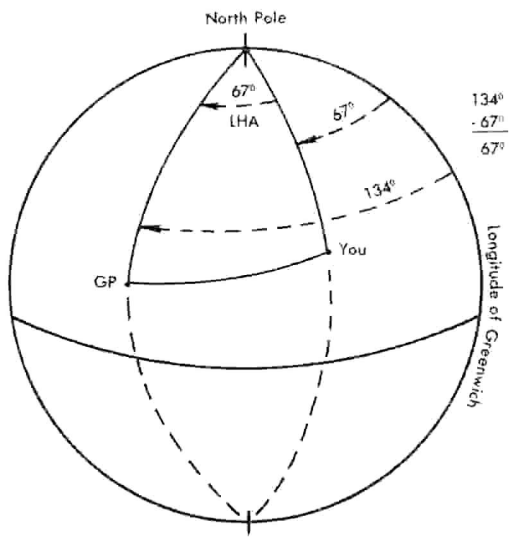


Figure 3.43: The eternal triangle, with the sun's GP and our guess position.

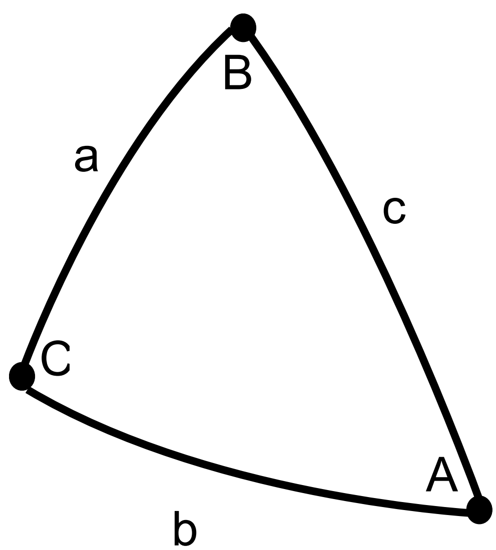


Figure 3.44: A generic spherical triangle.

$$\cos(c) = \cos(b)\cos(a) + \sin(b)\sin(a)\cos(C) \tag{3.27}$$

Note in these equations that lowercase letters refer to sides of the eternal triangle, and uppercase letters refer to angles that exists where two sides meet. With what we know from Figure 3.43 we can use Equation 3.26 to find the length of the side between the sun's GP and our DR position, by plugging in the variables we know and using our pocket calculator (that can do sin and cos), like this:

$$\cos(b) = \cos(66)\cos(53) + \sin(66)\sin(53)\cos(67) \tag{3.28}$$

or

$$\cos(b) = (0.407)(0.602) + (0.931)(0.799)(0.391) \tag{3.29}$$

or, finally,

$$\cos(b) = 0.536 \tag{3.30}$$

If you did this right, you should have the number 0.536 on your calculator display. Hit the inverse cosine button (INV-COS, or $\cos^{-1}$) to get 57.5°, or 58°. So, based on our AP, we are 58° from the GP of the sun, or about 3,480 nautical miles.

BUT WAIT! When we took our sextant measurement, we observed the sun to be at a zenith angle of 59°. We observed 59° and computed 58° based on our guessed AP. What gives! Well, think carefully about what these numbers mean. We know our AP isn't right; it came from your DR records and was rounded since you hate decimals. The only real truth about our position at this point is that the distance between you and the sun's GP is 59°. You know this because you just measured it with your sextant at 5 pm. This is real, but we were left with an unsolvable triangle, so we had to throw our AP into the mix and from this got that we are 58° from the sun's GP.

This means our AP is 1° in error, or 60 nautical miles off. Not bad for an AP (I guess!). In actuality then, we are farther from the sun's GP that our AP says. What we have thus far is shown in Figure 3.45, where we've labeled point "R" as our **R**eal position from the sun (this is where we really are, based on our sextant measurement). Remember that the "Y" point is still our guess that we put on our triangle at the very beginning. It is not right and we know it. We don't even know at this point if "R" and "Y" are on the same line leading to the GP of the sun, but we're working on it!

Before continuing, a word of discussion about the guessed position and all. Including the guess was really a practical matter. Why? Because if you're on a ship, you'd have some kind of map in front of you, onto which you could plot the points you are finding. The sun's GP, your AP, your real position "R," etc. But as you've seen, your distance from the sun's GP is in excess of 3,000 miles. Just plotting the two points "GP" and "R" would require a very large map. The normal working scale on small boats is about 1 nautical mile to 1 millimeter. A 3,600 GP sight would take a map 12 feet long![Schlereth (yellow), p. 23]. But you don't need to see the whole world at once. You'd rather work on a map that contains only the ocean in your immediate locality. For example, you don't want to have a map of

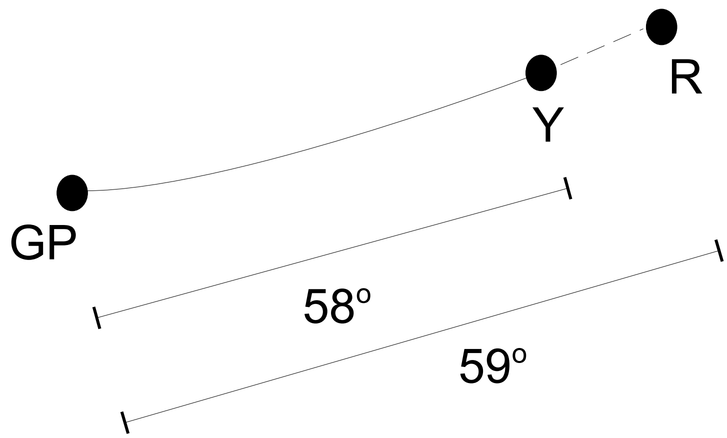


Figure 3.45: Our celestial navigation efforts thus far.

the entire Pacific Ocean in front of you, if your careful DR calculations put you just south of the Hawaiian Islands. This is not practical! Instead you get a map showing a convenient section of ocean surrounding your GP. You don't even need to plot the GP on your map, since our focus now will be simply to correct our AP relative to "R."

The initial zenith angle measurement of the sun, the starting point of all this, is a "whole big map" measurement. In this case, you concluded that you were on a circle 59° in radius, centered at the sun's GP. But now you know that you don't need to consider the full circle, because for one, you have your DR position, and second you know that the sun was to your west when you took your sextant reading. So you don't need portions of the circle to the north, east, and south of the sun's GP (because you aren't there!). Your next diagram might look like the one shown in Figure 3.46.

We can work this figure still further. Using the formulas above, we can find the angle between the meridian that contains our DR position and line connecting the GP and our DR position. Using Figure 3.44, we want to find angle A, which we can find using Equation 3.25. In particular

$$\cos(66) = \cos(58)\cos(53) + \sin(58)\sin(53)\cos(A), \tag{3.31}$$

which becomes

$$0.407 = (0.530)(0.601) + (0.848)(0.799)\cos(A) \tag{3.32}$$

or

$$0.088 = (0.678)\cos(A) \tag{3.33}$$

or

$$\cos(A) = 0.130. \tag{3.34}$$

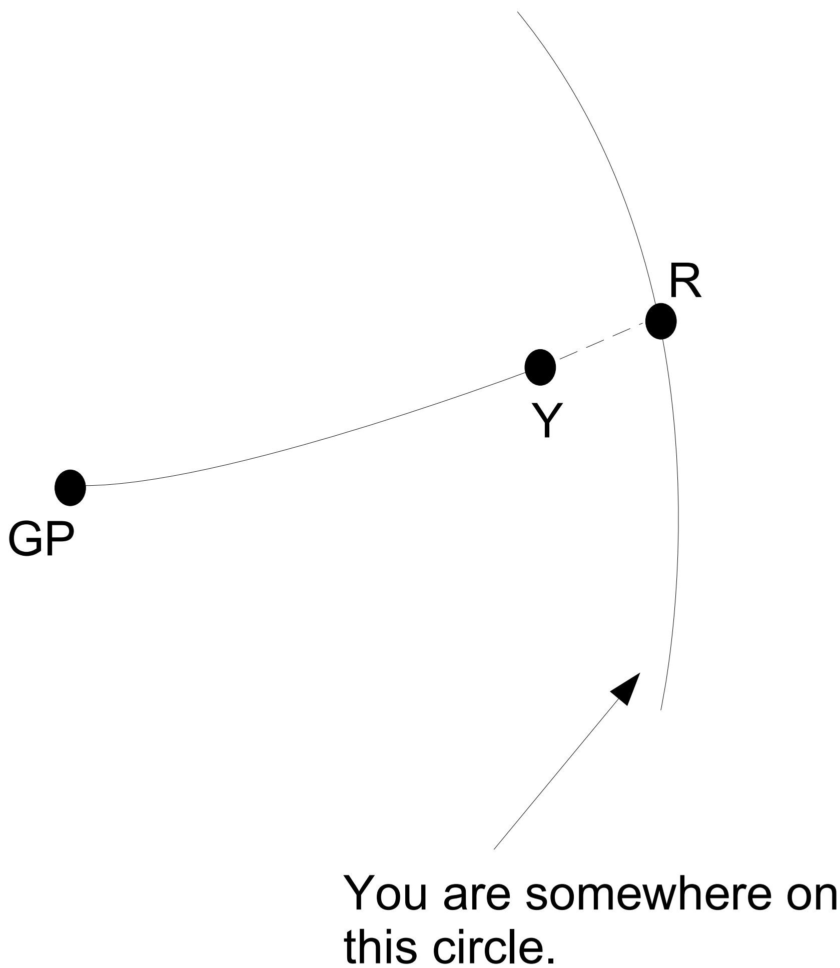


Figure 3.46: In this Figure GP is the geographic position of a celestial object. “Y” is our assumed position based on our best guess. “R” is our actual position from the GP based on our last sextant reading. Placing “R” and “Y” on the same line connecting to the GP is a guess as well at this point, but it at least gives us something to work with. The only thing we’re sure about in this figure is the distance between “GP” and “R.”

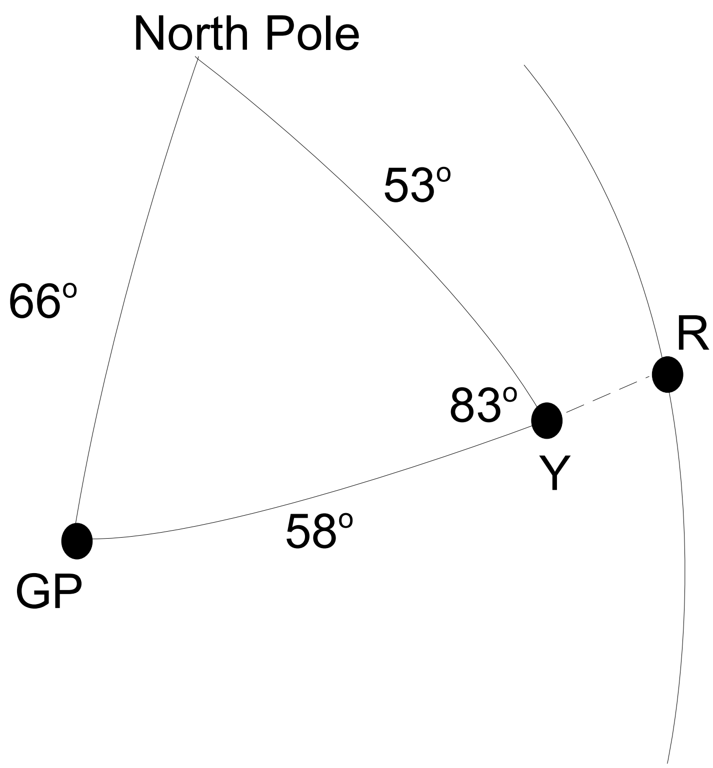


Figure 3.47: Still trying to find our position.

Taking the inverse cosine using our calculator gives angle $A = 83°$. So our triangle now is that in Figure 3.47.

Although still based on "Y," a point known not to be correct, this angle helps us in the following way. Let's draw a cross (coordinate axes) centered right on point "Y." In other words, let's forget looking at the whole, round earth. Let's zoom in, very close, and look at the "flat" patch of ocean on which we are floating. The vertical line will be the 67° longitude and the horizontal line will be our 37° latitude of AP.

Look carefully at this Figure 3.48 as it really brings together everything we've done thus far. It doesn't contain the sun's GP, 3000+ miles away (and it doesn't need to). It is a local zoom onto the patch of ocean we are on. On such a scale, our latitude and longitude lines are nice and straight. We have our "best guess" or point "Y" as our center of reference, and computed the angle of the line that leads from "Y" point to the sun's GP to be 83° (with respect to our line of longitude).

But we know point "Y" isn't correct, because the sextant measurement put our distance to the sun's GP at 59°, not 58°. So what do we do on our zoomed plot? Extend the line going to the sun's GP to the east another 1° ($59° - 58°$) as we did in Figure 3.46, and then draw a section of the circle with a 3000+ mile radius, as shown in Figure 3.49.

Now, let's use our actual position a little bit. Remember that the line connecting our actual position to the sun's GP is the radius of a big circle that we must be on (whose radius

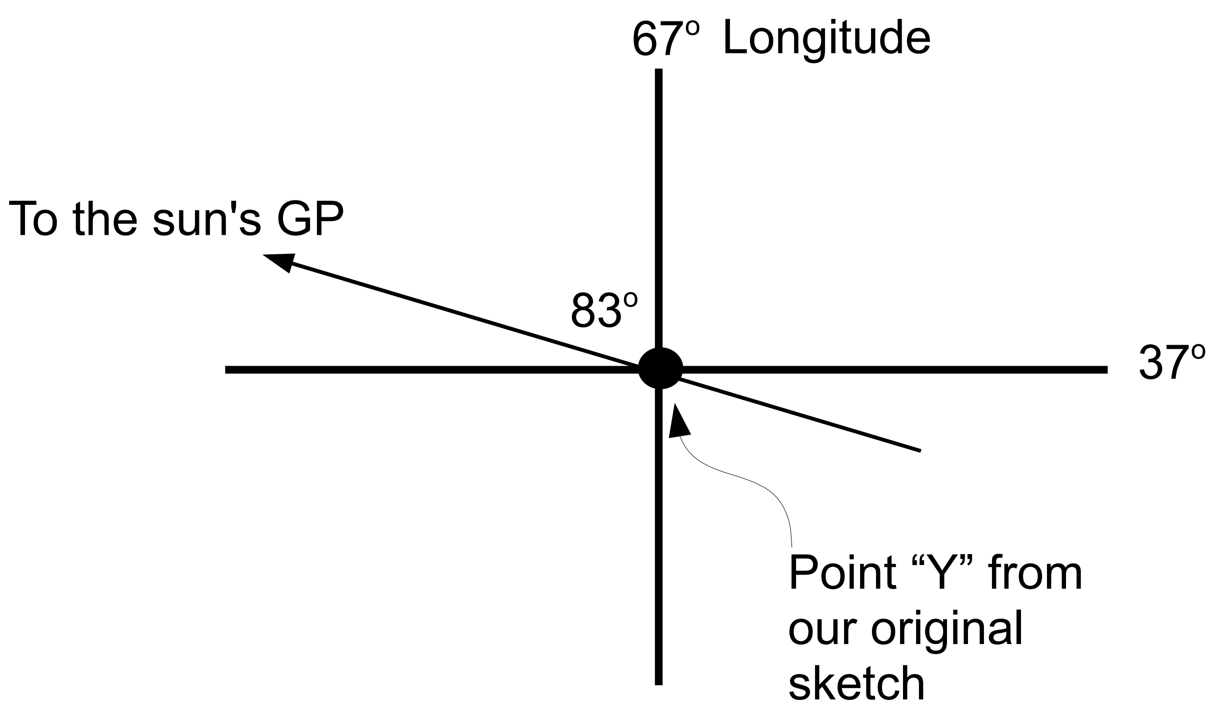


Figure 3.48: The point “Y,” our assumed position, zoomed in to just cover the portion of the sea near our assumed position. We also see that line leading to the sun’s GP.

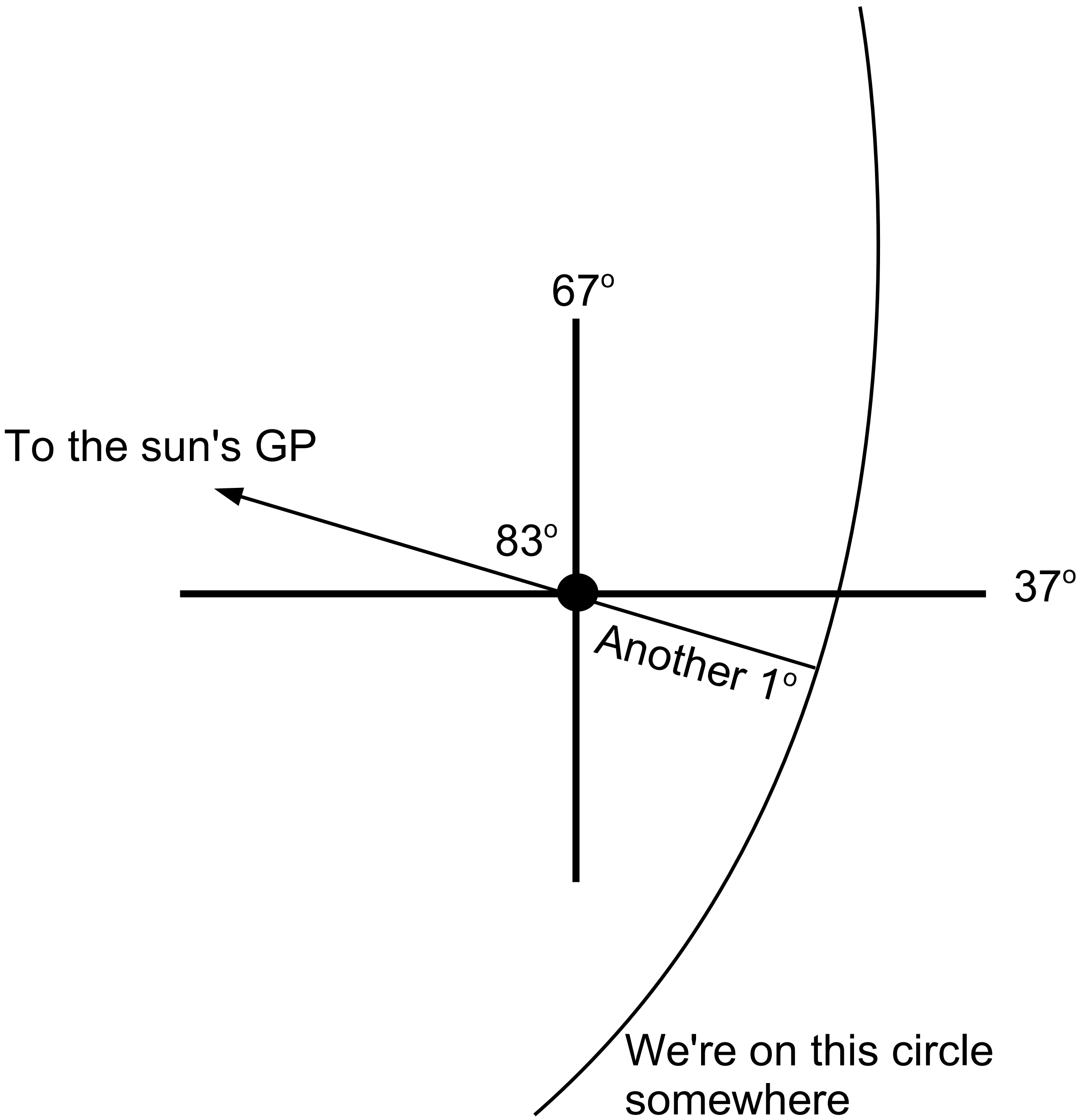


Figure 3.49: Line to the Sun's GP extended by 1° (the amount that we're off by). We are somewhere on the circle.

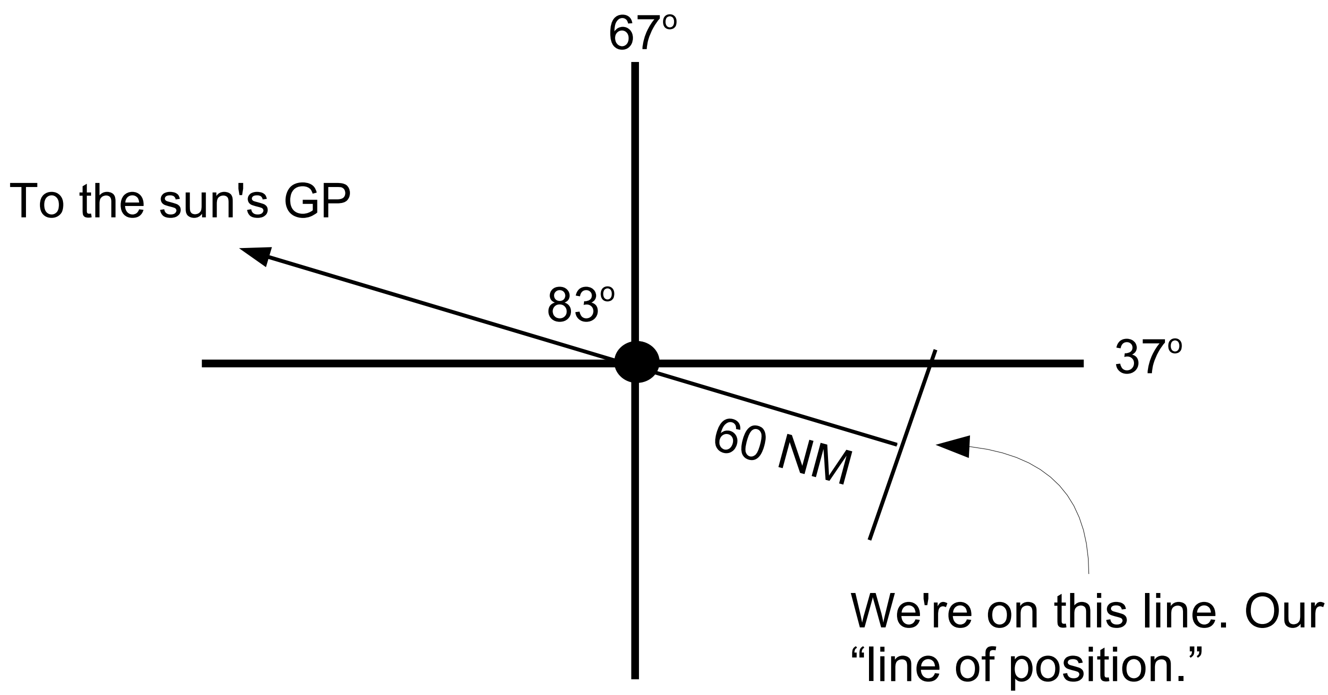


Figure 3.50: We're somewhere on the line, called our "line of position."

is the zenith angle we measured with our sextant). But we're trying to focus on the small patch of ocean that contains our boat. So forget the big circle; look at where the circle joins the line connecting "Y" and the sun's GP. Because the circle is so big, a small section that passes through our position can be approximated by a straight line as shown in Figure 3.50. This line is called our "line of position" or LOP for short. It is a tangent line to the huge circle with the sun's GP as its center. Notice also in this figure that we've replaced the 1° extension of the line leading to the sun's GP with 60 nautical miles (NM).

Believe it or not, we're done for now. There's nothing else we can extract from this analysis of our position. Although this still isn't the final answer, we're not totally lost anymore are we? We've narrowed our position down to being somewhere on a line, and we know we're on this line for two reasons:

1. Our zenith measurement of the sun (the 59°) is greater than the 58° we computed using the formulas and our AP (that originated from our DR records).

2. We know the sun is generally to the west.

Because of the above two criteria, we draw our LOP farther to the east of our AP by 1°, or the 60 NM. We don't know how long to draw the LOP in either direction, but we can constrain it some because the sun is to the west.

To find out exactly where we are on this line, we can do one of two things. First, we could repeat the entire procedure above, with another GP from another celestial object. Second,

since it is daytime and the Sun is the only object available to us, we can wait. Maybe for an hour or so. Let the sun move in the sky and let our boat drift a little bit. And over this hour, we'll keep *very careful* records of our speed and direction of motion. An hour later, the sun will have a new GP that we can look up, and we'll be able to advance our LOP relative to the current one.

So we are left to relying on the "unreliable" DR? Yes. But remember that DR is a measurement that accumulates error over many measurements. Imagine how lousy DR can be if you have to perform it during a storm, or just after a storm? But this isn't the case now. With careful work, there is no reason why short term DR, over an hour or so, cannot be a relatively accurate way of tracking our position.

Say over the hour of waiting, we do our DR bookkeeping and find our new position, relative to the old AP, as shown in in Figure 3.51. We have a new line connecting our new AP to the sun's new GP, and a sextant reading says we are actually 10 NM further east than our DR calculations indicate (so much for being careful over that hour!)

So now it's an hour later, and you still only have your position down to being on a new line. Hmmm..what next? How about see where the first LOP and second LOP cross? This is shown in Figure 3.52.

So...

...CONGRATULATIONS! You are no longer lost!

To check everything, you can wait some more, while paying careful attention to DR, and find a third AP. You can then construct a new line connecting your third AP to the sun's new GP, then derive a third LOP. Extending the third LOP into the direction where the first two intersect, you'll likely form a triangle with the other two LOPs, telling you that you are somewhere within the triangle. Not bad; over the vastness of the ocean, you have your position "fixed" to within a small triangle of area. With practice and a little luck, you could determine your position to a matter of a few square miles.

Finding your fix in this way gives you a truer and truer estimate of your actual position, in an iterative fashion. After finding the triangle formed by your three LOPs, you can assume that you are probably right at the center of the triangle. Use this point as your new AP, and start the procedure above all over again. The size of your ultimate triangle will likely grow smaller and smaller.

We're about to wrap up celestial navigation with a two more examples, but first something important to notice. We just went through many, many steps on the way to finding our location using the stars. So you might be thinking: why didn't early navigators just use this procedure to find where they were located? To answer this, remember what finding the GP of the sun was a critical component to all of this, and in order to do so, we had to look up the sun's GP in the Nautical Almanac, which requires that we know the time in Greenwich. It still all boils down to keeping good time while out at sea, and pre-1750's navigators could not do this.

Activity

Starting with a blank map, repeat the above exercise in celestial navigation.

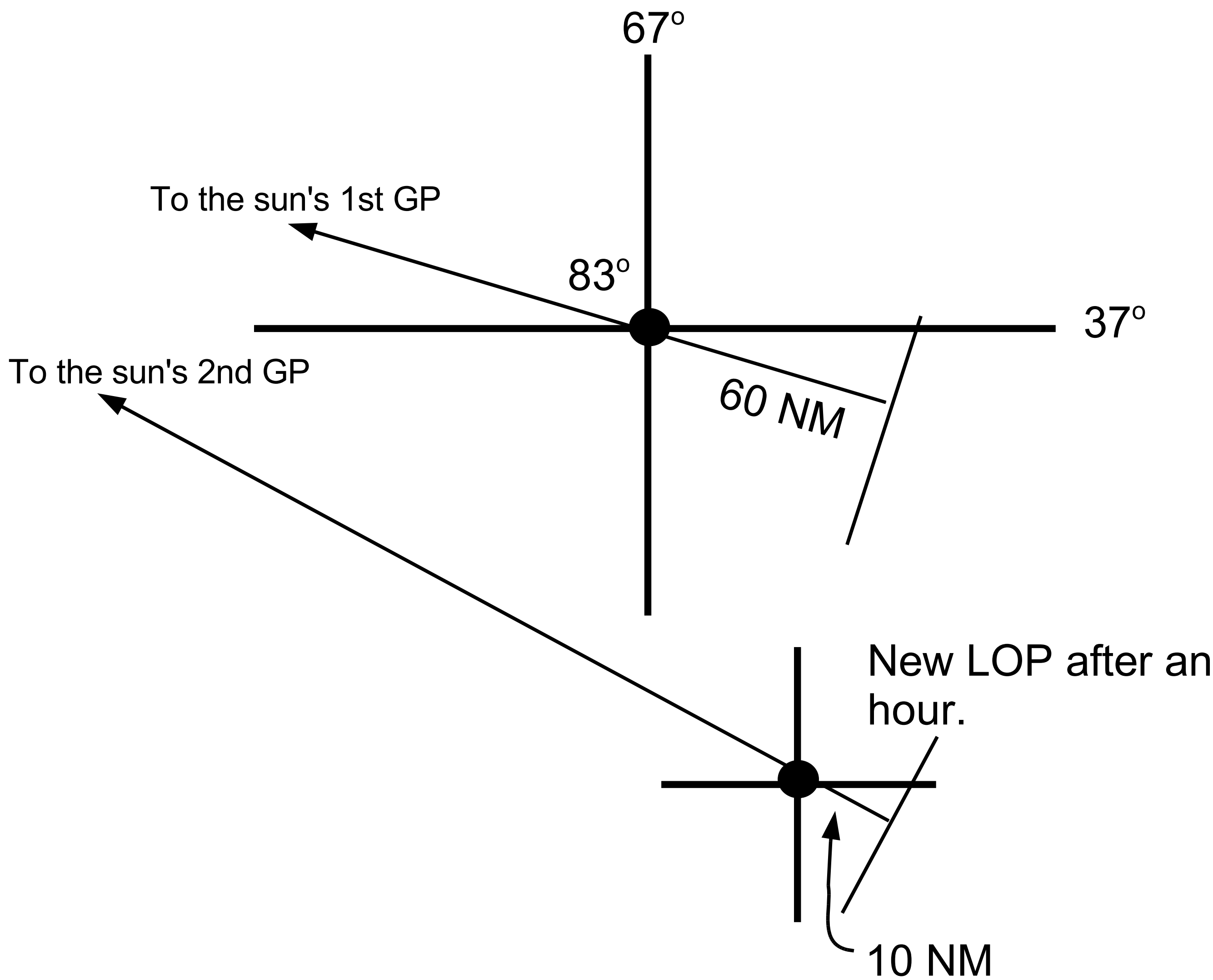


Figure 3.51: Our new position after waiting an hour. Note a new line connecting our position to the sun's new GP (an hour later).

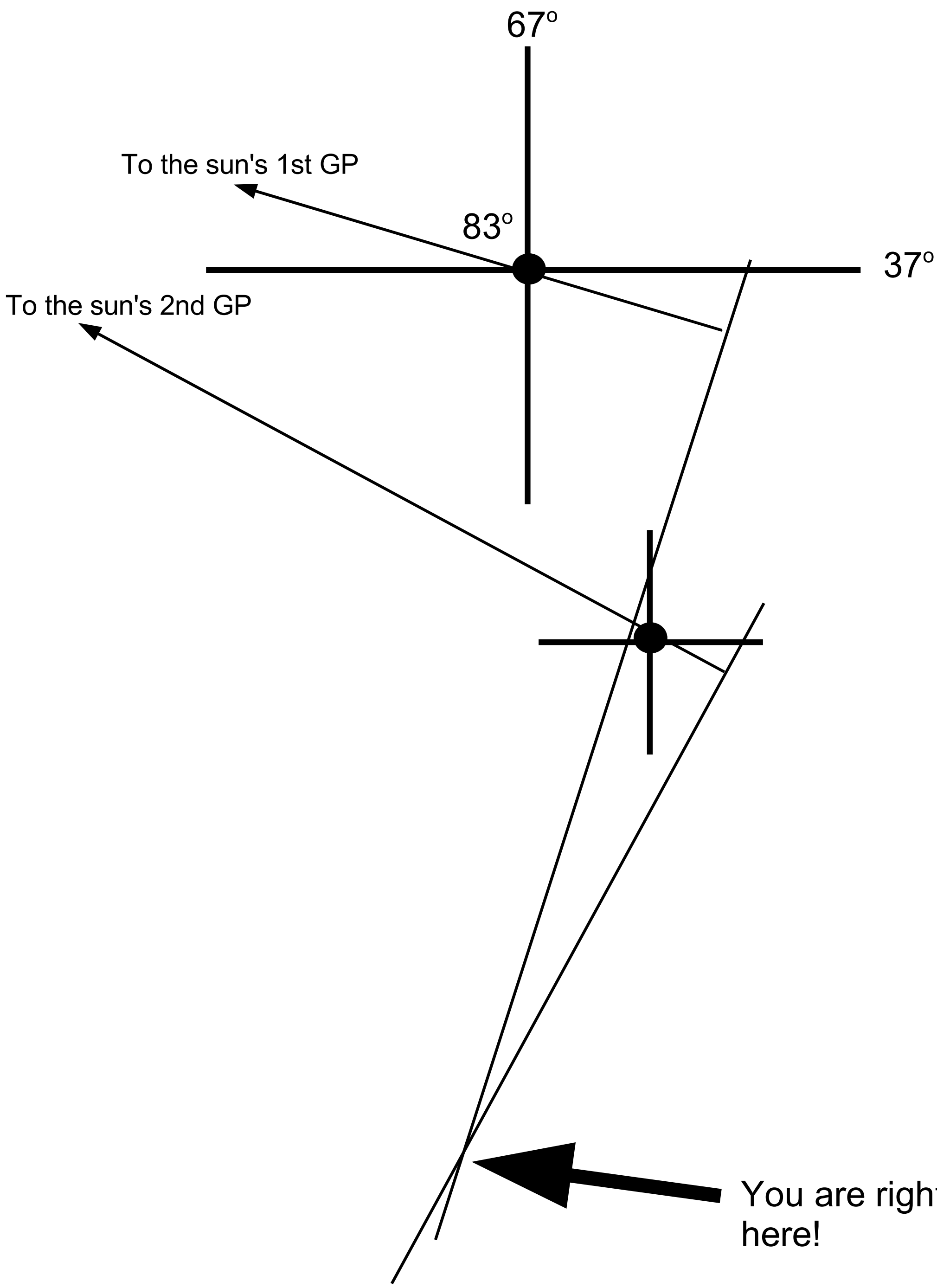


Figure 3.52: Our position is where the old and new LOPs cross. You're not lost anymore.

3.4.2 What did we just do?

The last section wasn't a theory of celestial navigation, it was *a technique* developed to handle the large circles of position derived from shooting the altitude of celestial objects. What exactly did we do, with the AP and intersecting lines and all? Let's take a final look at celestial navigation to tie it all together.

Take a look at Figure 3.53. This figure shows the full circle of position, as derived from a sextant measurement of the sun from your position at sea. The problem is that the sextant just gives you the radius of the circle of position that you are on, and nothing else. To find even your approximate position, you'd need to now narrow down where you might be (your angle) on the circle of position. This would be the angle between your position and the GP of the sun, which would reveal your location on the circle exactly, but you cannot find this angle from a single shot of the sun. Common sense allows you to at least narrow down this angle, since the sun was to the North-West when you shot it, meaning that you must be South-East of the GP. Further refinement of your location with a compass is only marginally possible, because a compass is clumsy and the earth's magnetic field deviates from true north, etc. So the best you can do with a single shot of the sun is to draw a tangent line to the circle, near your AP, and call it your LOP. But look at Figure 3.53. Suppose the radius of the circle is $3,000$ miles. The LOP you just drew is easily about $2/3$ of the radius, or $2,000$ miles long. So you have your position to $2,000$ miles. No good!

As mentioned above, if you had another celestial object to shoot, you would have two circles of position, and the two circles can intersect in two points, as discussed in Section 3.3.1. But this approach has two major problems. First, during the day, the only celestial object you can count on is the sun, so you must wait (an hour or so) and shoot the sun again. You hope your position doesn't change *very much* during your wait, or if it does, you track it carefully using DR. This would give you a second circle of position. Second, remember that circles of position of very large, and you won't be able to plot one, let alone two, on a single map. So what did we do? We recognized two facts:

1. During a short wait, the GP of the sun will not move very much, so the two circles of position generated will not be displaced very much.

2. Your AP indicates which of the two intersection points between the two circles is most likely to be your actual position.

Based on these facts, your two shots of the sun yields a picture like that shown in Figure 3.54. You see two closely spaced GP's (GP1 and GP2) and the arrows indicating the two intersection points of the two circles. You are located at one of these two points. Lastly, your AP, based on your DR records is indicated. It does not overlap with either of the two intersection points, but we didn't expect it to; your AP is just an approximation for your position. All it tells you at this point is that you are most likely at the lower right intersection point of the two circles.

So it appears as if your AP is "close" to your actual position, but it really isn't. If the radius of a circle of position is $3,000$ miles in Figure 3.54, then the distance between your

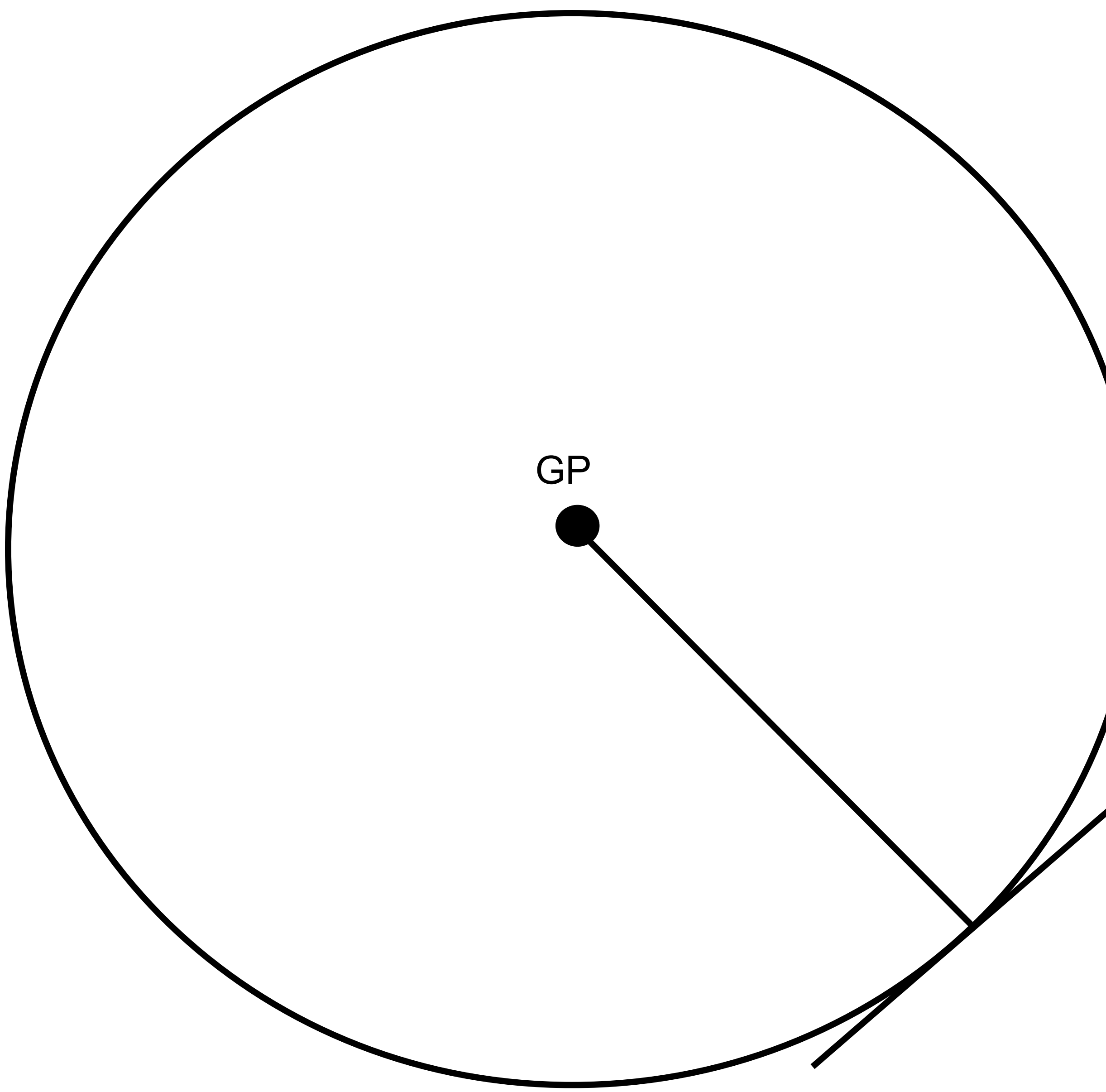


Figure 3.53: A GP from a sun sextant reading, and an attempted LOP.

AP and the lower right intersection point is about 750 miles. So we acknowledge then that we really need to focus our attention to the region near and around our AP. But before we do this, let's look at how our AP gives us an approximate angle between our position and the GP of the sun.

In Figure 3.55, we've drawn in two lines coming from GP1. One goes through our AP and onto the circle of position, and the other goes right to the actual intersection point between the two circles of position. In the realm of the vastness of the sea, the two lines are not that different in their directions. The drawing software used to generate the figure shows the lines to be different by about 10°. The two lines drawn for GP2 also differ by about 10°, as shown in Figure 3.56. If you figure that there are 360° in a circle and we have our location's angle to within 10°, we've eliminated over 97% of the circle from containing our position. So despite the reputation for DR to be so error-prone, it is quite useful in this regard! This is why the angle between our AP and GP (the 83° above) was such a reasonable starting point.

Next, let's zoom in 300x to the region near our AP. It would look like the view in Figure 3.57. Notice that the GP's are not on the map, but the lines leading to them are. This was the approach shown in Figure 3.48, and remember this line is the bottom line on the eternal triangle.

Finally, we draw tangents to both of the lines leading to the GPs, right where they touch their corresponding circle of position. This is shown in Figure 3.58, and look how close the intersection between the two tangent lines comes to the actual intersection between the two circles of position! As stated before, the AP and intersection point between the two circles was off by about 750 miles on the scale of the figure. The distance between the intersection point of the two circles of position and the intersection between the two tangents is now down to about 60 miles! In the vast ocean, we took our shots of the sun and now know our position to within 60 miles! This is awesome! Further shots of the sun, using this newly found position as a starting point will bring our calculated position even closer to where we actually are.

So this technique allows us to approximate the intersection point between the two circles of position by looking only at the region of the ocean near the vicinity of our AP. The tangent lines are approximations to the shapes of the huge circles of position, on the small sections of their arcs that fall within our tiny region of interest. Tangent lines are straight, so they're easy to draw and always have a direction that is perpendicular to the lines leading to the GPs of the sun. These lines incidentally are also the radii of the circles. Note also this technique only works if your original AP is at least near your actual position, as predicted by your sun shots. It is called the "intercept method," and was originally published by St. Hilaire in 1875.

3.5 How to find latitude

Sprinkled throughout the literature on navigation for the past 500 years is the idea that the pre-clock navigators could find their latitude at sea by observing the sun. Every sailor knew how to find latitude, and recall from Chapter 1, that a whole philosophy of sailing was

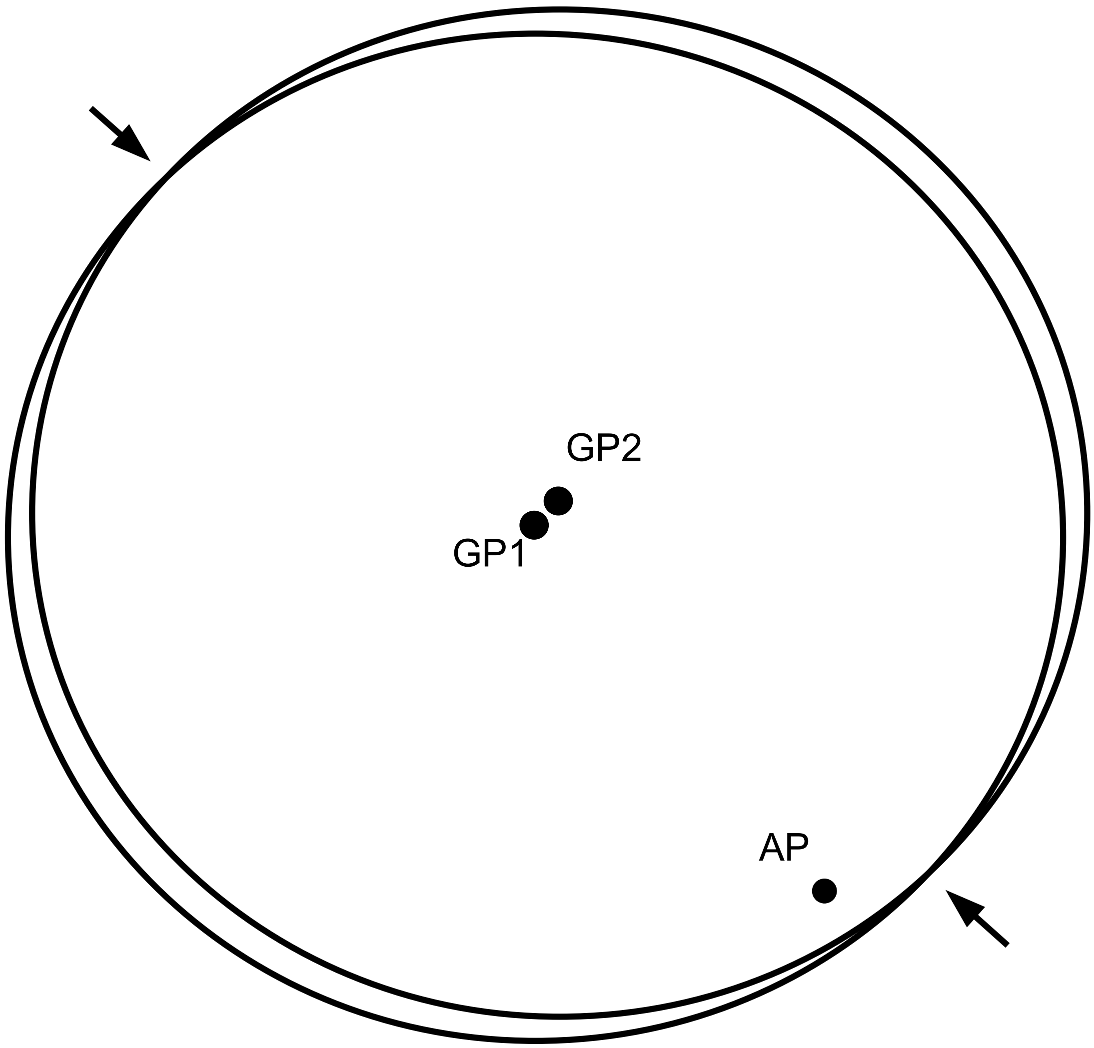


Figure 3.54: Two circles of position for two closely spaced GPs.

