Figures for Craig

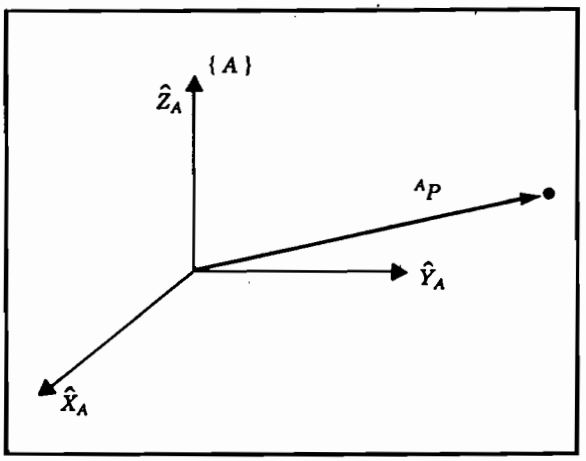

FIGURE 2.1 Vector relative to frame example.

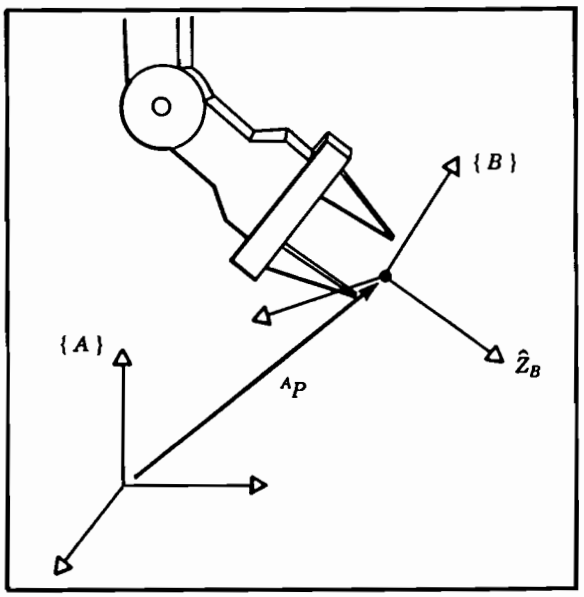

FIGURE 2.2 Locating an object in position and orientation.

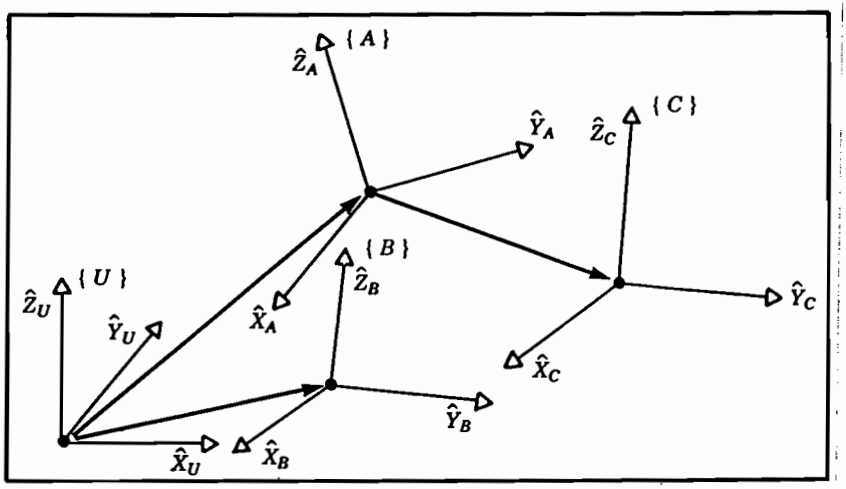

FIGURE 2.3 Example of several frames.

