Navigation basics and inertial navigation

Navigation is determining your own position and velocity. **Guidance** is the process of reaching a certain destination. Naturally these two topics are closely linked: you can't reach a destination if you don't even know where you are. In this chapter, we'll deal with the basics of navigation and also discuss inertial navigation. In the next chapter, we will discuss more navigation methods, and also deal with guidance.

1 Navigation basics

1.1 The reference ellipsoid

Before we can determine our position, we need to know with respect to what we want to know our position. One option is to approximate the Earth as a sphere with radius $R_e = 6378km$. There's just one downside. The gravity isn't always perpendicular to the surface for such a sphere. Instead, the **apparent gravity field g** satisfies

$$\mathbf{g} = \mathbf{G} - \mathbf{\Omega} \times (\mathbf{\Omega} \times \mathbf{R}), \tag{1.1}$$

where **G** is the Newtonian gravity field, $\mathbf{\Omega}$ is Earth's rotation vector and **R** your position with respect to Earth's center of mass. To solve this problem, we define the so-called **World Geodetic System Ellipsoid**. This ellipsoid represents the Earth surface. It is defined such that the average mean-square deviation between the direction of the apparent gravity and the normal to the ellipsoid is minimized over Earth's surface.

This ellipsoid has several important parameters. First of all, there are the **equatorial radius** R_E and the **polar radius** R_N. Second, there is the **eccentricity** $\epsilon = \sqrt{\frac{R_E^2 - R_N^2}{R_E^2}}$. We need these parameters to define the **meridian radius of curvature** R_M and the **prime radius of curvature** R_P. They are given by

$$R_M = \frac{R_E(1-\epsilon^2)}{(1-\epsilon^2 \sin^2 \Phi_t)^{3/2}}, \qquad \text{and} \qquad R_P = \frac{R_E}{(1-\epsilon^2 \sin^2 \Phi_T)^{1/2}}. \tag{1.2}$$

The definition of the **geodetic latitude** Φ_T will follow in the upcoming paragraph.

1.2 Indicating your position

To indicate our position, we still need a reference frame. We'll examine two of them, both having their origin positioned at the center of the Earth. In the **inertial frame of reference** $\mathscr{F}^i$, the Z_i axis points North, X_i points to a convenient star and Y_i is perpendicular to the other two axes. This reference frame does not rotate as the Earth rotates. On the other hand, the **Earth frame of reference** $\mathscr{F}^e$ does rotate along with Earth. Again, the Z_e axis points North, but now the X_e axis points through the Greenwich Meridian. The Y_e axes is again perpendicular to the other two axes.

To indicate the position of an aircraft, we can use **geocentric spherical coordinates**; relative to a spherical Earth. The aircraft then has a **longitude** λ (with East being positive) and a **(geocentric) latitude** Φ (with North being positive). Also, the distance from the center of the Earth is $R_e + h$, with $R_e = 6378km$ the radius of the Earth and h the height of the airplane. In this way, the position of the aircraft can be expressed by the numbers (λ, Φ, h).

We can also use coordinates with respect to an ellipsoidal Earth. This gives us the **geocentric spherical coodinates** (λ, Φ_T, h). h is the height above the reference ellipsoid. Φ_T is the **geodetic latitude**: the angle between the line normal to the ellipsoid surface and the equatorial plane. Finally, λ is still the longitude. (It remains unchanged.)

Let's suppose that we're flying with a velocity V_N to the North and V_E to the East. How will λ and Φ_T change? They will do this according to

$$\dot{\Phi}_T = \frac{V_N}{R_M + h} \qquad \text{and} \qquad \dot{\lambda} = \frac{V_E}{(R_P + h)\cos\Phi_T}. \tag{1.3}$$

1.3 History of navigation

Navigation became very important at the time when several European countries sent out ships across the world. The navigators on these ships could easily determine their latitude. All they needed to measure was the angle of the sun above the horizon at noon. However, finding the longitude was a lot harder. In fact, the problem was so hard that, in 1714 in London, an enormous prize of 20.000 pounds was promised for a solution.

John Harrison solved the problem. His solution was as impressive as it was simple: you just use a clock. The only problem was that, back in the early 18th century, clocks weren't accurate enough. But Harrison's clock, using springs, only had a deviation of about 0.7 seconds per day. With this clock, navigation was possible around the globe.

So how does Harrison's method work? First, you measure the local time, using the sun. Then, using your clock, you measure a certain reference time (like e.g. the London time). Then, by using the time difference, you can find the relative longitude, with respect to London.

2 Chategorizing navigation

2.1 Navigation system categories

Navigation systems must satisfy several **performance requirements**. They must have...

- **Accuracy** – The information that they give must be close to the actual value.
- **Integrity** – If the system can not give sufficiently accurate information, it must notify the user of this in time.
- **Availability** – It may not occur that the system is unexpectedly unavailable.
- **Continuity of service** – If the system stops working after, for example, 2 years, it's not really useful.

There are several ways to categorize navigation systems.

- A **primary means** navigation system is a system that meets the accuracy and integrity requirements, but does not necessarily meet the other two requirements all the time.
- A **sole means** navigation system is a system that, for a given phase of the flight, satisfies all four requirements. However, in real life, no single sole means navigation system exists. Currently, we always need combinations of primary means systems.
- A **supplemental means** navigation system is a system that must be used in conjunction with a sole means navigation system. An example is the GPS system: it is not integer enough. When the GPS system gives a wrong signal, it often needs 10 to 15 minutes before it can notify the user of this.

Another way to characterize navigation systems is as follows.

- **Positioning systems** measure the state vector of an aircraft without regard to the past. (Examples are using radio navigation or GPS.) The error of positioning systems is usually constant in time.
- **Dead reckoning systems** derive the state vector from a continuous series of measurements relative to an initial point. (The most important example here is the **inertial navigation system** (INS), using gyroscopes and accelerometers.) For such systems, the error accumulates over time.