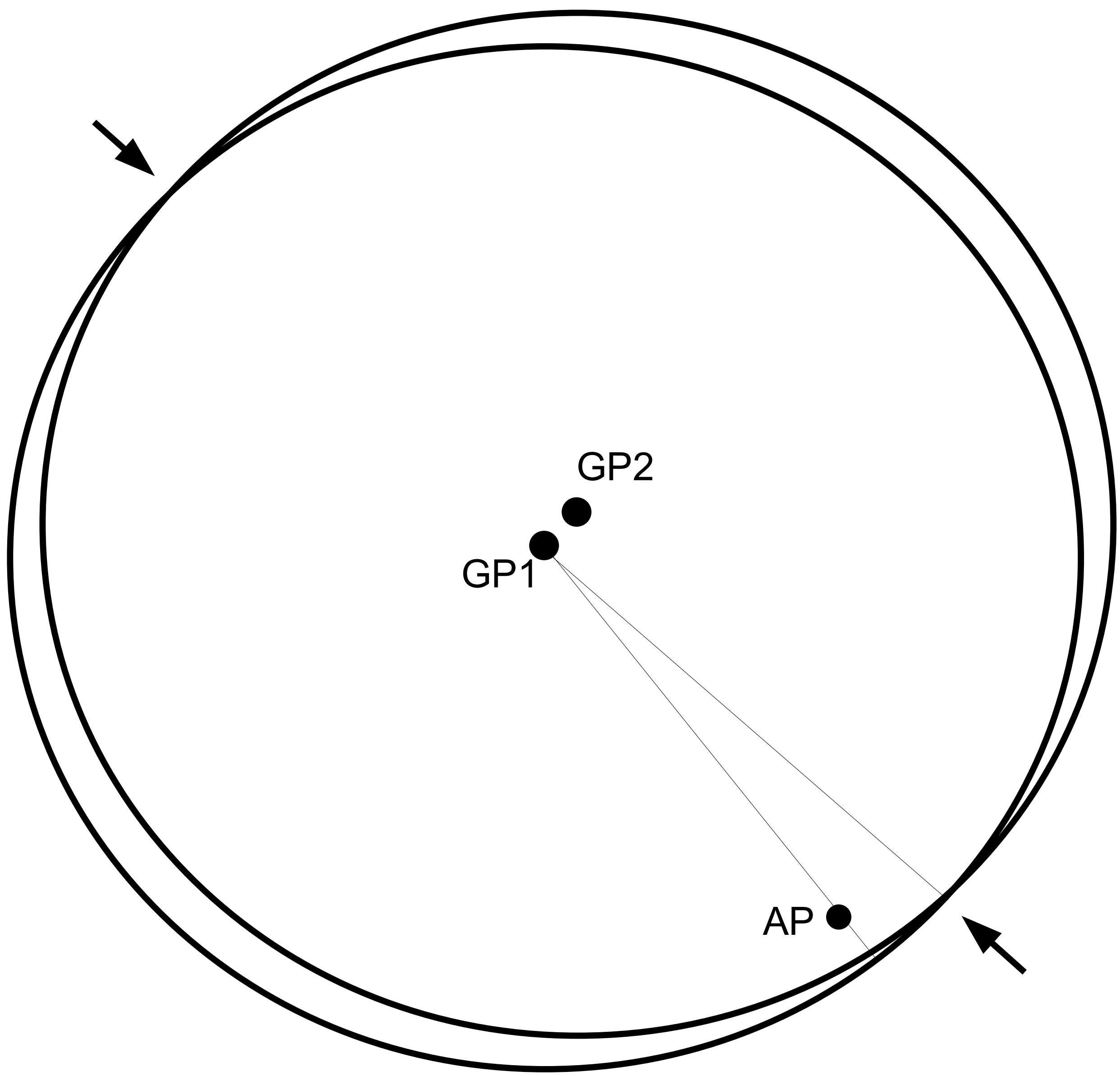


Figure 3.55: Lines from GP1 to both our AP and real position.

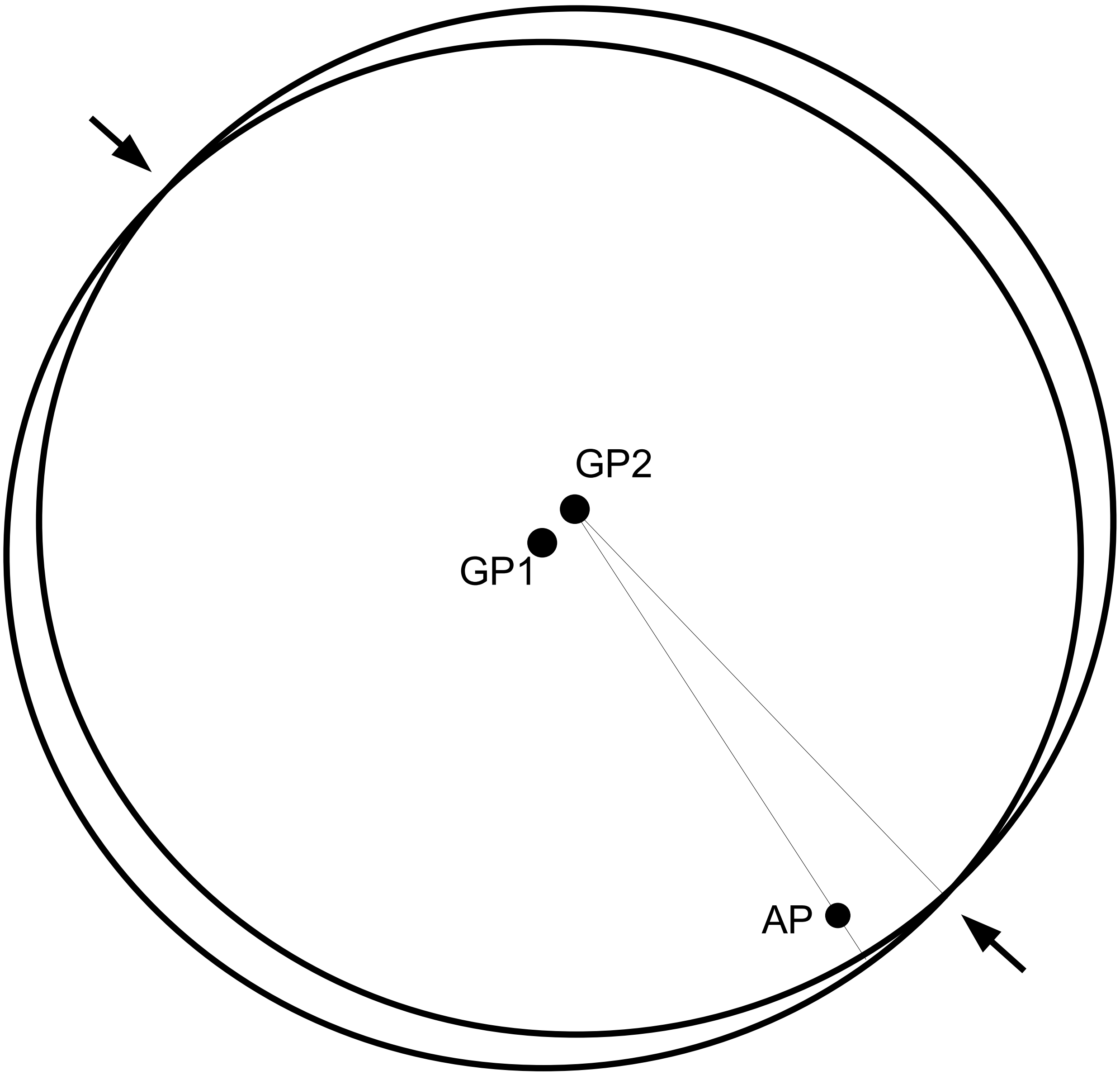


Figure 3.56: Lines from GP2 to both our AP and real position.

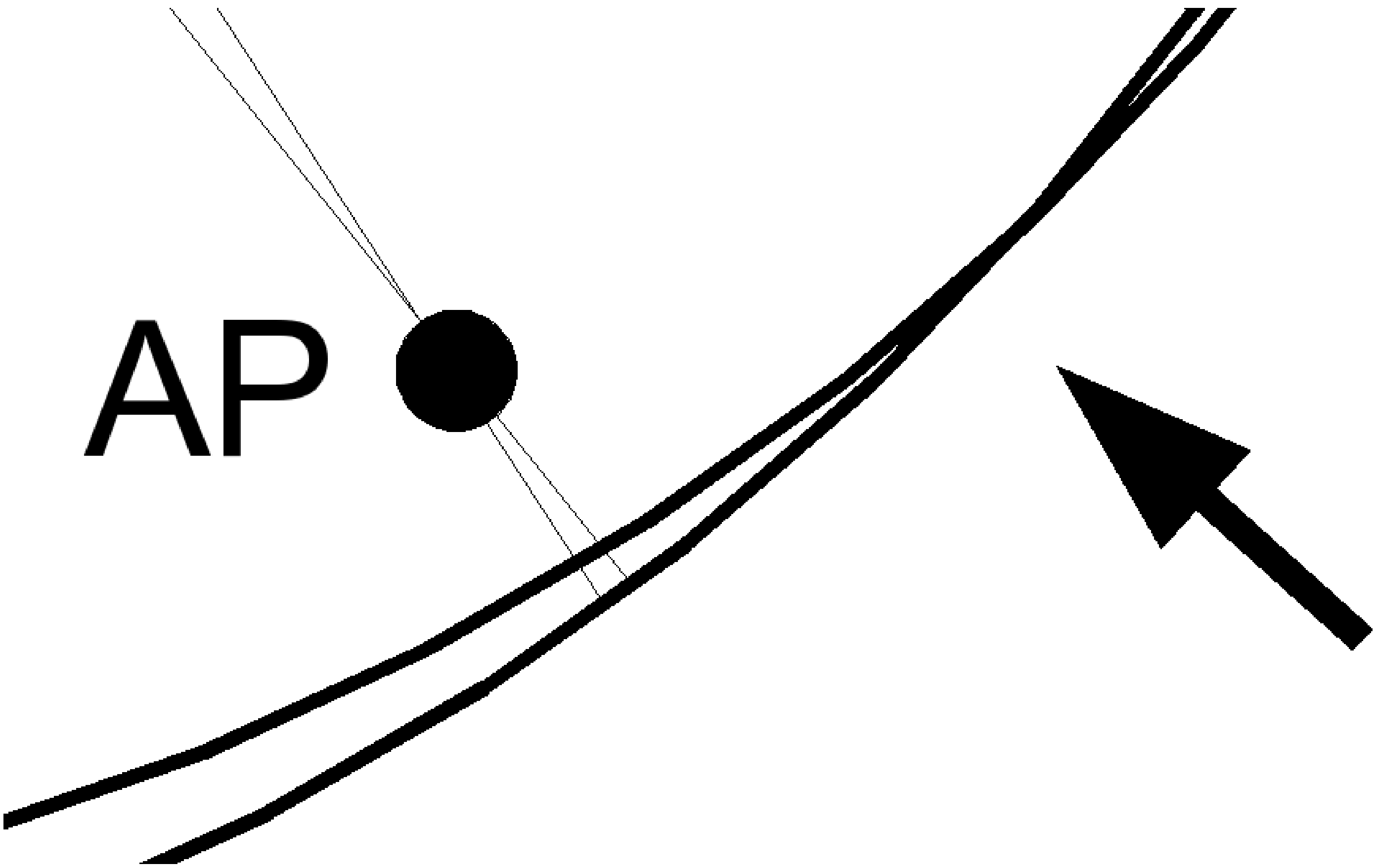


Figure 3.57: A 300x zoom to the region near our AP.

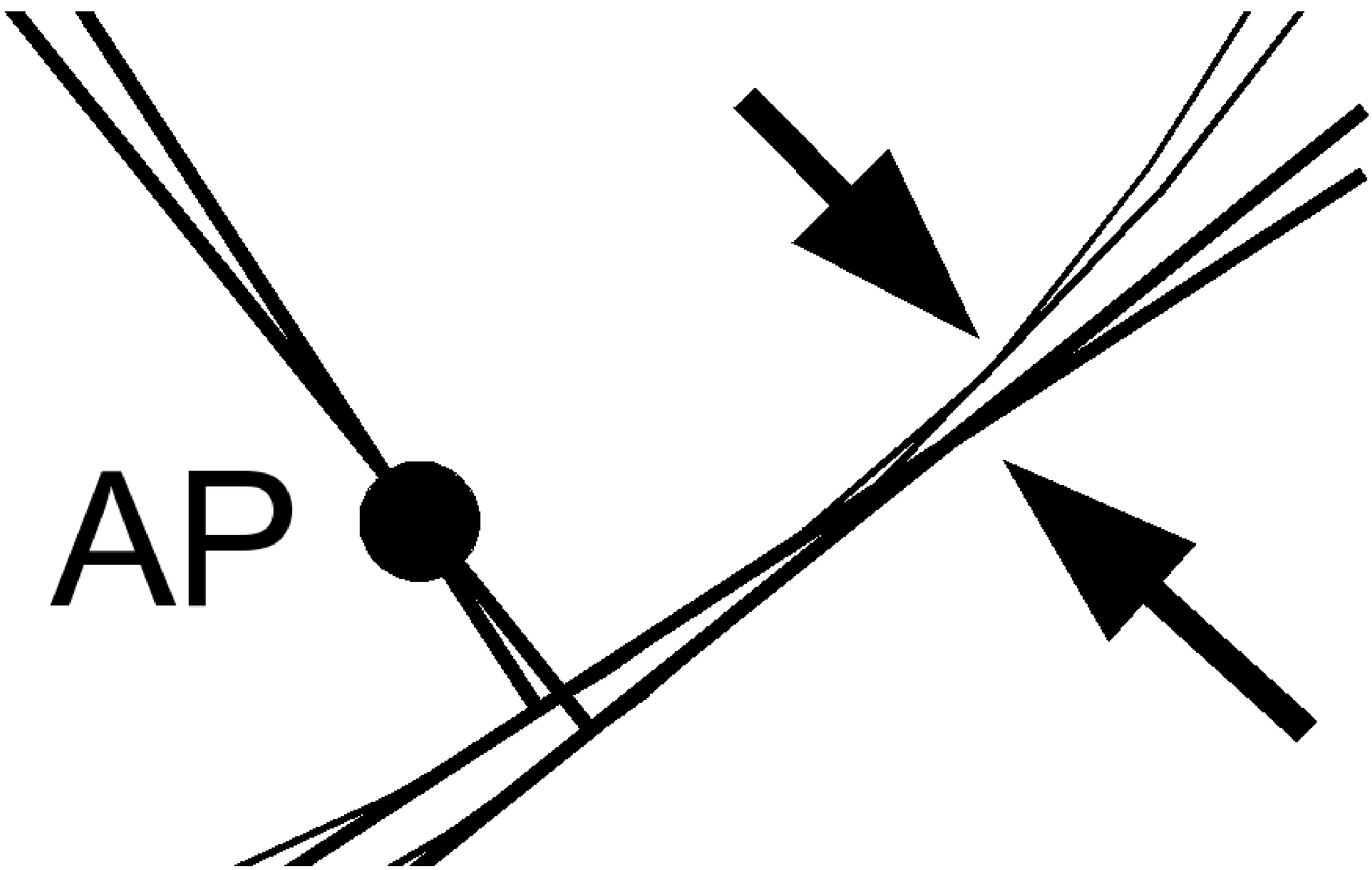


Figure 3.58: With the tangents (heavy lines) drawn, look how close their intersection is to the actual intersection between the two circles of position (thin lines).

developed around it, “running the parallel” as it was called. This could be risky (the Anson story), as at times it might be difficult to know if one should “turn left or right” when the proper latitude was reached (but this the same old problem with longitude).

Latitude is easy to find two ways. The first is by taking a shot of the sun at local noon and the second, is by taking a shot of Polaris, also known as the “north star.” (The Polaris shot only works, of course, if you’re in the northern hemisphere, but there are also guides, although not as convenient, in the southern hemisphere as well.) We’ll close this section by looking at both of these methods for finding one’s latitude.

3.5.1 Approximate Easy Latitude: A noon sight of the Sun

At all times of the year, the sun is directly over *some point* on the earth. In other words, the sun is directly over some GP on the earth. The GP is as high at $+23.5°$ in the northern hemisphere’s summer and as low at $-23.5°$ in the northern hemisphere’s winter, crossing the equator at $0°$ during the spring and fall equinoxes. The latitude of the sun’s GP is not very hard to calculate for all days of the year, and this was true as far back as at least the late 1600s. The oscillation of the sun’s declination between $\pm 23.5°$ might invoke your memory of the sine function from trigonometry, and indeed this can be used to compute the sun’s declination. In other words, tables of the sun’s latitude on each day of the year have been computable for a long time (prior to the Harrison’s time), having nothing to do with the hardship like that for computing the moon’s position over the course of a year. About the simplest approximate formula for the sun’s declination is

$$d = 23.45 \sin\left[\frac{360(284+N)}{365}\right], \tag{3.35}$$

where d is the sun’s declination (in degrees north (+) or south (-) of the equator) and N is the day of the year the declination is needed (1=Jan 1st, 2=Jan 2nd, etc.). The declination does depend on the time during the day, which doesn’t appear in the equation, so it can’t be entirely correct (it isn’t). But remember we are totally lost at sea and would be happy with just our approximate latitude. Also, 400 years ago, a table with some (any?) navigational information would have been desperately appreciated. A plot of Equation 3.35 is shown in Figure 3.59. Notice that the declination is at $-23.5°$ near the beginning and end of the year (dead of winter in the northern hemisphere), at $+23.5°$ in the middle of the year (mid-summer in the northern hemisphere), and has two equator crossings at $0°$ during the year (at the two equinoxes). Knowing the declination on a given day helps us in a way shown in Figure 3.60 (thanks to R. Echols for this discussion).

We can go out on the ship’s deck (at noon) and use a sextant to find the angle of the sun above the horizon. In Figure 3.60, this would be angle α. As long as we’ve been counting days at sea, the sun’s declination can be found from tables as angle d in the figure. Our unknown latitude is θ. A little careful geometry would tell us that

$$90 - \alpha = \theta + d, \tag{3.36}$$

or solving for θ, we’d get that

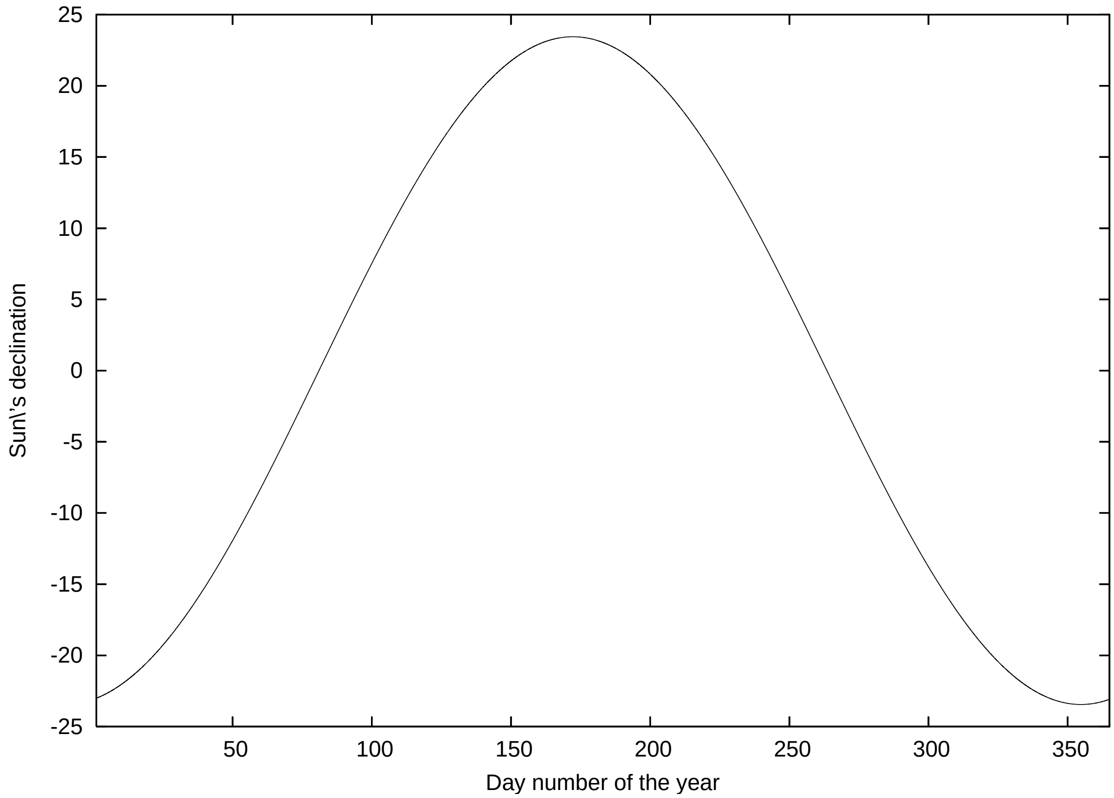


Figure 3.59: The sun's declination versus day number of the year. Notice it oscillates between $+23.5°$ and $-23.5°$ (the tropics), crossing the equation at $0°$ twice a year at the two equinoxes.

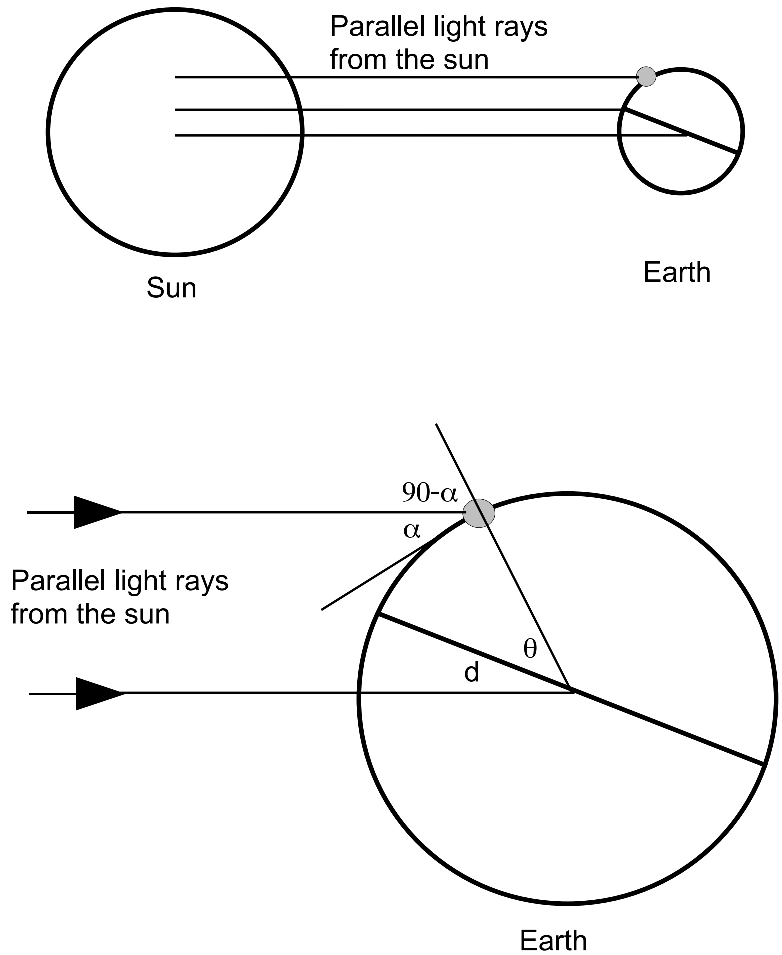


Figure 3.60: Finding approximate latitude from the sun's height above the horizon at noon. α is the sun's height above the horizon that we measure with a sextant, d is the sun's declination on that day, as found in a table, and θ is the latitude of our position (the small gray circle).

$$\theta = 90 - \alpha - d, \tag{3.37}$$

or

$$\theta = 90 - (\alpha + d), \tag{3.38}$$

So we can find θ, our latitude by subtracting the sum of α and d from 90°. That is, find the height of the sun above the horizon, add the sun's declination from the table, and subtract this number of 90° and that will be our latitude. This measurement is particularly fun twice a year at the equinoxes, when the sun is directly over the equator at noon, meaning $d = 0$; our latitude would then be simply $90 - \alpha$. That is, measure the height of the sun above the horizon, subtract it from 90°, and you'd have your latitude (this is true only during the two annual equinoxes, of course).

In either case, as discussed below, a night time sighting of Polaris (the north star) could be used to confirm any results found using the sun.

3.5.2 Better Easy Latitude: A noon sight of the Sun

In more modern times (post 1900), when the Nautical Almanac and marine chronometers (i.e. clocks) became refined and readily available, the height on the sun (at noon) could be used to more precisely find one's latitude. More precision requires knowledge of the time in Greenwich, essentially to correct for the *lack of any time* used in the procedure discussed in the previous section. At your local noon, the sun is at its highest point in the sky and is directly over the meridian on which you are located. This collapses the eternal triangle shown in Figure 3.42 to one where the distance between you and the sun's GP is along your meridian. That is, the sun's GP is on your meridian, and your distance from the GP (which comes from your altitude measurement) is more or less your latitude. To start, take a look at Figure 3.61.

As you can see the eternal triangle does not stretch across the earth's surface like a rubber sheet. Instead, it is contained between your position, the sun, and the sun's GP. In Figure 3.62, the important distances are shown. As you can see in this Figure, A is the declination (or latitude) of the sun's GP. B is the distance between your location and the sun's GP, and C is the latitude of your position.

Remember that C (your latitude) is what you'd like to find. The sun's latitude, A can be looked up in the Nautical Almanac, and B is what you get when you take the sun's zenith measurement. From the Figure you can see that $A = B + C$ so that $C = A - B$. So, when the sun is to your north, and you're in the northern hemisphere, you subtract the sun's zenith measurement from the sun's declination, and PRESTO! You have your latitude. You should be able to draw a similar figure for when you are in the northern or southern hemisphere, and/or the sun is to the north or south of you at local noon.

If you are paying close attention, you might be wondering about something. The sun's GP is needed from the Nautical Almanac. Doesn't this mean we need to know the time in Greenwich? Drat! Are we back to needing time at sea? Yes, but not accurate time. This

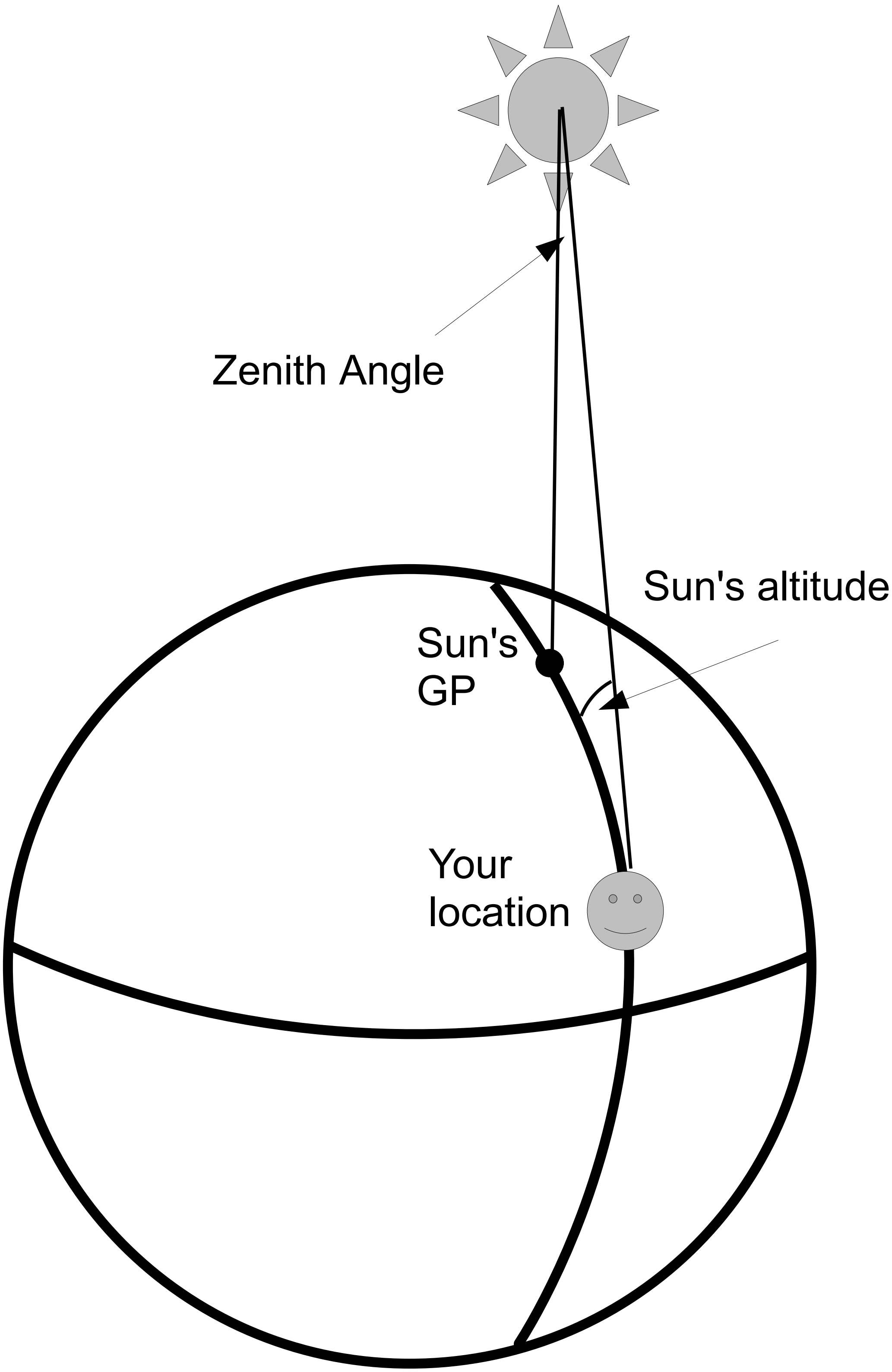


Figure 3.61: The Sun’s GP versus your position when observed at your local noon.

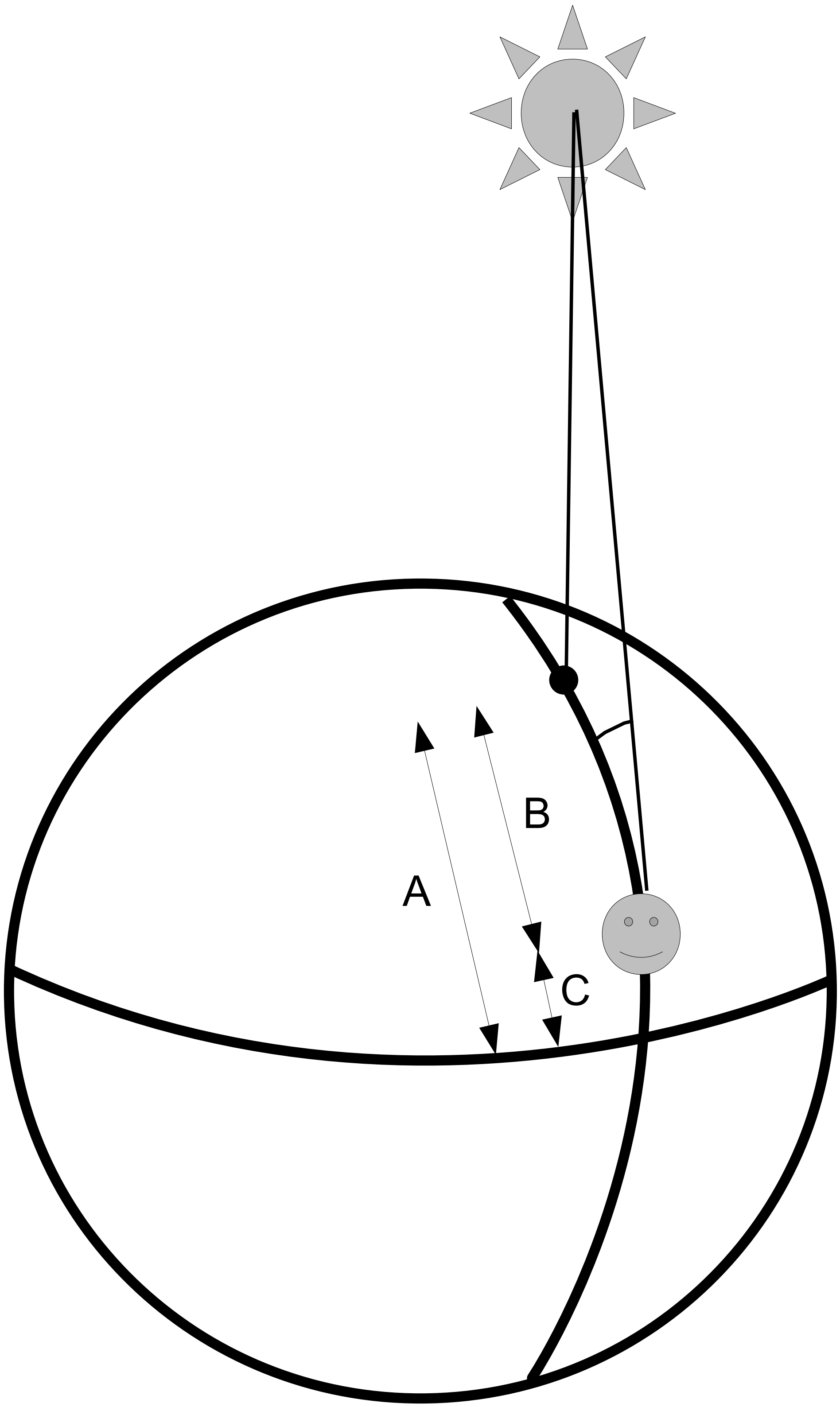


Figure 3.62: A is the declination (latitude) of the sun's GP, C is the latitude of your position, and B is the distance between your position and the sun's GP.

latitude procedure only uses one component of the sun's GP from the Nautical Almanac; the sun's declination. If you look at these numbers in the Almanac, they don't change very fast; at most by $1'$ *per hour* around the time of the equinoxes. In other words, the sun's GP doesn't wander toward or away from the equator (or north pole) along your meridian very fast. Refer again to Figure 3.21 and read down the "Dec" column; the numbers don't change very much over the course of an hour. Of course, when far out at sea for months at a time, even estimating the time in Greenwich would have been impossible; in fact it *was impossible* that's why we're all here! East-west position (longitude, which *was* your time) was simply a lost quantity. These days however, "approximate time" simply means you don't have to yell "mark" and take a precise time from your watch when you get the sun's altitude.

Activity

Try to meet at noon and do these two observations.

3.5.3 Easy Latitude again: A sight of Polaris at night

Viewers (at night) in the northern hemisphere have always had a special star, called "Polaris" or the "north star." It is a star that is almost directly over the northern tip of the earth's axis of rotation. This means if you extended an imaginary line straight up along the earth's north pole, the line would eventually run into Polaris.

The star can always be used to find the "north" direction (if you are lost at night and also lost your compass). It can easily be found using the two "pointer stars" on the edge of the Big Dipper's bowl. Just follow an imaginary line through these two stars, away from the "cup" of the Dipper, and you'll run right into Polaris (see Figure 3.35). Think carefully about how such a star can be used to get a rough estimate of your latitude. It works as follows.

If you are standing on the north tip of the earth's axis, Polaris would be directly overhead; you would have to look *straight up* to see it. Straight up means looking up at 90° relative to the horizon and that would be your latitude; 90°, since you are on the north pole. Suppose now you started to move south, toward the equator. Polaris would begin to get lower and lower in the sky. Your view to see it would have to change from looking straight up, to a little off of straight up; maybe 70°, instead of 90°. Your latitude would then be 70°. As you approach the equator, Polaris would be low, just beginning to dip and hide below the horizon, becoming completely hidden as you enter the southern hemisphere. Right at the equator, you would have to stare directly at the distant horizon to see Polaris; this is a view angle of 0°, which is your latitude on the equator. So getting your latitude from Polaris is as easy as finding its angle above the horizon.

There is one complication about using Polaris for a *more accurate* latitude determination (isn't there always?). The complication is that Polaris is *almost* over the north pole of the earth. In fact, if you are on the north pole, you'd have to look up at $89°8.5'$ or 89.14° to see it. For this reason, Polaris isn't directly over the earth's north pole (bummer), but travels in a small circle around it. The Nautical Almanac has tables that give you a correction for this; that is, a number you'd add or subtract from your measured altitude of Polaris, to correct

for this small circle in which Polaris travels. The corrections are given in terms of the "LHA of Aries." Remember Aries is the imaginary point on the celestial sphere that sets "zero" for right ascension. LHA stands for "local hour angle" and is the angle between your longitude (projected onto the celestial sphere) and the longitude of Aries ("local" refers to being relative to your local position). How does one find this? A simple diagram helps. To start, look up the GHA of Aries given the time in Greenwich. GHA stands for "Greenwich hour angle" and is the angle between the longitude of Greenwich (projected onto the celestial sphere) and Aries, as measured *west* of the prime meridian at Greenwich. The Nautical Almanac tabulates these angles.

Here's how LHA is found from GHA. Suppose at the time we need, the GHA of Aries is $43°39'$ and our last GHA (our longitude on earth) is $117°14'$ West. We could plot these on a circle as shown in Figure 3.63. This circle is what would be seen by looking down on the earth from above the north pole. The $0°$ mark is the prime meridian and Greenwich. You can see how the GHA of Aries is marked. The line pointing to it, if extended to the celestial sphere, would intersect Aries. Our longitude is where the $117°14'$ intersects the earth's surface. The LHA of Aries is how far *west* of our location do we need to sweep to find Aries. The answer in this case is $283°25'$, as shown in Figure 3.63.

So we look up a "Q" (or "wobble" correction factor) for Polaris, under this LHA and we get $+16'$. So suppose we took an altitude reading for Polaris and got $32°7'$. Our actual latitude would be $32°7' + 16'$ or $32°23'$. This is our full-out Polaris-determined latitude. The corrections vary from between $\pm 50'$ meaning your latitude could be off by as much as 50 nautical miles if you ignore this correction. Note that the correction requires knowledge of two issues that plague this entire class, the time at Greenwich to look up the GHA of Aries and our local longitude.

But remember, when lost at sea forget about the correction and remember

Your approximate latitude = the altitude of Polaris.

Knowing your latitude to 50 nautical miles would certainly be good enough for trips where one wants to "run the parallel." You simply try to sail east or west, while keeping Polaris at the same altitude, even if it required making nightly corrections.

Activity

Meet at night and find latitude from Polaris.

3.5.4 A Guess at History: Using Polaris + the Sun

Finally, just a guess about what sailors might have done with Polaris and the Sun without any idea about what time it is in Greenwich. Suppose you took an uncorrected Polaris sighting just as the sun was rising. It is a calm day at sea and you are pretty sure your latitude will hold (nearly constant) until noon. You take a sight of the sun at noon. In terms of Figure 3.62, you have C from your last Polaris sight, and B from your just-completed noon sighting of the sun. You can get the GP of the sun, A, using $A = B + C$ (see Figure 3.62).

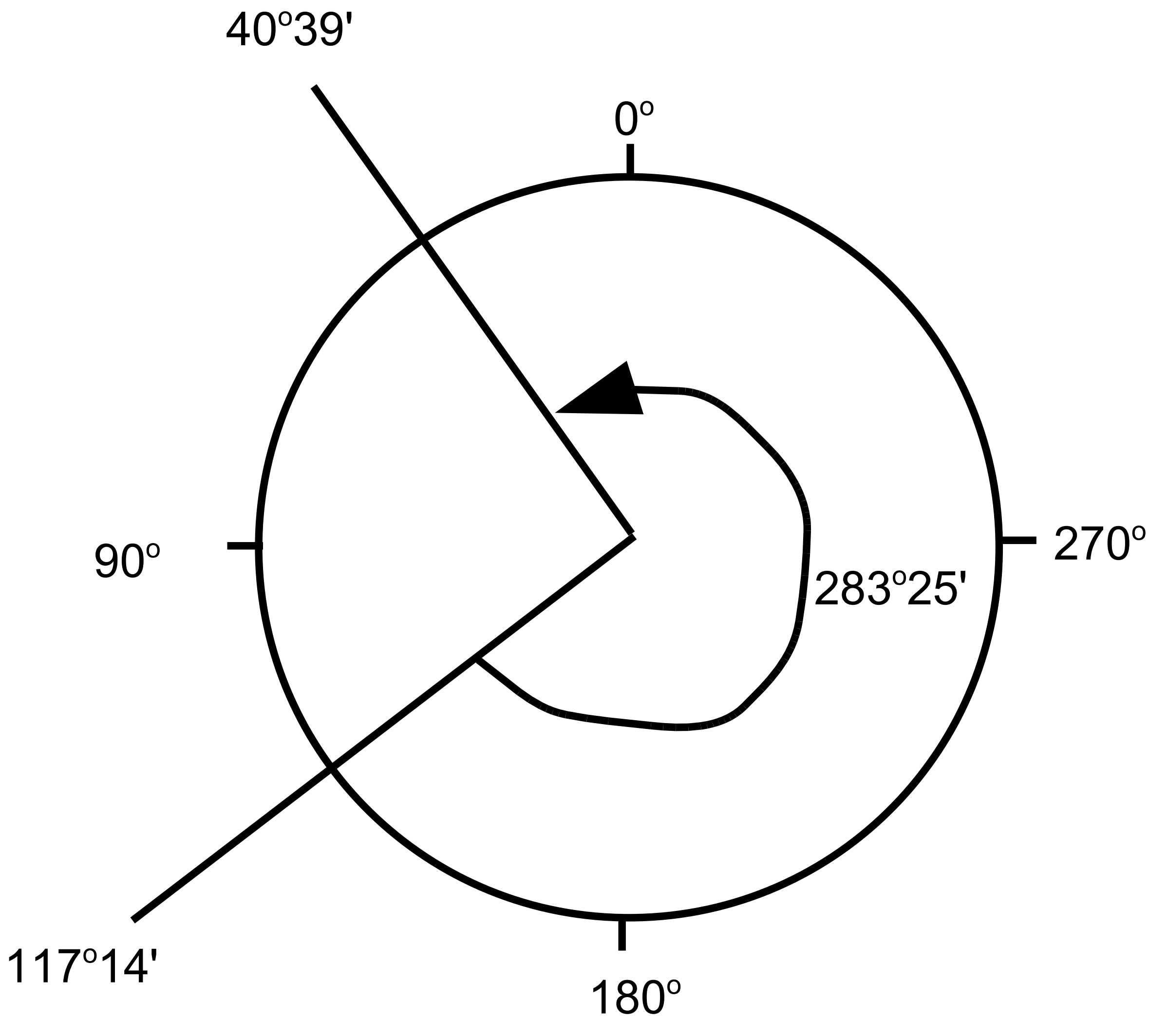


Figure 3.63: Finding the LHA of Aries to correct for the fact that Polaris is not directly over the north pole of the earth.