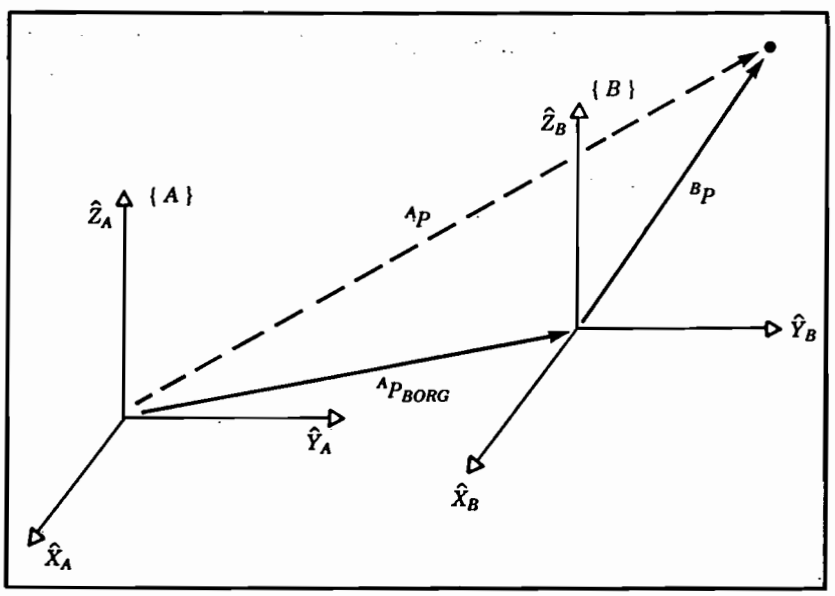

FIGURE 2.4 Translational mapping.

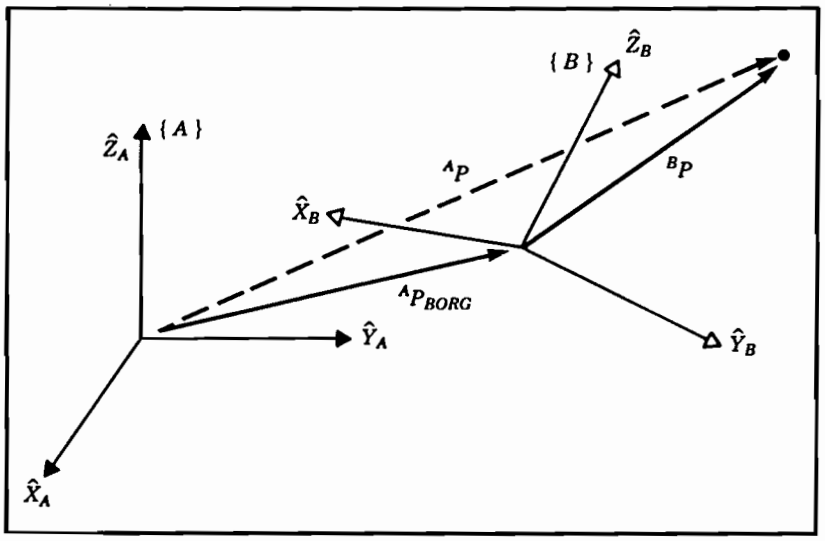

FIGURE 2.7 General transform of a vector.