2.2 Categorizing navigation errors

Let's suppose that we want to fly to some destination. If we fly directly towards our destination in a straight line, then it is called **direct steering** or **area navigation**. However, if we fly along preplanned airways, we are performing **airway steering**.

Of course, when we fly along airways, we don't exactly stay on the flight path. There will be deviations. There are several possible causes for these deviations.

- **Navigation sensor errors** – Errors when measuring your position.
- **Computer errors** – Errors when processing measured data.
- **Data entry errors** – For example, when the wrong position of a beacon is intered in the computer.
- **Display errors** – When humans read a display incorrectly.
- **Flight-technical errors** (FTE) – Due to reasons outside the navigation system.

The first four points of this list together form the **navigation system error** (NSE). If we add this up to the FTE, we get the **total system error** (TSE).

The NSE is a very important parameter. It is used to calculate the minimum distance between two airways or, similarly, between two runways. It can also be used to determine the risk of airborne collisions.

3 Inertial navigation

3.1 Inertial navigation system basics

An **inertial navigation system** (INS) use accelerometers to measure accelerations. These accelerations are then integrated along the axes to find the change in position. Advantages of such a system is that it is continuously avaliable, self-contained, autonomous, and accurate. However, it is also quite expensive and, since it is a dead reckoning system, its performance degrades over time.

An **accelerometer** is not much more than a box containing a mass and a spring. If the box is accelerated, the spring needs to exert a force on the mass to tug it along. (The mass has inertia.) The deflection of the mass from the equilibrium position gives an indication of the acceleration. However, what accelerometers measure is not exactly the acceleration. Instead, it is the **specific force** $\mathbf{f}$. To find the acceleration $\mathbf{a}$, we also need to take into account gravity. Thus, $\mathbf{f} = \mathbf{a} - \mathbf{g}$. To know in which direction $\mathbf{g}$ points, we have to know the orientation of the aircraft. For this, we can use gyroscopes.

One of the downsides of inertial navigation is the performance degradation over time. Let's examine this more closely. If an accelerometer has an **accelerometer bias** B then, after a time t, the distance will be off by $\frac{1}{2}Bt^2$. But the accelerometers aren't the only problems. We also have gyros. If we have a constant **gyro drift rate** W, then we will get a **platform tilt error** $\Delta\theta = Wt$. This will result in an acceleration error of $g\,\Delta\theta = gWt$. The distance error thus becomes $\frac{1}{6}gWt^3$.

3.2 Types of inertial systems

There are two important types of inertial systems. First, there are **stable (gimballed) platform** systems. Such a system have a stable horizontal platform. It consists of two loops. First, there is the **fast loop**. In this loop, gyroscopes measure the orientation of the platform. If it is not exactly horizontal, servo motors are used to turn the platform back to a horizontal position. Second, there is the **slow loop**. This loop contains accelerometers. These accelerometers are used to find the position of the aircraft. Of course this position is what we wanted to know. But the position data is also used to precess the gyroscopes, such that they follow the Earth's curvature. (If we don't, then apparent drift and transport wander will start to play a role.)

The second type of system is the **strapdown** system. Now, the INS is mounted on the airplane. Gyroscopes measure the orientation of the INS. Then, computers use this data, together with the output of the accelerometers, to determine the position of the aircraft. Strapdown systems thus do not have a real stable platform. But in the computer, there is some kind of virtual 'stable platform', known as the **analytic platform**. And, just as in the stable platform system, the strapdown system also consists of two loops. But again, these loops are only present in the computer.

Since the 1980s, when computers have become widely used, strapdown systems have become the dominant type of inertial systems. They have reduced mechanical complexity, improved reliability, lower power consumption, lower volume/weight and lower cost.

3.3 Schuler tuning

The downside of inertial navigation systems, is that they require tuning: you need to set the initial orientation. This is not very difficult, when the system is standing still. But now let's suppose that we accidentally do give the platform a horizontal acceleration a. This will result in a platform, tilted by an angle

$$\theta = \arctan\left(\frac{a}{g}\right). \tag{3.1}$$

The effect of this tilt angle is somewhat surprising. Because of the (unknown) tilt angle, the system senses a component of the gravity for which it does not compensate. It is thus interpreted as an acceleration. It now seems as if the aircraft is (for example) accelerating: the system will say the aircraft flies faster than it really does. Because of this, the slow loop of the platform will compensate more for the effects of the curved Earth. The tilt angle of the platform thus decreases. After a while, this will result in a tilt of the platform in the other way. This will then cause the aircraft to decelerate. And the whole process happens again, but now in the opposite way. This repetitive process is called **Schuler tuning**.

We can derive equations for the effects of errors. A platform tilt angle $\Delta\theta$ (or similarly, a constant accelerometer bias $B = g\Delta\theta$) will give a distance error of

$$D_e(t) = \frac{B}{\omega_s^2}\left(1 - \cos\omega_s t\right), \tag{3.2}$$

where $\omega_s = \sqrt{\frac{g}{R_e}}$ is the **undamped natural frequency** of the Schuler loop. Similarly, the distance error due to a constant gyro drift rate W is

$$D_e(t) = WR_e\left(t - \frac{1}{\omega_s}\sin(\omega_s t)\right). \tag{3.3}$$

So, both errors result in an oscillation. This **Schuler oscillation** has a fixed period $T_s = 2\pi\sqrt{\frac{R_e}{g}} =$ 84.4 minutes, irrespective of where you are on Earth. And surprisingly, the Schuler oscillation doesn't only occur for the real platform of a stable platform system. It also occurs for the analytic platform of a strapdown system.