You then use the Nautical Almanac to look up what Greenwich time that day the sun is to have the GP you just found. This is using the Almanac in "reverse." Intead of looking up the sun's GP for a given time, you first find the sun's GP then look up the time in Greenwich that corresponds to it. This might be your best guess at the time in Greenwich. Also, it is noon for you, so you know your local time. You could then approximate your longitude from the difference of these two times (in hours) times 15° per hour. You also have your DR to compare with.

This is just a guess at navigational history; the author doesn't know if this was done or not. In either case, remember you started with an uncorrected Polaris sighting (so your latitude C is off by a little bit) and you assumed that your latitude stayed constant for the hours between your last Polaris sighting and your next local noon. Seems at least plausible, but there is still no historical evidence that "pre-clock" navigators were anything but *utterly ignorant of their longitude.*

Chapter 4

Time and Time Keeping

4.1 Introduction

Whether stated or not, time has been central to our discussion thus far. Finding longitude, at sea or elsewhere, is all about knowing the time. Even with the wonderful way in which celestial navigation allows us to use the stars to find our location, one must still know the time. The longitude prize was one by finding a way of keeping accurate time (at sea). In this chapter, we'll dispense with longitude, sailing, and scurvy, and try to look at time as its own entity. So then, what is time?

It would be nice if we could find a good definition of time, but they're all kind of abstract. The dictionary that comes with Apple's Macintosh computers defines time as "the indefinite continued progress of existence and events in the past, present and future regarded as a whole." Huh? Dictionary.com defines it as "the system of those sequential relations that any event has to any other, as past, present, or future; indefinite and continuous duration regarded as that in which events succeed one another." These definitions aren't helping.

What does "time" mean to you? Perhaps you're always feeling like there's not enough of it? The incessant clicking of a clock? Is it always getting "later and later?" Are you getting "older and older?" Or do you find yourself wishing time away, as in "I wish it was Friday?" or "I can't wait until next summer?" Or do you desperately cling to it, as in "I wish this day would last forever?" Time is a funny entity in our lives, and it affects us all in undeniably critical ways. As you've also seen, time played a critical role in solving the longitude problem. In fact, keeping time *was the solution.*

Time has a strong psychological meaning too. Fun things "go by fast" and boring things "go by slow." We wish some things could "hurry up" and others could "slow down." It can cause us great stress (not enough time), or allow us to completely relax (a vacation with lots of free time). There was a recent paper[C. Stetson, plosone.org, 12/2007, Issue 12] which claimed that time seemed to slow down during a frightening event. Is it some increased temporal resolution we have when we are frightened, or is it just the way we remembered something stressful? (It was actually a function of how the test subject recalled the event; a "richer encoding of memory" made the event appear to last longer.)

Time is something like an an asymmetrical arrow drawn on our existence. In a car, we

can go forward and backward, but time always points forward. It never stops or reverses, so it is not the same as space in that regard. History (the past) is unchangeable, and the future is not certain. We really only have slight control over a singular instant of time, which is "right now." There, you missed it. That "right now" just passed. No one can say for certain what will happen even one second from now.

The Riddle of Time

[Jespersen, p. 3-5]

- It's present everywhere, but occupies no space.
- We can measure it, but we can't see it, touch it, get rid of it, of put it in a container.
- Everyone knows what is is and uses it every day, but no one has been able to define it.
- We can spend it, save it, waste it, or kill it, but we can't destroy it or even change it, and there's never any more or less of it.
- We can see distance and feel weight and temperature, but we cannot apprehend time by any of the physical senses. We cannot see, hear, feel, smell, or taste time. We know it only though consciousness, or through observing its effects.
- "Now" is constantly changing. We can buy a good meter stick, or one-gram weight, or even a thermometer, put it away in a drawer or cabinet, and use it whenever we wish. We can forget it between uses–for a day or a week or 10 years–and find it as useful when we bring it out as when we put it away. But a "clock"–the "measuring stick" for time–is useful only if it is kept "running." If we put it away in a drawer and forget it, it "stops," it becomes useless until it is "started" again, and "reset" from information available only from another clock.
- We can write a postcard to a friend and ask him how long his golf clubs are or how much his bowling ball weighs, and the answer he sends on another postcard gives us useful information. But if we write and ask him what time it is–and he goes to great pains to get an accurate answer, which he writes on another postcard, well, obviously before he writes it down, his information is no longer useful or valid.

What's really important though, at least for the longitude problem, is not some agreeable definition of idea of time, but the ability to *measure* time and to *keep* good time. This chapter will address this idea: How does one go about keeping track of time?

4.2 Measuring Time

One way of measuring time is to use something periodic; something that happens over and over again, that we can use as a time keeper. Every time this "thing" starts over again,

we know the same amount of time went by and we can make a measurement. So what's convenient and periodic? How about a day?[Feynman, p 5-1]. A day seems to happen over and over again. Using a day as a our time keeper might work to keep track of how old we are, how long we've been at sea, or how long seeds in our garden have been planted. But then we might ask, are days really periodic? Are they regular?[Feynman, p. 5-2]. Summer days seem longer than winter days. Is there a way to test if days are really periodic?

To test for "constant-day-ness" you'd need something to "slice up" a day or something that would allow you to observe the passage of a day, piece by piece. How about an hour glass? From the moment the sun rises, you could start an hour glass, and when the top "tank" is empty, you could tip it–again and again–until the sun sets. Count the number of times you've tipped the hour glass, and compare your counts after several seasons. What sort of count is convenient? Is 10,000 tips a day convenient? 2? 50? Let's say that 24 was your favorite number and you decided on 24 tips a day, and you called each tip the passage of "an hour."

Now, what if the count of "24 tips a day" varied? Between some sun rises and sun sets it was 22 tips, others it was 30? What then? Is the hour glass no good? Maybe the sunrise/sunset combination is not a good start/stop combination? (See, there's time again, constantly eluding us, even as we try to measure it; this is what poor John Harrison went through with his clocks.) Well, we know the sun's highest position in the sky is unique for us; why not use that? Rest assured, with a good enough hour glass, you would find the time between successive meridian passages of the sun will be the same from day to day. (Note: But don't put much faith into an hour glass; they stink as time keepers.) So we haven't really proven anything, but we've found a periodic time standard. Certainly though, our work isn't done with a reliable clock that has a resolution of one day. We need to do better in order to time events or durations in our everyday lives.

4.2.1 Short Times

Let's try to figure out how to time fractions of a day. The sun's highest point gives us a day; the hour glass (ahem) gives us the hour, 24 of which form a day. How do be split up an hour? Let's try a pendulum, which is a weight tied to the end of a string, that swings back and forth. As you'll learn below, a pendulum that can be built fairly easily, can have more-or-less any "swing time" we'd like it to have, and it can keep this swing time fairly well. Suppose then we decided to build a pendulum that had $3,600$ swings between turns of the hour glass. We could call each swing of the pendulum a "second." And once again, if our pendulum consistently swung back and forth $3,600$ times between turns of the hour glass, we'd begin to have confidence in the pendulum. So we'd have a day, hour, and second well under control. But how much faster can we go? How much more can we slice up our time base?

For anything shorter than a second, mechanical objects (i.e. things that move) don't work very well. Nature doesn't provide anything obvious to use either. To test this, try to accurately time something lasting a tenth of a second. What would you use? A blink of your eye? A snap of your fingers? To go shorter than a second, you'd need to use electronics.

Details on electronic timers will be discussed below, but for now, trust us that we can *easily* slice up a second using electronics (switches, wires, light bulbs, etc.).

To test our ability to keep track of seconds, we'll follow our current trend of taking our last acceptable timer, and slicing it up (we sliced up the day into hours, and hours into seconds). Let's slice up the second now. We'll build a series of electronic timers, each 10 times faster than the one before it. That is, each new [faster] timer is "calibrated" against the previous [slower] one[Feynman, p. 5-2]. But as we go less than about a second, we won't be able to manually observe the time ticking by. We need some kind of electrical device to show us what our fast electronic time keeper is doing. Luckily for us, devices called "oscilloscopes" have been around since about the time the television was invented. But look at the quagmire we're in! We are at the point where we need a machine to observe our machine to tell us how well we're able to keep track of time!

Rest assured though, it all works. Electrical timers these days have no trouble keeping time down to 0.000000000001 seconds, or 1 picosecond (one million millionth of a second). What about smaller times? A femtosecond, which is 14 zeros in front of a 1 is coming of age, but beyond that, no one really knows. Does time exist if we cannot measure it? Does it make sense to talk about times so short that we cannot measure them?

Here is an odd idea to end this section. One of the big physics theories of the day is the "Planck time" or 0.000000000000000000000000000000000000000000005 seconds (a 5 with 43 zeros in front of it). It is the time it takes a photon, traveling at the speed of light, to travel one "Planck length" which is about 0.0000000000000000000000000000000001 m (a 1 with 34 zeros in front of it). The number comes from a unique way other fundamental constants we behold and trust can be combined to give time. Currently, no one thinks the laws of physics apply for time scales shorter than this. If you're interested, the formula for the Planck time is

$$t_{planck} = \sqrt{\frac{hG}{2\pi c^5}}. \tag{4.1}$$

The reason why it is hard to ignore the Planck time is because its formula involves constants that are so sacred in physics. In the formula, h is Planck's constant, which is fundamental to all of quantum physics, G is the gravitational constant which is fundamental to describing the universe, and c is the speed of light. There is rarely a theory that pretends to describe the very very small (h), and the very very large (G), but there you have it. Your conclusions about t_{planck} are as good as anyone else's at this point.

So here's where we're at with short times. We can accurately time down to somewhere in the picosecond to femtosecond range, but that is all. Current physics-research literature (circa 2008) is beginning to discuss the attosecond (0.000000000000000001 seconds) range more and more, but for now that is all. So in principle, we could build a clock that kept track of days, hours, minutes, seconds, tenths of seconds, etc. all the way down to picoseconds or so. And that's the best we can do for now. Perhaps technology will someday allow us to keep time down to the Planck time. Who knows? What would this do for us? But keep things in perspective. In Harrison's time it was tantalizing to keep track of "boring old" seconds accurately, and when this was possible, finding longitude suddenly became a solved

problem. As you'll see later, keeping time down to the nanosecond (0.000 000 001 second) is required for GPS to work. We'll find some use for attoseconds someday.

4.2.2 Longer times

After discussing some ideas on how to keep track of short times, how about keeping track of long times? How about times longer than one day? Well, this is easy. Just start counting days. As long as someone is around to count the days, we'd do just fine[Feynman, p. 5-3]. We'd find periodicity in our world at about 365 days (1 year), and nature keeps track of years by tree rings, erosion sediment, flooding of the Nile, Monsoons in south-east Asia, etc.

But here we go again with the question of keeping track of time. What about really long times? Hundreds or thousands of years? Just as the hour glass and pendulum helped us with short times, we now need a periodic event to keep track of longer times. What would this be? What sort of "clock" could we build that kept track of thousands of years, for instance?

For very long times, we invariably turn to radioactivity. Some atoms exhibit what is called "radioactivity." Think of it as an atom that glows. The glowing can be the atom spitting out light or a particle, like an electrons, neutrons, or protons. With each particle spit out, a little bit of the original sample disappears; the atom that spit out the particle does so at the expense of turning into some other atom. This is called "radioactive decay." Radioactive decay follows a simple trend too: for equal time intervals over which the sample is allowed to decay, equal fractions disappear.

As an example, suppose we prepare a brand new radioactive sample, say 5 grams of element X. We then let it glow for 5 minutes and notice that 1/8 of it has disappeared, meaning we'd only have 4.375 grams of it left. If it glows for another 5 minutes, 1/8 of the 4.375 grams would disappear, leaving 3.828 grams. After more and more time, when half of the original 5 grams is gone, we say that material X has reached its "half-life." If we followed this sample a bit longer, its decay would look as shown in Figure 4.1.

From this figure, you can see that about half of the sample has decayed after about 25 minutes. This is precisely how radioactive decay can be used as a long-term clock: you measure when one-half of an original sample has decayed. This is called the "half-life" of a sample, sometimes written as $t_{1/2}$. The bars in Figure 4.1 follow an exponential decrease in height as time goes on. This is typical of radioactive decay as well: the amount of sample at any given time exponentially decreases. In Figure 4.1, the exponential function describing the height of the bars turns out to be $5e^{-t/37}$, where 5 is the amount of the original sample, and 37 is the "mean life-time" of the sample, sometimes written as τ. Half-life and mean-life are related as $t_{1/2} = (\ln 2)\tau$. You can think of τ as sort of the "average" overall useful life of the sample.

So, back to keeping track of long times. Carbon-14, for example, has a half life of $5,000$ years. So say we made a fresh batch of it; 1 gram of C-14. We could set it somewhere and leave it alone. If someone came back to it and found 0.5 g, of our original batch, we could all agree that $5,000$ years has just gone by. There! We've just times the passage of $5,000$ years! There are many such radioactive elements we could use for our "long" time keeping. Helium-3 has a half life of about 12.3 years. Cesium-137 has a half-life of

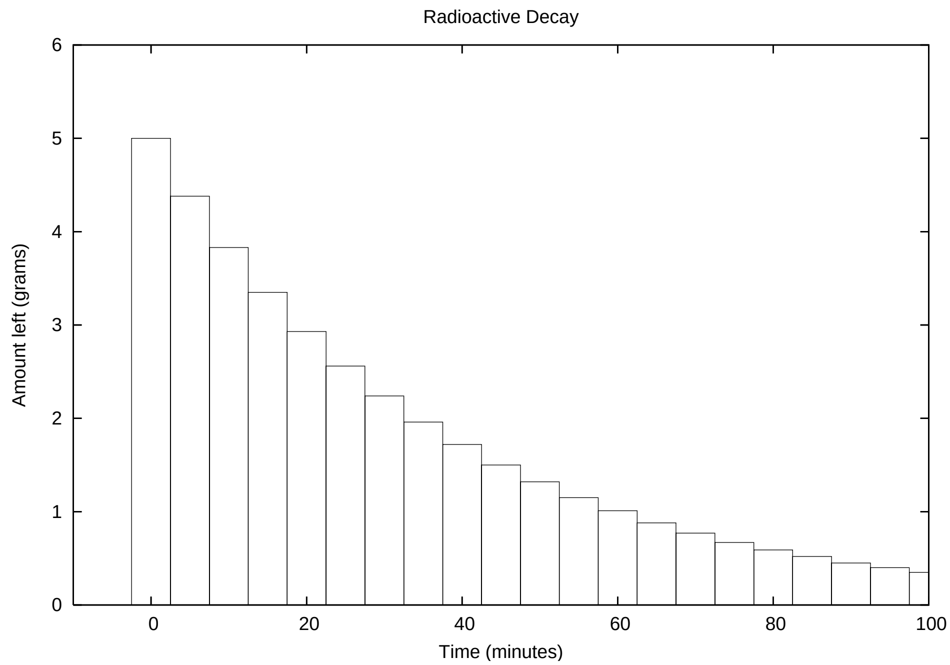


Figure 4.1: Amount left of a hypothetical radioactive sample as time goes on.

about 30 years. The other issue about radioactive decay, is that it's relatively insensitive to environmental parameters. So for instance, Carbon-14 from deep within the earth (and whatever conditions exists down there) has a half life of $5,000$ years, and so does the Carbon-14 in a newly planted tree (Carbon-14 exists in the atmosphere; it's created by collisions between when cosmic particles hitting the air surrounding the earth). So radioactive decay is a pretty reliable clock for keeping track of "long times." Dinosaur bones, the earth, and prehistoric mammals are all dated using this radioactive decay concept.

4.2.3 A summary of long and short times

This interesting summary of time scales was taken from a 2006 issue of Scientific American called "From Instantaneous to Eternal," by David Labrador.

ONE ATTOSECOND (a billionth of a billionth of a second)
The most fleeting events that scientists can clock are measured in attoseconds. Researchers have created pulses of light lasting just 250 attoseconds using sophisticated high-speed lasers. Although the interval seems unimaginably brief, it is an aeon compared with the Planck time—about 10^{-43} second—which is believed to be the shortest possible duration.

ONE FEMTOSECOND (a millionth of a billionth of a second)
An atom in a molecule typically completes a single vibration in 10 to 100 femtoseconds. Even fast chemical reactions generally take hundreds of femtoseconds to complete. The interaction of light with pigments in the retina—the process that allows vision—takes about 200 femtoseconds.

ONE PICOSECOND (a thousandth of a billionth of a second)
The fastest transistors operate in picoseconds. The bottom quark, a rare subatomic particle created in high-energy accelerators, lasts for one picosecond before decaying. The average lifetime of a hydrogen bond between water molecules at room temperature is three picoseconds.

ONE NANOSECOND (a billionth of a second)
A beam of light shining through a vacuum will travel only 30 centimeters (not quite one foot) in this time. The microprocessor inside a personal computer will typically take two to four nanoseconds to execute a single instruction, such as adding two numbers. The K meson, another rare subatomic particle, has a lifetime of 12 nanoseconds.

ONE MICROSECOND (a millionth of a second)
That beam of light will now have traveled 300 meters, about the length of three football fields, but a sound wave at sea level will have propagated only one third of a millimeter. The flash of a high-speed commercial stroboscope lasts about one microsecond. It takes 24 microseconds for a stick of dynamite to explode after its fuse has burned down.

ONE MILLISECOND (a thousandth of a second)
The shortest exposure time in a typical camera. A housefly flaps its wings once every three milliseconds; a honeybee does the same once every five milliseconds. The moon travels around Earth two milliseconds more slowly each year as its orbit gradually widens. In computer science, an interval of 10 milliseconds is known as a jiffy.

ONE TENTH OF A SECOND
The duration of the fabled "blink of an eye." The human ear needs this much time to discriminate an echo from the original sound. Voyager 1, a spacecraft speeding out of the solar system, travels about two kilometers farther away from the sun during this time frame. A hummingbird can beat its wings seven times. A tuning fork pitched to A above middle C vibrates four times.

ONE SECOND
A healthy person's heartbeat lasts about this long. On average, Americans eat 350 slices of pizza during this time. Earth travels 30 kilometers around the sun, while the sun zips 274 kilometers on its trek through the galaxy. It is not quite enough time for moonlight to reach Earth (1.3 seconds). Traditionally, the second was the 60th part of the 60th part of the 24th part of a day, but science has given it a more precise definition: it is the duration of 9,192,631,770 cycles of one type of radiation produced by a cesium 133 atom.

ONE MINUTE
The brain of a newborn baby grows one to two milligrams in this time. A shrew's fluttering heart beats 1,000 times. The average person can speak about 150 words or read about 250 words. Light from the sun reaches Earth in about eight minutes; when Mars is closest to Earth, sunlight reflected off the Red Planet's surface reaches us in about four minutes.

ONE HOUR
Reproducing cells generally take about this long to divide into two. One hour and 16 minutes is the average time between eruptions of the Old Faithful geyser in Yellowstone National Park. Light from Pluto, the most distant planet in our solar system, reaches Earth in five hours and 20 minutes.

ONE DAY
For humans, this is perhaps the most natural unit of time, the duration of Earth's rotation. Currently clocked at 23 hours, 56 minutes and 4.1 seconds, our planet's rotation is constantly slowing because of gravitational drag from the moon and other influences. The human heart beats about 100,000 times in a day, while the lungs inhale about 11,000 liters of air. In the same amount of time, an infant blue whale adds another 200 pounds to its bulk.

ONE YEAR
Earth makes one circuit around the sun and spins on its axis 365.26 times. The mean level of the oceans rises between one and 2.5 millimeters, and North America moves about three centimeters away from Europe. It takes 4.3 years for light from Proxima Centauri, the closest star, to reach Earth—approximately the same amount of time that ocean-surface currents take to circumnavigate the globe.

ONE CENTURY
The moon recedes from Earth by another 3.8 meters. Standard compact discs and CD-ROMs are expected to degrade in this time. Baby boomers have only a one-in-26 chance of living to the age of 100, but giant tortoises can live as long as 177 years. The most advanced recordable CDs may last more than 200 years.

ONE MILLION YEARS
After traveling for a million years, a spaceship moving at the speed of light would not yet be at the halfway point on a journey to the Andromeda galaxy (2.3 million light-years away). The most massive stars, blue supergiants that are millions of times brighter than the sun, burn out in about this much time. Because of the movement of Earth's tectonic plates, Los Angeles will creep about 40 kilometers north-northwest of its present location in a million years.

ONE BILLION YEARS
It took approximately this long for the newly formed Earth to cool, develop oceans, give birth to single-celled life and exchange its carbon dioxide–rich early atmosphere for an oxygen-rich one. Meanwhile the sun orbited four times around the center of the galaxy. Because the universe is 12 billion to 14 billion years old, units of time beyond a billion years aren't used very often. But cosmologists believe that the universe will probably keep expanding indefinitely, until long after the last star dies (100 trillion years from now) and the last black hole evaporates (10^{100} years from now). Our future stretches ahead much farther than our past trails behind.

4.2.4 Heisenberg: The best we can do

The last three sections, about keeping track of short and long times had a theme: reliability. We need periodic events that are reliable, so we can trust them as accurate time keepers. Radioactive decay is very reliable, as is the time between meridian passages of the sun (although our earth is spinning slower as time goes on). Electronic timers are very good, but hour glasses cannot be trusted. There is a profound law in science that tells us a bit about how precisely we can expect to measure time and it's called the "Heisenberg Uncertainty Principle."

When we know what time it is, call it t, we might be concerned about how well we really know this time to be correct. Say we timed a race and a runner did 1 lap in 120.8 seconds. Do we mean exactly 120.8 seconds, or do we mean between 119 and 121 seconds? The 0.8 seconds might not be good enough to judge a tight race. So how well do we know and trust the 120.8? What if we were on trial and the 0.8 seconds was the difference between a guilty and not-guilty verdict?

Such a question begs then, what is the "uncertainty" in our measurement of $t = 120.8$ s? The uncertainty is called Δt, and suppose the manufacturer of the stop watch we used to time the race put a sticker on the back of the stop watch that said that Δt for the watch was 0.01 seconds. The manufacturer knows this because to be safe (and avoid lawsuits) they tested their stop watch against a more precise clock and determined that their design had some random error in it at the level of 0.01 seconds. (Note: there's that circular theme about time again; testing one time standard against another, just like the day was tested with the hour glass). They couldn't figure out what was causing the 0.01 second random error, but they decided to sell the stop watch anyway because most people won't worry about such a small timing error. So any time from the watch would have to be written as 120.8 ± 0.01 seconds. This means if you read 120.8 s, it could actually mean 120.79 or 120.81 seconds. If you don't like this, then buy another watch.

So this is what "uncertainty" is for time. It's how well we trust the time we are stating. It turns out that there is an ultimate precision to which we can know any time, and it *is not* zero. In other words, we can never exactly know what time it is. This is stated by the Heisenberg Uncertainty Principle to be

$$\Delta t = \frac{\Delta E}{h}. \tag{4.2}$$

In this equation Δt is our uncertainty in time, h is Planck's constant 0.000 000 000 000 000 000 000 000 000 000 000 663 and ΔE is the uncertainty in the time keeper's energy (of all things). You can think of ΔE as the ultimate form a clock has (i.e. what it looks like, what it's made of, etc.)

This law doesn't apply to macroscopic objects, like a clock sitting on your desk, but it does tell us that there is a limit to how well we can (someday) end up knowing time. The equation states that Δt (the smallest uncertainty in time) is actually equal to some number that isn't zero! For the sake of this book, this means we'll also never know exactly what our longitude is, but surely we'll never need to know our longitude to, for example, with width of our ship!

4.2.5 Time in Science

In science time always stands alone as a variable. Science describes many things we can measure and observe, but they aren't all entirely unique; many of them are interrelated. For example, you've probably heard of such things such as force, current, voltage, charge, mass, resistance, momentum, velocity, position, acceleration, torque, stress, strain, pressure, to name a few. Momentum though, comes when you mix mass and velocity. Current comes from mixing charge and time. Force mixes mass and acceleration. But time stands alone. It does not come from any mixture of other parameters; it is fundamental.

Time is also our critical "fourth dimension." Space, or the precise location of something, serves as the first three dimensions. The first is your left-right location. The second is your up-down location, and the third is your up-back position. In mathematics, we call this your x, y, and z coordinates (three dimensions). But your location can change with time. You might be walking or moving somehow. Even if you are standing still, the earth is still moving through the universe. Changing one's position in space always has to do with time, the fourth dimension. So to fully specify where something is located, you would have to deal with four dimensions: x, y, z, and t (t for time).

For scientists, time can create quite a headache when solving a problem. Learning how something "evolves" in time creates untold difficulty. Remember the job of predicting the moon's position (Chapter 2), given it's erratic behavior? The very word "prediction" carries the connotation that time is involved; you want to state, with confidence, that something will happen in the future. This is hard to do and no one can be sure their answer will be correct. To make a prediction, one must be so confident as to declare that they will peer into the future, and state what is going to happen (e.g. where will the moon be?).

Go out and look at the moon. A simple one minute shot with a sextant tells you a lot about where the moon is at that instant. Next, state with scientific certainty, exactly where it will be 15 minutes from now. How would you do it? You would probably need a Ph.D. in physics to figure this out, but this is the difficulty that the dimension of time adds to your analysis. Imagine next the job of compiling the Nautical Almanac for a given year. You must predict the GP of all the planets, the sun, many stars, and the moon for each hour of each day for an entire year! All to predict not just where a ship might be, but also when it might be at that position.

You might think that anything Albert Einstein toyed with must be some of the most bizarre and esoteric ideas out there. Well guess what, he struggled to describe time for many years of his life. You might know that he is most famous for publishing a theory called "relativity," which gave an entirely new and bizarre meaning to space and time. At first, It was very hard for people to believe, and took many years for the "rest of us" to even prove his theories orrect (thus far, Einstein's theories have never been shown to be wrong). One of his basic arguments began with the seemly simple task of having two people agree as to what time it is. You might think two people can simply look at their watches, then at each other, and agree that it is, say exactly 12 noon, when the sun is high in sky. It turns out that it isn't quite that simple (and this doesn't have anything to do with slightly different watch settings either). If one person is moving, then the two parties will not agree that it is

12 noon (this will be discussed further in chapter 4).

Finally, what about the job of "keeping time?" We discussed short times, and long times, but about about good old "human time?" Practical time, the kind we need to find our ship or get to class on time. How do we keep track of this? We all know that we can look at a clock to know what time it is. But what is a clock? Do you know how a clock works? What machinery, theory, or science goes into a machine that can keep accurate time? Now, add the reliability requirement that a clock must work 24 hours a day, 7 days a week, for years and years at a time, if it is to be useful to us.

4.3 Time Keeping with a Clock

This is a critical section in our understanding of how the longitude problem was solved. Why? Because we'll finally study how accurate time is kept. Not just estimating time as in counting "alligator-one...alligator-two...alligator-three" or even time the way an hour glass (falling grains of sand) would keep it. Here we'll look at methods of correct, proper, and accurate time keeping. As far as we're concerned here, time is kept with a **clock**, and a clock is an artificial, mechanical device made by human beings. It is a collection of parts, numerals, dials, wires, and other "stuff" that *we make*, that when put together and turned "on" will start and (hopefully) continue to keep accurate time.

With your thoughts perhaps stuck on the words in the past two chapters, imagine now the possibilities of a working clock for navigation. We could wait until the sun was directly overhead at high noon in Greenwich. We could set the dials on the clock to 12:00, and start sailing, say west. Two hours later, when we've lost site of land, our clock would read 2 pm, and clocks back in Greenwich would also read 2 pm. Suppose a few days go by and we get hit by a horrific storm would hit, and batter us around for the next 14 days. We wouldn't have a sun or Polaris sighting due to cloud cover. Suddenly the skies would clear. We'd shoot the sun at noon then read our clock. It reads 5:45 pm. Without being there, we know the clocks in Greenwich also read 5:45 pm. We'd know that we are 86.25° west of Greenwich. How nice! This is the gift that John Harrison gave to the world: accurate, portable, and convenient time keeping.

Now let's think a bit more about clocks. Of all demands put upon a human-made instrument, none have been more stringent that those placed upon a mechanical clock[Andrews p. 3]. A clock has to keep going 24 hours a day, 7 days a week, 365 days a year, for several months (if not years) without attention. This is a requirement that goes far beyond that which is expected of almost any other mechanical device. Furthermore, during its "on time," we expect the clock to keep good time, or it will be no use to us.

But just what is a reliable clock? We might think of 99% as being good enough for most things in life[Andrews, 1992]. But not a clock. A clock good 99% of the time would lose almost 15 minutes a day. You'd be quite late for things in your life, and for longitude, this would put you off by 225 nautical miles at the equator. Upping the ante to 99.99% would cause a clock to lose 8.84 seconds a day. If you set your wristwatch and it kept time like this, it would be off by almost 5 minutes after a month. Further, 99.99% would not qualify for the longitude prize offered nearly 300 years ago, which demanded a clock that kept time

to within 2 seconds a day, and this also meant 2 seconds a day out a sea, where at times conditions were so wretched that *even the rats would die.*

4.4 Nature's Clocks

So we make clocks. They are artificial inventions of our own making. But to build one, we need a starting point. We need some parts; something to build a clock with, and not just the box in which to hold it all. We need to design the "guts" of the clock. But what on earth could these parts be? What does it mean to keep time? Well, we'd like to produce a machine that somehow went "tick-tock-tick-tock" with the same length of silence between each tick and tock. This length of silence is sometimes called "latency" of a clock, and is the key "time keeping" mechanism. The actual sound made by the "tick" or the "tock" is typically made by two mechanisms inside of the clock coming into contact with one another. The contact point might be where another mechanism, that keeps track of the ticks and tocks was activated, such as the advancement of a clock hand, for instance.

The latency of the clock is where most of our design attention must go. This is the critical time keeping period of the clock. It will set the overall consistency and regularity of the clock. We must focus most of our attention here, to achieve the 99.99% that we are after. And remember, 99.99% isn't a terribly good clock; it would be a good first attempt at a longitude clock, but probably wouldn't win the longitude prize.

We must address the desired resolution of our clock, also critically set by the latency period. The smaller the silent period, the finer the resolution our clock can have. If the silence period was, say 1/10th of a second, then we could potentially build a clock that would keep time to one 10th of a second, probably pretty good for most purposes. But, we have to be very careful with our design, and work to ensure that the latency period always came out to be 1/10th of a second, to within 99.99%. So the silent period could be as long as 0.10001 seconds or as short as 0.09999 seconds, under any "normal" environment for the clock.

Lastly, we must think of powering our clock. How is going to keep the tick-tock-tick-tock mechanism going? Old mechanical clocks were wound up. Modern day clocks are plugged in, or run on batteries. And however we power our clock, we can't let it mess with our latency mechanism.

As a starting point, we can turn to a few items found in nature that can provide us with many natural time keeping mechanisms; something about which we can design our tick-tock-tick-tock mechansim. These mechanisms cannot be called clocks themselves, but are certainly starting points for some kind of accurate time reference, and are the topic of this section. Nearly all of them can be traced to exist in all of the clocks that Harrison built, or even all clocks that one can buy today, in modern times.

4.4.1 The Pendulum

Suppose you took a piece of string and tied one end to the hook in the ceiling and on the other end, you tied a heavy weight, like a rock, or a piece of metal. The hook in the ceiling

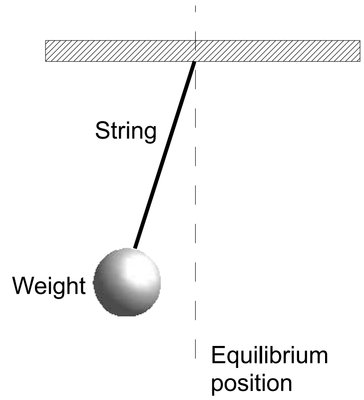


Figure 4.2: A simple pendulum, that can swing back and forth.

is commonly called the "pivot point" and the heavy weight the "bob." This bob, string, weight system is called a "pendulum," as shown in Figure 4.2. If you allowed the bob to sit, with the weight exactly below the pivot point, the system would be in in equilibrium. Nothing would be moving, or start to move. If you pulled the weight back, however, while keeping the string taught, and then let it go, the weight would begin to swing back towards its equilibrium position. Because it is now moving, it will not stop abruptly at its equilibrium point, but would begin to swing back and forth about it, in a repetitive motion called an "oscillation."

People have long known about pendulums and have studied them carefully. It is known that if you pull a pendulum bob back a small distance, it will return to your hand in a time, T, where T is "the period" of the pendulum, defined as

$$T = 2\pi\sqrt{\frac{L}{g}}. \tag{4.3}$$

T, the period, is defined as the time it takes the bob to swing from its release point, stop on the other side of equilibrium, then back to its release point again. This is "one period" of the pendulum. In the equation above, g is the acceleration due to gravity near the earth's surface, or 9.8 meters per second, per second, L is the length of the string, and 2π is just 2 times 3.1415... or about 6.283. So, as an example, suppose the length of our string was 2 meters (about 6 and a half feet). The period of the pendulum would be

$$T = 2\pi\sqrt{\frac{2\text{ m}}{9.8\text{ m/s}^2}} \tag{4.4}$$

or T would be about 2.84 seconds. This means if we pull the weight back and let it go, it would return to our hand again in about 2.84 seconds. Such an interval is the perfect time base for building a clock. In other words, the pendulum is a natural time keeper! Using this as our core mechanism, we might be able to design a clock with a latency period of 2.84

seconds. Dropping the length of the string to about 25 cm (10 inches) would give us a 1 second latency period; perfect for a clock that is designed around counting seconds.

Look carefully again at Equation 4.3 to see just what properties of a pendulum the period will depend on.

- The 2π is just a number (6.283...).
- The parameter g is a property of the earth. This is the acceleration of gravity, near the surface of the earth. For most purposes, it's always 9.8 m/s^2 or about 32 feet/s^2.
- What makes a pendulum so nice for a clock, is that its period seems to only depend on L, the length of the string you choose.
- Curiously missing from Equation 4.3 is the amount by which you initially pull the pendulum back. In other words, a 2 meter pendulum will always return back to its release point in 2.84 seconds, if you pull it back 2°, 10°, or even up to about 20°. There is only one precaution, in that the equation for T fails somewhat if you pull the pendulum back to far when starting it, but for pulls less than about 20° or so, the equation holds up just fine.
- Also missing from Equation 4.3 is the mass of the weight tied to the end of the pendulum. In other words, you could tie a small pebble or an elephant to the end of a pendulum, and it won't affect the period; the elephant and the pebble both make a pendulum with the same period.

There is an interesting historical note about pendulum clocks that dates back to 1656[Andrews, p. 35]. Christiaan Huygens worked on building pendulum clocks around this time and developed a theory to handle the period of a pendulum for large initial pulls. He even tried working on such a clock for solving the longitude problem, but failed.

So we have found a simple, and natural time-base, but there is some bad news as far as the longitude problem goes. Obviously there will be no stable pivot point for a pendulum on board a rocking ship, so a pendulum clock simply won't work on a ship (too bad).

Second, the earth's acceleration due to gravity varies some over the surface of the earth, and at different altitudes above the earth's surface. And, if g varies, so will the period of the clock. Here's what can cause g to vary.

- In general, g decreases with height above the earth, so a pendulum-based clock would run a bit faster at higher altitudes.
- Even at sea level g varies, largely due to the composition (density) of earth directly below the location of the pendulum. The variation can be one part in one million, but enough to vary the latency of the clock.
- Due to the shape of the earth (it's not a perfect sphere), g can be as small as 9.73 m/s^2 at the equator and as large as 9.83 m/s^2 near the poles (g drops to about 8.5 m/s^2 up where the space shuttle orbits).

Third, the pendulum has to swing through the air. Air resistance (and air pressure) not only provide a frictional force that tends to stop the swinging motion, but it can vary too. The "highs" and "lows" discussed during the nightly weather report are examples of variations in air pressure. Variations can cause the pendulum to swing erratically, and the air resistance will stop your 2 meter pendulum from swinging, typically in under 2 minutes.

Lastly, and most importantly, L will vary due to temperature. This is because all materials contract when cooled and expand when heated. This means a grandfather clock would run faster (smaller T) in the winter (or at night) when temperatures are cooler. In this case, L decreases a bit since all materials contract when cooled, and T depends on $\sqrt{L}$. The same clock would run slower (larger T) in the summer (or during the day), when the temperatures are higher. In this case L increases and T would increase as well, again since T depends on $\sqrt{L}$.

So a pendulum, even for use on a land-based grandfather clock, is full of problems. Most designers would quit, but Harrison did not, and here we point out, for the first time, the genius of his clock design, and his drive to solve problems as they are presented to him.

Activity

Pendulum lab. Create pendulums, time periods with a stopwatch. Estimate error in pendulum timing. Time a short walking trip with the watch, then with the pendulum. Investigate effects of mass, length, and initial amplitude on period.

Harrision's Gridiron Pendulum

As mentioned, all materials expand when heated and contract when cooled. This is all encompassed in a relationship that looks like this:

$$\Delta L = \alpha L \Delta T \tag{4.5}$$

What this means is as follows. If you have an object with a length L, such as a pendulum, its length will change by the amount ΔL if the object is subjected to a temperature ΔT.

Just a few words now about the Δ symbol. In science, Δ always means change, and is defined as $\Delta X = X_{end} - X_{start}$, where "end" and "start" refer to the starting and ending values of the quantity X. So if you had a pendulum at $T = 20$°C with a length of 2 m, it would have $T_{start} = 20$°C and $L_{start} = 2$ m. Suppose the temperature then increased to 25°C, and this caused the length of the pendulum to increase to 2.1 m. Your T_{end} would be 25°C, and L_{end} would be 2.1. All told then, ΔT would be 5°C and ΔL would be 0.1 m.

What about the α in Equation 4.5? This is called the "expansion coefficient" and is different for different materials an object might be made from. As some examples, lead has an $\alpha = 0.000029$ per degree C. Aluminum has $\alpha = 0.000023$ per degree C. Brass has $\alpha = 0.000019$/°C. Copper is 0.000012/°C and lastly, steel has $\alpha = 0.000011$/°C. So, if we know the original length of some material and the temperature change it has undergone, we can predict how its length will change. And, as we've seen with length change comes a

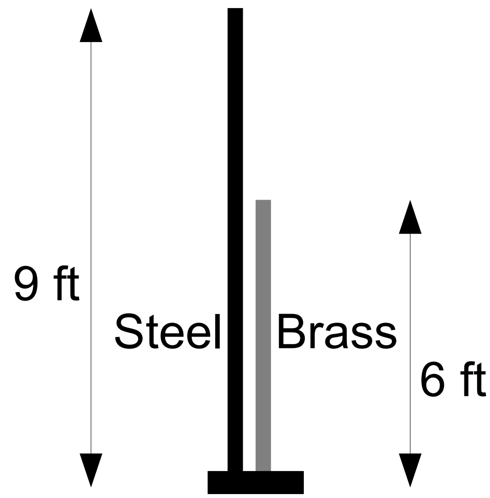


Figure 4.3: Harrision's gridiron pendulum. The spacing between the top tips of the two bars will remain constant regardless of temperature.

change in period of a pendulum. As an example, a 2 m copper rod would expand by 0.24 mm if the temperature rose by 10°C.

Harrison wanted to design a pendulum clock whose period was insensitive to temperature changes. This means a pendulum whose length was somehow immune to the natural expansions and contractions that all materials undergo when their temperature changes. As you might expect, large changes in temperature were expected on a ship at sea, even at a similar latitude, never mind traveling between say, the cold North Atlantic and the balmy equator.

Here's what Harrison did. He looked at the expansion coefficients and noticed that brass expands more than steel by about three to two (or more accurately about 100 to 62)[Gould, p. 41]. Further, Harrison knew from Equation 4.5, that an object's length change (ΔL) is proportional to its original length, L. So if you had two pieces of steel, one twice as long as the other, the longer one would have a ΔL that is twice the other for the same temperature change, ΔT. Harrison thought, why not combine brass and steel in such a ratio to cancel the effects of thermal expansion? His plan was as follows. Say you took a steel rod 9 feet long, and a brass rod 6 feet long, and laid them side by side. Next, you connected their lower ends. Their upper ends would be about 3 feet apart, as shown in Figure 4.3.

Now suppose that the temperature begins to rise and the materials naturally want to expand. As mentioned, brass expands faster than steel by about a factor of 3 to 2. But the expansion is also proportional to the length of each rod. Harrison cleverly made the length of brass 6 feet long, 2/3 of the length of the steel rod at 9 feet long. That is, the lengths were always chosen in the inverse ratio of the expansion coefficients. In Figure 4.3, as the temperature rises, the steel rod will begin to expand, lowering its lower end. But fixed to this lower end is the brass rod which is also expanding, its top part reaching ever higher. Due to each rod's length, the slower expanding steel is "quickened" due to its longer length,

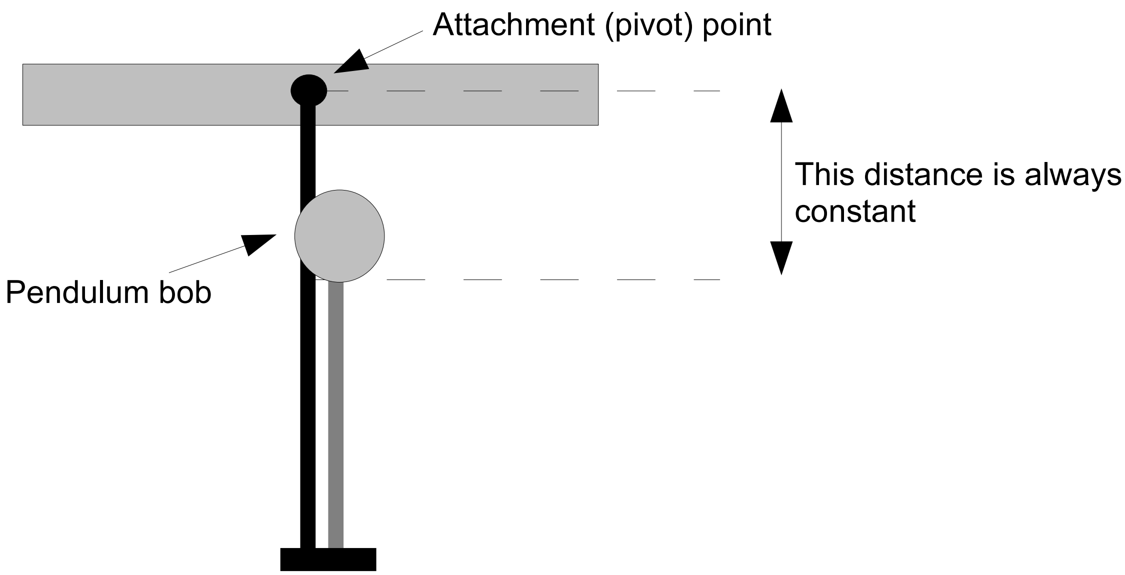


Figure 4.4: A pendulum whose length is insensitive to temperature.

and the faster expanding brass is "slowed" by its shorter length. Thus the distance between the top tips of each rod will forever be exactly the same, regardless of temperature. The (theoretical) completed pendulum, insensitive to temperature is shown in Figure 4.4, where a pendulum bob is attached to the upper end of the brass rod. It will always remain at the same distance from the top attachment (and pivot) point of the steel rod, making a pendulum with a constant length. (The figure is not drawn to scale. One can imagine very small diameter steel and brass rods, relative to a very heavy pendulum bob.)

In practice, the gridiron pendulum can be shortened and folded, using alternating brass and steel rods. A more practical illustration of a gridiron pendulum is shown in Figure 4.5. In this figure, steel is in black and brass is in gray. You should study this figure and convince yourself that the distance between the pendulum bob and the top pivot point will always be a constant. Lastly, if you ever happen to see a grandfather clock (even today's models), take a good look at it. The swinging part will most likely resemble Figure 4.5.