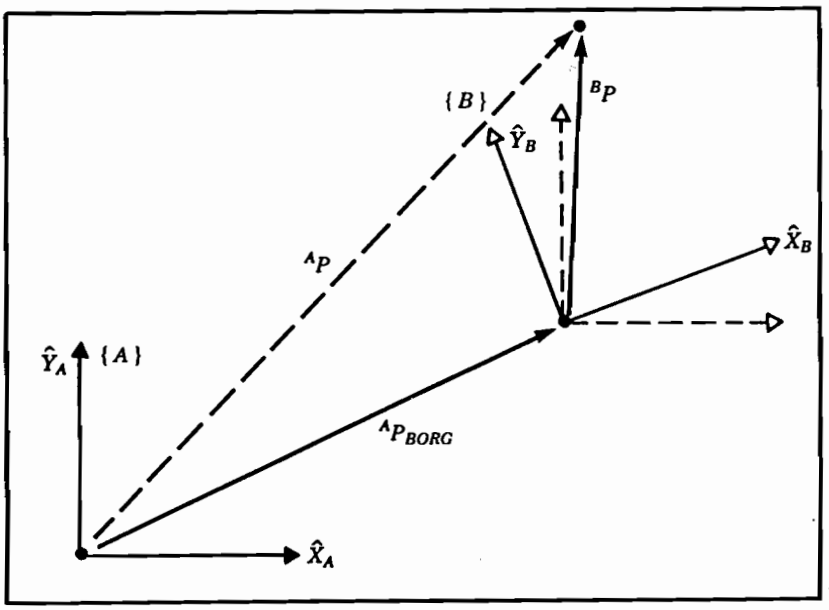

FIGURE 2.8 Frame $\{B\}$ rotated and translated.

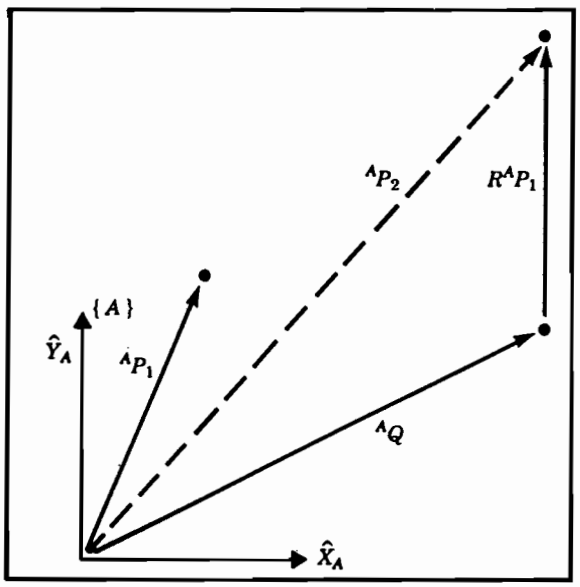

FIGURE 2.11 The vector $^AP_1$ rotated and translated to form $^AP_2$.

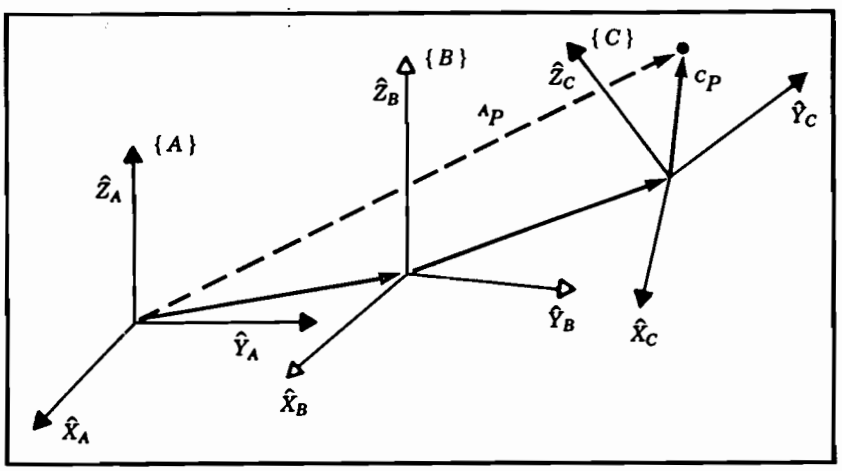

FIGURE 2.12 Compound frames: each is known relative to previous.