Lastly, a technological advance that would have made Harrison sweat with excitement. Invar is a Nickel-Steel alloy invented in 1896, 120 years after Harrison's death. It has a thermal expansion coefficient that can be as low as 0.00000062/°C, which is a whole factor of 100 smaller than any other metal. To compare with the copper rod discussed above, a 2 m Invar rod would expand by 0.0024 mm if the temperature rose by 10°C.

4.4.2 Mass on a Spring

Suppose you took some mass (a weight), kind of like what you tied onto the end of the pendulum in the last section, and instead attached it to the end of a spring. Then, you attached the other end of the spring firmly to the ceiling. This is shown in Figure 4.6.

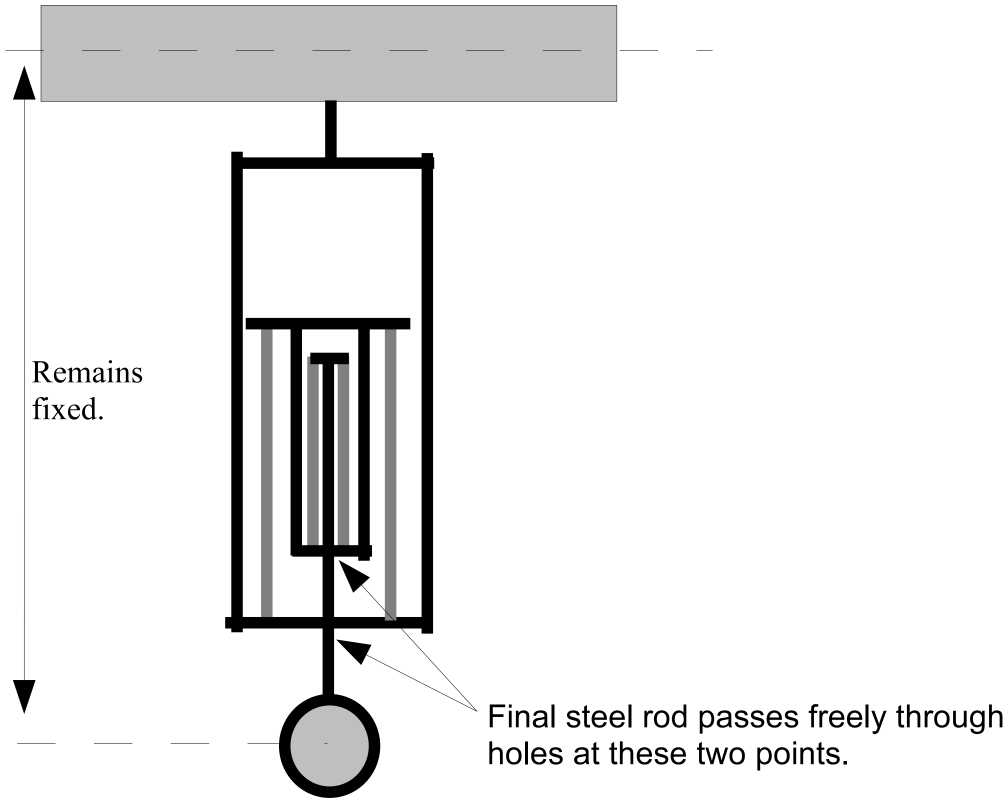


Figure 4.5: A practical gridiron pendulum. Steel is in black and brass is in gray.

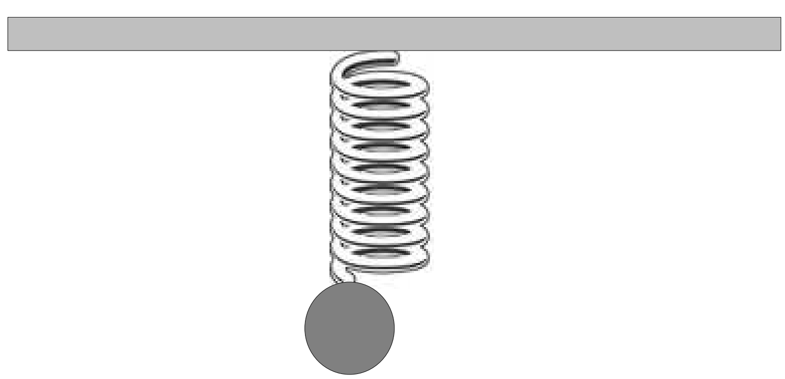


Figure 4.6: A natural time keeper: the mass/spring system.

If you let the system hang at rest, the position of the mass will be the equilibrium position of this system (much like the pendulum's equilibrium position).

Suppose now that you pulled the mass down by a small amount. This will stretch the spring out a bit. Then, you let go of the mass. What will happen? The spring will pull it back up, and get it moving in the upward direction. Because it is now moving up, the mass will pass right through its equilibrium point (just as the pendulum did) and keep going up until it starts to compress the spring to a point where the spring is able to stop it, then get pushed down again. The mass will keep going down, once again past its equilibrium point, until it begins to stretch the spring out again, and eventually stops, almost exactly at the point where it was originally released. The cycle, or the mass-spring oscillation, will now repeat itself.

Like the pendulum, this up-down motion is the "period" of the mass-spring system, and also like the pendulum there is an equation that tells you the period of the mass-spring system. The period is defined as how long it will take the mass to go up from its release point, stop, and return to its original release point. The equation for the period is

$$T = 2\pi\sqrt{\frac{m}{k}}. \tag{4.6}$$

Great! Another system with a natural period. The mass-spring system is another natural time keeper!

In this equation, 2π is 6.283 and m is the mass (in kilograms) of the weight attached to the free end of the spring. So *unlike* the pendulum the period *does* depend on how heavy an object you attach to the spring. A small pebble and an elephant will certainly cause the mass-spring system to have a different period. Larger masses will have larger periods and vice-versa.

The variable k is a parameter that takes the "stiffness" of the spring into account. Some springs are very stiff, like the coils attached to each wheel on a car to make the ride smoother, or the big ones that make opening a garage door easier. Other springs are very easily bent, like a slinky, which can barely hold its own shape, or the springs that pop the toast out of a toaster, that you can easily push down with your hands[Bloomfield, p. 62] . Stiff springs can have a very large number for k and will have a shorter period, since they want to "snap back" so quickly. Something like a slinky will have a very small k and take "forever" to pull the mass back and forth and have a larger T. This parameter isn't really that important to us and the longitude problem; just keep in mind what it means and what it is for.

As for some typical numbers, suppose you had a spring with $k = 10$ N/m (Newtons per meter) and you tied a 2 kg mass to the end of it. It will have a period of

$$T = 2\pi\sqrt{\frac{2}{10}} = 2.8 \text{ seconds}. \tag{4.7}$$

So when you pull the 2 kg mass down, it will go up, turn around, and return to the original release point in about 2.8 seconds. There are two areas of concern, however.

- The first is is that the overall maximum stretching or compressing of the spring must stay small. You don't want the mass to oscillate wildly on the end of the spring and

you don't want to pull on the spring so hard that you start to permanently deform the spiral structure of the spring. In other words you don't want to compromise the spring's natural elasticity (ability to stretch and compress).

- The second is that the top attachment point must stay fixed. Too bad again for using a mass-spring on a ship. If the attachment point started to move up and down, while the mass was trying to move up and down, you'd get a "feedback" or "coupling" between the motion of the attachment point and the motion of the mass itself, causing chaos in the otherwise uniform motion of the mass that you'd like to maintain.

So the simple mass-spring arrangement shown in Figure 4.6 won't hold up on a ship, but the spring is very valuable item in many clock-building ideas. In fact, Harrison used two on his first sea-clock, called H1.

Activity

Spring lab. Time periods with stopwatch and estimate errors in spring. Time a short walking excursion with watch, then with spring. Investigate effects of stiffness and mass on period.

Springs in Harrison's First Sea-Clock

John Harrison started his first attempt at a sea-clock (and the longitude problem) in 1729. The clock is called "H1" and is on display at the National Maritime Museum in Greenwich, England. This section contains a discussion on some of the inner workings of H1.

Harrison well knew the natural time keeping abilities of pendulums and mass-spring systems, and knew the simple designs presented above would not, by themselves, work aboard a rocking ship. So he would start, at a very fundamental level, with a design that would be immune to the rocking motion of ship. Here's what he did.

Suppose you had two identical brass balls, that had exactly the same mass and shape, that were mounted on the ends of a metallic brass bar. The bar was, in turn, mounted on a pivot, exactly at its center, as shown in Figure 4.7.

The pivot is a fixed point, and the rod can rotate freely about the pivot. Collectively, we'll call the balls, rod, and centered pivot the "balance." As shown in Figure 4.7, the balance is stationary (not rotating). Now, see if you can answer this question: Suppose you gently grabbed the top (or bottom) ball between your two fingers, and slowly rotated the entire balance about its pivot 30° clockwise. So you reconfigured the system to not be exactly balanced vertically, as shown in Figure 4.7. This rotation is shown in the right of Figure 4.8.

As you approached 30°, suppose that you gently let go of the ball, without giving it a push or nudge one way or the other. What would happen to the balance? Since the balance is now "unbalanced" would gravity pull the top ball down, causing the balance to rotate clockwise? Or would the entire system remain stationary after you let go? The correct answer is that the balance would remain stationary. It would not begin to rotate just because it is not oriented vertically anymore. Why?

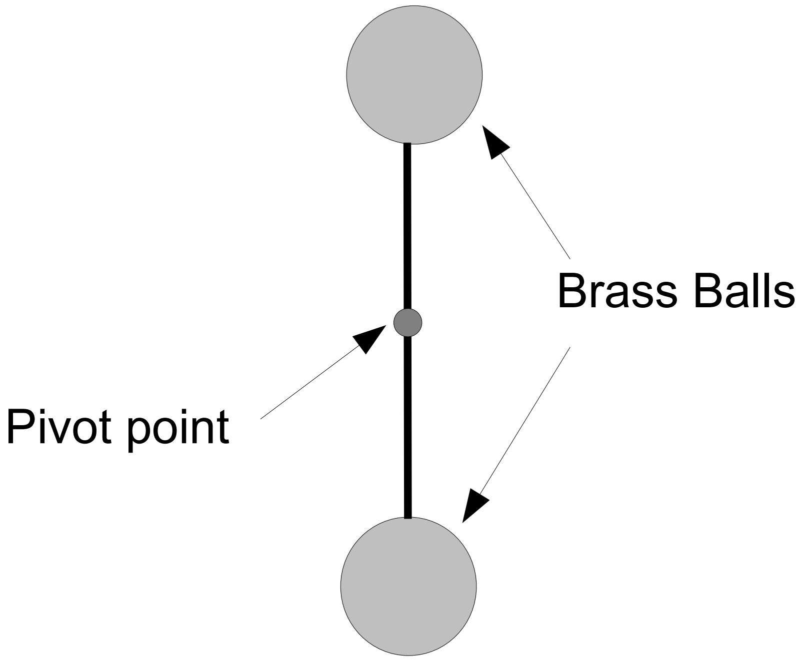


Figure 4.7: Harrison's balance that ultimately would be immune to the rocking motions of a ship.

If you consider carefully what the system consists of you should be able to convince yourself that the system will remain stationary when set in any angular position. The pivot point is exactly centered between two objects with identical weights. If one ball rotates toward the right, the other rotates toward the left. If one is pushed down, the other moves up, etc. In more of a "physics sense" the only thing really happening is that gravity always pulls both balls straight down, as shown in Figure 4.9.

Each ball affects the overall motion of the balance in the following manner. If ball A "had its way" it would cause the balance to rotate clockwise. Ball B would cause the balance to rotate counterclockwise. But each ball "having its way" is called the "torque" a ball exerts on the balance. Torques are forces that cause things to rotate. So the balance is really a tug-of-war between the two balls. Ball A would like the balance to rotate clockwise, and ball B would like the ball to rotate counterclockwise. Which will win?

The torque a given ball can create is related to three things. The first is the distance from the pivot, and the second is its weight. As we've said, both balls are exactly the same distance from the pivot, so no neither wins there. The second is the weight of the ball. Again, the balls are exactly the same, so neither wins there either. The last item is the angle gravity makes with respect to the brass bar connecting the two masses. Since the system is perfectly symmetric, this angle is the same for both balls.

The conclusion to all of this is that neither ball can "have its way." No ball can individually influence which direction the balance will rotate, so the balance always stays stationary, no matter what angle at which it is set. So in Figure 4.10, the balance will not begin to

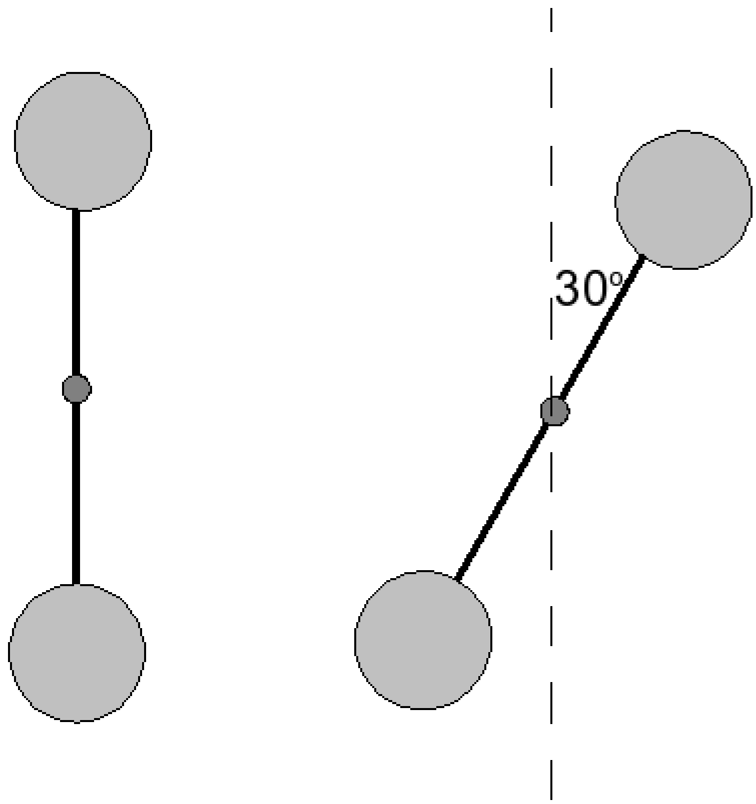


Figure 4.8: The balance rotated 30° clockwise (right figure). If let go, would it rotate or stay stationary?

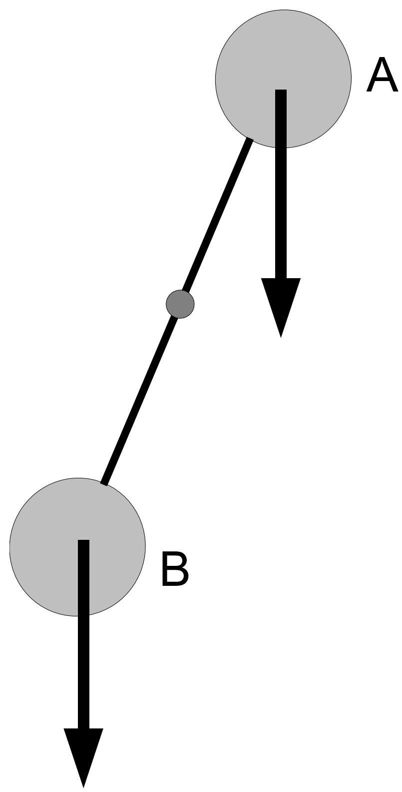


Figure 4.9: Gravity pulls both balls straight down, as shown by the heavy arrows.

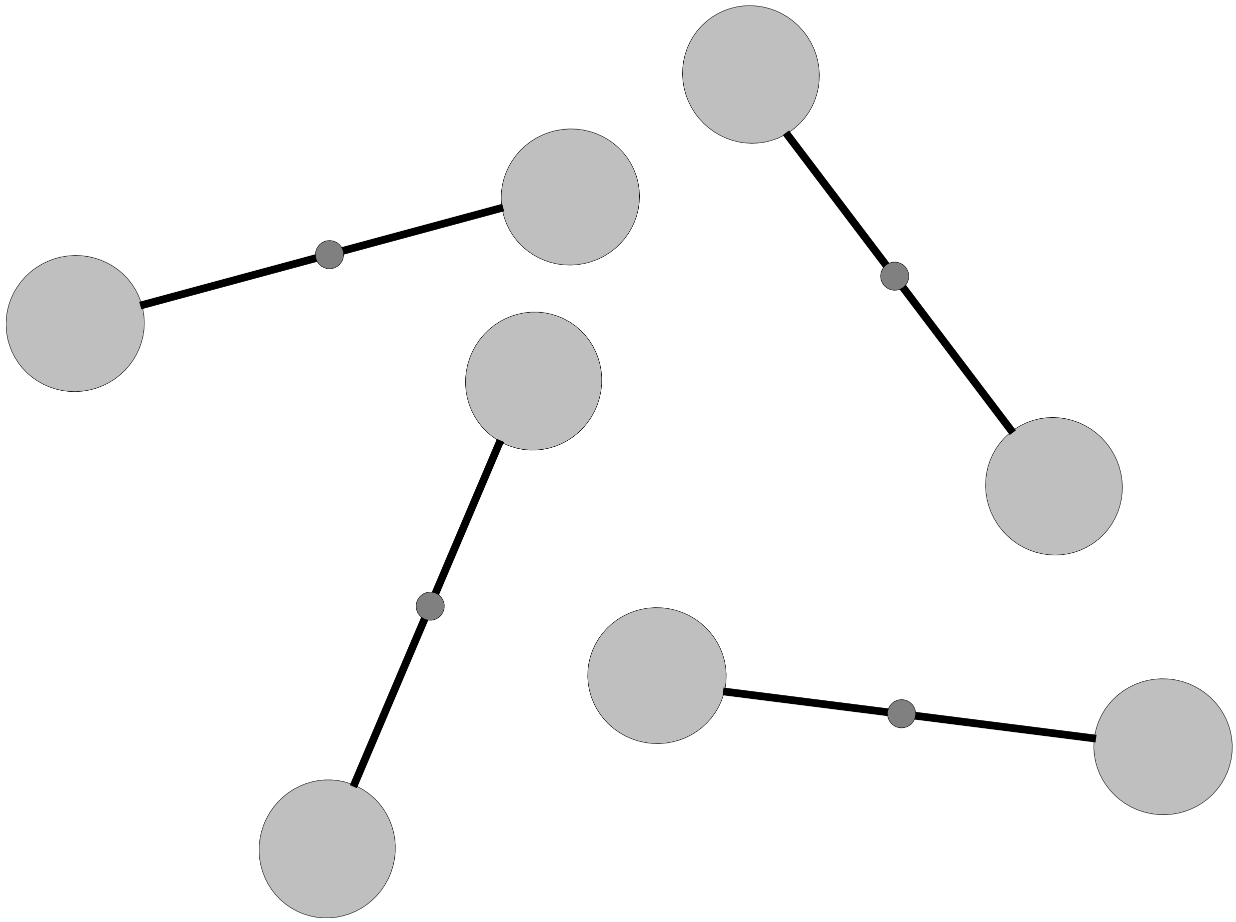


Figure 4.10: The balance will not rotate, for any orientation.

rotate if left in any of the configurations shown (or any configuration not shown for that matter!). Physically, each balance is said to have "no net torque." This means there is no resultant or net torque that would cause the balance to rotate on its own.

If you've ever taken a physics course, you probably studied a system called the "Atwood Machine." Harrison's balance is essentially the circular version of an Atwood Machine. But this "propeller" motion is not needed for any clock applications. You can forget about it as far as the longitude prize goes.

Hopefully you are beginning to see why Harrison started with such a balance. On a rocking ship, the pivot point of the balance, which is connected to the clock, which is connected to the ship (sitting on a table, or the floor, etc.) will get tossed around with the ship. Up down, left, right, up, and back, or any simultaneous combination of these directions. But these motions will not affect the orientation of the balance, because the balls can only move due to the torques they themselves impinge on the system (which are zero). Any external up or down jerk due to the ship is equally suffered by all elements of the balance. It will not upset the balance of torques of identical balls at identical distances

with identical angles you'll always find in this system. If the balance is set at 32.835° during 16 foot swells in a hurricane aboard a ship in a hurricane. It will remain at 32.835°.

Let us now look at the core time keeper Harrison designed into his first sea-clock. He didn't just use a pendulum or a spring, he used them both as shown in Figure 4.11. He constructed two torque free balances near one another. The balances were connected near their ends with springs.

Figure 4.11(A) is how the time keeping mechanism would look at rest. To start the clock, the top two balls are pulled apart, as shown in Figure 4.11(B). This compresses the bottom spring and stretches the top spring. When released, the stretched top spring will pull the top balls together, and the compressed bottom spring will push the bottom balls apart. The pair of balances move to the configuration shown in Figure 4.11(C), and then back again. In other words, the two balances continually oscillate, as mirror images of each other.

Since the balances are each torque free, they do not have any sort of natural period of their own, and their period comes solely from the connecting springs. A detailed analysis of two springs in parallel like this means one period of the system will be

$$T = 2\pi\sqrt{\frac{m}{2k}}, \tag{4.8}$$

which looks a lot like the period for a single mass-spring system, except for the 2 in front of the k. Each spring is just like the simple mass-spring system described above, except that the attachment point for each spring is not a fixed point, but on the *other balance*. This is not problematic for a spring, and it will keep oscillating; springs can contract and expand and "do" their regular oscillation in any orientation (gravity doesn't affect their oscillation).

To to visualize the time keeping mechanism in Harrison's first clock, you must imagine a spring with masses (the balls) attached to both ends. The masses oscillating back and forth, in unison, on both sides. They travel in mirror image motions with respect to one another. They both travel away from each other expanding the spring, then back toward each other, compressing the spring. The time it takes them to make such a round trip is given by the equation for T above. The balances then are nothing more than mechanisms to hold the balls as they oscillate, as dictated by the springs.

The motion of the two balls, as insisted on by the expansion and compression of the spring, is called the "normal mode" of the spring. A normal mode of an oscillating system (such as the masses and springs here) is the pattern of motion in which all parts of the system move, sinusoidally with the same frequency. Frequency is just the mathematical reciprocal of the period, T. If the system has an oscillation period of 2 seconds, it oscillates back and forth at a frequency of 0.5 Hertz (Hertz = "per second". The frequencies of the normal modes of a system are known as its "natural frequencies" or "resonant frequencies.")

Lastly, we mention a possible problem with Harrison's balances. The literature is vague on this[see Andrewes, p. 211, bottom paragraph], but during H1's trial to Lisbon, Harrison's clock seemed to be keeping very good time for most of the voyage. At some point however, the clock abruptly was off by around 2 minutes or so. Harrison did not know what happened, but here's what the problem could most likely be traced to. The problem is not easy to grasp without some basic knowledge of physics, but we'll try to describe it here.

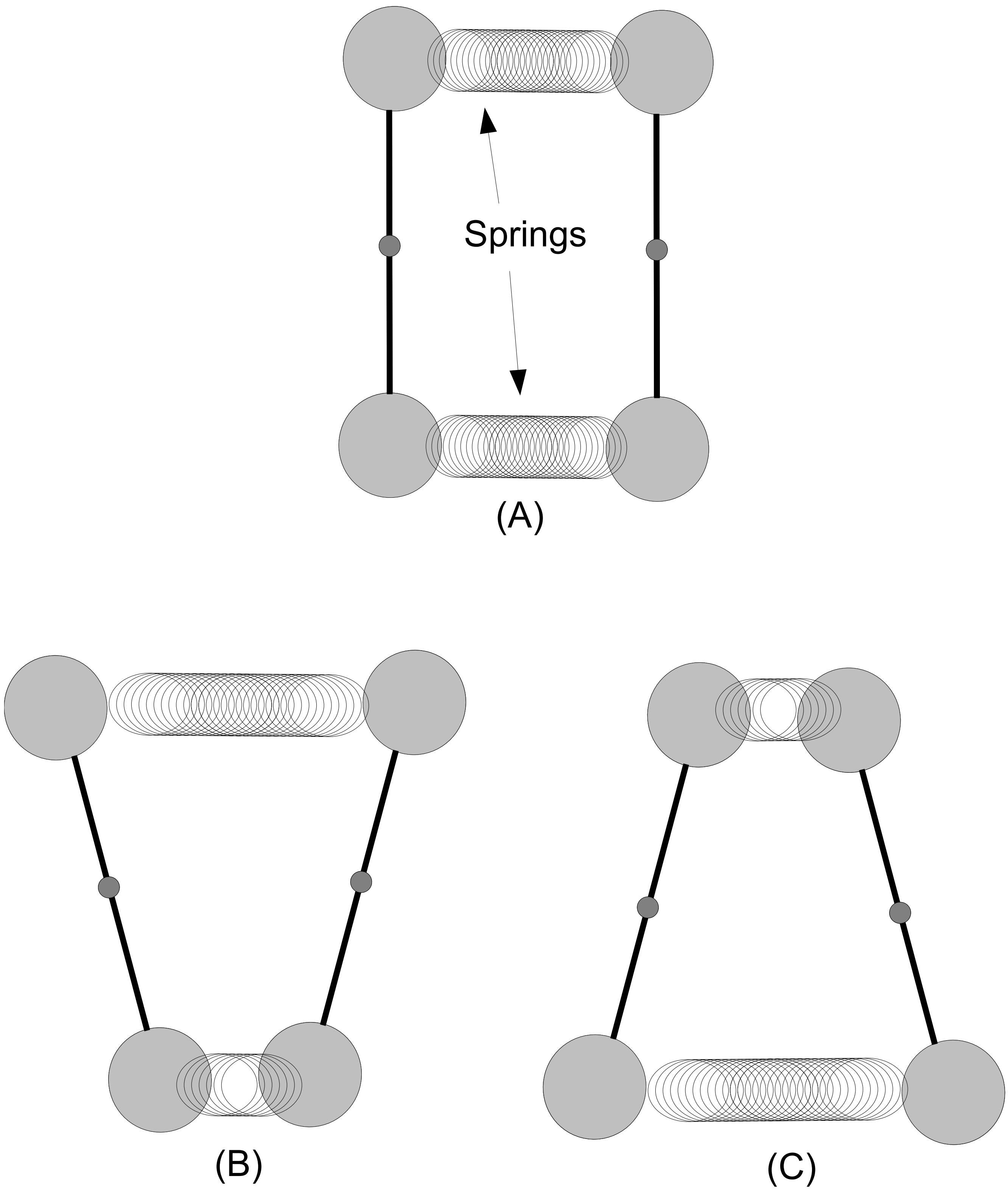


Figure 4.11: The core time keeping mechanism Harrison used in his first sea clock.

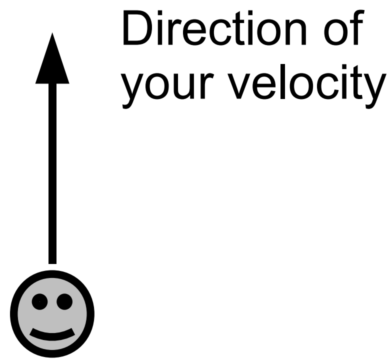


Figure 4.12: The arrow is your velocity. You are going due north.

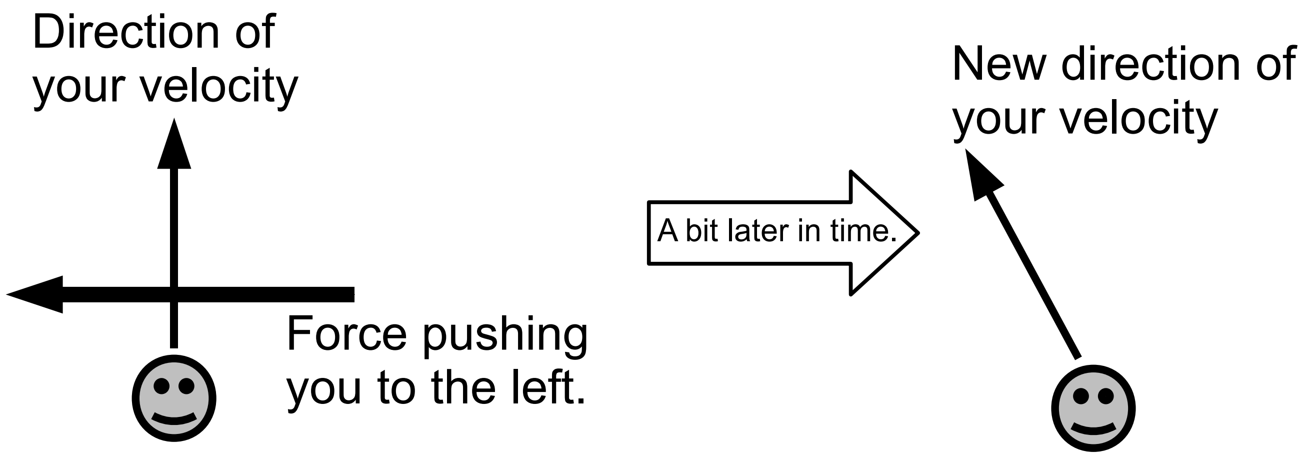


Figure 4.13: A force that will cause you to turn left.

Imagine that you are on a ship traveling due north, as shown in Figure 4.12. The arrow is meant to show the direction of your velocity. (Arrows in physics are called "vectors," so we'll use this word from now on.)

If no other forces (pushes or pulls) touched you, you would continue moving north at the same speed, forever. This is one of Isaac Newton's famous laws stating "a body in motion persists in that state of motion unless acted on by an external force." In other words, you would persist in your northward direction unless something pushed you in another direction, like the wind, or the boat's rudder, etc. So to turn left for example, a force would have to push you to the left, as shown in Figure 4.13.

In the left portion of Figure 4.13, there is a large leftward force pushing you to the left. The force could be the wind, the rudder, or a tug-boat. If you let a little time go by, to let the force "do it's job," your velocity vector will look like it does in the right part of Figure 4.13. What has changed? Your velocity vector is now pointing to the left a little bit. The force has turned you to the left!

So what does this have to do with Harrison's clock? Well, suppose your ship got into a long persistent turn to the left as shown in Figure 4.14. Turns are really just sections of circles, so your turn would start as shown in Figure 4.14A.

You are going north with some velocity, v, and there's the force, F pushing you toward the left. In order to make the turn, the force must always reorient itself toward the center of your turn, as shown in Figure 4.14B. Notice that only if the force points toward the center of the circle can it always nudge your velocity vector leftward, allowing you to stay in the turn. Such a force is called a "centripetal force" (centripetal means "center seeking") and is always responsible for keeping any object in a circle.

The strength of the centripetal force required to keep you in the turn depends on how "tight" of a turn you are trying to make. A quick, sharp turn requires a larger inward force. A slow, methodical turn does not require such a large force. And this is the potential problem with Harrison's balanced-based clock.

Suppose H1 was sitting on a ship as shown in Figure 4.15 and the ship was making a turn to the left. Notice how the left balance is more left than the right balance. Or in other words, the left balance is technically in a tighter turn than the right balance. This means the left balance will experience a larger centripetal force than the right balance, and this will throw off the time keeping mechanism of the clock. This is because the left balance will experience an overall greater force (spring + centripetal) than the right balance.

The effect might not be much, and maybe it was just bad luck that Harrison's first clock was affected by this strange combination of forces. In is not unconceivable that at some point during the voyage the ship engaged in a long sustained turn, or even a tight swift turn for a prolonged period. In either case, this is a problem with balance clocks "on the move" and would have been discovered sooner or later as a fundamental problem with using the balances in a sea-clock.

4.5 Core Elements of Mechanical Clocks

In Section 4.3, we began with the idea that we would finally look into a clock and see how one works. We took a look at Harrison's clock and some of the clocks that nature provides for us. All of this discussion was very much *mechanical*. Springs, gears, pendulums, and the like. Things that move in precise ways that keep good time. Let us close this section with a final, more general look, at just what is inside of a mechanical clock. These were the first invented (1600s) that had any precision to speak of, and the ideas used in their construction are still used today. Harrison was really a master of design. He took mechanical clocks to the extreme, making the best there were, and ushered in a new era of clockmaking.

4.5.1 Sustained Energy

By a "mechanical clock," we mean a device that keeps time without using electricity. In our times, this means no batteries or anything that plugs into the wall. Just what then, would keep a clock running? In Harrison's time, only two things existed that would cause a clock to run: a falling weight, and compressed spring.

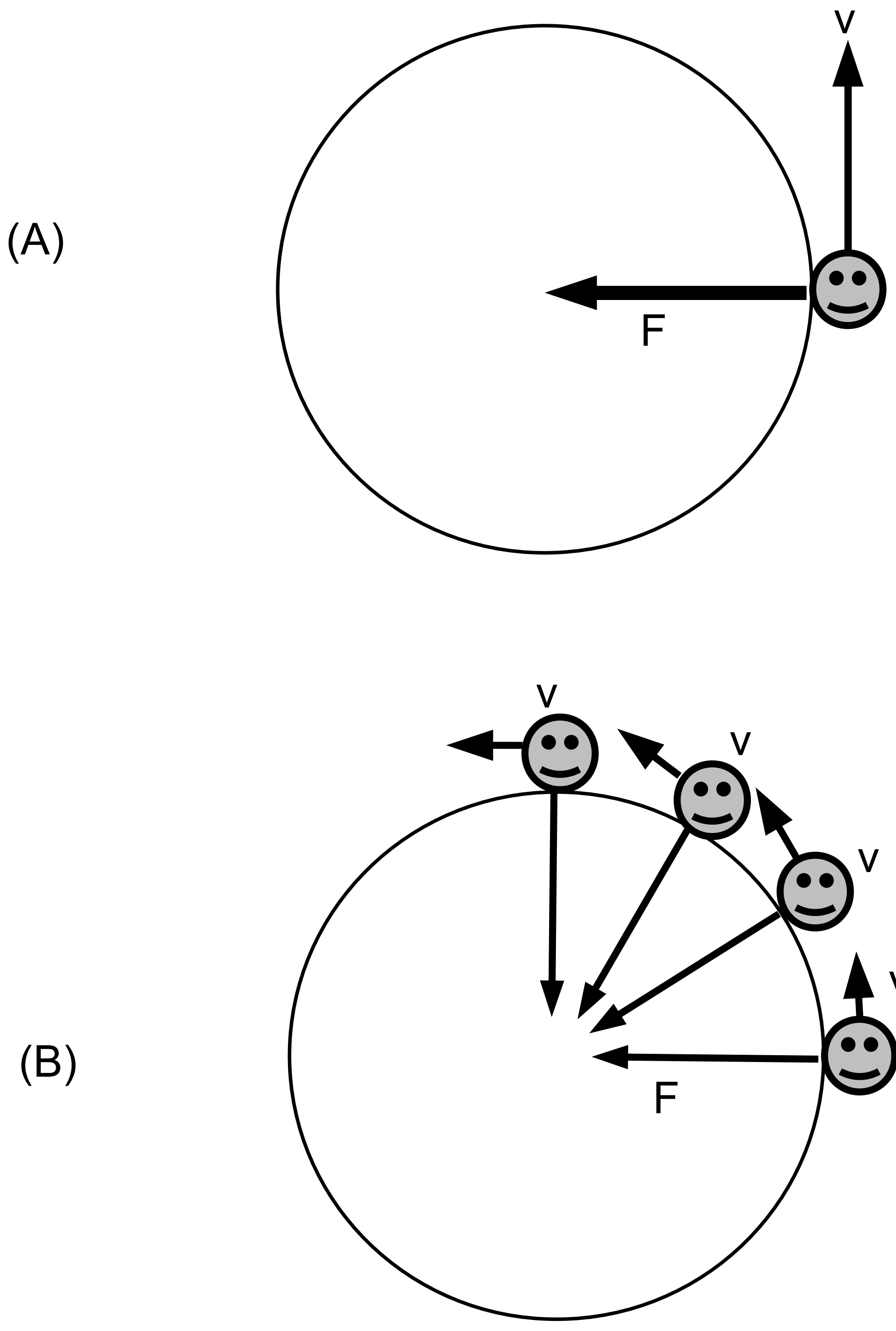


Figure 4.14: Turns are all just sections of a circle. Here the force (F) that will cause you to turn left by nudging your velocity vector (v) to the left.

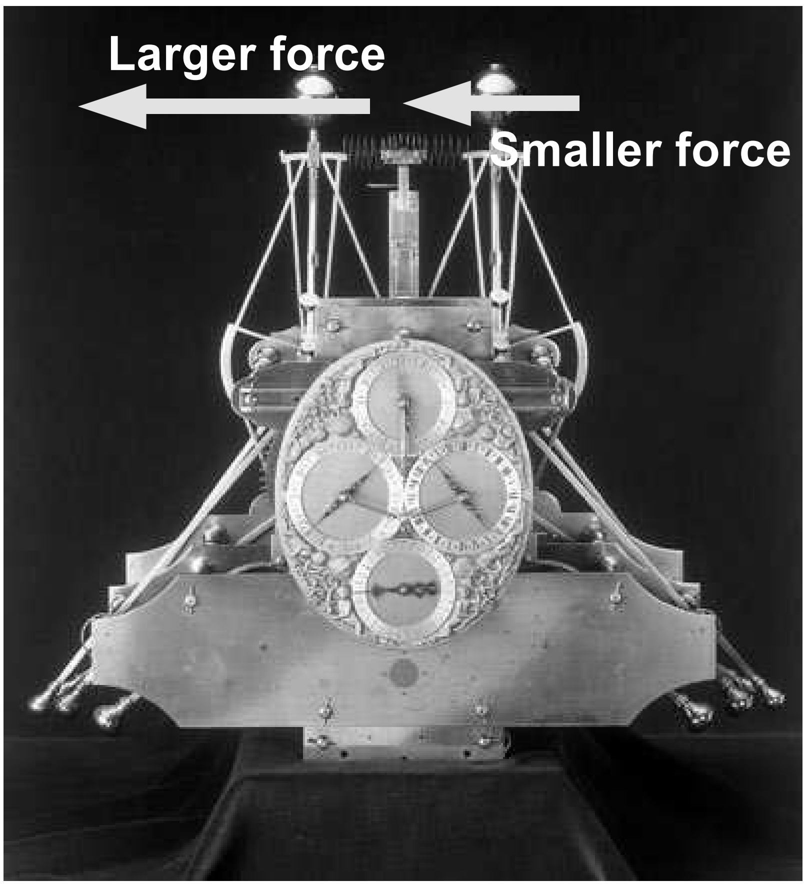


Figure 4.15: As H1 sits on a ship, differing strengths of the centripetal force could create unequal forces on the balances, causing the clock to gain or lose time.

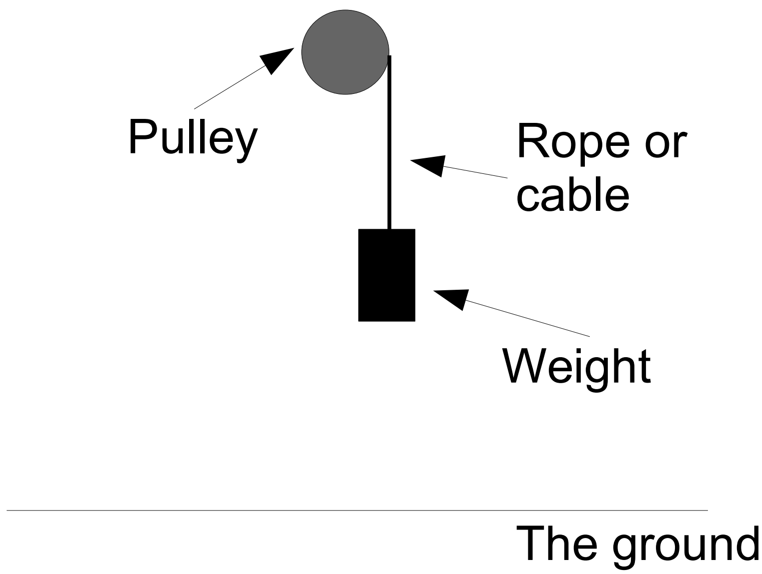


Figure 4.16: The start of a gravity-fed clock: a pulley attached to a heavy weight with a rope or cable. As the weight falls, the pulley turns.

A falling weight is what you might expect. Something heavy that is always pulled toward the ground by the earth's gravity. A clock that works with a falling weight might very well be called a "gravity clock." Figure 4.16 shows the simplest start toward constructing a gravity clock: attaching a weight to a pulley.

As you might guess, the pulley (or clock) is "wound" when the weight is as high as it will go, almost touching the pulley. This is stored or potential energy in its purest form. Only when the weight is released, does it start to fall, turning the pulley (via the rope wrapped around it) while it speeds up as it falls. Obviously the energy is gone (and the clock stops) when the weight hits the ground. So where's the clock? Well the turning pulley's rotational rate (a frequency) can be accurately predicted by elementary physics, but this is the not the nature of the clock. The falling weight is merely an energy storage device; it can be used to run a clock as it falls. Many early mechanical clocks were driven by the use of falling weights, and you can often see them in today's classic grandfather clocks. In fact, giving a weight some room to fall is one reason why grandfather clocks are so tall. We'll hold off on the "where's the clock" answer for just a bit. First, let's mention another energy storage device: the spring.

We encountered springs earlier in this chapter when discussing a natural oscillator; something that moves back and forth with a predictable, and stable frequency. But think for a moment about a compressed spring in the context of energy storage, as shown in Figure 4.17. In this figure, the hand is pushing on the block, which pushes on the spring, compressing it. What would happen if the hand is quickly pulled away? The spring would rapidly expand again, shooting the block up into the air. Rapid expansion? Shooting blocks? This all requires energy. Yes, it does, and it all came from the compressed spring. The lesson here is

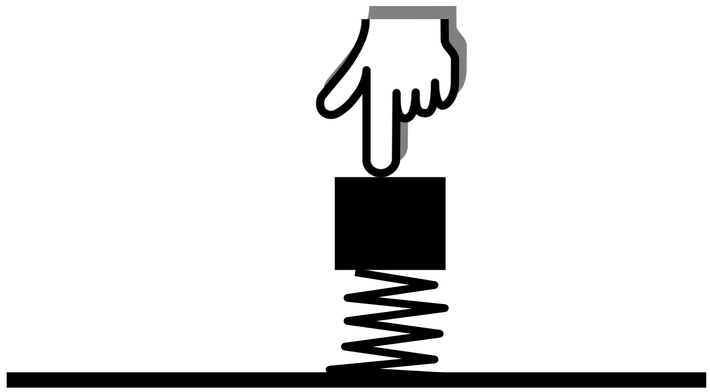


Figure 4.17: The spring, with the block on top of it, is being compressed by the push from above. The compressed spring is storing energy.

that compressed springs store energy, which can be released, by letting the springs expand again. The spring's energy comes from deforming the "wirey" structure of the spring, from its natural manufactured length; the falling block's energy comes from gravity. Either way, both devices store energy.

For springs, the conventional coiled spring, as shown in Figure 4.17 isn't the only game in town. Anything that can be compressed or squeezed, that also decompress or re-expand when released has the same energy storing capabilities as a spring (a sponge, a rubber band, a tennis ball, elastic cord, etc.). Imagine what would happen if you took a drinking draw and wrapped it around your finger. As long as you hold it wrapped the straw stores energy. You know this because as soon as you let go, it will unwrap, flying off of your finger as it returns to a straightened configuration again. It stored energy while it was wrapped, and released it when you let go. It is a sort of "circular spring."

The straw was mentioned because when springs are used to give energy to a clock, it is called a "mainspring," and is usually a circular coil, like the straw around your finger. "Winding" the clock is a process whereby the circular spring becomes more and more tightly wound into a circle, storing as much energy as possible. An example of a mainspring is shown in Figure 4.18.

This spring is made from a thin, flexible metal strip, which is wound like the straw around your finger. One end is attached to the axis of a gear, while the other end is attached to the outer perimeter of the gear. As the mainspring expands, its fixture on the end of the gear causes the gear to turn. So just like the falling weight turns a pulley, now we see an expanding, circular, mainspring can cause a gear to turn.

4.5.2 Escapements: Controlling the Energy Release

If left alone, the weight in Figure 4.16 would fall to the ground and it would all end. Similarly, the mainspring in Figure 4.18 would quickly unwind, ending the gear's motion. The spring and weight are merely energy storage mechanisms; something to keep the clock running.

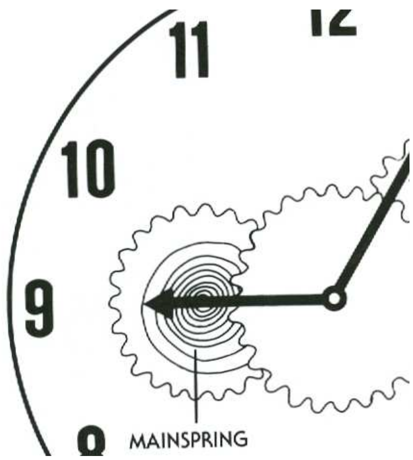


Figure 4.18: A coiled mainspring provides energy to the gears of a clock[From Brackin].

The next key step in time keeping is to *control* the release of the energy stored by these systems, so it all doesn't get released at once. This is done with an ingenuous device called an "escapement," a word undoubtedly derived from the word "escape" as in allowing energy to escape. Harrison did not invent the escapement, but he took them to previously unseen levels of perfection and sophistication. The parts needed to add an an escapement to a clock are illustrated in Figure 4.19.

First of all, you'll see in Figure 4.19(a) that the perimeter of the otherwise smooth pulley has been outfitted with teeth, making it into a gear. So when the weight falls, the gear turns. Second, in Figure 4.19(b), a pendulum has been attached to a rigid, "upside-down-V-shaped"' construct. The pivot point is right at the apex of the upside-down-V (which is called a "pallet" or "anchor.") A few overlapped oscillations of the pendulum mechanism are shown in inset (c).

Now, suppose that the pendulum and gear-pulley are combined as shown in Figure 4.20. Can you visualize what would happen if the mass was allowed to fall and the pendulum allowed to oscillate? A snapshot of the oscillation at three different times is shown in Figure 4.21.

Let's start at part (a) of Figure 4.21. The pendulum was swinging to its leftmost extreme position, when something interesting happened: one of it pallets hit and meshed into the groove between two teeth on the gear-pulley (called the escape wheel). This is shown at the position of the little arrow in part (a). This stops the pendulum and and escape wheel and if the escape wheel is stopped, the falling mass stops too. Next, as pendulums do, it begins to swing toward the right. As it does so, the pallet and escape wheel's tooth separate from each other and the pulley begins to turn freely again as the weight begins falling again. So in part (b) of this Figure, the escape is freely spinning and the weight is freely falling. But not for long! Within one-half of the pendulum's period, the other pallet meshes into a gear tooth stopping the wheel and weight once more as seen in part (c). So the escapement has done it! The energy release of the falling weight has been controlled with a precise period:

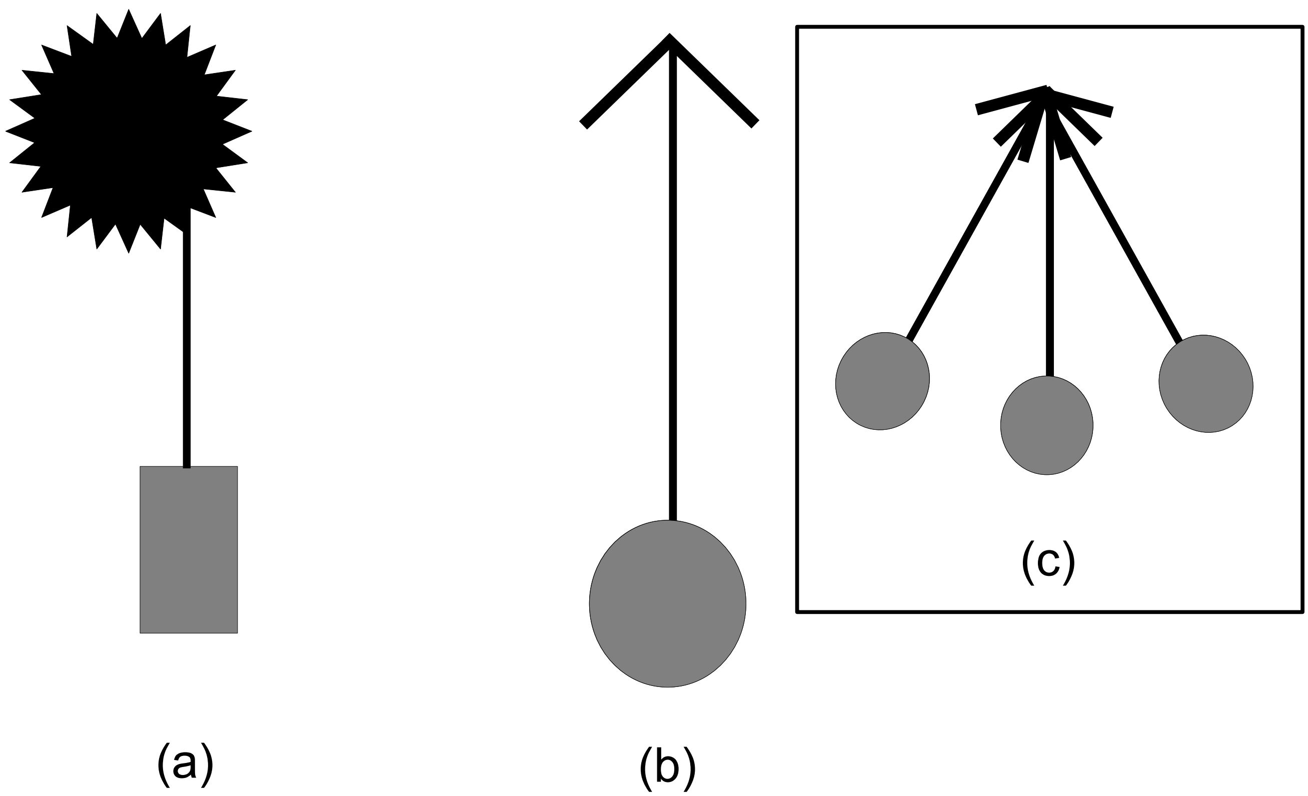


Figure 4.19: Parts needed to add an escapement to a clock.

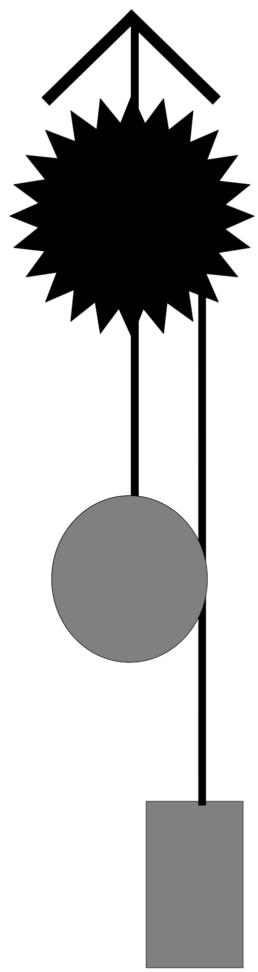


Figure 4.20: A simple escapement assembly.

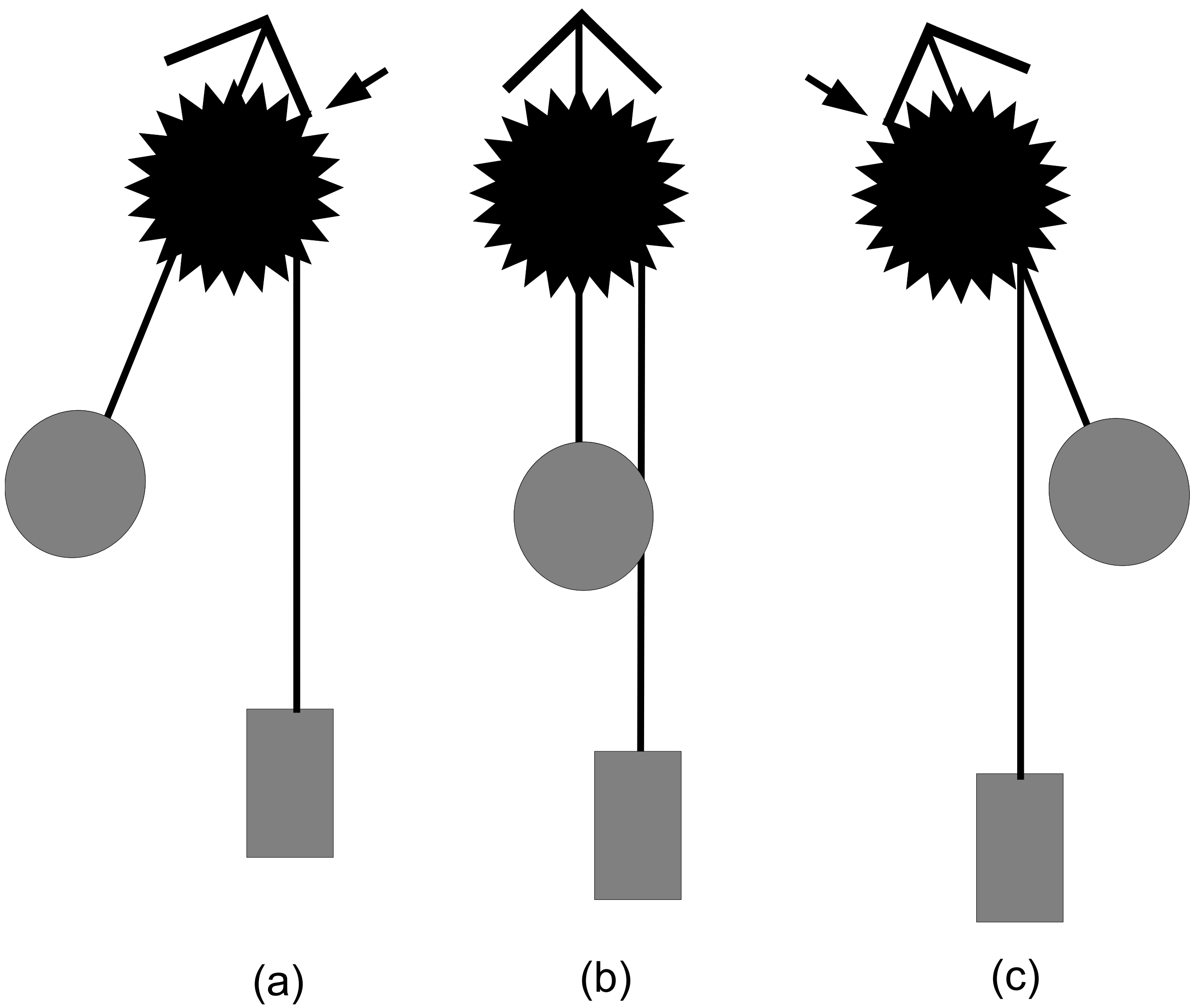


Figure 4.21: Three snapshots of a simple escapement assembly at work.

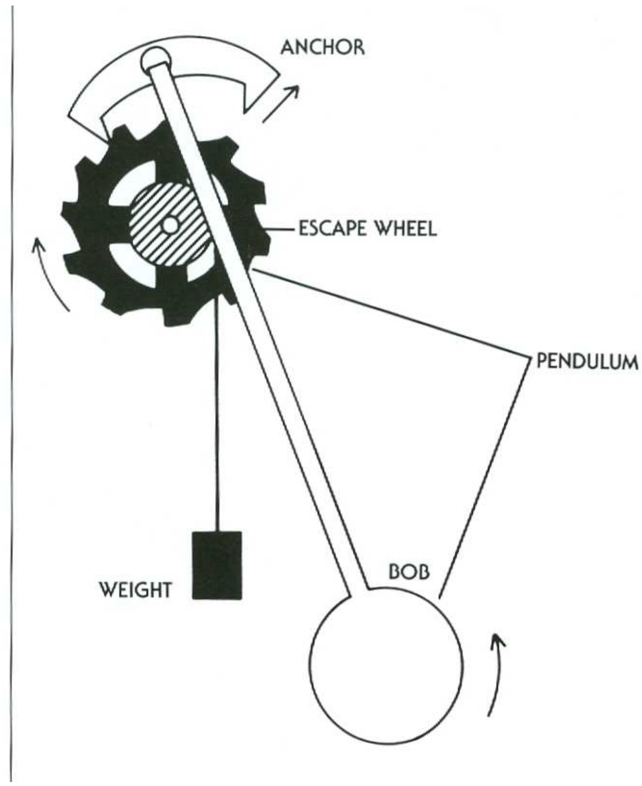


Figure 4.22: Note the shape of the teeth on the escape wheel. From Brackin p. 34.

that of the pendulum itself. The wheel is only allowed to spin only for brief intervals of time: when the pallets are not in contact with any gear teeth.

As you might guess, the escape wheel is meshed with other gears, in the right ratio to drive a second and minute hand around a clock face. In a nutshell, this is how a mechanical, gravity driven, clock works. Lastly, the hits between the pallet and gear teeth cause the familiar "ticking" sound of a clock.

There is an interesting oddity about the design of escape wheels. The gear teeth are *never* even and symmetric as shown in Figures 4.19, 4.20 and 4.21. They always have a antisymmetric shape. An example is shown in Figure 4.22, and another from Harrison's H1 is shown in the top of Figure 4.23.

Owing to friction and air resistance, all pendulums will eventually come to a stop. This would cause the escapement mechanism, and periodic energy release, to cease in operation. Although it is very difficult to describe, the antisymmetric teeth, on one end (or the other) of the pendulum's swing, cause the escape wheel to give the pendulum a little nudge, just as the pallet lifts away from the escape wheel's teeth. This nudge directs a small amount of the stored energy back into the pendulum, to keep it swinging.

There is one last escapement mechanism worth mentioning: the "verge and foliet escapement." This is shown in Figure 4.24. It works like this. As the escape wheel turns (notice the hugely antisymmetric teeth), the pallets connected to the top and bottom of the verge

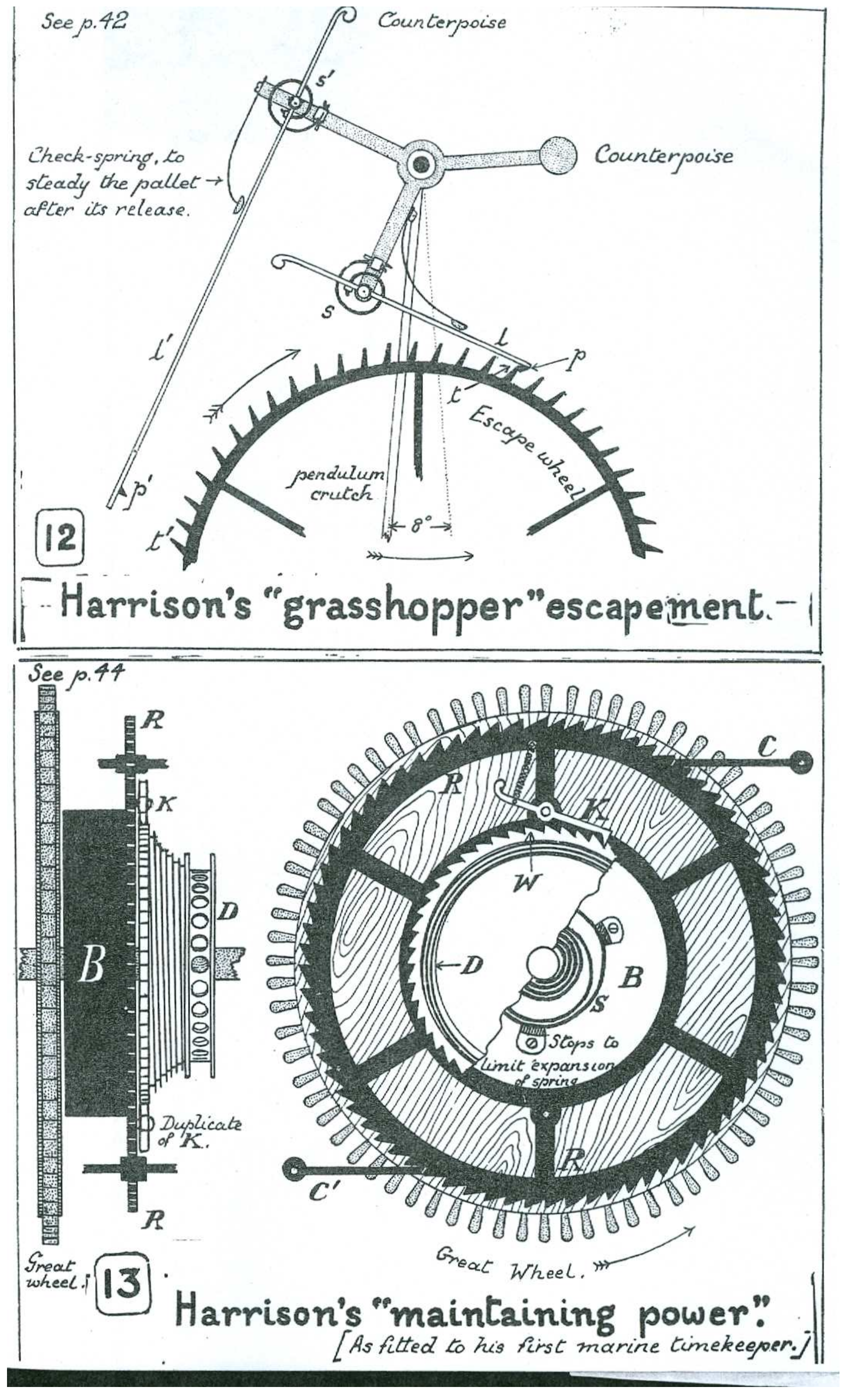


Figure 4.23: Note the shape of the teeth on the escape wheel. From Gould.

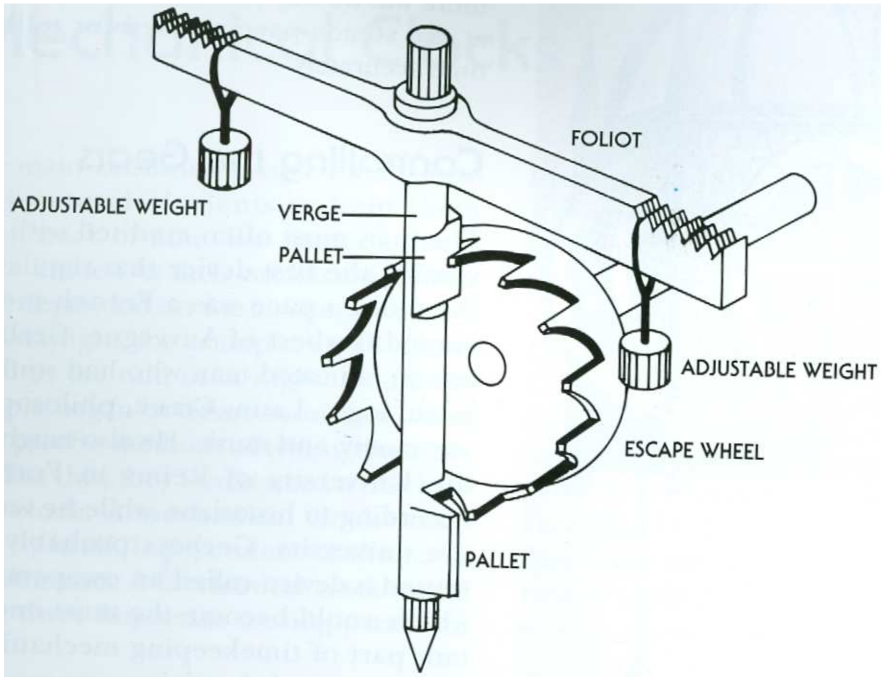


Figure 4.24: Note the shape of the teeth on the escape wheel.

take turns coming into contact with the teeth of the escape wheel. Once in contact, the escape wheel will push the given pallet back out as the wheel continues to rotate. But pushing a pallet back out causes the entire verge to rotate, forcing the other pallet to eventually make its own contact with the escape wheel. It gets pushed out and the verge rotates in the other direction, forcing the original pallet back in contact with the escape wheel again. The periodicity of the clock is set by the time it takes the pallets to engage, then re-engage the escape wheel.

The T-bar on the verge is called a foliot. It's length and mass, as well as the position of the weights determines how long it takes a pallet to reach the escape wheel. The more massive the foliot and the weights, the more time it will take between pallet engagements. The verge then serves as the axis of another gear (not shown), which can drive second and minute hands on a clock. In motion, you would see the T-shaped verge/foliot pair turning back and forth.

The common property of the escapement mechanisms shown is that they have very strong orientation dependencies (hence potential problems if deployed at sea). Falling weights must be still and vertical, and a pendulum's pivot must be stationary. The verge and foliot need not be vertical, but a single orientation must be found for its falling weight on the escape wheel.

To solve this problem, a coiled spring, as mentioned, can also be used as the energy storage device. The spring will uncoil in the same manner independent of its orientation and could be used in place of the falling weight in any of the mechanisms shown above. But

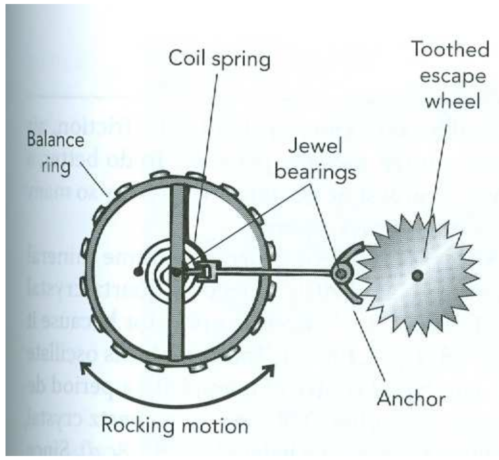


Figure 4.25: Note the spiral hairspring that causes the balance wheel to go a few degrees clockwise, then a few degrees counterclockwise, and back again. From Bloomfield, p. 299.

what sort of escape mechanism also has no orientation dependence? You guessed it! Another spring!

Remember from our discussion of springs that when a spring is pushed or pulled from its natural length, it will oscillate back and forth at a precise frequency. Any elastic material will respond in this manner. If a thin strip of metal is shaped into a spiral, you'll get what's called a "hairspring" and a "hairspring escapement" as shown in Figure 4.25.

In this arrangement, notice the spiral hairspring. One end is anchored and the other end is attached to the center of the balance wheel. If the balance wheel is turned slightly (clockwise or counterclockwise), the spring will coil, or uncoil a bit. This is analogous to one pushing or pulling on a more common "straight" spring. Either way, the hairspring has been displaced from its manufactured equilibrium position a little bit, and like all springs, it will want to spring back to equilibrium. In doing so, it overshoots, stops itself, and then goes back again; it begins to oscillate with a well defined period. But since it is also attached to the axis of a small wheel (the balance ring), the ring will begin turning a small amount clockwise, then back counterclockwise, and back again. The ring will "rock" back and forth. The balance wheel is attached to an anchor, which, like before, engages, and disengages an escape wheel, controlling its release of energy. The hairspring will oscillate independent of its orientation, and if the escape wheel is also powered by a coiled spring, you'll have an entire driven escapement system that will work in any orientation. This is the design behind most mechanical wristwatches and what went into Harrison's H4, which ultimately won the Longitude Prize.

A few more notes about Harrison's work. Harrison was a master carpenter, and so he made his gears and escapement mechanisms out of wood. He used oak wood (which is very strong), with the grains of the wood running parallel to the direction of the teeth to maximize their strength. Further, he practically invented the roller bearings, which mated all of the wheels and gears to the axis about which they spin. Also, the wood he chose for the

bearings and anchor tips is a wood called "lignum vitae," which is a very hard, dense wood (so dense that would sink in water). More importantly, the wood provides its own natural lubricating oils which do not dry out. Indeed in the 1700s, Harrison's wheels and escape mechanisms were virtually frictionless. He was a true master of his materials. Harrison also had a way of winding the sea clocks (putting more energy back into the mainsprings), without interrupting their time keeping. This would be critical for sea clocks, but not so much for a grandfather clock in one's house. [Gould] describes how this worked, but for now, it defies this author's understanding!

4.6 More Bizarre Clocks

4.6.1 Kepler's Laws

Time keeping using the planets and stars is now all but abandoned. No one counts days by observing the sun at its meridian passage. Maybe hearing that it's the "summer solstice" triggers a fond summer memory from a year ago, but that's about it. No one counts years based on seasons. Nevertheless, the planets hold some remarkable time keeping abilities that come from their very placement and structure in the solar system. The laws that summarizes these time keeping properties are called "Kepler's Laws" and they are as follows.

Law 1: The Law of Orbits

The law of orbits states that all planets (in our solar system) move in elliptical orbits, with the Sun at one focus. Here's what this means. The planets do not move in circular orbits. You probably know this because the earth is closer to the sun in the northern hemisphere's winter and farther in the summer. Also, the earth is the same distance from the sun on the first day of spring (March 20) and first day of fall (September 21). A path around the sun that is not equal at all times cannot describe a circle with the sun at the center. Instead, the path is a stretched circle, or ellipse, with the sun not at the center, but at a focus, as shown in Figure 4.26.

This figure shows two focal points, one containing the sun, the other empty. The earth would orbit along the path outlined by the solid, dark line. The focal points are needed for the old high-school trick used to draw ellipses: the sum of the distances from one focus, to the dark line, then to the other focus is always the same for a given ellipse. Notice how the earth at mid-winter and mid-summer would fit at the nearest and farthest approaches of the elliptical path. The equinoxes fit just "above" and "below" the sun in the figure.

Law 2: The Law of Areas

The law of areas states gives us the first hint of time keeping abilities of planets. It states that a line connecting a planet to the Sun sweeps out equal areas in equal times.

Figure 4.27(A) shows an imaginary line connecting the Sun and the earth (or any planet). As the earth moves, imagine this line staying connected to the earth and "sweeping out"

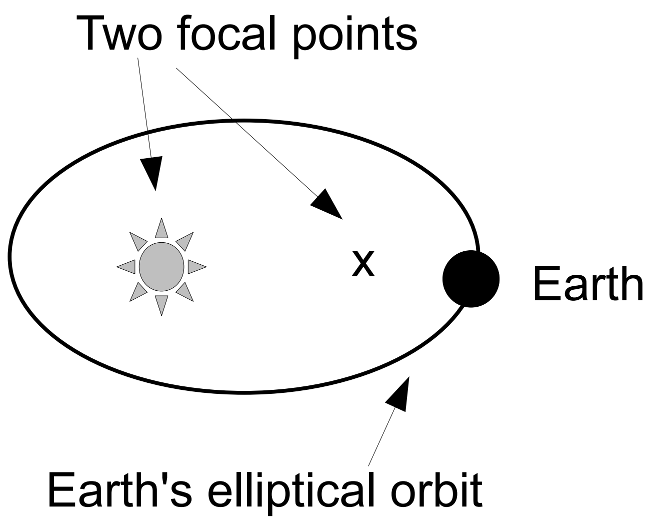


Figure 4.26: Kepler's first law: Planets (like the earth) move in elliptical orbits with the sun at one of the two focal points.

some area as shown in Figure 4.27(B). Suppose the time between the two positions of the earth in Figure 4.27(B) is the same as the time between the positions in Figure 4.27(C): the areas shown would be the same; "equal areas swept out in equal times." Sure the area in Figure 4.27(B) is long and narrow, but the area in (C) is short and fat. All told, the gray triangles have the same area.

Is this useful for time keeping? Not directly. No one is going to build a "Kepler" clock that uses this law to keep time, but it's a hint at a natural time-keeping mechanism in the planets. If you lived above the solar system and had the view of the earth's orbit as shown in Figure 4.28, then you could conceivably record the earth's position and draw out the triangles the imaginary lines sweep out. When two triangles had the same area, you'd know the same amount of time went by. Wildly different in composition, but fundamentally, not different from how a pendulum swings. Planets sweep out equal areas in equal times, and a pendulum swings out and back in a constant time. Pendulums just happen to be more convenient for clock building here on earth.

Law 3: The Law of Periods

Planets all have a period, or the time it takes them to go around the sun one time. For the earth, this period, T is one year. But the period of a planet is also related to its orbital radius, or how far it is from the sun. Planets that sit farther from the sun take a longer time to go around once, and hence have a larger T. For instance, Jupiter is just over 5 times farther from the sun than the earth is, and its period is 12 times longer than the earth's. Neptune's orbital period is about 184 years, and it was discovered in 1846. So it is just finishing the first complete orbit since it was discovered[R. Brown].

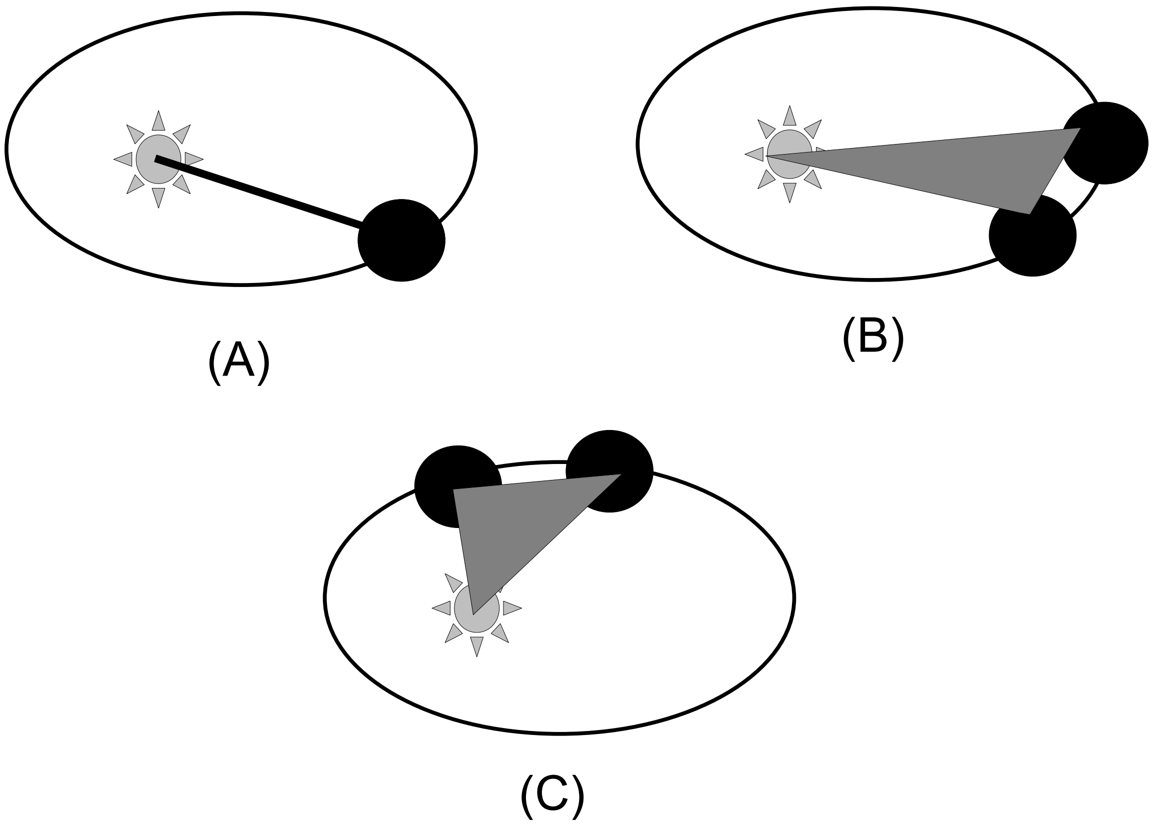


Figure 4.27: Kepler's second law: A line connecting the sun and a planet (A) sweeps out equal areas in equal times (B) and (C).

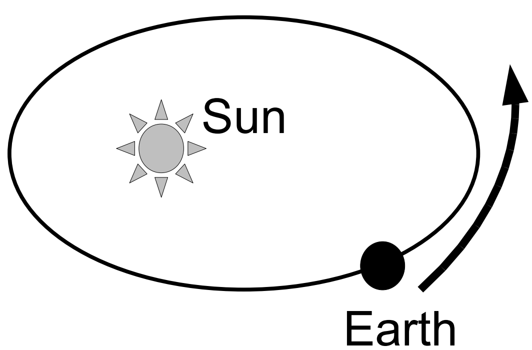


Figure 4.28: If you had this view of the earth/sun system, you could use Kepler's area law to keep time.

This relationship between periods and orbital distance is just what Kepler's third law is all about: The square of the period is proportional to the cube of the orbital distance. We are sorry this isn't a simple relationship, such as "the period is proportional to twice the orbital distance." You must square the period and cube the orbital distance for this law to work. In other words, for any planet with an orbital period T and orbital distance r,

$$\frac{T^2}{r^3} = \text{ a constant.} \tag{4.9}$$

In fact, for all planets in our solar system, the constant is 3×10^{-34} in units of $(\text{year})^2/(\text{meter})^3$. The constant involves the mass of the sun, which all of our planets revolve around. This is why the constant is, well, "a constant." In some other system, where a planet is orbiting around some other "central body," the constant would be different. So where's the time keeper? Well, step out of the solar system again and look down on it. You'd see 8 planets orbiting the Sun, sort of like gears of a clock. All of them will move such that $T^2/r^3 = 3\times10^{-34}$ y^2/m^3. This mechanism would be good for keeping track of years. The Earth's period would trigger the "year dial" on your clock. Venus could almost trigger the "half-year" dial with a period of 0.62 of the Earth's. Jupiter could almost count decades with its 12 year period. If you wanted an orbiting body, like an artificial satellite, to have a particular period, T, you could use this third law to find the radius of the orbit that would be required. This would be the radius to which you'd have to launch it, in order for it to have the period you desire.

4.6.2 Atoms

This section has a good name to it. Use an "atom" to keep time; "atomic time" if you will. An "atomic clock." It sounds good, so atoms *must* do a great job keeping time. It turns out that atoms actually do an excellent job keeping time; atoms are used to create the most accurate clocks known. In fact, an "average" atomic clock loses about 1 second in every billion years! To see how they work though, you need to know a little bit about atoms and light.

You probably know that atoms are the small constituents that make up everything around us. Water is H_2O, two hydrogen atoms connected to a single oxygen atom. Gasoline for our cars consists of long chains of carbon atoms with hydrogens sticking off to the sides. The air we breathe is 70% of pairs of nitrogen atoms, and 30% of pairs of oxygen atoms. To understand how atomic clocks keep such good time, we need to understand a little about the structure of atoms.

What is an atom?

You probably remember from high school that an atom looks like a little solar system. In the center is a positively charged nucleus, made of protons and neutrons. Orbiting around the nucleus are one or more negatively charged electrons, as shown in Figure 4.29. This is called the "Bohr Model" of the atom, after Neil's Bohr, who first developed a theory on atomic structure around 1900. His model has mostly been shown to be wrong, but it does have some successes in that it happened to correctly predict a few properties of the

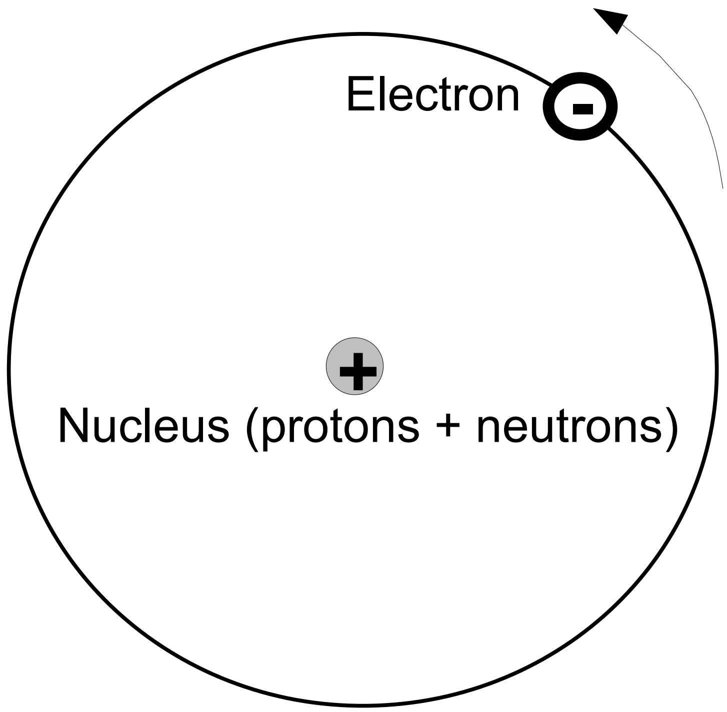


Figure 4.29: Model of an atom, showing how absorbing light energy can result in a larger orbit.

Hydrogen atom. This was considered a great success back then, but it fails miserably for any other atom. The trouble is that Bohr's model is a "classical" model for an atom. Classical refers to the assumption that atoms are made of little discrete particles; an electron like a little ball, orbiting around a slightly bigger ball (the proton). We now know that atoms are "quantum" by design, which is beyond the scope of our work here. Nevertheless, this model is a good starting point for discussing atomic structure. We'll be able to use it to guide us in understanding how atoms can be used as clocks.

In the Bohr model, for example, a hydrogen atom has a single-proton nucleus with a single electron orbiting around it. A sodium atom is similar, except that it is a bigger, more involved atom than hydrogen. It has 11 protons and neutrons packed together, forming the nucleus, with 11 electrons orbiting around it. For the sake of atomic clocks, we'll always be able to ignore the nucleus. Atomic clocks exploit the *electrons* moving around the nucleus, and the only property you need to know about these electrons is that they hold energy in their orbits.

Electrons in their orbits have energy that comes from two sources. The first is in their kinetic energy, or energy they have because they are moving. The speed an electron in a hydrogen atom must have to stay in a stable orbit around the nucleus is given by

$$v = \sqrt{\frac{ke^2}{mr}}, \tag{4.10}$$

where k is a constant, $k = 9 \times 10^9$ N m^2/C^2, e is the charge of an electron, $e = 1.6 \times 10^{-19}$ C, m is the mass of an electronr $m = 9.11 \times 10^{-31}$ kg, and r is the radius of the electron's orbit, $r = 5 \times 10^{-11}$ meters. Putting these numbers into the equation for v gives a speed of about $2,249,039$ m/s or about 5 million miles per hour. You might think this speed would require an enormous amount of energy, but it does not, because the electron has such a small mass. Nevertheless, the electron is moving, and it has energy because of this motion.

The second source of electron energy is in its separation from the nucleus. If you look carefully at Figure 4.29, you'll notice the positive nucleus sitting close to the negative electron. Another high school fact you might remember is that opposite charges attract each other or "opposites attract." This means the electron would love to simple fall freely into the proton. It doesn't though, because of its orbital motion. It's motion at the exact speed discussed above keeps it from falling into the proton, just as the earth's motion around the sun prevents it from falling into the sun. This "held apart-ed-ness" is another form of energy, called "potential energy." The word "potential" is used because the electron could *potentially* slam into the proton, but it doesn't because of its orbit. Incidentally, to prove that energy exists between a positive proton and negative electron, we could design an experiment that let the electron fall toward the proton. If this was allowed to happen, the atom would begin to glow, releasing the energy it had.

Finally, one last point about atoms and the energy in their electrons. On a whim the energy of an electron can be changed by shining light on the atom. Light? Yes, light, like the light from the Sun or that which is illuminating the page of this book. Light is a form of energy, and if you make light shine onto atoms, the electrons in the atom can absorb the light's energy. This is illustrated in Figure 4.30.

In Figure 4.30(A), you see an atom with light shining down upon it. If the light is absorbed by the atom, the light will disappear; in its place you'll see the same atom, but the electron will be in a larger orbit, farther from the nucleus, as shown in Figure 4.30(B). In other words the absorption of light by an atomic electron causes it to jump up to a larger and more energetic orbit. The energy from the light is added to the energy the electron already had.

The reverse is also true: an electron in a large, energetic orbit may "fall" to a smaller orbit, releasing energy (i.e. light) in the process. So that's it on atoms for now. They have a nucleus with one or more electrons orbiting around in stable orbits, that have a certain energy associated with them. If an atom absorbs light, the electron can gain the light energy and jump to a larger orbit. An electron can also fall to a smaller orbit, releasing energy in the form of light, in the process. We'll bring all of this together a bit later. Let's discuss light a little bit now.

Light

When we say light, we *do mean* the good old light that comes from the sun, or the light that's probably illuminating the room in which you are sitting. This is just the "stuff" that can give energy to electrons. There are a few catches to using light to "excite" atomic electrons, which we'll discuss in a moment. How light is needed for atomic interactions requires you

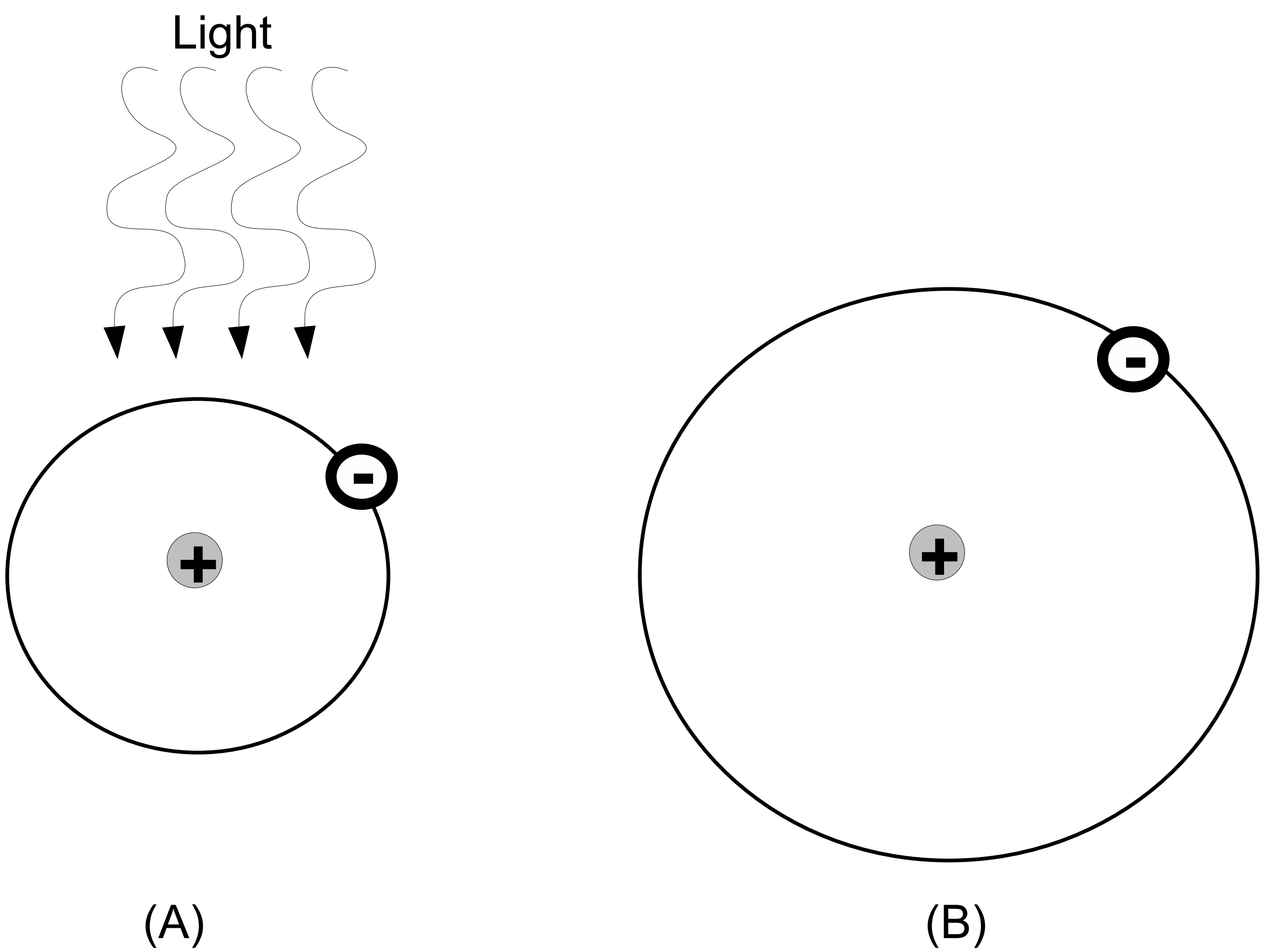


Figure 4.30: Elementary model of an atom.