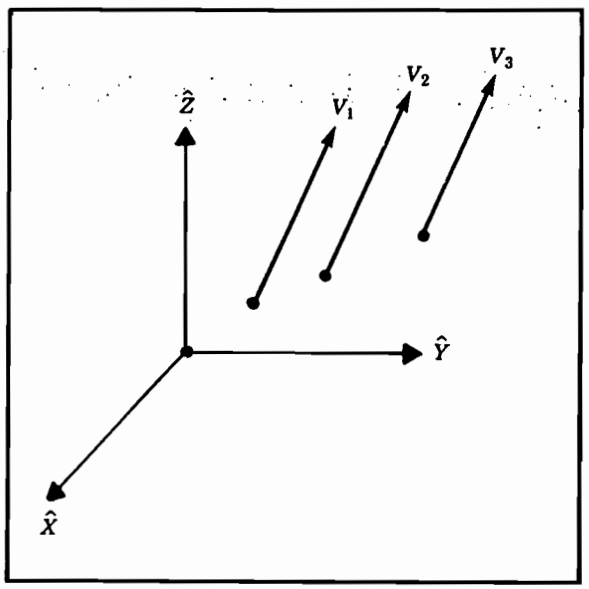

FIGURE 2.21 Equal velocity vectors.

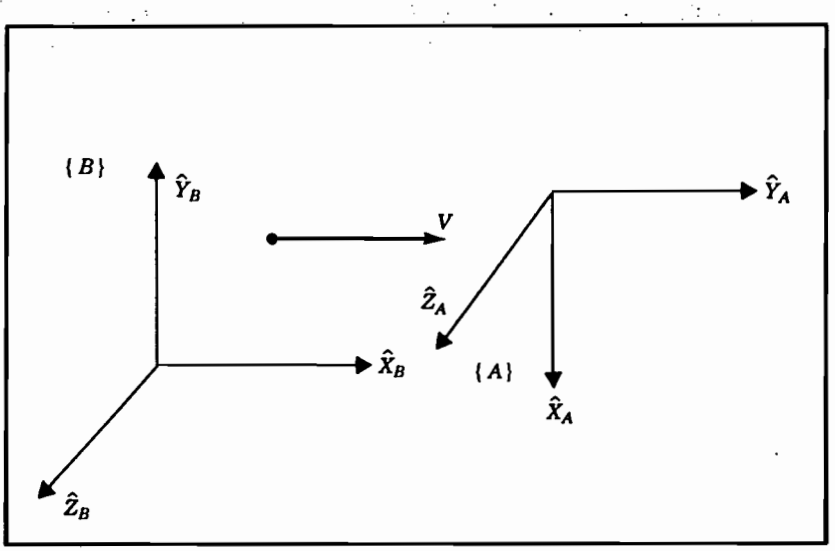

FIGURE 2.22 Transforming velocities.

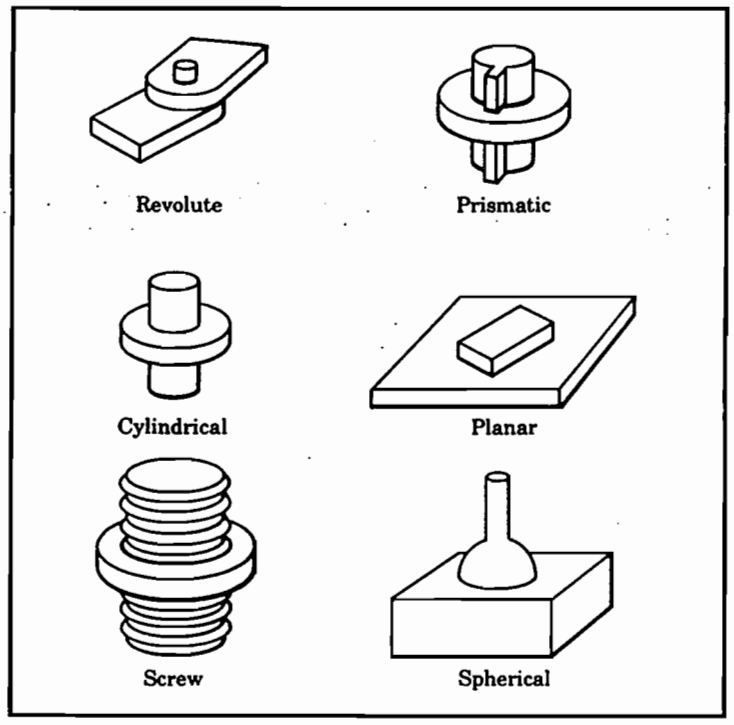

FIGURE 3.1 The six possible lower pair joints.