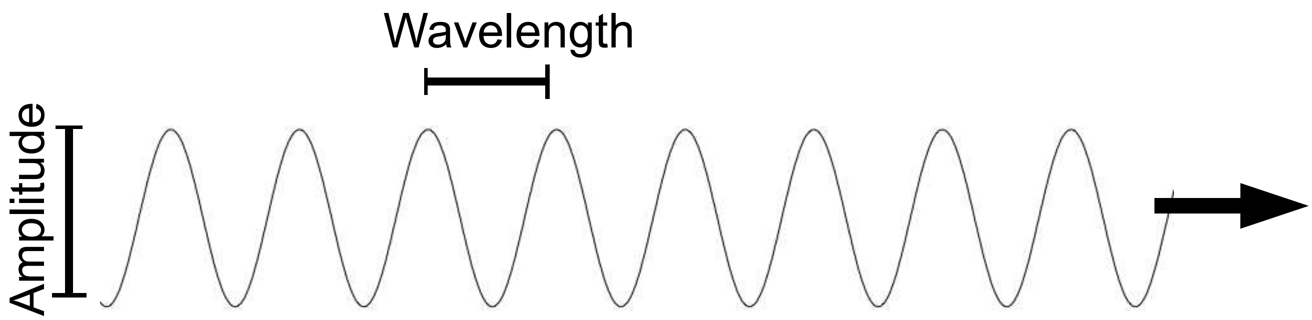


Figure 4.31: Two properties of a wave: its amplitude and wavelength.

to know a few properties of light itself.

To begin, light is a wave. A wave just like one in the water; it has high points, low points, and sort of "wriggles" along and you can certainly surf on the crest (or high point) of a wave. Sound (music, speech, noise, etc.) is also a wave. High points in a sound wave correspond to higher pressure; a low point to lower pressure. These pressure differences are what stimulate your ear-drum so you can hear. A speaker that produces sound pushes and pulls on the air producing these pressure differences. These are a few general examples of waves, but there are also very specific properties of waves, as shown in Figure 4.31.

Suppose Figure 4.31 shows a wave of light (or water, or sound) traveling toward the right. The height of the wave, is called the amplitude. This is the maximum distance between a high crest and a low valley. If the wave is light, the amplitude would be how "strong" or bright the light is. If it were a sound wave, the amplitude would be how loud the sound is (the volume). The distance between two crests (or valleys) is called the "wavelength" of the wave, and this is a very important property, for it defines what type of effects the light (or sound) might have in our environment. Here are some examples.

If a light wave has a wavelength between 400 and 700 nanometers, the light will be visible to humans. These are the wavelengths are eyes are sensitive to. Blue light is around 400 nm, and red light is around 700 nm. Green light is at about 520 nm, etc. Light having wavelengths of about 1 cm are microwaves. These are used to heat our foot and communicate with satellites. Your cell-phone also uses centimeter wavelengths to "beam" your [important] conversation to your carrier's network. Light with a wavelength of about 10 nm is an X-ray, like the doctor might use to examine your bones. Light with a wavelength of about 3 meters is used for FM radio broadcasts; 300 m is used for AM radio.

Another important parameter of light is its frequency, which is very important to us because it is used directly to drive atomic clocks. To understand frequency, take a look at Figure 4.32, and imagine you are stationary, and watching a light wave travel by you at the position shown.

Suppose you are looking straight down, and count how many times per second a crest (high point) on the wave passes you. Say you watch carefully and count 10 crests go by in 3 seconds. This light would have a frequency of $10/3 = 3.3$ "per second" or 3.3 Hertz (light

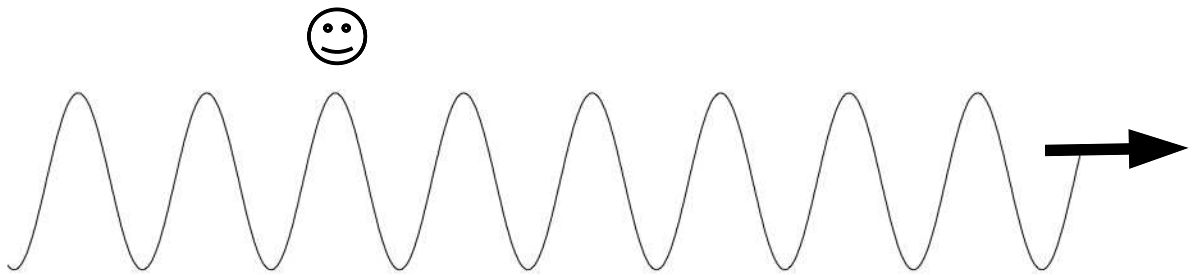


Figure 4.32: The frequency of this wave would be how many times per second a crest passes by you.

of this frequency is very unusual, but is emitted by lightning bolts). Pay attention to the units of the numbers here: 3.3 "per second." Per second of what? 3.3 *crests passing* per second. This is the frequency of light. Visible light has a much higher frequency. If you were watching the wave in Figure 4.32 go by, and it was the red beam from a laser pointer, you'd see about 4.6×10^{14} crests go by per second. This is about 460 million million crests per second!

Light is a very pure form of energy. Either light is present, carrying its energy, or it doesn't exist. The energy comes from the mechanism that forces it to oscillate (go up and down) as it travels like a wave. Frequency is a very important parameter for this because, as mentioned, the frequency of a light wave measures how quickly the wave goes up and down. The faster it goes up and down, the more energy it has. In fact, the energy in a light wave is given by

$$\text{Energy} = h \times \text{Frequency}, \tag{4.11}$$

or more succinctly, if energy is E and frequency is f, then

$$E = hf, \tag{4.12}$$

where h is known as "Planck's Constant," $h = 6.63 \times 10^{-34}$ J·s. Planck's constant was discovered (by a man named Planck) back in the early 1900s, when Bohr was working on his atomic theory. Planck's constant was correct and is still widely used today when describing atoms and such. This equation clearly shows that the larger f is, the more energy light carries. That is, the higher the frequency of light, the more energy it has.

Lastly light is very fast as it travels. Too fast for us to notice in everyday activities. It has a speed which is $300,000,000$ meters per second, or $186,000$ miles per second, or about 1 foot every nanosecond. At this speed, it takes light 8 minutes to travel the 93 million miles from the Sun to the Earth. There is a key relationship between the speed, frequency, and wavelength of any way, which is:

$$\text{Speed} = \text{Wavelength} \times \text{Frequency}. \tag{4.13}$$

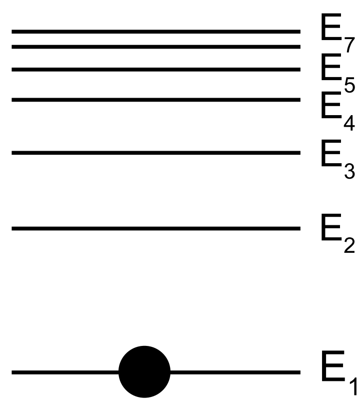


Figure 4.33: A sketch of an atom: the energy levels of the electron.

Atoms and Light

We're almost ready to discuss how atoms are used to keep time. First, though, a short discussion on what happens when we now combine atoms and light. As discussed, the orbits of electrons typically aren't that important. When atoms are studied or discussed, people rarely draw the orbital paths of electrons. The energy of their orbits, however, are very important, so mostly atoms are drawn like that shown in Figure 4.33.

In this figure, you can see a series of horizontal lines. Each line represents a possible energy the electron in an atom can have. The lines are labeled E_1, E_2, etc. for different possible energy levels of the electron. E_1 is the lowest energy level, and as shown, the electron is apparently in this energy level at the moment. This level would correspond to the smallest orbit, and the smallest value of energy the electron could a have in the atom. The energy levels get higher in energy as you go up to E_2, E_3, etc. In reality, the energy levels go up forever, so you could have something like $E_{93847593874 95}$, which would be a very large, highly energetic orbit. When you get to ∞ at E_∞, the electron has so much energy that it has left the atom.

Notice two more things about Figure 4.33. First, the spacing between the energy levels is not the same. The energy levels get closer as you go higher. This is a real effect and is a consequence of the electron's kinetic energy being related to it velocity squared, as in $KE = \frac{1}{2}mv^2$. Second, notice that the energy levels are discrete, that is, they have the obvious [nonuniform] spacings between them. In other words, the electron can have energy E_1 or E_2, but nothing in between. There is no $E_{1\frac{1}{2}}$ for instance. This is the "quantum" nature of atoms; electrons in an atom can only have particular (or quantized) amounts of energy.

So here's how atoms and light mix. It was stated before that if you shine light onto an atom, the electron can absorb the light and go to a higher energy level, but here's the catch: the energy in the light must exactly match the spacing (or gap) between the energy level the electron is in, and the next highest level. This is illustrated in Figure 4.34.

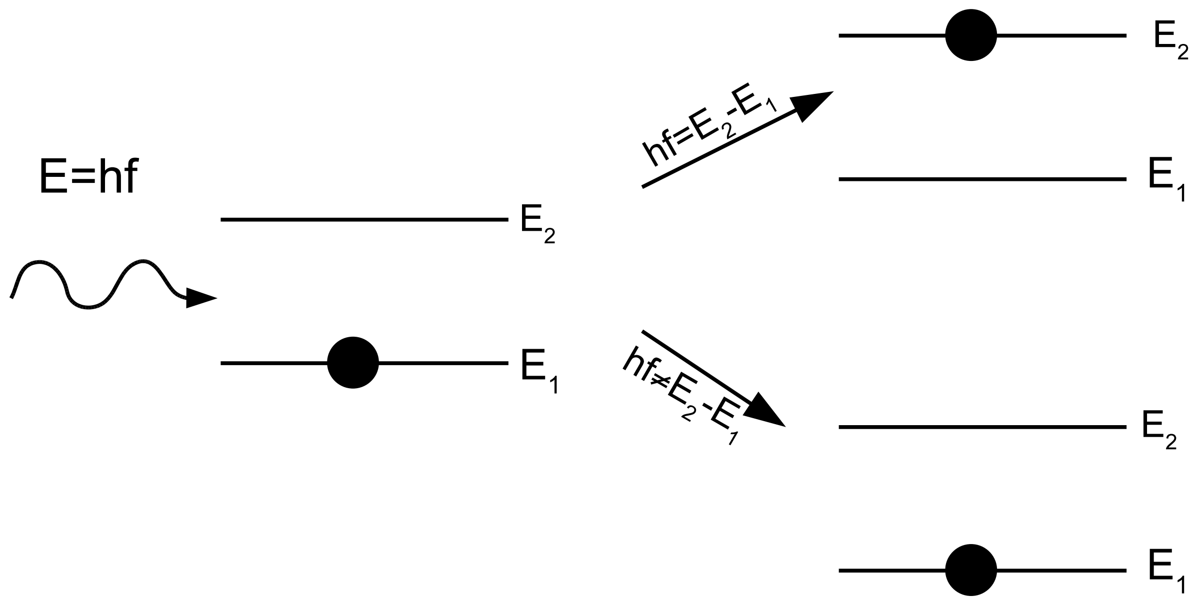


Figure 4.34: In order to cause an electron to jump to a higher energy level, the energy of the light must exactly match the energy spacing between the two electron energy levels. That is, $hf = E_2 - E_1$.

To the left is an atom with its electron in the lowest energy level E_1, and light is irradiating the atom from the left. If the energy of the light, hf, exactly matches the spacing between E_1 and E_2, or if $hf = E_2 - E_1$, then the electron will absorb the light and will jump from E_1 to E_2. The light will now be gone, but you'll be left with an electron in a more energetic level. If the light's energy level does not match $E_2 - E_1$, then the electron will remain in level E_1 and the light will pass right on by.

So we are in a position to revise an earlier statement that said "you can shine light on an atom and cause an electron to jump to a higher energy level." This should be corrected to "if you can shine light on an atom, an electron will jump to a higher energy level if the light's energy exactly matches the difference in the electron's energy levels." It turns out that this is not all that easy to do. If you set out with the explicit goal of exciting an electron in a particular atom, from one energy level to another, some quite sophisticated laboratory equipment would be required.

Atomic Clocks

At last then, with a background on atoms, light, and atoms+light, we can now discuss how atoms and light are used for time keeping. First off, review Figure 4.32, light has a given frequency. Frequency is implicit in timekeeping. For light, it's how many times per second a crest goes from "up to down to up again." This is like the "tick-tock-tick" of a clock, or the "out-and-back" motion of a pendulum or mass-spring system. The major flaw in

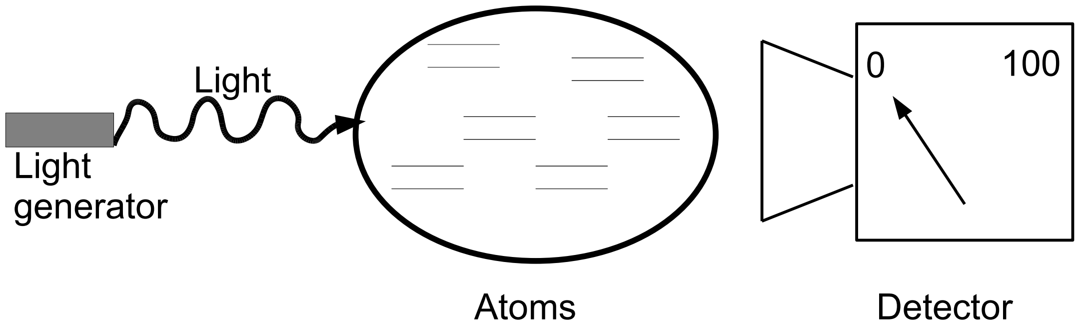


Figure 4.35: Potential arrangement for an atomic clock. If light gets absorbed by the atoms in the round container, it won't register on the detector. If light doesn't get absorbed, it will pass through the atoms and hit the detector. The frequency of light is constantly adjusted to keep the detector reading as low as possible. In this figure, light is apparently being strongly absorbed in by the atoms, since very little is hitting the detector.

the mechanical (pendulum or spring) clocks are that they are susceptible to environmental factors, like friction, pressure, and temperature. Light waves, however, are entirely unaffected by these parameters. In other words, red light will have a frequency of 460 million million crests per second (see above) in the cold of Antarctica or the heat of Death Valley. Under the ocean or in the parched Sahara Desert. It doesn't matter. So it is the precision oscillation of light waves that are at the heart of atomic clocks. Here's how they work.

Suppose you had an arrangement as shown in Figure 4.35. Light is generated by a "light generator" and is aimed at a glass cell containing a bunch of atoms (the atoms are denoted by two little energy levels). (Note: In actuality, the "light generator" will be a laser that generates visible or near visible light, or more commonly a microwave generator that generates light with a wavelength of about 1 cm.)

On the other side of the glass cell, is a light detector. The more light that hits it, the larger number the needle will point to. Now remember that if light going into the glass cell is absorbed by the atoms, causing an electron to jump to a higher energy level, then the light would disappear; it wouldn't make it to the detector and the detector would read a low number (as it is in Figure 4.35).

Now, like a pendulum or mass-spring system, the light generator is a (partly) mechanical device and so it is susceptible to temperature, pressure and humidity changes. Suppose it started to get hot and the frequency of the light it produced no longer matched the energy spacing of the atom. What would happen? More and more light would go right through the glass cell and hit the detector. That is, the light would cease to be absorbed by the atoms. The atoms would become more and more transparent to the light. In this case, the detector needle would read a higher number, as shown in Figure 4.36 since more light is entering it.

But remember, atomic energy levels are mostly immune to the temperature changes affecting the light generator, so their energy levels (and gaps between them), stay constant. So if the amount of light absorbed changes, we'll know that something must be shifting with

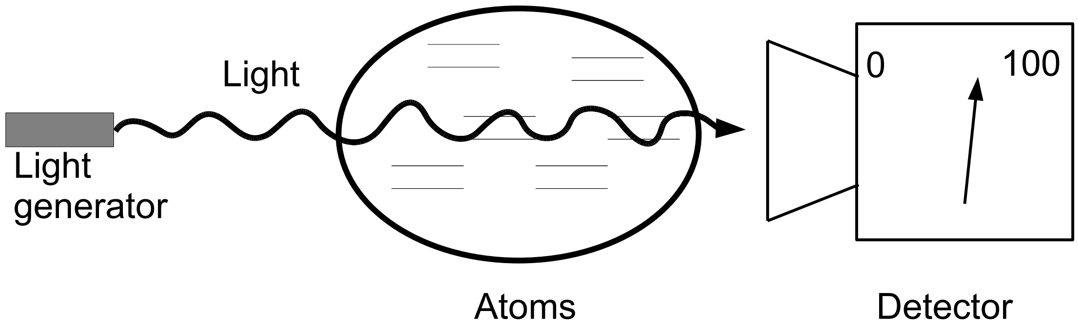


Figure 4.36: In this figure, light is not being absorbed strongly by the atoms, since a lot of light is going right through the atoms and entering the detector. The frequency of the light must be wrong, relative to the energy spacings of the atoms.

the light generator, not the atoms, and the light generator's frequency should be adjusted. In particular, increased light entering the detector is a clue that the frequency of the light is wrong, or has become wrong for whatever reason. What we need is a way for the detector to send a signal back to the light generator automatically, in order to adjust its frequency until the detector goes back to reading a low level signal again. We can do this by connecting the detector and light generator electrically, using a "feedback loop" as shown in Figure 4.37.

The feedback loop allows the detector to tell the light generator to adjust its frequency. It works like this. The light generator turns on and shines light into the atoms. If a lot of light hits the detector then we know the generator is not outputting light of the proper frequency; perhaps the frequency is too high. The detector then tells the generator to lower its frequency and try again. If this doesn't work, the detector says "make another adjustment and try again." It keeps doing this until the detector sees a low light level hitting it, then it tells the generator to stop adjusting its frequency. Everything is fine (for now).

Suppose after a while though, the light level on the detector rises again (it always will). The detector might tell the light generator to lower its frequency, and after doing so sees more light hitting it. "Oops!" it thinks, "wrong direction." It then tells the light generator to raise its frequency and in doing so, it sees the light level hitting it begin to drop again. This "back and forth" goes on as long as the system is running. Initially, large adjustments in the light generator frequency will be needed. But after a while something of an "equilibrium" (or agreement) is reached between the light generator and detector, so only small changes or "tweaks" are needed and the detector reading stays at a low level.

So although all of the action is between the generator and detector, remember that at the core here are the atoms, and their immutable energy levels. They are regulating everything. If the light generator's frequency drifts *relative to the atomic energy levels* the detector is forced to make adjustments. The detector is only happy when low levels of light are hitting it, meaning the light generator's frequency matches *the energy levels of the atom.*

This is the heart of an atomic clock: the atoms themselves serve as a check on the frequency of the light produced by the light generator. If anything drifts relative to the atomic

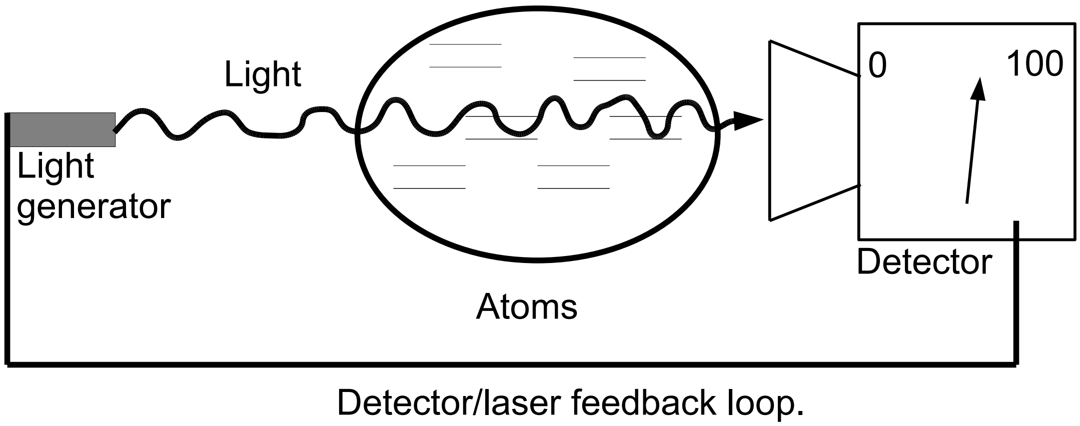


Figure 4.37: A feedback loop between the light generator and detector. Since light is getting through the atoms and entering the detector, the detector must now tell the light generator to adjust its frequency. The pair will make continual adjustments to minimize the amount of light passing through the atoms and entering the detector.

energy levels, an adjustment is made. So the atoms play an analogous role to Harrison's gridiron pendulum and its ability to compensate for temperature changes. If *anything* causes the light generator to shift the frequency of the light it is producing, relative to the atomic structure, an adjustment is quickly forced on the generator to compensate for the shift.

All told, an atomic clock produces a very stabilized light frequency. Also, feedback loops like this are not that uncommon. Think for a minute how the thermostat and heater (or air conditioner) in your house maintains the comfortable temperature that you have set. Using atoms to regulate a light generator's frequency is entirely analogous to how your heater knows when to come on relative to a desired house temperature you have set.

Time keeping and atomic clocks

So where does the timekeeping come from with the atoms and the light generator? We'd like a hand that counts seconds perhaps? The heart of the time keeping is the ultrastable frequency of the light generator the atomic clock maintains. Think of the frequency as the "beating" of the clock or analogous to the swing of a pendulum. The pendulum is mechanical and can be directly connected (with gears, etc.) to a second hand. What do we attach to the atomic clock? The light generator is an electrical device, containing electrical circuits and other elements that are able to generate the light in the first place. The most common atomic clocks use Cesium atoms with a precise absorption frequency of 9,192,631,770 Hz or about 9.1 billion times per second (think of this as a pendulum that goes back and forth 9 billion times per second!). If you look back at Equation 4.13, you'll see this corresponds to a wavelength of light of about 3.1 cm. These are microwaves. Before the advent of atomic clocks, the second was defined as 1/60th of 1/60 of 1/24th of a day. Now one second is

defined as the amount of time it takes a microwave generator, tuned to excite the Cesium atoms, to oscillate up and down $9,192,631,770$ times.

The key in generating a useful time keeper is to send part of the stable frequency signal, from the light generator, into a "frequency divider" circuit. These do nothing more than take in an electrical frequency and divide it by some number and output the result. They are somewhat straightforward to build, and are usually covered in a college-level electronics class. Think about the Cesium clock mentioned above. Suppose the atom-stabilized frequency was divided by a "nice round" number like 1 billion. The divider would then output a frequency, or "heartbeat" of about 9.2 seconds. And it would be *very stable* because remember that the origin of the selected frequency is the atomic energy levels of Cesium atoms. In fact, the 9.2 second frequency would just about be the most stable 9.2 second time-based ever created by humans! It is not uncommon for atomic clocks to lose 1 second over a million years. Remember that the eventual 9.2 seconds here is "just" a time-based that can be used for a clock. We've seen other ways of doing this. A pendulum, for example, with a length of 20.5 m would oscillate back and forth with a period of 9.1 seconds too, but as you know, it would be terribly susceptible to temperature changes and air resistance (not to mention quite large).

A picture of an "atomic beam clock" is shown in Figure 4.38. In this figure the "source" is where the gaseous atoms come from. Magnet A causes the electrons to jump into a particular energy level. The microwave cavity shoots 1 cm-ish light waves at the atoms. Only the atoms that jumped up in energy from magnet A will absorb the microwaves. Magnet B causes atoms that absorbed the microwaves to turn into the detector. Notice how the detector is connected to a "servo mechanism," that itself is connected to the microwave generator. If the detector sees a drop in atoms reaching it, it will adjust the frequency of the microwaves. The frequency divider and digital display are also shown coming out of the microwave generator.

A more technical representation of a clock that uses Rubidium (Rb) atoms (instead of Cesium) is shown in Figure 4.39. The functionality is the same, only Rubidium has an oscillation of $6,834,700,000$ times per second. This means the "tick" and "tock" of an Rb clock has approximately 0.14 nanoseconds between them $(1/6,834,700,000)$, and a second based on Rb clocks would be defined as the time it takes the light generator to oscillate up and down $6,834,700,000$ times.

If you study this figure carefully, you'll see many of the items that have been discussed thus far. Inset (a) shows the energy level diagram for an actual Rb atom. Although more complicated than hydrogen, an electron can visit any of the quantized horizontal lines and jump to and from another with the absorption or emission of light. Inset (b) shows the output of the light detector. The lowest point on the "dip" is where the Rb atoms are maximally "soaking up" the light put out by the light generator; this is the stability point on which the clock seeks to "lock on." Following the arrow to the right, the phase sensitive detector is a bit of electronics that closely watches this dip and ensures the system stays there by adjusting the light generator's frequency. We note also how a crystal oscillator sits between the displayed time (the clock) and a multiplication circuit. The light generator needs to output light at a frequency of $6,834,700,000$ Hz, but the crystal oscillator cannot

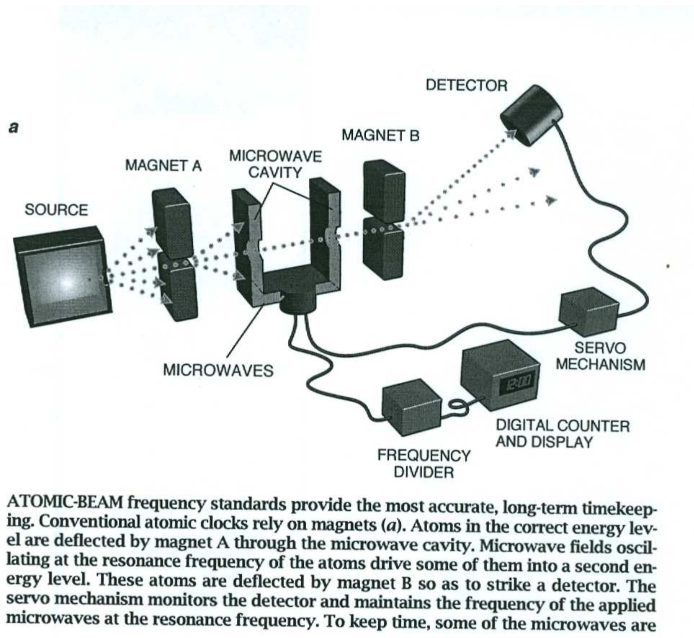


Figure 4.38: A schematic for an atomic beam clock. From Itano, SciAm, July 1993, p. 62.

run this fast. Perhaps it oscillates stably at 6 Hz (6 times per second). The "multiply by M" box multiplies this 6 by a factor of one billion or more, to the frequency needed to cause absorptions in the glass cell containing the Rb atoms. The other branch, with the 6 Hz on it, is used by the clock to tell the time (with maybe an additional division by 6 for a 1 second timebase?).

So how good are they?

To close on atomic clocks, a short discussion on how good they are at keeping time. There are many numbers out there on this. Here are few we found. A modern Cesium-beam clock is good to 1 second in $10,000,000$ years[Jespersen], or about 1 microsecond per day[Hood, p. 60]. Another figure puts it at 1 second in $3,000,000$ years[Barnett, p. 160]. You can compare these numbers to Harrison's H4, which won the longitude prize. It lost 1 minute and 54.5 seconds in 147 days, during the trip to the West Indies[Andrewes, p. 143]. This is about 0.7 seconds per day, or about 256 seconds per year.

4.6.3 Electronics

Mechanical devices seem to be the most obvious solutions for timekeeping. Pendulums, springs, masses, gears, etc. all come to mind when one considers what it might take to build a clock. Atoms are clearly the best at keeping time, but a little overkill. What about a clock for "the rest of us?" It turns out though, that for ease, precision, miniaturization, and

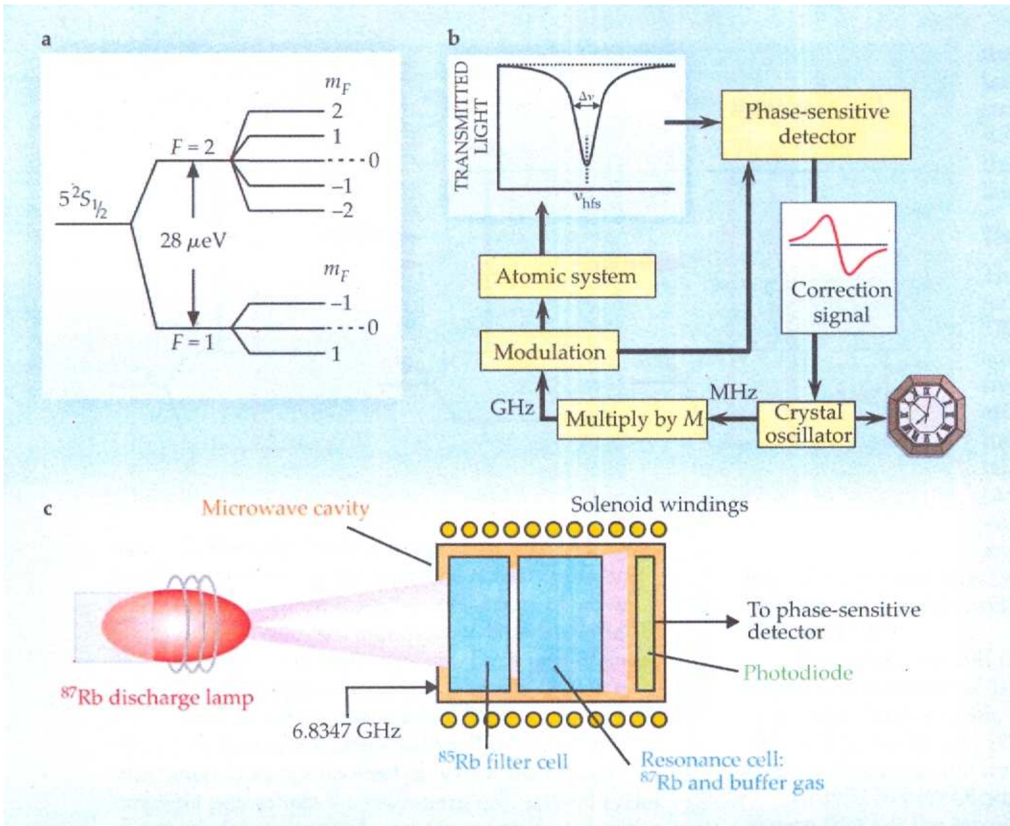


Figure 1. A generic rubidium vapor-cell atomic clock. **(a)** The ground state of ^{87}Rb is split into two hyperfine levels, characterized by the quantum number F. The transition between the two $m_F = 0$ sublevels, unchanged to first order by stray magnetic fields, is used to stabilize a crystal oscillator's frequency. **(b)** The frequency of a voltage-controlled crystal oscillator, scaled up by a factor M, is locked to the hyperfine-splitting frequency $\nu_{hfs} = 6.8347$ GHz using phase-sensitive detection. The control voltage corrects the crystal oscillator's frequency. **(c)** The clock's atomic system consists of a few simple elements.[15] Optical pumping by lamplight creates a population imbalance between the hyperfine states of the ^{87}Rb atoms in the resonance cell, so that in steady state few atoms are in the $F = 1$ state and the transmitted light intensity is relatively high. When microwaves at ν_{hfs} are applied to the resonance cell, there is a net transfer of atoms from $F = 1$ back into $F = 2$ and a decrease in transmitted light intensity. The correction signal of the clock is proportional to the derivative of that microwave-induced resonant decrease in transmitted light intensity.

Figure 4.39: Schematic for a Rubidium-based atomic clock from Camparo, Physics Today, 11/07, p. 34.

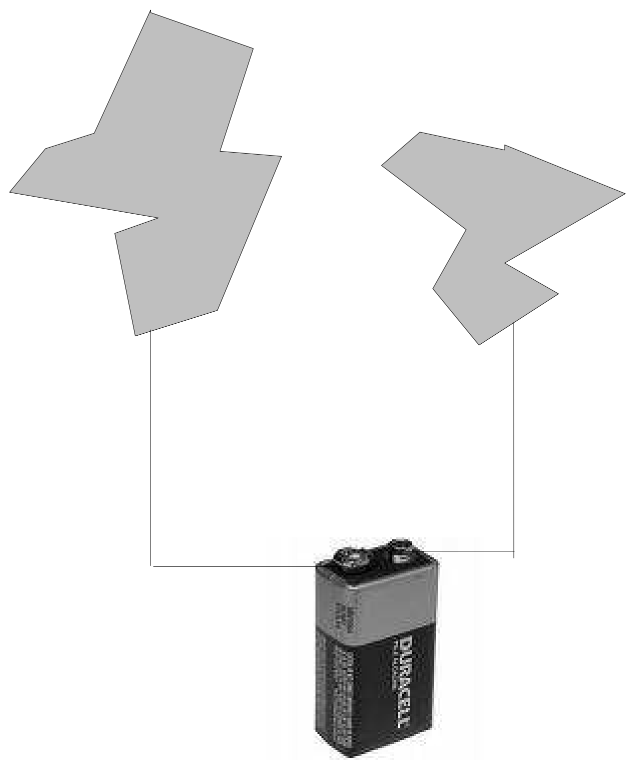


Figure 4.40: A simple capacitor: two metallic objects brought near each other, then connected to a battery.

everyday-reliability, electronics is the way to go. This section will look at three important areas in electronic timekeeping: RC circuits, resonant crystals, and digital electronics.

Resistors and Capacitors

Suppose you're interested in electronics, and sign up for "Electronics 101." During lesson one, you'll learn about batteries, switches, wire, and perhaps light bulbs. In lesson two, you'll learn about resistors and capacitors; and learn them well, because these two items are so important that they'll come up again and again, even after a lifetime of electronics training.

Capacitors are very simple by design. Just take two metallic objects, place them near each other (but not touching) and connect each to the opposite poles of a battery, as shown in Figure 4.40. After a very brief time, positive charge will flow from the positive terminal of the battery onto the metallic object to which it is attached. Likewise, negative charge will flow from the negative terminal to the other object. The situation will now resemble that shown in Figure 4.41.

Equal amounts of positive and negative charges will appear on each metallic object, and the opposite charges will try to get as close to each other as they can. They are constrained by the physical boundaries of the object they are on, and can only get as close as the spacing

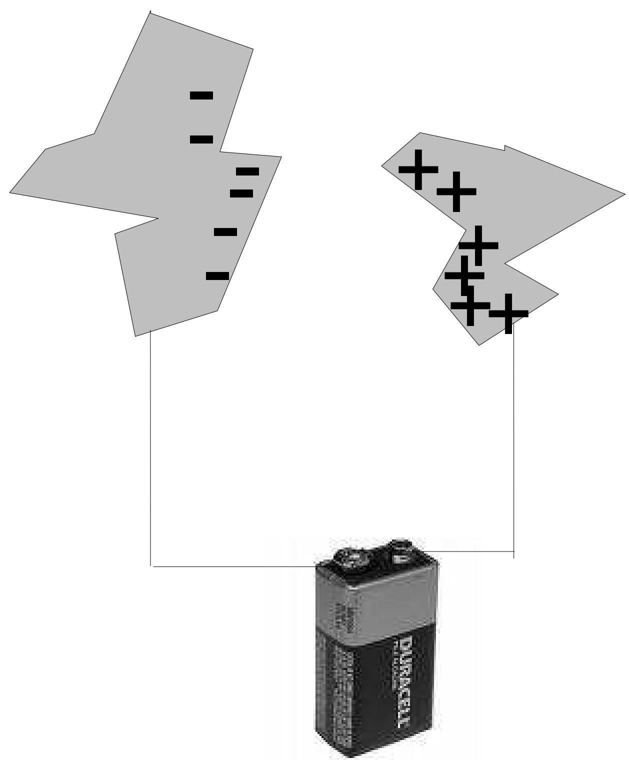


Figure 4.41: A charged capacitor.

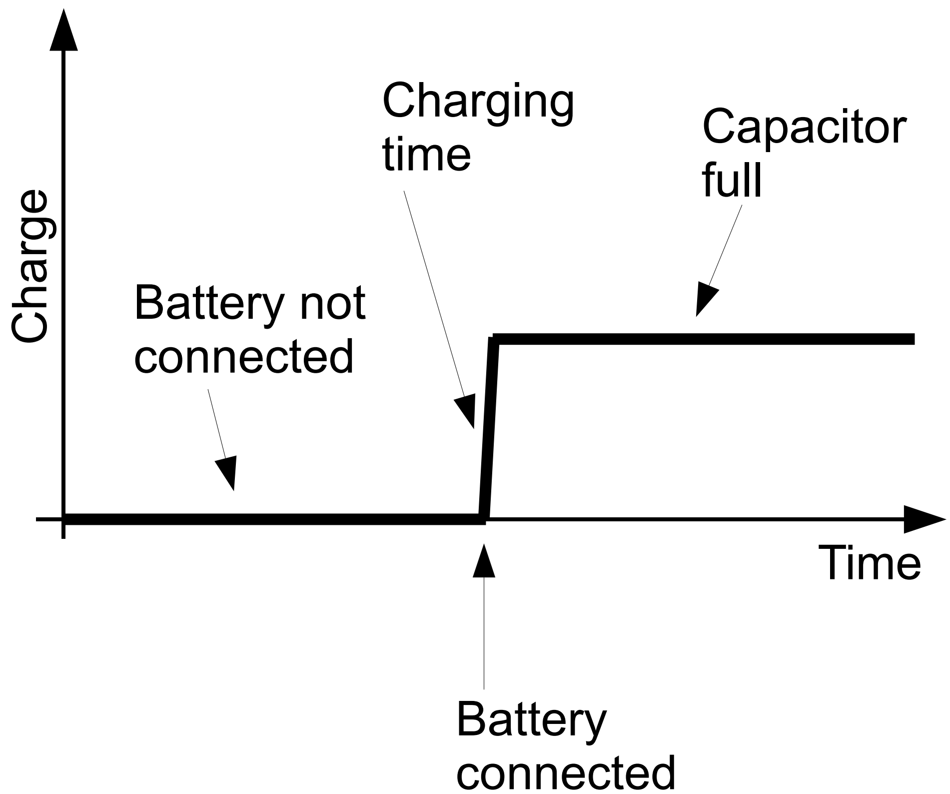


Figure 4.42: The charge as time evolves for a charging capacitor.

between objects. If the battery is removed, the charges will remain on the metallic objects, and this is the purpose of a capacitor: to store charge. A capacitor does for charge what a bucket does for water.

Although there is no obvious time keeping mechanism in a capacitor (yet), it is helpful in that it can only store a certain maximum amount of charge before becoming "full" or "fully charged." Just as a bucket can only hold so much water, a capacitor will only hold so much charge. Even if a battery remains connected, no more charge will flow out onto the metallic objects. As we'll see below, this charge limit is something like the maximum swing amplitude of a pendulum.

Suppose we graphed the amount of charge on the capacitor in Figure 4.41 as time went on. At first there is no charge on the capacitor. When the battery gets connected, charge quickly fills the capacitor to its maximum capacity and the "action" is over. A graph of this would look something that shown in Figure 4.42.

As you start where the "Charge" and "Time" axes cross and slowly go to the right, you see the interval where the battery is not connected. The heavy line is at "zero charge" since no charge is on the capacitor. Next, we come to the time when the battery gets connected. Quickly, but not instantly, the charge on the capacitor grows to its full capacity. This charge-growth stage is called "charging the capacitor." Finally, the capacitor is full, and nothing

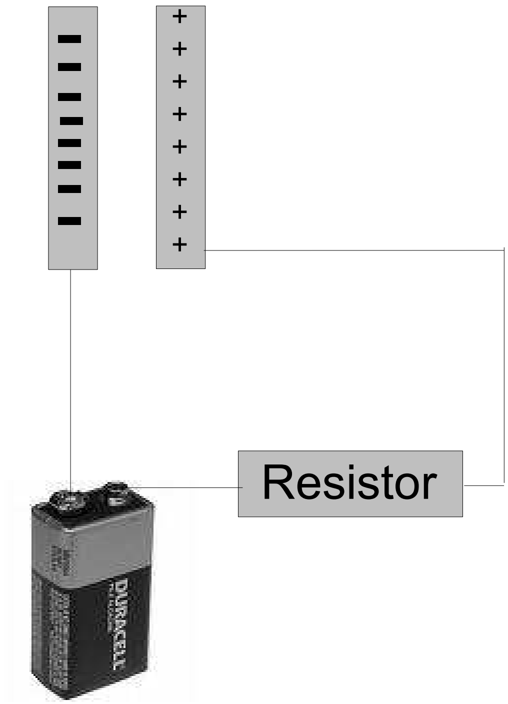


Figure 4.43: Slowing the charge rate using a resistor.

else changes. It refuses additional charge from the battery and the charge level stays at the same height. So a capacitor stores charge.

What does a resistor do? Simply put, a resistor slows the flow of charge. The higher the resistance, the slower the charge will travel. As for an analogy, think of a bunch of marbles as the flow of charge. Compare dropping the marbles through the air versus through a big tub of pancake syrup. The syrup does for the marbles what a resistor does for charge. For electronic timekeeping, we need to combine the slowing effect of resistors and the storage effect of capacitors to start electronic timekeeping. It works as follows.

Suppose then, you returned to Figure 4.41. It was mentioned that when the battery gets connected, the capacitor will accept the flow of charge to a point; then it will be full. Filling a capacitor with charge takes the briefest of time periods. On the order of nanoseconds (0.000000001 seconds) or less. Suppose that after charge begins to leave the battery, you forced it to flow through a resistor first, on the way to the capacitor, increasing the time it takes the charge to fill the capacitor. The circuit for this is shown in Figure 4.43.

Notice how all charge that flows from the battery needs to pass through the resistor first. In a case like this, the charging graph might now resemble that shown in Figure 4.44. Notice how the charging region is much more gradual than in Figure 4.42. The resistor has slowed the charging process. It turns out that we can choose any appropriate values for the resistor and capacitor to tailor the charging curve to any value we might need.

Suppose we chose the resistor and capacitor so that the capacitor went from zero charge to

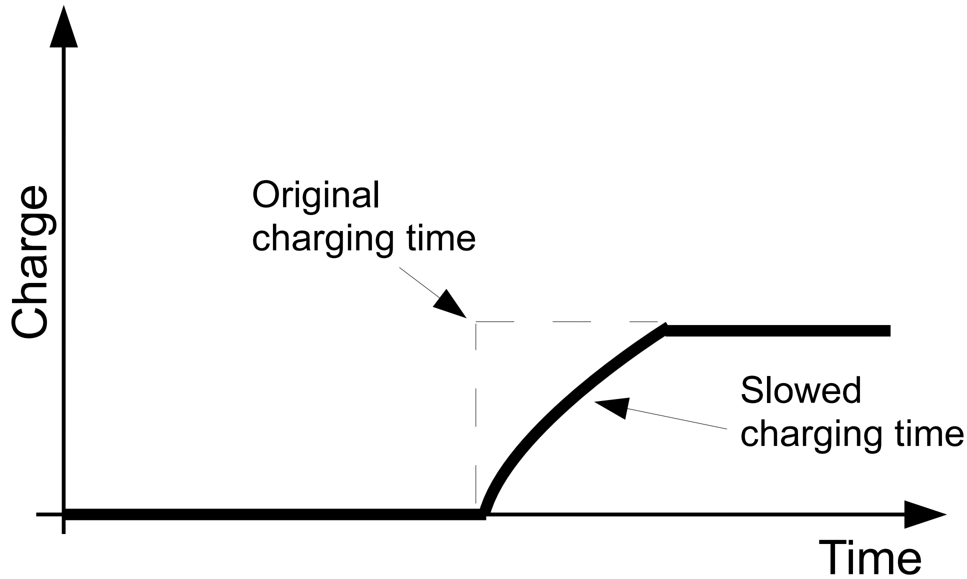


Figure 4.44: A slowed charging cycle for a capacitor.

fully charged in exactly one second. In other words, we'd always know that with the specially chosen capacitor and resistor, the amount time that will elapse from when we connect the battery to when the capacitor is fully charged, is exactly 1 second. In other words, we've just built an electronic, one second time base!

So where can a ticking clock come from? We need one more element. We'll call it a "charge monitor," but it is actually a transistor, which is an electrically controlled switch. The charge monitor circuit is shown in Figure 4.45. Notice that Point A is connected to the positive side of the capacitor, and point B the negative side.

Assuming the capacitor started with no charge on it, here's what happens. The battery is connected and charge begins to flow to the capacitor, at a controlled, one second rate, due to the resistor. Point A of the charge switch is constantly monitoring the charge level on the capacitor. When the capacitor reaches its "full" state, the charge switch does two things. First, it tells a second hand on a clock face somewhere to advance by one, since we know one second had just gone by. Second, it does something very clever. It internally connects Points A and B together. What does this do?

We know capacitors can store charge by keeping them separated on two different metallic objects. When the objects are brought together though, either by allowing them to touch, or by connecting them with a conducting wire, the positive and negative charges quickly neutralize each other, and the capacitor returns to a zero-charge state. So, when the charge switch internally connects points A and B together, the capacitor quickly becomes neutralized. So what happens now? The capacitor is "empty" and starts accepting charge from the battery again. In other words, the one-second charging cycle starts all over again! After another second, the capacitor will be fully charged, will advance the second hand, and

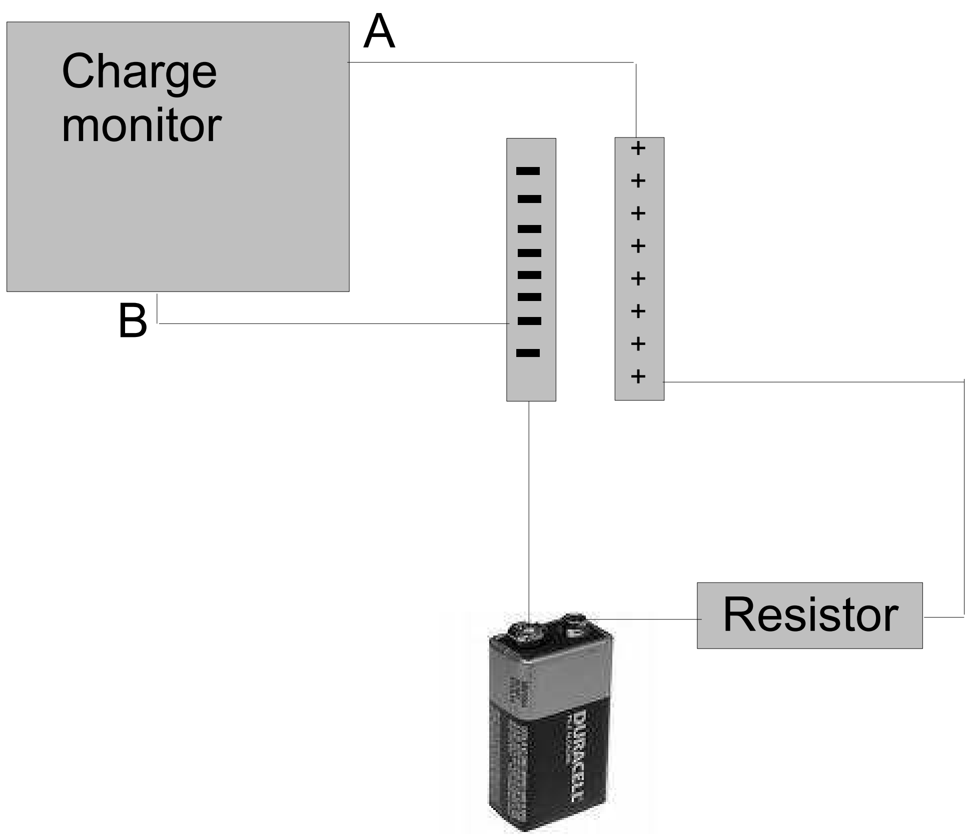


Figure 4.45: A circuit that will make a resistor-capacitor circuit “tick.”

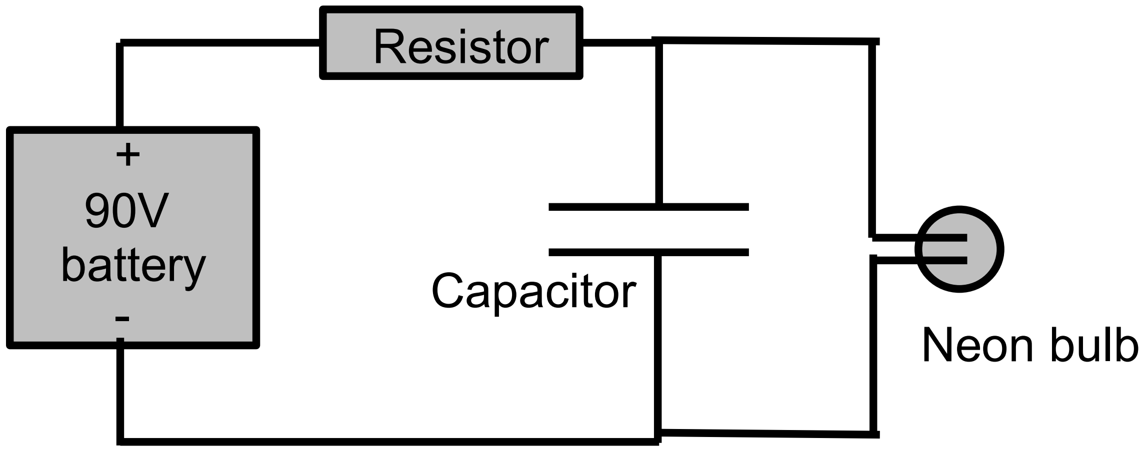


Figure 4.46: A circuit that will flash the neon bulb with fairly decent regularity. It's called a "relaxation oscillator."

will become quickly discharged again. We've made a "ticking" clock out of resistors and capacitors.

4.6.4 Electrical circuits that have a period

Remember that in order to keep time, you have to have something that moves or "beats" with a stable period, like a pendulum or a mass-spring system. The basis for an electrical clock is some kind of circuit that "beats" as well. In the case of electronics however, the beating would be a periodic surge in current or voltage, that could then drive a digital display on a clock, for example. When a circuit "beats," it is said to "oscillate." Here are two electrical circuits that oscillate.

Relaxation Oscillator

Figure 4.46 shows a simple electrical oscillator. Assume the neon bulb stays off, conducting no electricity, unless it sees a voltage across it of 90 Volts. At 90 V, it will turn on, drawing as much charge as is available to it. The neon lamp is like the "charge monitor" in Figure 4.45. We know this circuit will oscillate because the neon bulb will flash on and off; this would be our time keeper. Can you describe why this circuit would oscillate?

Digital Electrical Clock Pulser

There is another example of a relaxation oscillator that can be built quite easily. Let's talk a little bit about digital electronics first though. In digital electronics, all voltages are either +5 V or 0 V. In computer lingo, the +5 V might mean "true" or binary 1 and the 0 V might mean "false" or binary 0. The true and false are sometimes called "logic states." In this digital world, there is a clever device called an inverter that, as you might guess, flips one of

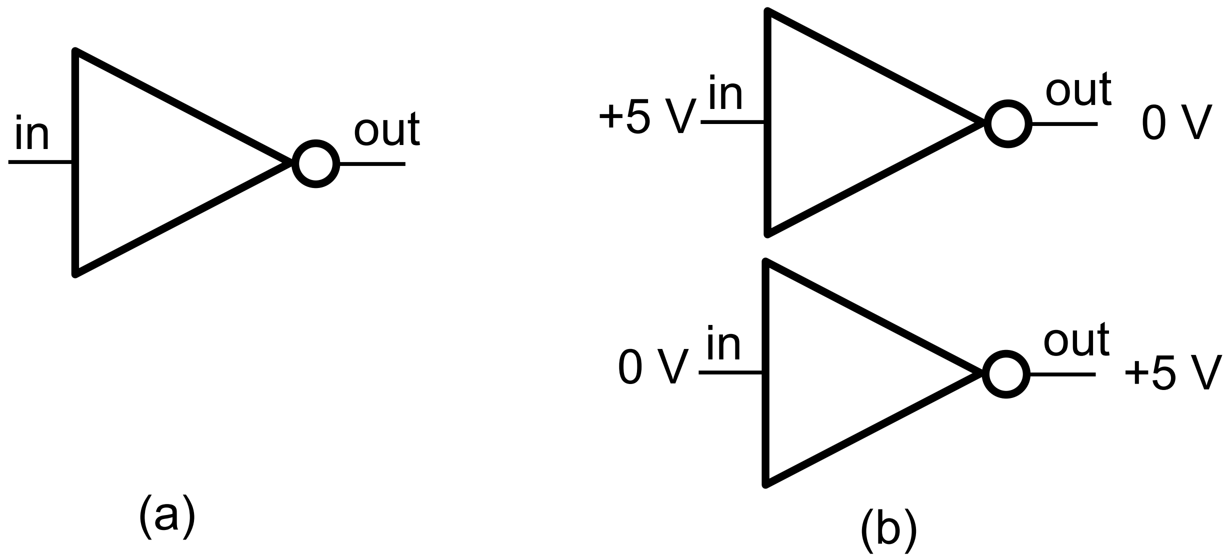


Figure 4.47: A logic or digital inverter. It flips a +5V to 0 or a 0 to a +5V.

these logic states into the other. That is, it converts a +5 V to a 0 V or a 0 V to a +5 V. That is, an inverter converts a true to a false, and a false to a true. The electrical symbol for it is shown in Figure 4.47(a) and the state-inversion process in shown in Figure 4.47(b). It turns out that flipping states like this is very useful in making an electronic clock. The rate at which a true becomes a false, then back to a true again would represent one period of the electronic clock.

Given the simple functionality of an inverter, what do you think would happen if the output of an inverter was wired to its own input, as shown in Figure 4.48? Suppose you built the circuit and flipped it on, and as luck would have it, 0 V was first to appear on the input of the inverter. The 0 V would be sent through the inverter, becoming a +5 V on the way out, which is fed back into the input, which would get converted into a 0 V on its way out. This would get fed back into the input, becoming a +5 V on the output, and on and on. See where the oscillation comes from? If we got a third wire to monitor the output of the inverter, we'd see a rapid succession of +5 V, 0 V, +5 V, 0 V, +5 V, 0 V, in other words, an electrical oscillator! Or, if we wired an LED onto the output, it would flash, just as the neon bulb did for the relaxation oscillator above. Each flash would be a like the swing of a pendulum. This could be considered an electric clock, and for a typical inverter the frequency, would be about 1 million times per second (1 Megahertz).

To make the frequency a little more useful as an everyday time keeper (perhaps to time seconds) a few components could be added to the circuit as showin in Figure 4.49.

As you can tell, there are two inverters now, that are somewhat tied together, meaning the output of one seems to feed the input of the next (read from left to right). And notice the output of the right inverter is fed back around to the input of the first (thus the output is still tied to the input of the system, as was done in Figure 4.48). What sits in between

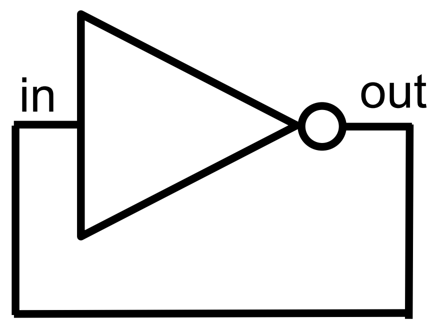


Figure 4.48: Feeding the output back into the input of an inverter give you a basic electrical oscillator.

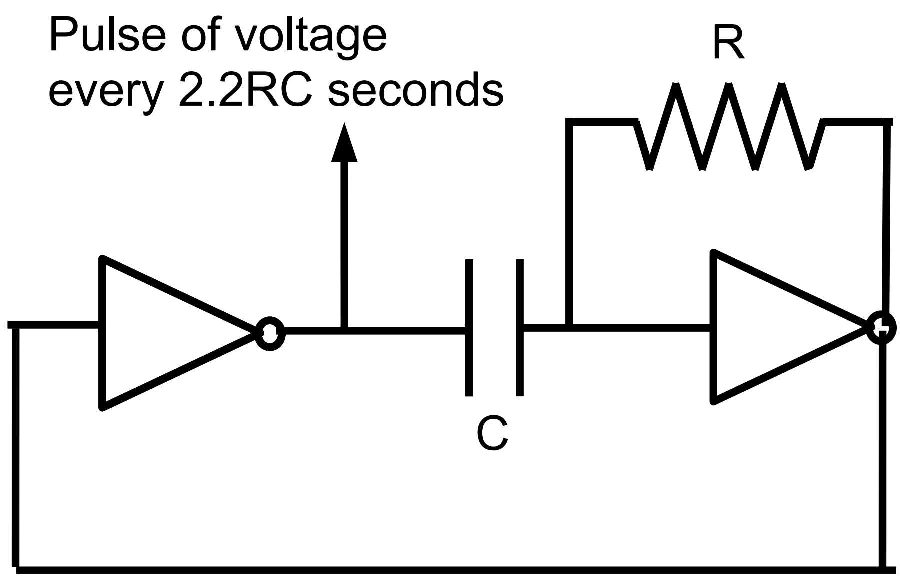


Figure 4.49: With R=10 kilo-ohms and C=100 μF, this circuit will have a period of about 2.2 seconds.

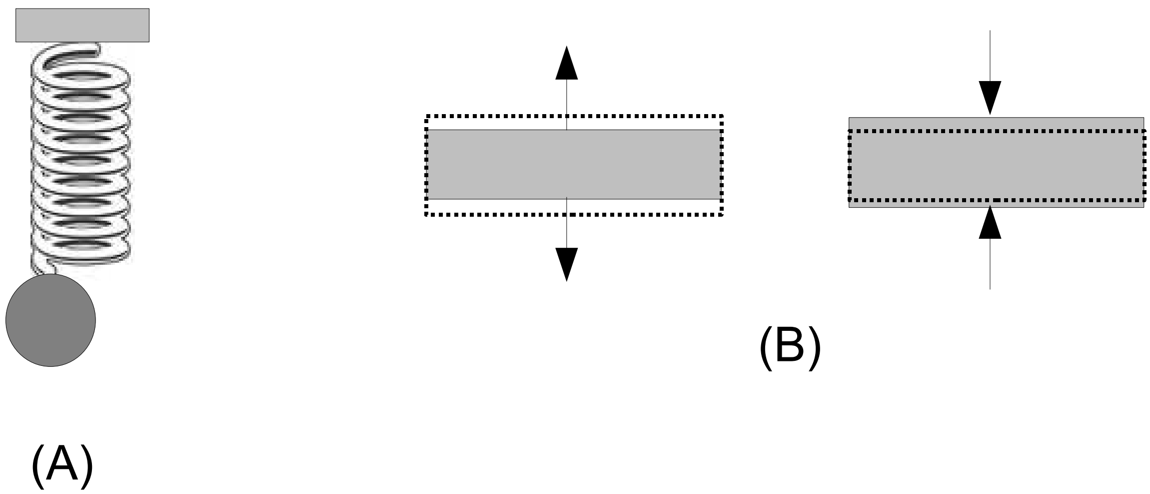


Figure 4.50: How the expansion and compression of a crystal (B) is like the mass-spring system (A).

the two is, you guessed it, a resistor/capacitor circuit, like the one discussed above. Suppose when you power on this circuit, a logic +5 V (or logic “1”) happens to appear on the input of the leftmost inverter. This means 0 V (false) will appear on its output. So the left plate of the capacitor is at 0 V and if you follow the wire, the right edge of the resistor will be at +5 V. This means the RC circuit will begin charging with a timed curve like that shown in Figure 4.44. When the capacitor voltage (tied to the input of the rightmost inverter) reaches +5 V, the rightmost inverter input will be +5 V (true), and its output will be 0 V (false). This will cause the input to the left inverter to be 0 V (false) as well (follow the wire), causing the output of the left inverter to go to +5 V (true). Now the left plate of the capacitor will be at +5 V (true) and the right side of the resistor will be at 0 V (false); just reverse from how it all started! The capacitor will now discharge slowly until its voltage (and connection into the right inverter) drops to below +0 V (false), starting the timing cycle all over again. The conclusion: if you attach an LED to the output arrow shown, it will flash on and off. Congratulations, you’ve just created a visible, flashing electronic clock!

Activity

Build relaxation oscillator. Use stopwatch to find period. Try to find error in the circuit’s period. Time walking excursion with stopwatch then with circuit.

4.6.5 Crystals

For electronic clocks, nothing beats using a crystal, like a quartz crystal. Here’s why. Remember the mass-spring system from Section 4.4.2? Well it turns out that a crystal is essentially the electronic analog of a mass-spring system. To see why, take a look at Figure 4.50.

In Figure 4.50(A), you see a reminder of the mass-spring system. When the round mass is pulled down slightly, it will oscillate at a fixed period of $T = 2\pi\sqrt{m/k}$, as discussed in Section 4.6. Figure 4.50(B) doesn't look like much, but think of the rectangular boxes as a small quartz crystal. In appearance, it might look like a small cube of glass (it's made from the same material as glass, fine, white sand, or Silicon Dioxide). Now, believe it or not, if the edges of the crystal are pulled as in Figure 4.50(B) [left], it will snap back to its original position, just like the mass-spring does after it is pulled. If the edges of the crystal are pushed on (the crystal is compressed), as in Figure 4.50(B) [right], they will want to expand back out to their original position, just like a spring would if it were compressed. Also just like the spring, as the edges try to return to their original position after being pushed or pulled, the crystal will overshoot its equilibrium position and keep oscillating back and forth (just like the mass-spring system or the pendulum).

Now this all might be a bit hard to believe. A small and very hard, glass-like crystal oscillating in response to being pushed or pulled? Yes! But the stiffness of a crystal is very very high. Think of a crystal as a very, very stiff spring. It is like a spring with a very large k. This means the period, T is going to be very small. For example, the T of a typical quartz crystal is 0.00003 seconds. So if you tap on a quartz crystal, it will oscillate back and forth $32,768$ times per second, or it will take 0.00003 seconds to expand out, then contract again. The range of its motion will be something like 10 micrometers, or 10 millionths of a meter. So you'd never see a crystal vibrating, but interestingly enough, the crystal can still push on air and $32,768$ Hz is only a factor of two outside of the range for human hearing. So for a crystal that oscillated somewhat slower, it might be shown to be vibrating by *listening* for it!

So the oscillation is pretty fast and is normally called a "vibration" instead. Also, the amplitude is pretty small. Compare these parameters to a pendulum, or mass-spring system that might oscillate up and back maybe 1 time per second, with very visible and measurable amplitude of several centimeters or more. And, because the crystal isn't sliding across anything, or moving through much air, it will vibrate for a very long time. Also, the thermal expansion coefficient of a crystal is very small, so it isn't greatly affected by temperature changes. For a clock, a crystal gains or loses less than a tenth of a second per year[Bloomfield, p. 300]. Wouldn't John Harrison be envious!?

So how is a crystal used in a clock? Well, it turns out when the crystal is compressed or expanded, the natural electrical equilibrium of the crystal's internal structure (electrical by design: protons and electrons) gets upset and a voltage will appear across the crystal when it is not in equilibrium. The farther from equilibrium, the higher the voltage will be across it. This is called "piezoelectricity" and is very common. So if electrodes are placed across the vibrating crystal, a voltage can be tapped off of it, as shown in Figure 4.51. In this figure, the dotted horizontal lines are equilibrium positions of the crystal. For the compressed and expanded forms of the crystal, a voltage appears.

Now just like the atomic clock, the period of the voltage will be very stable, since it originates from the crystal itself. Crystals typically keep time to about 0.001 seconds per day, or about 1 second per week[Itano, ibid.] Further, since a crystal in this context is an electrical device, it can be monitored with an electronic circuit; in particular, one that

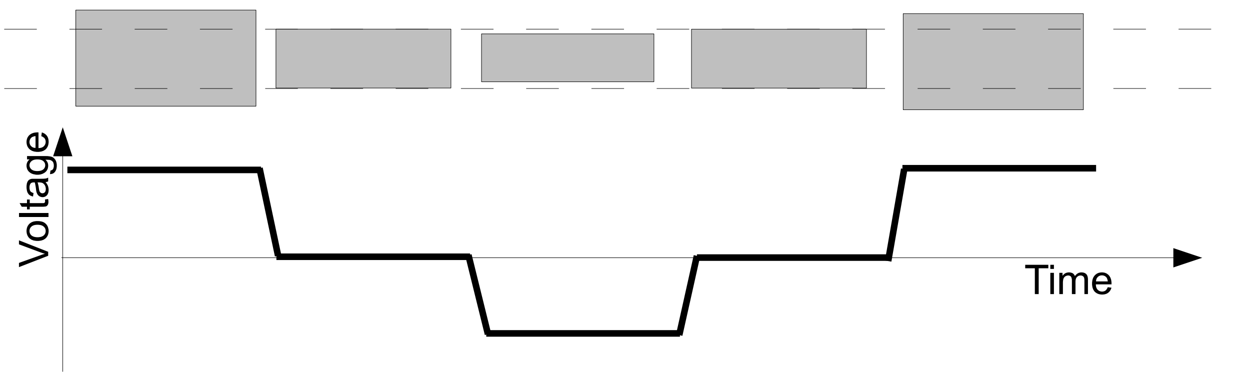


Figure 4.51: Observing the voltage across a crystal as it oscillates. The dotted lines are the equilibrium positions of the crystal surfaces.

divides the frequency, say by $32,768$ to give a period of one second, which can then drive an electronic counter that would count seconds. Crystals are small in size (4 mm x 0.3 mm) and use very little power; they're ideal for portable clocks and wrist watches. This is indeed what happened in the 1970s: LEDs and electronics miniaturization were "all the rage" back then, and wrist watches with big red digits on them began appearing for about $500. Just how a crystal is used in a circuit is as follows.

In use, a crystal doesn't directly drive an electrical oscillation; it sits in a circuit that already oscillates, in a manner that allows it to *regulate* the oscillation, to some ultra-stable mode. Here's an example. Remember the simple inverter circuit shown in Figure 4.48 above? It turns out that the frequency of this circuit is not be all that stable. The 1 Megahertz oscillation would be the average frequency, but could easily vary by up to 25% on either side of this number. Why the instability? For one, a 0 V might be sent back around before the inverter has "recovered" from the +5 V that formed the 0 V in the first place (called "pile-up"). This would result in missed conversions. The sequence might then look like 0,5,0,5,5,5,0,5,0,0; the string of 5's would be like a pendulum getting stuck at one extreme of its motion from time to time. Second, an oscillation forced upon an inverter like this would be terribly susceptible to temperature variations and local electrical noise. It is something of an electrical clock, but it would need to be stabilized. This is where a crystal could be used.

To stabilize this oscillation, we would put a resistor and a crystal into the circuit as shown in Figure 4.52. Here's the way this would work. Suppose a +5 V coming out of the inverter corresponds to some electrical current flowing through the wire. As it travels down the wire, it has a choice of paths when it gets to point "A" in Figure 4.52. When encountering a junction, current always splits between paths according to how much resistance it encounters, always preferring to find a "path of least resistance." If, for example, the resistor has twice as much resistance as the crystal, then twice as much current will flow through the crystal than the resistor.

When the circuit is first turned on, it is impossible to predict what fraction of current is

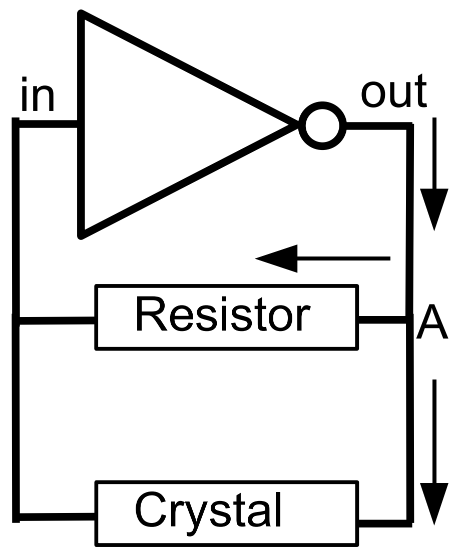


Figure 4.52: Add a crystal to stabilize the oscillation.

going down what path. We really have no idea what the crystal is doing at this point, but certainly the amount of current (large? small?) going through the resistor is enough to keep the oscillator going, as described above (however unstable it may be).

The instability of the circuit turns out to be good thing for the crystal, and a funny thing happens after the circuit has been on for a little while. Whenever the unstable oscillation happens to “beat” at the frequency of crystal was made for ($32,768$ Hz for example), the resistance of the crystal suddenly gets very very small; much, much smaller than that for the resistor. This is a very dramatic electrical effect.

This means that just about all current coming out of the inverter will decide to flow through the crystal, rather than the resistor. And remember, this is current pulsating at the precise frequency for which the crystal was cut. This effect is so strong, that other unstable periods simply are not allowed to exist anymore. The crystal has “bullied” them out of the circuit, and these circuits are sometimes called “crystal disciplined” circuits. As long as the current oscillates at the right frequency, it will see the ultra-low resistance in the crystal branch, and as mentioned, electrical current likes to see a low resistance through which to travel. Suddenly then, the oscillation in the inverter circuit gets very stable, with a period set by the crystal itself. In this stable mode, the frequency in the circuit is said to be “on resonance” with the crystal, and the crystal will be happily compressing and expanding to it fullest extent.

Finally, one can purchase crystals for about $5.00 that are manufactured to “like” oscillating between 10,000 times per second and up to 3 million times per second. Also, crystals are sold by their known stability as well. One can buy, for example, a crystal known to be stable to 10 ticks per million over the course of a year. This means, over the course of operating for one year, 10 ticks per million might be bad (out of synchronization). This is something like losing 1 millisecond per month. Their uses are mainly in watches, clocks for

microprocessors, and video displays.

Chapter 5

Modern Ideas on Time and Navigation

5.1 Introduction

Believe it or not, in the early to mid 1990's, the "Problem of Longitude" reached a near pinnacle in popularity. Really? A navigational problem from 300 years ago peaking in popularity in the 1990s? Yes really! In fact, in 1993, an entire conference on longitude was held in Cambridge, MA. Called the "Longitude Symposium," over 500 people from seventeen countries attended. It is indeed an odd topic for an international conference; some of the attendees (in their writings) allude to jokes being thrown about as they told people about the conference to which they were traveling. Why would we need a conference dedicated to finding longitude? A couple of years later, the Dava Sobel book on Longitude appeared, pushing the popularity of this topic to its peak, allowing the general public to read about the story. An intriguing event happened at the conference, as outlined in the conference proceedings, called "The Quest for Longitude." Although the proceedings, (in book form) is now out of print, and expensive to purchase, it is the absolute authority on this topic, and required reading for anyone interested in this topic. It is simply wonderful to read. The event was as follows.

During one of the concluding sessions, a physicist in attendance rose and pulled a small black box from his pocket. He pressed a button on it, and read some numbers from its small screen:

$$\textbf{Latitude: } 42°22.546', \textbf{ Longitude: West } 071°6.904'$$

"No big deal," you're thinking. "So the guy had a GPS (Global Positioning Satellite) receiver, that anyone can buy now (2008) for under $200; my friends and I geo-cache with one almost every weekend." Well, back in 1993, these handheld devices were just beginning to appear, and were very expensive. This was an amazing feat, amplified more so by the nature of the conference.

During my research on this topic, I continue to be thoroughly impressed by the writings of William H. J. Andrewes, the chair of the Longitude Symposium. I wish then, to quote

some of his words from the introduction to the conference proceedings, in response the the "button pushing" event just described.

> *With this simple act, he[the physicist with the GPS] demonstrated how advances in science and technology had reduced the problem of finding longitude–once the most crucial challenge facing every seafaring nation–to a small "black box" that provides a precise answer almost instantaneously on a miniature screen. Like so many products of modern technology, the Global Positioning System receiver conceals the nature of the problem being solved: the theoretical complexity, the technical difficulties, and the untold thousands of hours of thought and labor extended to tell us, so simply exactly where we are. Technology has become such an integral part of our lives that we tend to take it entirely for granted. It has raised our expectations, changed our priorities, and left us perilously ignorant of how helpless we would be without it. Traveling overseas, we now complain when delayed for an hour: we have forgotten that once there were problems finding continents.*

Quotes like this help to remind us how far science and technology have allowed us to go. Although time and longitude are old, and completely solved problems, there are some modern ideas about them, and we shall cover two of these ideas in this chapter.

The first idea is about time. When you think about it, what more do we need to know? We know how to keep it accurately, and we know that time is critically linked to finding longitude. Time goes forever forward, and our computers can even reset their times internally when day-light-savings rolls around. We can keep track of small times and large times, etc. What else is there? Well, as you probably know, there's always a danger in assuming you "know something." Whether it's arrogance or a way of stifling creativity, we should never assume we "know all there is to know." And time is no exception.

Enter Albert Einstein, in 1905, perhaps the most brilliant scientist ever. In that year, with the longitude problem readily solved, and reliable clocks and watches being mass-produced, it's quite possible that no one was thinking about time *at all*. But Einstein put forth a view on time that continues to jolt the scientific community even today. In fact, few people ever have or will ever fully understand time at the level Einstein did. The first part of this chapter will discuss Einstein's view of time.

The second modern idea is about the Global Positioning System. By looking at a black box, you can know where you are on the earth to within a meter, just by pressing a button. How can this possibly work? Something about satellites? GPS, for sure, over-solves the longitude problem and takes all of the "fun" out of figuring out your location. Remember the chapter on Celestial Navigation? One author of a book used as a reference in that chapter (Hewitt Schlereth), joked about being careful to read the time as you mark the sight of a celestial object. He said something to the effect of "as you sight the object, read the time from your wristwatch (or GPS unit)." (Don't get it? Read the quoted sentence again, slowly.) It seems that all electronic devices these days have a clock built into their displays, and GPS units are no exception. Indeed the GPS unit, all at once, removes one's need for a watch, celestial navigation, the time in Greenwich, and in some cases, even a map itself! To

be safe though, the same author also sarcastically references dead batteries or dropping the GPS unit into the water.

The operation of a GPS system probably ranks as one of the most complex and multi-faceted systems ever conceived. It draws from the basic to the most esoteric ideas in physics and mathematics, and somehow still works. In the second part of this chapter, we'll delve a bit into the inner workings of this system.

So, Einstein and GPS; two modern twists on time and navigation. Let's take a look.

5.2 Albert Einstein

Although there is no useful science between the two, we find it curious that the longitude problem and one of Einstein's theories were both centered around the same basic problem:

Knowing the time at two different places, at the same instant.

In the longitude problem, of course, you needed to know the time at Greenwich and the time at your location at sea. Einstein's thoughts on time were based around simply how one expects to know two times at two different places at all. Here's how Einstein's ideas on time go.

Activity

Mechanical universe on length contraction and time dilation.

5.2.1 Einstein's Ideas on Time

You probably have a pretty good idea about what a clock is at this point. Something that goes "tick-tock" and keeps track of hours and minutes. There are a variety of ways to make a clock (electronic, mechanical, or even atomic). Since there are essentially an infinite number of ways to make a clock (Wikipedia lists about a dozen escapements alone), for the sake of the discussion here, we'll make our own clock. Let's make a clock out of two mirrors and a pulse of light. It will look like that shown in Figure 5.1

Here's the way it works. By some means, we shine a pulse (or flash) of light into the space between two mirrors that face each other. As shown in Figure 5.1(a), a pulse of light is traveling upward, toward the top mirror. In (b) it strikes the upper mirror, giving us a "tick." Since mirrors reflect light, our pulse's direction is now abruptly reversed, and it begins to travel back down to the lower mirror. As before, when the pulse of light strikes the lower mirror, our clock goes "tock." So, the "tick" and the "tock" serve as our timebase. To track this, perhaps we have some kind of detector embedded in the upper and/or lower mirror that updates an internal counter each time the light hits it. The internal counter is where we'd read the time. As for the period of our clock, we know that time traveled is distance divided by velocity, or $t = D/v$. Here D is the distance between the mirrors and $v = c$, with c being the speed of light. So for example, if the mirrors are 1 foot apart, for

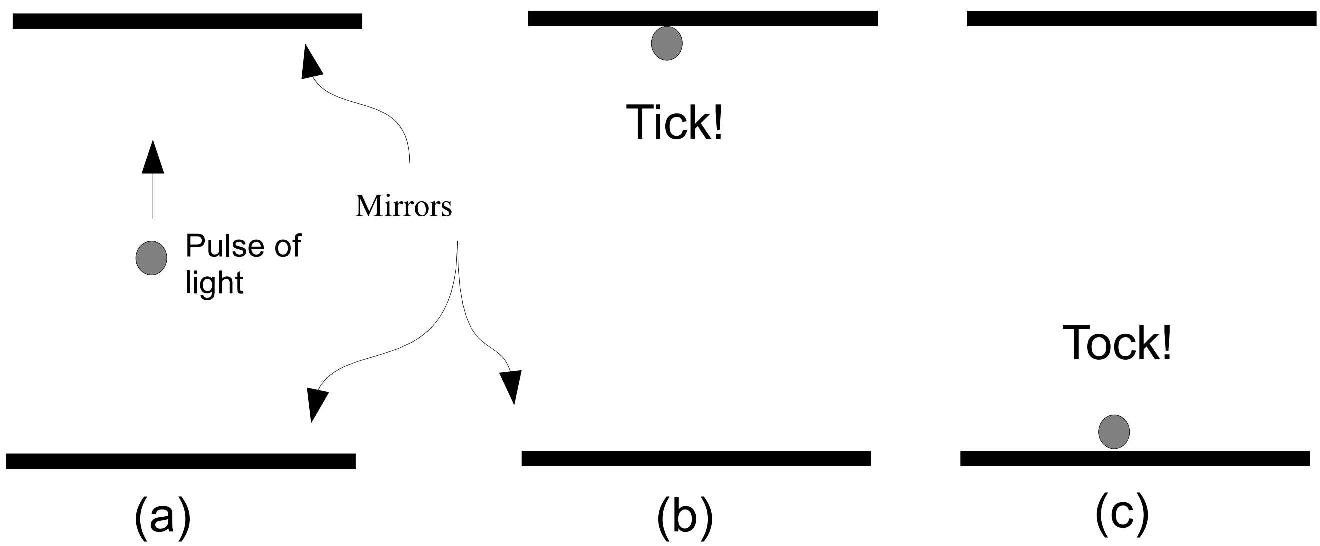


Figure 5.1: A clock we've invented that "ticks" and "tocks" when a pulse of light bounces back and forth between two mirrors.

example, and there's a detector only in the upper mirror, then our clock would have a period of about 2 nanoseconds (since light travels about 1 foot every nanosecond).

A couple of notes before proceeding. First, our clock doesn't actually make a "tick" or a "tock" sound (light reflecting from a mirror is perfectly silent). We are just trying to emphasize how our clock keeps time. Second, bouncing light between two mirrors isn't as crazy as it sounds. This is exactly what goes on inside of a laser, and microwave-based lasers, or MASERS, are still routinely used as ultra-precise clocks.

Now, suppose we make two identical clocks and start them so that they are exactly synchronized. In other words, they both "tick" and "tock" at precisely the same moments (One could be the clock at Greenwich and one could be the clock that goes out on our ship). Suppose we wanted to mass produce our clocks, but first wanted to verify that our manufacturing process was reliable, and that we are able to make two identically accurate clocks. So, we call in two horologists to verify that our two clocks are indeed identically synchronized.

Horologist #1 (or H1 for short) is standing exactly midway between both clocks, as shown in Figure 5.2. At a particular instant, the clocks are in their "tick" position. H1 sees both clocks in this "tick" position and declares them to be synchronized. But our competitor is a bit skeptical, and wants to know just how H1 knows both clocks went "tick" at the same time. The reason given is because H1 looked at the clocks with his or hers own eyes. He or she declares that they saw the light hitting the upper mirrors of both clocks at the same instant.

But what does seeing mean? It means light traveled from each clock, into H1's eyes, carrying with it the scene of the light hitting the upper mirror (this is what vision is, after all). This is not a trivial issue either. Nothing can travel faster than the speed of light,

which is indeed a large speed, but it is not infinitely fast. This means that any information exchange, between two parties, must always take some amount of time; that is, it is not possible to relay information to someone or something *instantly*. In the case of viewing the clock reading from H1, it takes a certain (non-zero) amount of time for the light coming from the "tick" or "tock" event to reach H1's eyes.

So, what if H1 moved a bit to the right? Could they still declare the clocks to be synchronized? The answer is "no," because as mentioned, it would take a finite amount of time for information to travel from one the clock to H1's eyes. The visual information from the right clock will arrive a bit sooner than from the left clock, because it has a smaller distance to travel. It would appear as if the right clock is "ticking or tocking" a little earlier than the left clock. See how confusion on this matter is already setting in? Admittedly, the fact the person is standing closer to one clock or the other could be addressed, but this question of "simultaneity" of events is one of Einstein's postulates[Eisberg, p. 17]:

> *Two instants of time, observed at two different points, are simultaneous if light signals emitted from the two different points arrive at the midpoint simultaneously.*

Indeed, declaring something to be simultaneous is a bit difficult. Think even of what it means to declare that "the train arrived at 7 o'clock"[Eisberg, p. 16]. This statement is actually an abbreviation for "the train arrived and the little hand of a nearby clock pointed to 7 simultaneously." In our everyday lives, declaring two things to be simultaneous that are near each other, requires no further discussion. If you see a train pull in and the clock on the other side of the station says 7 o'clock, you would be happy. Even if the horologist (H1 above) was a bit closer to one clock or the other, he would observe the closer clock to tick in the range of nanoseconds (0.000000001 s) earlier than the other. Again, no big deal for our everyday lives, which is why all of Einstein's simultaneity arguments seem so useless. But as we probe deeper into the matter and invent more and more sophisticated machines, these things begin to matter.

So, according to Einstein, to fully classify the clocks as simultaneous, H1 *must* stand at the mid-point between the clocks. So let's say they do so, and declare the clocks to be synchronized. Great! Einstein still wondered though, what would happen if one of the identical clocks was moving, while the other one was sitting still. This sounds silly, for who cares about a moving clock versus a clock sitting still? They're still "ticking and tocking," right? Well yes, but here's where we might get into trouble.

Let's ask horologist #2, or H2, to examine our clocks. H2 agrees to do so, but is very busy and can only look at the clocks from a moving car, as they drive by, on to their next appointment.

Let's say the clocks both go "tick" just as H2 drives by the midpoint between the two. At this instant, the visual light from both "tick" events begins traveling toward H2's eyes. While the visual signal is traveling (taking its non-zero time to reach H2), H2 advances toward the right in the car, moving at speed v (for velocity) to the right. As shown in Figure 5.3, the light visual from the right clock will get into H2's eyes first. Confirming its "tick." A bit later, the "tick" visual from the left clock would arrive. H2 reports that the clocks are not

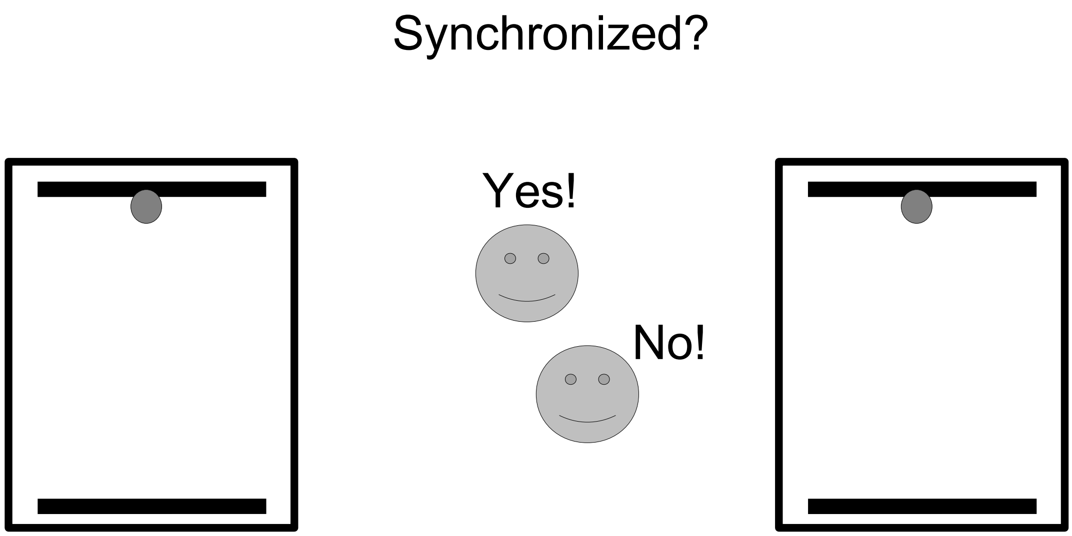


Figure 5.2: Our two light clocks examined by an observer standing at the midpoint between the two clocks.

synchronized. We plead with H2 to drive by and look again. They do so, and again declare the clocks to be unsynchronized for the same reason. Due to H2's motion, it is difficult to agree that the two "ticks and tocks" of the clocks are simultaneous.

So there you go. You called in two experts and they cannot agree on the simultaneity of your clocks. One reports that the clocks do agree and the other states they do not. Once again, this might sound silly, but there is something very peculiar about clocks that move.

5.2.2 Clocks on the move

Let's look more closely now at our light-based clocks. One is shown in Figure 5.4.

The mirrors are separated by a distance D, and light travels at speed c, which is 3.0×10^8 m/s. This means the time between "tick" and "tock" is the time it takes the pulse of light to travel between the mirrors which will be a time t equal to $t = D/c$. As stated above, if $D = 1$ foot, then t will be about 1 nanosecond, for example. As long as you are at rest with respect to the clock, you will observe a period of 1 nanosecond between each "tick" and "tock."

Since there seems to be something funny about moving clocks, let's now put the exact same clock in motion toward the right, with speed v. A time-lapse picture of the moving clock might look like that shown in Figure 5.5.

The light pulse continues its up/down "ticking and tocking." But, since it is moving past us, we do not see the light going up and down in a vertical line (as shown in Figure 5.4, but rather in a diagonal line as shown in Figure 5.6.

The diagonal line traced out by the light pulse of the moving clock is longer than vertical line the light pulse travels through in the stationary clock. This is a funny distortion of the

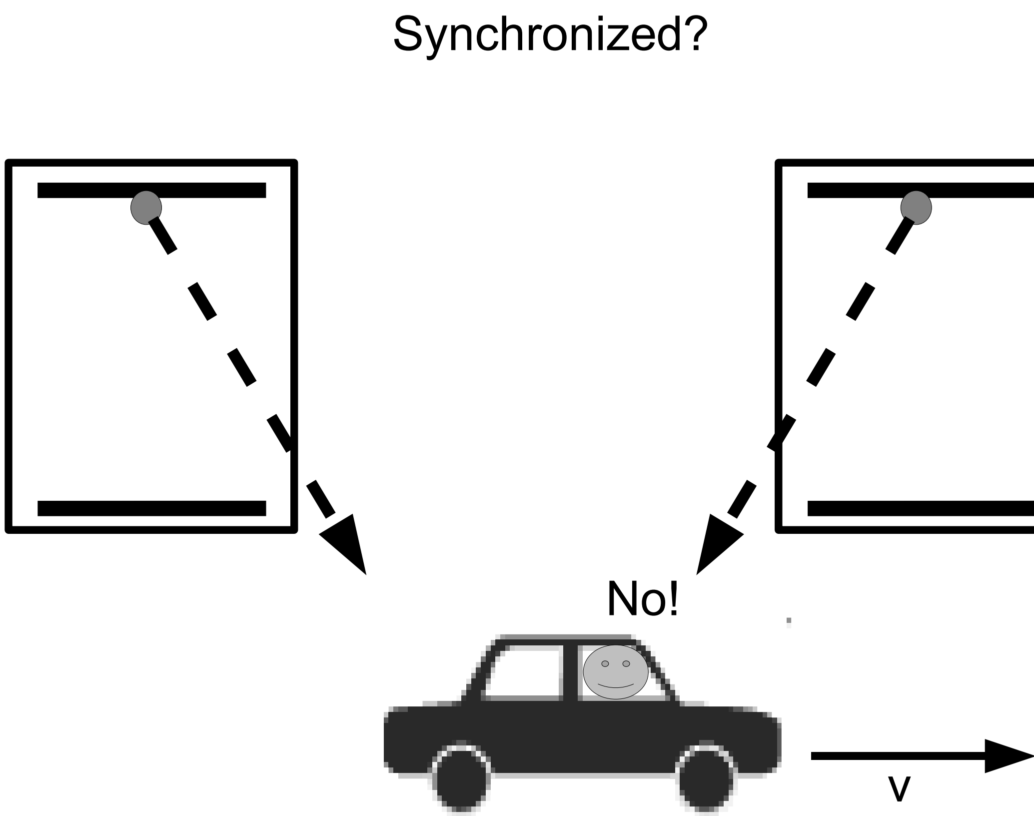


Figure 5.3: Our two light clocks examined by an observer moving past them in a car at speed v (for velocity).

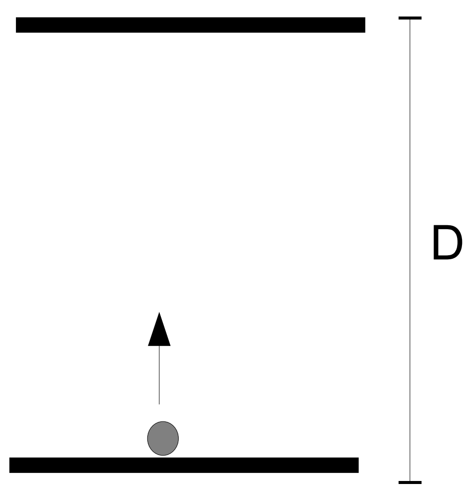


Figure 5.4: Our light clock.