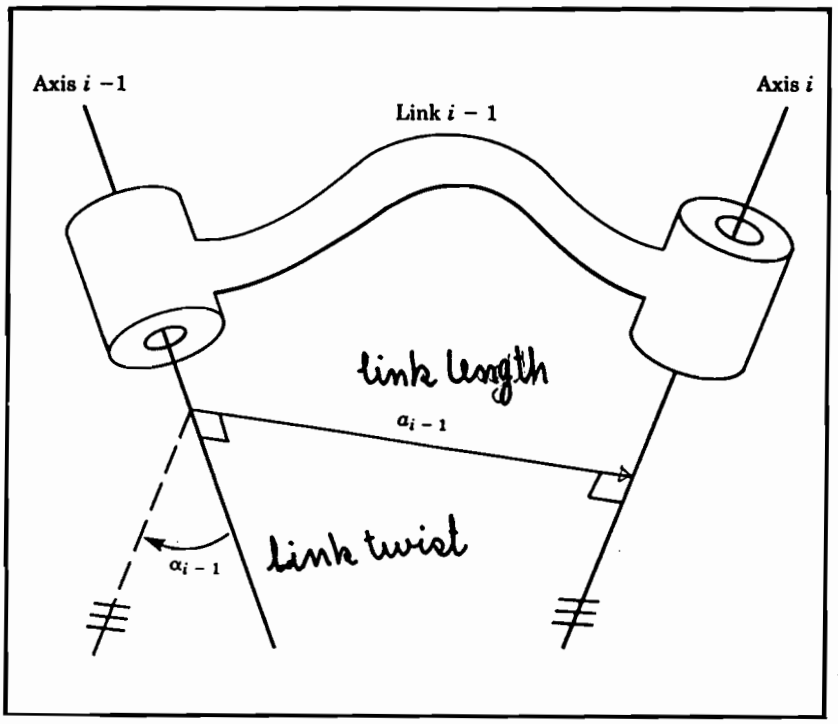

FIGURE 3.2 The kinematic function of a link is to maintain a fixed relationship between the two joint axes it supports. This relationship can be described with two parameters, the link length, a, and the link twist, α.

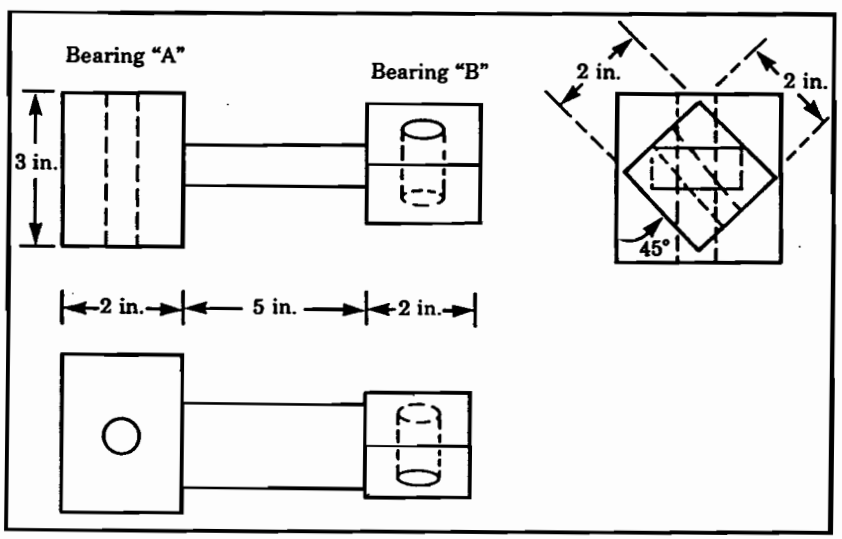

FIGURE 3.3 A simple link which supports two revolute axes.