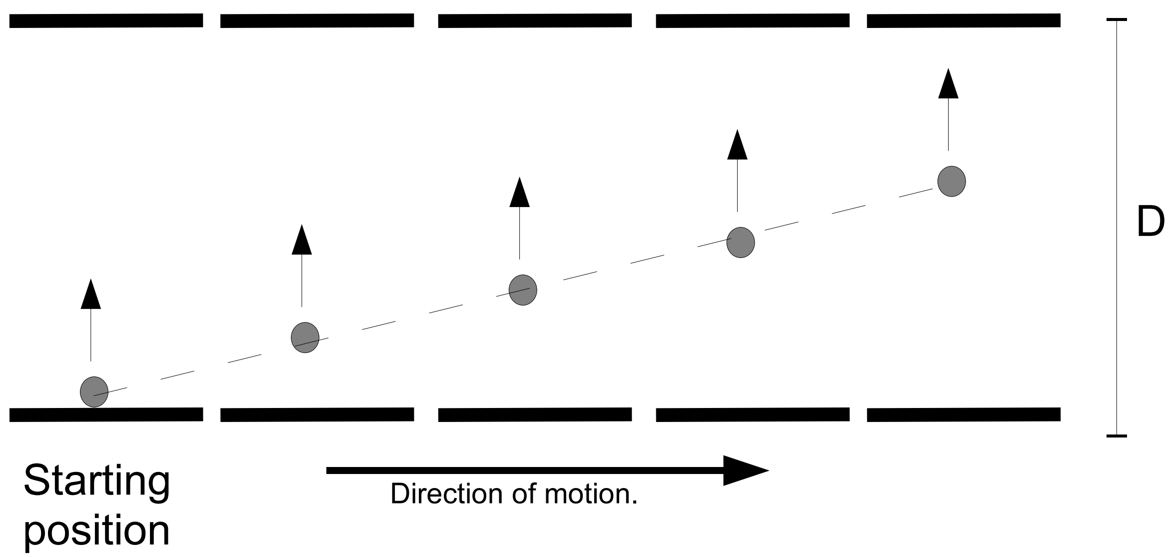


Figure 5.5: Our light clock on the move.

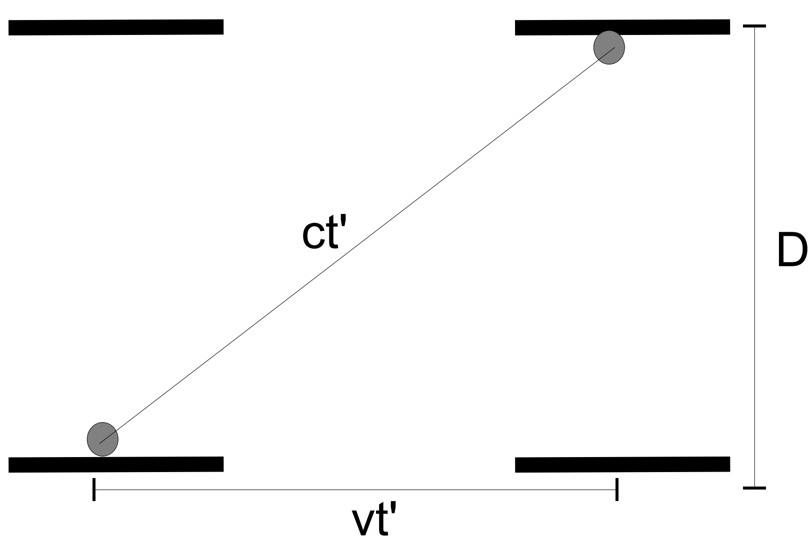


Figure 5.6: Our light clock as seen as it moves past us.

light pulse's apparent motion, owing to the fact that the clock is in motion. Notice that all told, the distances involved with the moving clock form a triangle (a right triangle). The diagonal (or hypotenuse) is set by the time it takes the light pulse to travel from the bottom mirror to the top mirror. Since light travels with a speed c, the length of this hypotenuse will be ct'. We are using the symbol t' here for the time as kept by the moving clock, because we should probably not claim at the onset that it will be the same time as t, the time of the stationary clock, since there appears to be some oddities with moving clocks. Continuing, the lower, horizontal line is how far the clock has moved during the time it took the light pulse to go from the bottom mirror to the top mirror, vt'. Lastly, the vertical distance is still the same distance between the mirrors, D.

For the stationary clock we were able to obtain an equation for its ticking time to be $t = D/c$. Let us see if we can find an equation for t', the ticking time of the same clock, but when its moving. In high school, we learned of the Pythagorean Theorem, which states that the sum of the squares of the legs of a right triangle is equal to the square of its hypotenuse. Let us apply that here to get

$$c^2t'^2 = v^2t'^2 + D^2. \tag{5.1}$$

We can solve this for t' to get

$$t'^2 = \frac{D^2}{c^2 - v^2}. \tag{5.2}$$

So, we have an equation for t', the ticking time of the moving clock. It is certainly different than that from the same clock when it wasn't moving ($t = D/c$). To check the equation thus far, let us set $v = 0$ to make the clock at rest again. If we do so, we'll get

$$t'^2 = \frac{D^2}{c^2 - 0^2}, \tag{5.3}$$

or just

$$t' = \frac{D}{c}. \tag{5.4}$$

So when the clock is at rest, $t' = t$ as we might expect it should. So we are getting somewhere. If only the second horologist could have just stopped their car! It appears as if the equation for our moving clock tick time, t' has some meaning, since it agrees with t when $v = 0$. To compare t' and t (the moving and stationary times), we can notice that if $t = D/c$, then $D = ct$, and we can put this in for D in Equation 5.2 to get

$$t'^2 = \frac{c^2t^2}{c^2 - v^2}. \tag{5.5}$$

This puts t and t' into one equation, so we can see how the two compare. A c^2 can be factored out of the bottom and canceled with the c^2 on the top like this:

$$t'^2 = \frac{c^2t^2}{c^2(1 - v^2/c^2)}. \tag{5.6}$$

and

$$t'^2 = \frac{t^2}{1 - v^2/c^2}. \tag{5.7}$$

Lastly, taking the square root of both sides, we get that

$$t' = \frac{t}{\sqrt{1 - v^2/c^2}}. \tag{5.8}$$

We apologize for the rather "mathy" few pages, but Equation 5.8 is our big result here, and a very famous one put forth by Einstein. It relates the ticking time of the stationary clock, t, and that of the moving clock t'. How do they relate? Well, if you look at the term v^2/c^2, you'll notice that this is a fraction that is always less than 1. Why? Because c is the speed of light, the fastest speed possible, and v for any object will always be less than c, so v^2/c^2 will always be less than 1. This means $\sqrt{1 - v^2/c^2}$ will also be less than 1 and greater than 0, since we are subtracting "a number less than 1" from 1 itself. Lastly, this means that

$$t' = \frac{t}{\text{a number less than 1}}. \tag{5.9}$$

Now if you remember your division, if you divide a number like t by a number less than 1, it will make t/a-number-less-than-1 a number larger than 1 as a result. Thus, t' is bigger than t! This means that:

The ticking time of the moving clock is longer than the ticking of the stationary clock!

or

The moving clock appears to tick slower than the stationary clock!

This is one of Albert Einstein's most famous results, and is called time dilation (dilation means "to become longer."). And remember this dilation result arose simply because the dilated clock was put into motion! Before proceeding, let's answer a few common questions that often come up when one first learns about time dilation.

So this means that if I see a clock that is moving, it is ticking slower than if it were standing still? Yes. To the best of your observing ability, a moving clock will appear to tick slower to you, than if it was standing still.

Say my friend is riding by me on a bike. Is his watch running slower than mine? Yes. As far as you can tell, your friend's watch is running slower. Your friend, however, won't think his or her watch is running slower.

Wait. Why won't my friend think their watch is running slower if I do? Because your friend isn't moving with respect to his own watch; their watch is moving with them on the bike. But guess what? Your friend will think *your watch* is ticking slower.

So we can't agree then? No, your observations won't agree, and this is the heart of Einstein's theory on time. Two people cannot agree on their observations on the rate of time if one is moving with respect to the other. You would need to use Einstein's time dilation equation (above) to translate the tick time of your friend's watch with your own, or vise versa. Using the equation is the only way you will agree.

Suppose my friend and I identically synchronize our watches. They then get on their bike and ride down to the end of the street and back. When he gets off, are you telling me that his watch will now be behind mind in time? Yes.

How big of an effect is this? For any attainable human speed, the time dilation effect is very, very small. Say a bike moves at 9 meters/s (about 20 miles per hour). The ratio of v^2/c^2 will be $9^2/(3\times 10^8)^2$ or about the size of 1 versus the number 1 quadrillion (1,000,000,000,000,000). This makes $1-v^2/c^2$ about 0.9999999999999991, and the square root of it about .9999999999999995. This means that t' is 99.99999999999999% the same at t. In other words, you won't notice the time dilation effect Thus, the effect is so small that you won't notice it, and nothing you can buy at your local hardware or sporting goods store can measure it either. Even for a 747-400 going 600 miles per hour (270 meters/s), t' will be 99.99999999999999849% the same as t.

Who cares about such a small effect? In everyday life, we don't. But as we know, advances in technology have taken us to smaller and smaller limits of space and matter. Our technology exploits the laws of quantum mechanics each day (transistors, integrated circuits, lasers, biochemicals, etc.) We are able to probe atomic structure and see what atoms are made of. Electrons move around the nuclei of atoms at nearly half the speed of light; this makes $v^2/c^2=0.25$, leading to $t'=1.15t$ in the time dilation equation. This is a 15% difference in time!

Has time dilation actually ever been observed using everyday equipment? Yes. Here are three examples:

- The first "macroscopic" test of time dilation came in 1971 with the Hafele and Keating experiment. To quote the abstract from their publication, in Science Magazine, Vol. 177, July 14, 1972, p. 166-168:

  > *During October 1971, four cesium beam atomic clocks were flown on regularly scheduled commercial jet flights around the world twice, once eastward, and once westward, to test Einstein's theory of relativity with macroscopic clocks. From the actual flight paths of each trip, the theory predicts that the flying clocks, compared with the reference clocks at the*

> *U.S. Naval Observatory, should have lost* 40 *nanoseconds during the eastward trip, and should have gained* 275 *nanoseconds during the westward trip. The observed time differences are presented in this report.*

- As silly as this sounds, a lot of "last generation" scientific equipment can be found on eBay, including ultra-precise atomic clocks. (New ones cost around $6,000 and are as big as a shoe box.) People can and do buy these and run them at home. A person at LeapSecond.com, left one clock at home and brought another on a camping trip. When he got back, the clocks were off by 22 nanoseconds. This was a dilation caused by an altitude difference between the clocks, but the point is that very small time discrepancies are routinely measured in "everyday life." He took his kids and said upon return "that was the best 22 nanoseconds I've ever spent with my kids."
- In outer space there are a plethora of "cosmic rays" flying around. Around here, many are ejected from the Sun in the form of protons (Hydrogen atoms with their electrons stipped off) and neutrons flying around at very large speeds. When a proton encounters the earth's upper atmosphere, it will collide with the nuclei of the N_2 and O_2 molecules floating around up there. When this happens, the proton and nuclei break apart, and one byproduct of this collision is a small particle called a muon. Now, muons are very unstable, and will eventually "dissolve away" (on average) in about 2.2 microseconds. This means, as they come out of the collision at about 99.94% of the speed of light, they would travel about $2.2 \times 10^{-6} \times 0.9994c = 660$ m, or about 2,000 feet before dissolving.

  The puzzling thing about muons though, is that we detect their presence down here on the surface of the earth, $10,000$ m (31,000 feet) below their collision point. How is this possible? Their lifetime should only permit them to travel 600 m or so. The reason is time dilation. They are traveling at such a high speed relative to us on the earth, that the "internal clocks" (as we observe) on the muons slow down, giving them a time-dilated lifetime of about 65 microseconds as they travel at 99.994% the speed of light. This time dilated lifetime allows them to travel nearly $20,000$ m (or more) on average, allowing a strong flux of muons to make it to the earth's surface. In the Cal Poly Physics Department, we detect muons in our advanced laboratory at the rate of about 15 per hour.

5.2.3 It moves so its life will slow

With the basics out of the way, we will close this section with a few more remarks. First, we obtained our time dilation equation using the funny light clock described above. But light clock, or not, any periodic entity will run slower when it moves. The pulse of light moving between the mirrors might very well be a pendulum bob as it swings (driving an escapement), or the mass on the end of a spring that is bouncing up and down. Both would appear to swing or undulate over a longer path, resulting in a longer swing or spring time.

Your heart beats rhythmically. It's not human-made, and not a mechanical clock; would it beat slower? Yes. Imagine you get into a car (or one of the 1971 jets) and had someone

observe the outer wall of your heart move in and out as your heart beats. It will follow some "longer path" just as the light pulse did. The entire rate of your body would slow; your heart-rate, the rate at which your brain thinks, growth of your cells, etc. You would appear to age slower to someone at rest on the ground. You may have heard of the "twin-paradox," where two identical twins are born. One is put on a rocket at sent into outer space for 20 years, then comes back. When the twin returns, the one from the rocket is 5 years old and the twin who stayed on earth is 40. Is this true? Yes. The twin that went on the rocket would be younger. Moving through space at a high speed causes time dilation of any time periods associated with the twin in flight, at least observed by the stationary twin back at home. How much depends on how fast the rocket is able to go, setting the critical v^2/c^2 parameter, and on how long the journey lasts. Strange, but true. This is the theory of time that Einstein gave to us.

5.2.4 Einstein and Harrison

Finally, one last thought on John Harrison. He fought the Board of Longitude, which was stacked with astronomers, so unfairly biased toward an astronomical solution to the longitude problem. But Harrison knew that time and a timekeeper would ultimately be the best solution, and so he dedicated his life to this end, eventually winning the prize. He was right.

Wouldn't he be pleased to know that time, whose prominence Harrison fought so hard for, remains an ever elusive parameter, even today. Perhaps the skeptics among us are all like the "board of longitude" refusing to believe strange predictions about time, and Einstein was the Harrison, convinced of its truth. To date, no experiment has ever disproved any of Einstein's predictions.

Measuring time continues to confound us, as it did for him way back then. Harrison was able to build a clock out of wood that could survive on a ship and keep time to within 2 seconds a month in 1760. Around 300 years later, atomic clocks that are able to keep time to within 1 nanosecond in a million years have exposed ever more flaws in our thinking about time. Even asking "what time is it?" can be difficult to answer depending on the circumstances. Harrison simply wanted to be able to keep track of the time in Greenwich from far out at sea. Could he even conceive of a nanosecond back then? Or that the very act of even asking for the time at stationary Greenwich from a moving ship would someday be tied to one of the grandest theories of our time?

5.3 The Crown Jewel in Navigation: The Global Positioning System

5.3.1 Introduction

The Global Positioning System (GPS for short) is a system that allows anyone to instantly know their latitude, longitude, and altitude to within about a meter, no matter where they

are on earth. There are two parts to the GPS, the receiver and "the rest of it all." The receiver is a $200, handheld gadget you can buy from Amazon.com. You turn it on, wait a few seconds, and your location will be displayed on a small screen. "The rest" is what this section is about; how something like this could possibly work (hint: clocks and time are essential). But before proceeding, we would like to point out two curious similarities between ultra-modern-day GPS and the times of John Harrison.

The first is in the magic of the handheld GPS unit. You simply turn the thing on and your latitude, longitude (and altitude) pop up on a small screen. Anyone even mildly technically minded, must think "wow" when this happens; "how does this work?" Back when Harrison was showing people his clocks, he had a hard time convincing them to believe in clocks. Sure he could show them the gears, and escapement, and mainspring, but people were deeply religious. It was a sacrilege to trust a machine to something like timekeeping, especially since "things from above" like planets, the moon, and the stars also told time, as set by their Creator. This is part of the partisanship Harrison had to deal with in defending his clocks; people simply didn't believe they could work.

But look how far technology has taken us. You press the button on a GPS receiver, read your location, and you believe it; so do airline pilots, ship captains, truckers, and families playing geocache games. We believe it partly because it is correct, time and time again. No one these days would discount the readings from a GPS unit, failing to believe it due to some higher authority. As written by Andrewes at the start of this chapter, technology has certainly "changed our priorities," at least since the time of Harrison.

The second similarity, is in the basic informational mode of GPS. GPS sends its signals using electromagnetic radiation or "radio waves" if you will. As you know, radio waves beam invisibly across space sending information (like music, commercials, cell phone calls, etc.) far and wide. A ship at sea is often in constant radio communication with land. If radio existed in the 1700's, a captain could have "radioed" Greenwich, asking for the time. Also, one of the first radio-based global navigational aids, called the "WWV" radio, broadcasts the time in Greenwich over the radio, at regular intervals, all over the globe. All you need is a short-wave receiver from Radio Shack to get the time in Greenwich.

Back in the days of the Longitude Prize, there were many "crackpots" with similarly described theories on how the longitude problem could be solved. One was particularly "out there" and it worked as follows. Suppose a ship was about to set out across the Atlantic. On board, you would put a dog that you would promptly injure mildly with a knife. Before setting off, you would soak some cloth with the dog's blood. The cloth would stay on shore and the dog would go with the ship. Anytime you immersed the cloth in some magical potion, called the "powder of sympathy," the dog would yelp out in pain. To keep time aboard the ship, a trustworthy person on land need only to dip the cloth into the powder at fixed intervals and the dog (on the ship, thousands of miles away) would yelp, signaling to those on board, the time in Greenwich. This was based on some "unknown" way of communicating across vast distances, with some kind of fast, invisible signal. Although complete nonsense, the so called "powder of sympathy" sounds like a desperate wish for something like today's radio waves.

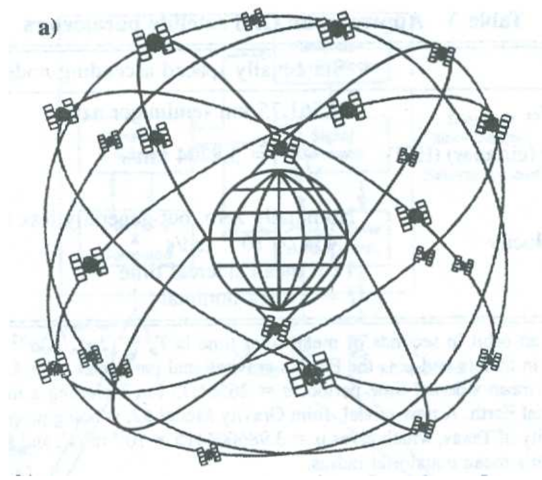


Figure 5.7: A snapshot of the 24 GPS satellites. From Parkinson, GPS: Theory and Applications, Vol I, p. 39]

5.3.2 The GPS System

GPS is a space-based navigational system. It consists of a system of 24 satellites positioned around the earth so that, from almost any location, at least four satellites will be above the horizon[Ashby, Physics Today, May 2002] at any given time. A snapshot of the satellites at a given instant of time is shown in Figure 5.7.

For obvious reasons, the satellites form a "birdcage" around the earth, and are sometimes called the "GPS Constellation." We would have liked to say "visible" but you can't necessarily see the satellites. (The best time to see an artificial satellite is just as the sun is setting and a satellite is heading west where the sun's rays can reflect off of one and into your eyes; they appear as fast moving "dots of light" high in the sky.)Stated another way, at least four satellites are always in your "part of the sky," and not obstructed by the Earth itself. This is important because if you are to "find yourself" anywhere on earth, it is necessary to be able to send you information about your location, via radio waves, and this is what the 4 satellites in your part of the sky are always doing. All you have to do is receive and interpret their signals, and you can find out where you are. (This is what the $200 GPS receiver does.)

Why satellites from space? Covering the entire earth is a lot easier from above. Satellites have an unobstructed view of any point on earth, so being on high mountains, low canyons, between skyscrapers, or next to a brick wall, isn't detrimental to GPS's operation. As long as you have a clear view of the sky, the satellites can talk to you. Contrast this to ground based radio stations or cell phone towers, where reception can fade in and out depending on the local landscape (buildings, mountains, etc.). "Can you hear me now?" We also maintain that the view of the ground is "unobstructed" even with clouds, because while clouds are opaque to visible light (that we see with) they are transparent to the radio waves used by GPS. The satellites are placed in orbit at about $20,000$ km above the earth, with an orbital

period of 11 hours and 58 minutes (half a day). Twice a day, a given satellite can be found at the same location on the celestial sphere.

Again a tie with history and the longitude problem. If the satellites were somehow made visible, and 4 of them were passing overhead at any given time, with a period of half of a day, sighting them would make excellent world-wide timekeepers. Anytime the same satellite was spotted at a given location, you'd know another 11 hours and 58 minutes went by. Again, a curious analogy with a crackpot longitude theory from Harrison's day. An idea was proposed to anchor ships all across the Atlantic that would fire cannons, at known times, to a height of $6,500$ feet. On board your ship, you could sight and hear them firing to figure your time. A wishful idea back then to install some kind of earth-wide machine to help with navigation; just what GPS is! Back then, anchoring ships at regular intervals to the ocean floor would sound about as nutty as building satellites and launching them into space with rocket ships!

A key functioning element of GPS is that the precise positions of the satellites are always known. Also, each satellite carries one or more atomic clocks on board, and all 24 satellites are synchronized to have exactly the same time. The GPS satellites are in constant communication with ground-based stations throughout the world. At these stations, computers are constantly calculating where the satellites should be based on the physics of satellite orbits. Also, a computer calculates what time it should be by taking the average time of 60 or more land-based atomic clocks (50 cesium clocks and a dozen hydrogen masers)[Ashby, Physics Today, May 2002]. Orbital and time corrections are constantly sent up to the GPS satellites, to keep them up-to-date. When done, the satellites agree with ground-based clocks down to about 20 nanoseconds (0.000000020 seconds)[Ashby, Physics Today, May 2002].

In the next section, we'll look at how the satellites are used to find your location. So we leave this section with the following information: GPS consists of 24 satellites, 4 are always available to you, no matter where you are on earth. Each satellite knows its position very well, and precisely what time it is. Lastly, the clocks on all of the satellites are ulta-synchronized, meaning the time that one tells is the time they all tell.

5.3.3 Finding your location

Your location on earth can be fully specified by three variables: your latitude, longitude and altitude. For sake of simplicity, we'll call these three variables x, y, and z. Next, a rather funny things happens, and we apologize to the reader for bringing this up again: the GPS system doesn't trust that you really know what time it is! Again with the time! They couldn't keep track of it when Harrison was around and now that we can, it can't really be trusted! Sure you have your watch on, and your GPS receiver probably has a very accurate crystal controlled clock, but these won't suffice to find your location within a meter. So then, we have to add another "unknown" to the list, time, which we'll give the variable t. So all told, GPS needs to tell you four parameters: x, y, z, and t.

Note that this is our first result about GPS. Why do you need 4 satellites? Because you have four unknowns. You might remember from high school algebra, that you always need as many equations as unknowns. So two unknowns require two equations. Four unknowns require four equations. But "equation" here is a strong word at this point. Equation can

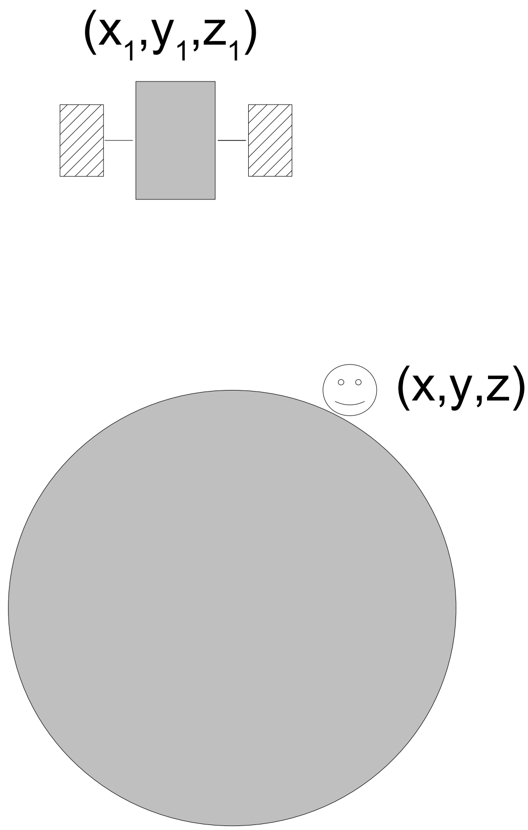


Figure 5.8: A basic GPS configuration: you, a satellite and the earth.

also be taken to mean "bits of independent information." So to solve for all four unknowns, you need four "bits of information." So this is why 4 satellites are needed.

When decoding the radio waves sent by GPS, the first thing your GPS receiver does is learn what time it *really is* from the satellites, which remember, are synchronized by ground-based atomic clocks. This immediately gives you one of the unknowns, t. All GPS satellite transmission packets are encoded with the current time as far as the satellite knows it. Once your GPS receiver knows what time it is, it can start finding your position, and here's how it does it.

Figure 5.8 shows the basic GPS layout. It shows a single GPS satellite and the Earth. Let's call the satellite "Satellite #1" and say it is located at position (x_1, y_1, z_1). These could be the satellite's declination, right ascension, and altitude, on the celestial sphere. As mentioned, above, let's call your position (x, y, z). Mathematically, the distance between you and Satellite #1, d_1 is given by the "distance formula," or

$$d_1 = \sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2}, \tag{5.10}$$

and this is shown by the dotted line in Figure 5.9.

The interesting "twist" is that there is another quantity to which this distance, d_1 is

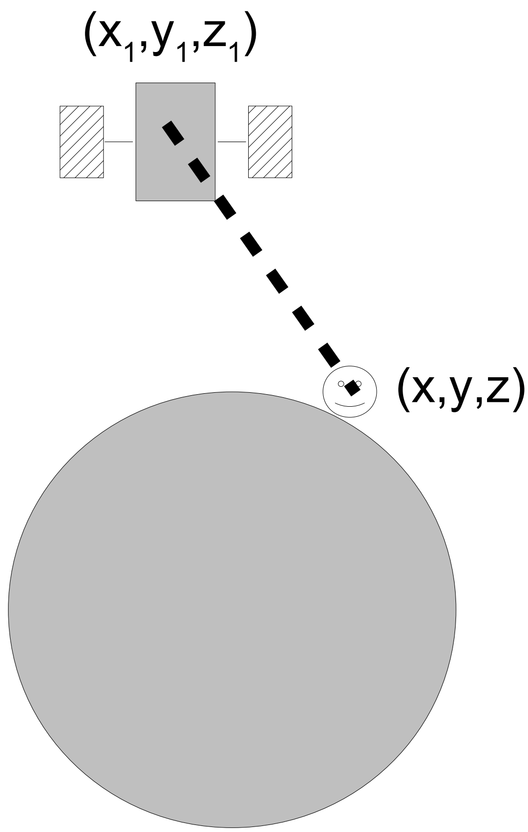


Figure 5.9: The dotted line is the distance between you and the GPS satellite. It can be computed using the distance formula from high school algebra.

equal. A radio wave from the satellite, which travels at the speed of light, takes a certain amount of time to propagate between the satellite and your GPS receiver. For Satellite #1, we'll call this time Δt_1. As a rough number, the signal from a satellite directly overhead at $20,000$ km would take about $20,000$ km$/c$ or 2.0×10^7 m$/3 \times 10^8$ m/s $= 0.066$ s or about 66 milliseconds to travel to your receiver. (Your eye blinks in about 100 ms.)

Since distance=speed $\times$ time, it is also true then that the distance between your GPS receiver and Satellite #1 is $c\Delta t_1$. But this distance is also given by d_1 (above), so we can write that

$$d_1 = c\Delta t_1, \tag{5.11}$$

or

$$\sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2} = c\Delta t_1. \tag{5.12}$$

So the question now is, where do we get Δt_1, the time it takes the radio wave to travel from a satellite to your receiver? This is where the utility of an atomic clock on board each satellite comes into play.

Remember that as a result of the earlier transmission, your GPS receiver's clock is now synchronized with the time on board the GPS satellites. Let's call this time t_S for "time on the satellite." After the first contact with a GPS satellite this number is stored inside of your GPS receiver. Suppose next, a satellite sent you a signal that contained (among other things) the time at which that very signal left the satellite. Let's call this time t_{L1} for "time a signal left Satellite #1." When you receive this signal, a little computer inside of your GPS receiver can compute Δt_1, the time it took the signal to get to you to be

$$\Delta t_1 = t_{L1} - t_S. \tag{5.13}$$

So the radio-wave-travel-time is now known. The "other things" in the transmission are the precise coordinates of the satellite, or x_1, y_1 and z_1. This means the unknowns in Equation 5.12, are reduced down to only 3, x, y, and z. We have one equation now, and we need two more. Where can we get them? By receiving and decoding the signal from two other satellites! Using the identical procedure as above, we can compute Δt_2, and Δt_3, to form this set of equations:

$$\sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2} = c\Delta t_1, \tag{5.14}$$

$$\sqrt{(x-x_2)^2+(y-y_2)^2+(z-z_2)^2} = c\Delta t_2, \tag{5.15}$$

and

$$\sqrt{(x-x_3)^2+(y-y_3)^2+(z-z_3)^2} = c\Delta t_3. \tag{5.16}$$

Now we have three equations and three unknowns. Solving these three equations is a routine computational job, and x, y, and z will readily "pop out" as your latitude, longitude, and altitude. A tad bit more of electronics paints this information onto the little screen

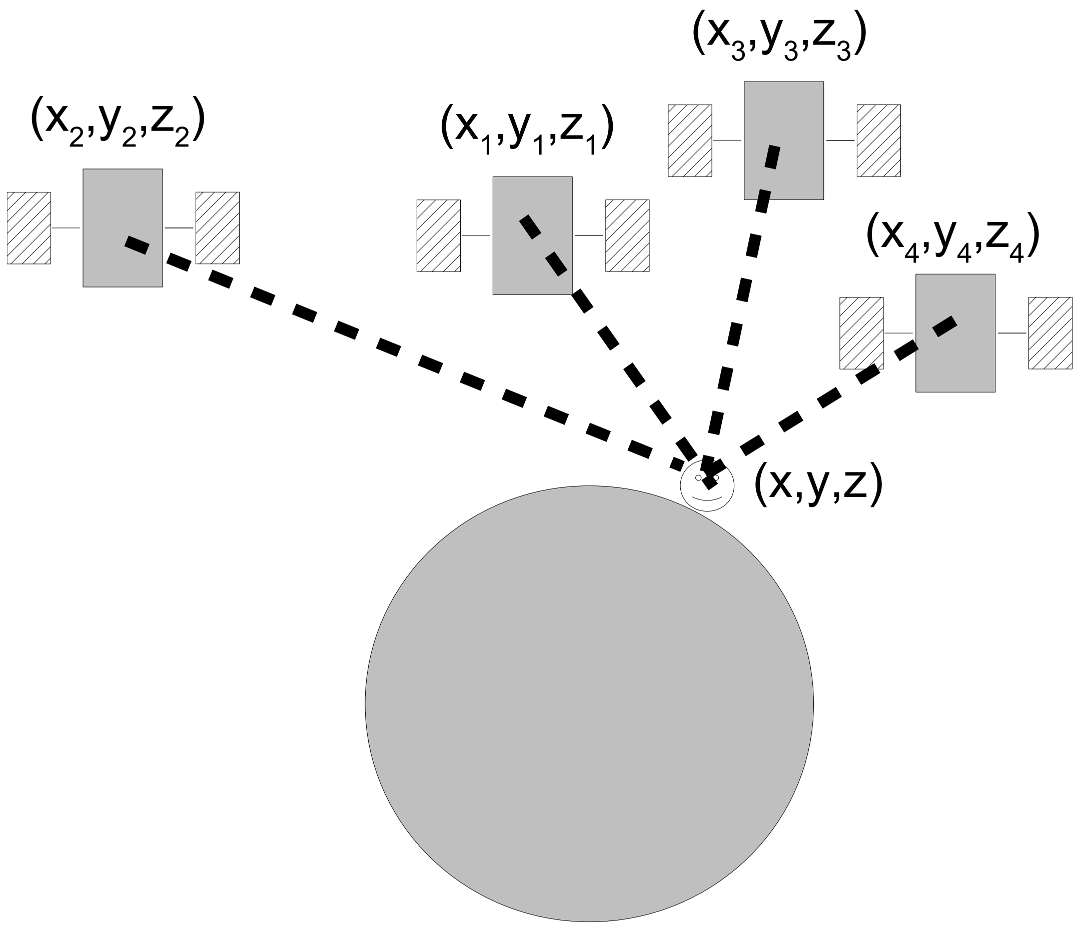


Figure 5.10: Using four GPS satellites to synchronize the clock in your GPS unit, then calculate your position.

on your GPS receiver, and PRESTO! You know where you are! The full satellite distance picture is shown in Figure 5.10.

Lastly, the equations involving the $x's$ and the square roots above are equations of circles. The circles pass through two points, a given satellite and your exact location. Three circles are needed to determine your location. Does this sound familiar? It should! It the same principle used in celestial navigation, from Chapter 2! That's right! The geo-location idea for celestial navigation is the same as for GPS!

5.3.4 Why Atomic Clocks?

GPS could only work with the high resolution and stability offered by atomic clocks for two reasons. First, the high resolution is needed because apart of finding your location is timing how long it takes a radio wave to travel from a satellite to your receiver. This gives you the $c\Delta t$ terms above. Radio waves travel at about 1 foot every nanosecond. If this

timing was off by even, 5 nanoseconds, (0.000000005 seconds) then your position could not be found to about 5 feet. Modern GPS demands better resolution that this. Mechanical clocks are good to about a second and would give you your distance to within $300,000,000$ meters. Basic electrical clocks are good to somewhere in the millisecond (0.001 seconds) or microsecond (0.000001 seconds) range. This would give your position to within $300,000$ meters, or 300 meters, respectively. You need something that can time reliably down to the one-ish nanosecond range. Atomic clocks can do this.

Second, stability is needed because a clock up on a satellite cannot always be adjusted. Current GPS clocks are updated by the ground stations about twice a day, so once set, the clocks need to stay stable for a good 12 hours. That is, their "ticks" and "tocks" must remain precisely spaced for the duration. The clock can not run too fast, or too slow. As an example, if your system can tolerate a 5 nanosecond error buildup before its next update, then this is the same as saying you are willing to have a error in up to 5 feet in your location. Half a day is about $50,000$ seconds, so the atomic clock has to be stable to $50,000$ seconds/5×10^{-9} seconds $= 10^{13}$ seconds per second. This means for every 10^{13} "ticks," (ten million million ticks) only 1 of them may be unstable. Current atomic clocks have 1 faulty tick in 10^{13} over 1-10 days[Pakrinson, Global Positioning System: Theory and Applications Volume 1].

Rubidium (Rb) clocks are used in the current GPS fleet. As you might guess, the clocks must be designed to be completely maintenance free. We find some of the technical difficulties interesting to ponder. Here are a few excerpts from Camparo, Physics Today, 12/07, p. 36-37:

> The best Rb clocks ever built are now flying on GPS satellites. Those clocks are so stable that in three hours they accrue less than half a nanosecond of time error. In the same three hours, general relativistic effects due to the satellite's elevated location, if left uncorrected, would cause a satellite clock and a terrestrial clock to accrue a little more than 5 microseconds of time offset. In other words, GPS Rb atomic clocks can sense those general relativistic effects at the fourth significant figure. (See the article by Neil Ashby, PHYSICS TODAY, May 2002, page 41.)
>
> Rubidium clocks aboard satellites operate under different conditions than do terrestrial clocks, and the need to adapt the clocks to those conditions has led to a range of physical and chemical research. For example, the inner glass surface of a discharge lamp experiences a harsh environment, with energetic ions etching the surface and alkali-metal atoms chemically interacting with the exposed glass. If that chemistry were to lead to lamp failure in a terrestrial clock, one would simply ship the device back to the manufacturer for a replacement lamp. However, no one services a failed clock on a GPS satellite, so the lamps must operate unattended and reliably throughout the satellite's many-year mission. In the 1980s the question of a lamp's lifetime motivated
>
> Robert Frueholz and his colleagues at The Aerospace Corporation to study alkali–glass physical chemistry.[9] Frueholz and colleagues used electron spectroscopy and secondary-ion mass spectrometry to show that Rb forms an oxide on the lamp's inner glass surface and penetrates into the glass matrix, quite likely taking the form of a rubidium silicate. The complete depletion of Rb through those mechanisms would lead to a lamp's failure. Using differential scanning calorimetry, the researchers then determined the rates at which various glass types consume Rb. As a result of their work, clock manufacturers can now determine how much liquid Rb needs to be deposited into a lamp to guarantee its operation over any specified mission lifetime.

So we get a glimpse at the minute technical details that go into making such a system work.

Lastly, we have been bragging about the stability of atomic clocks since they were first mentioned. It turns out, however, like most things, even atomic clocks degrade over time. Harrison fought build-up on his gear bearings by using naturally oily wood. His escapements were nearly frictionless with jewels as their tips. He fought temperature effects by mixing different metals. Atomic clocks designers see their clocks lose 1 part in about $30,000,000,000,000$ each day. That is, a single tick in 30 billion is bad each day, and *no one knows why.* The effect is called "linear frequency aging."

5.3.5 Closing Remarks

Just a few closing remarks. First, it seems conceivable that only three GPS satellites could be used to generate the three GPS equations shown above. It is the author's understanding though, that in practice, the satellite that is used to give you the time is *not* also used to compute your distance. Owing to small variations in the satellites' atomic clocks, this is a way of giving four unbiased pieces of data (the time and three position coordinates). In other words, GPS only relies on a given satellite for one bit of information. Either the correct time, or its coordinates. Never both.

Second, hopefully you see the reason why GPS is only possible with atomic clocks. Suppose you wanted GPS to find your position to within 1 meter (as it does). Light takes 3.3 nanoseconds (that's 0.000000003 nanoseconds) to travel 1 m. If the Δt's used in the above calculations above are off by more than 3.3 nanoseconds, GPS will fail to resolve your position to within this distance. Accurate timing, to many significant digits, is crucial.

Third, as abstract as you may think Einstein's relativity is, consider that modern day GPS cannot work without it. For one, there is relative motion between the satellites themselves and the satellites and the clocks on the ground. And remember from above that time dilation is all about relative motion. A GPS satellite travels at about $4,000$ m/s[Ashby] relative to the ground (this satellite speed is not negotiable; the laws of physics require such a speed to keep the satellite in a stable orbit). In this case, the clock in the satellite runs at "only" 99.999999991% of the rate of the clock on the ground. This is a small time lag of only 0.000000009% slower. Although small, this lag can accumulate over time. Suppose you decide to use your GPS receiver in a location for about 1 minute. The clock in the satellite speeding by overhead would be off by about 5.4 nanoseconds in this time, which could put your location off by over 5 feet, or over 1.5 meters! Computer software in the GPS receivers are programmed by the manufacturers to compensate for such a relativistic effect.

Lastly, think again what you need to do in order to find your location with GPS. One step; push a button. Could a machine have been built to find your location so easily in Harrison's time? Maybe! Suppose Harrison made two H4's and integrated them both into one "machine." One clock read the time in Greenwich, and the other the local time, updated manually once-per-day at local noon. Perhaps through another combination of gears, the difference in the two clocks could be computed and with another set of gears, this difference could be multiplied by 15. (This is not as far fetched as it sounds. Early calculators were entirely mechanical, using gears and the like to perform basic mathematical operations.) With a difference and multiplication by 15, you'd have a machine with a "longitude" dial on it! Once per day, by merely setting the local clock, you would have your longitude! This is about as easy as pressing a button, but what a job Harrison would have, not only to build this device, but to get people to believe the readouts on such a "longitude machine."

Chapter 6

Could you do what Harrison did?

6.1 Introduction

It is not known exactly why Harrison became so committed to the longitude problem. Most likely it had something to do with his passion for clocks, since it was well established that keeping time at sea would allow one to find longitude. Even so, as he began, he had to know the odds were stacked highly against him. He was a self-educated carpenter from the English countryside. He was not at a big university nor at a scientific lab and had no connections to like-minded individuals. He would have to enter the mainstream of the society in London at that time and deal with all of the ups and downs that it had to offer. Did he know he would spend his entire life on this quest? Did he expect or want his first clock to win the award? Was he intimidated by the "important" people he encountered? Would they take him seriously? Did he think a single person could solve the longitude problem? Did he think that he would eventually meet the King of England? Did he visualize himself defending his ideas in front of the board of longitude? What was Harrison's attachment to "the machine" at a time when God (i.e. astronomy) was the rule of the land? Finally, why did the longitude prize go unclaimed for so long? Weren't there any other serious contenders?

As a young, unbound, and impressionable college student student, I would like to ask you to pause and wonder for a moment if **you could do what Harrison did?** Could you identify a world-wide problem? Then, whether you knew something about it or not, would you be willing start working on a solution to it? Would you dedicate your life to solving the problem?

6.2 Problems

You might think that the list of problems facing our world is endless and that you do not know where to start. It turns out that a few organizations have realized that stating and prioritizing world-wide problems is a huge step toward solving them. Here are a few resources that you might consider drawing from.

6.2.1 Xprize.org

If you are looking for a problem with a scientific solution and a hefty prize (just like the longitude prize), then look no further than the X-Prize foundation a www.xprize.org. They are currently looking for solutions to landing a robot on the moon, developing a 100 mpg car, decoding the human genome, and delivering payloads to the moon.

6.2.2 The Copenhagen Consensus

This is a steering committee that is helping world government prioritize world-wide problems. You can find information at copenhagenconsensus.com, and see the book "Solutions for the World's Biggest Problems" by Bjorn Lomborg. If you can find this book, the introduction alone is well worth reading to give you some perspective on solving "big" problems. The current prioritized list is:

1. Hunger and malnutrition
2. Unsafe water and lack of sanitation
3. Diseases
4. Lack of education
5. Conflicts
6. Living conditions of children
7. Climate change
8. Living conditions of women
9. Deforestation
10. Air pollution
11. Corruption
12. Land degradation
13. Financial instability
14. Subsidies and trade barriers
15. Loss of biodiversity
16. Terrorism
17. Vulnerability to natural disasters

18. Population: migration
19. Arms proliferation
20. Drugs
21. Lack of intellectual property rights
22. Lack of people of working age
23. Money laundering

We find it interesting that these prioritized solutions don't necessarily have technological solutions. One might think a 100 mpg car or clean, unlimited energy might be high on this list, but they don't appear at all. Perhaps these technological problems are based only in highly developed societies such as our, in the United States. For example, people without enough food to eat are not concerned with being able to drive or leave their web-server on 24/7.

6.2.3 The Bill and Melinda Gates Foundation

See gatesfoundation.org. A foundation set up by the well-known founder of Microsoft, that focuses mainly on health and education issues.

6.2.4 Obvious problems

Many of these types of problems are identified by simply looking out of your window or by reading the newspaper. What can you find?

6.3 The Assignment

The requirements of this assignment can be broken down into three parts, as follows:

Part I: Introduction The first is a written statement of the world-wide problem that you'd like to solve. Clearly state the problem and why you perceive it as being a problem. This can include "official evidence" as to why your choice is considered a problem as well as your own personal feelings and observations. In general, you should probably not invent your own problem. Like the longitude story, find a problem that we've all already acknowledged is a problem. This will make it easier for you to justify yourself and your efforts.

Part II: Your solution Discuss in detail how you are going to solve this problem. You can certainly summarize your ideas and "blue sky" thoughts, but you must also provide concrete steps that will be taken that will solve your problem.

Part III: Detailed Budget Analysis You must justify all costs associated with your solution. Do this in an Excel spreadsheet so you can actually create columns and add numbers together, to arrive at a grand total. In this solution you are proposing, you can't just say things like "give me \$100 million and I'll create a 100 mph car." You need to justify every cent of that \$100 million. Do you need a room full of computers and a nice office? Fine! Put them on your list. A small office in San Luis Obispo is about \$2,000 per month and a well equipped computer costs about \$1,000. Now what?

Chapter 7

Map Worksheets