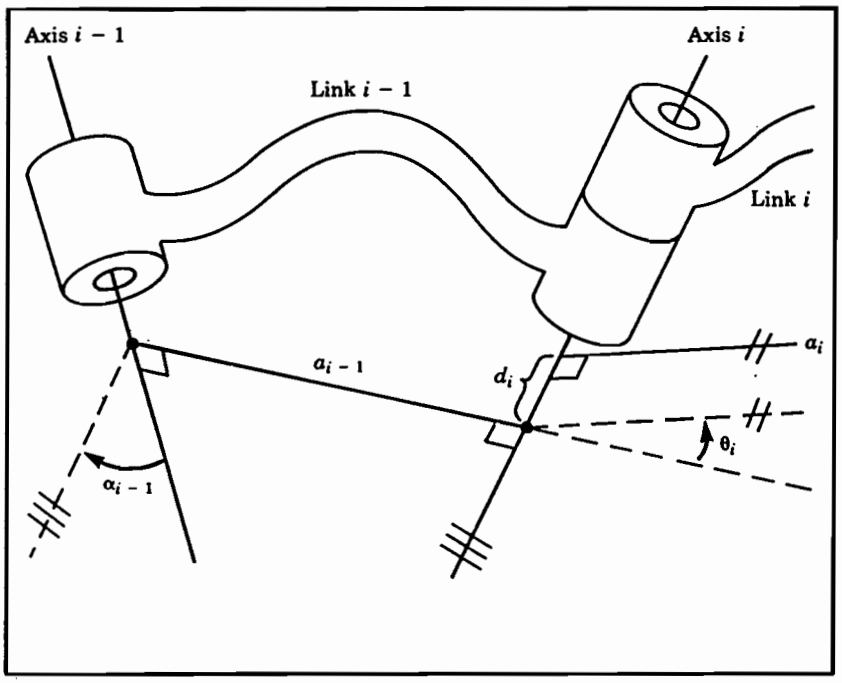

FIGURE 3.4 The link offset, d, and the joint angle, θ, are two parameters which may be used to describe the nature of the connection between neighboring links.

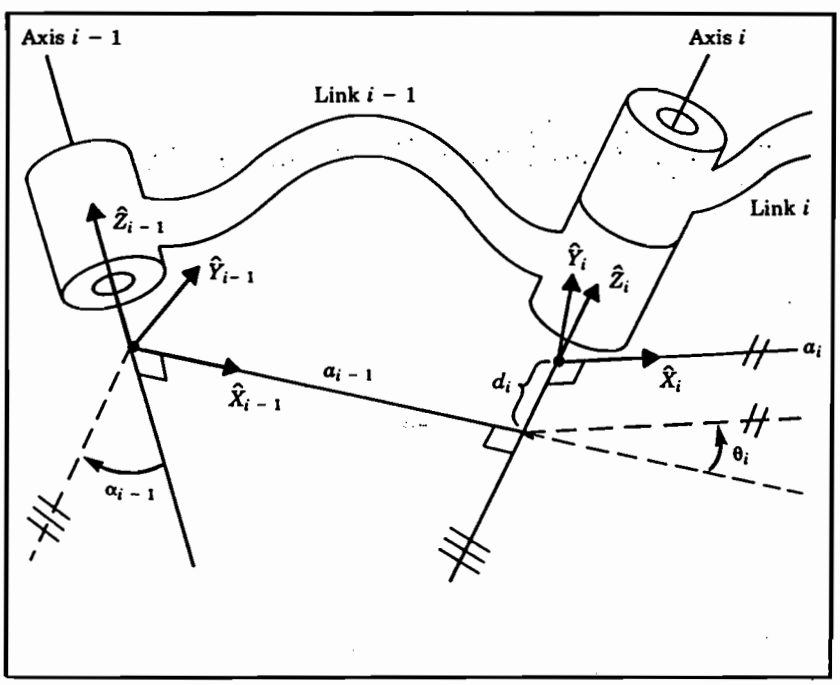

FIGURE 3.5 Link frames are attached so that frame $\{i\}$ is attached rigidly to link i.

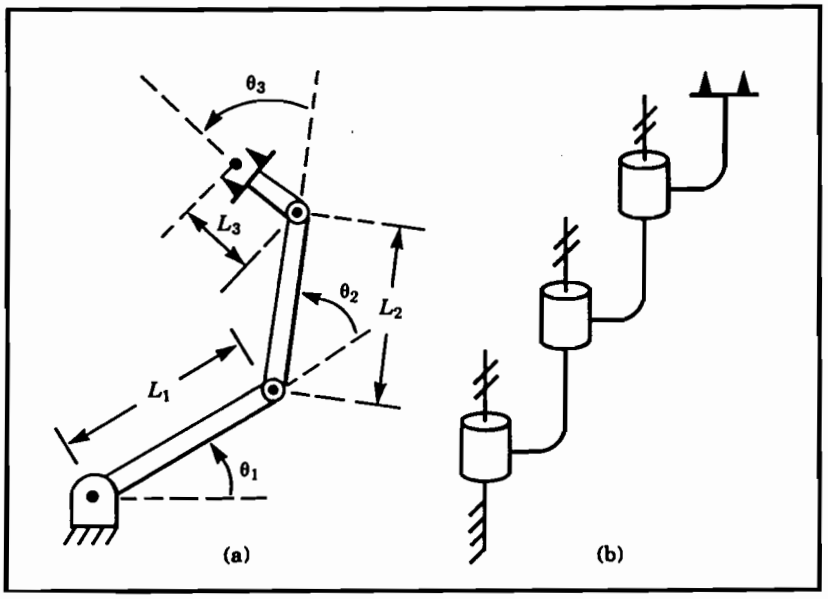

FIGURE 3.6 A three-link planar arm. On the right we show the same manipulator by means of a simple schematic notation. Hash marks on the axes indicate that they are mutually parallel.

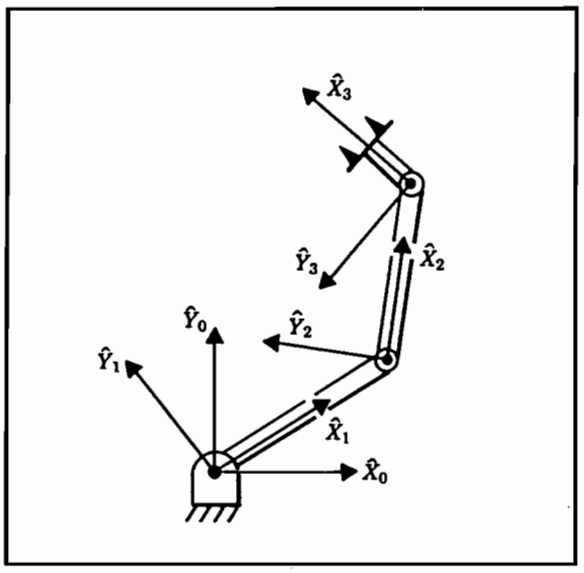

FIGURE 3.7 Link frame assignments.

i	α_{i-1}	a_{i-1}	d_i	θ_i
1	0	0	0	θ_1
2	0	L_1	0	θ_2
3	0	L_2	0	θ_3

FIGURE 3.8 Link parameters of the three-link planar manipulator.

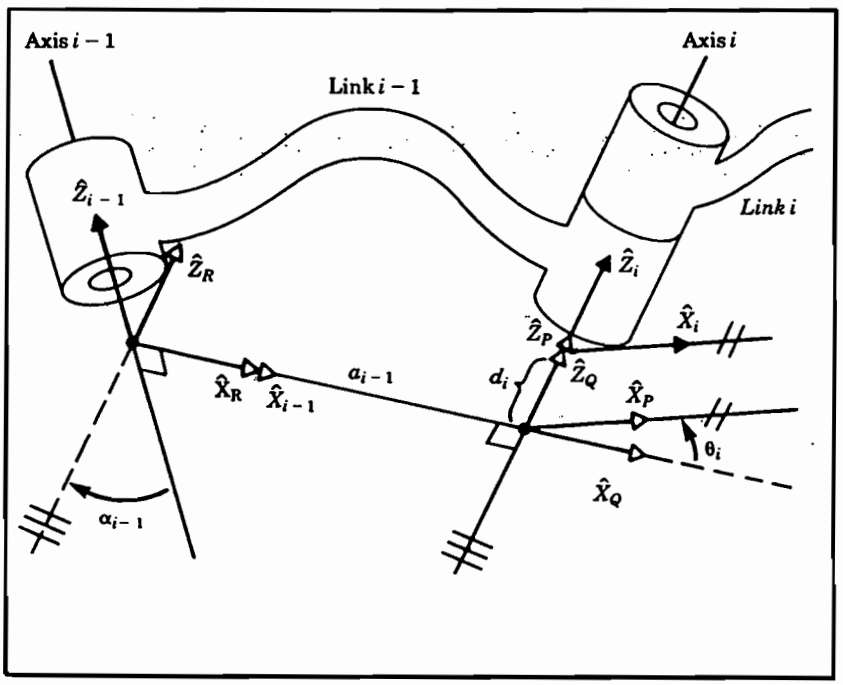

FIGURE 3.15 Location of intermediate frames $\{P\}$, $\{Q\}$, and $\{R\}$.

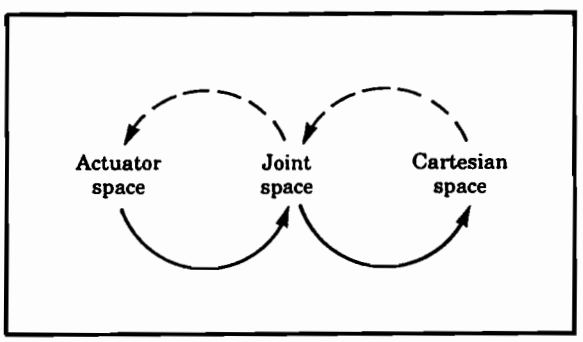

FIGURE 3.16 Mappings between kinematic descriptions.

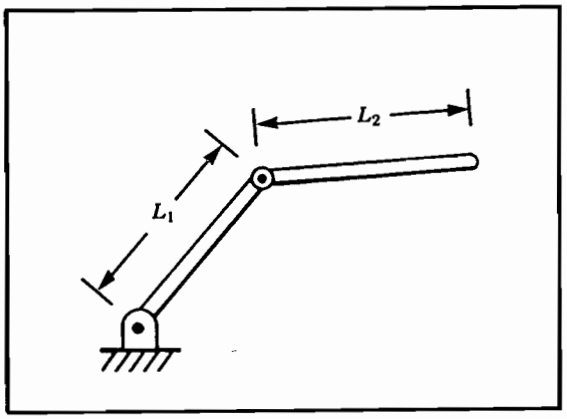

FIGURE 4.1 Two-link manipulator with link lengths l_1 and l_2.

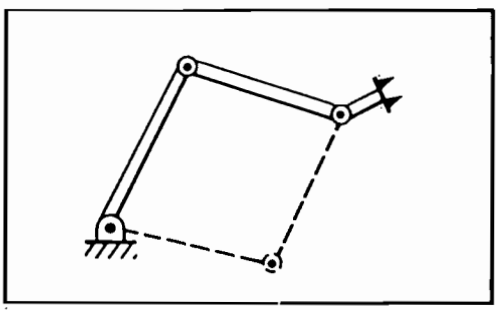

FIGURE 4.2 Three-link manipulator. Dashed lines indicate a second solution.

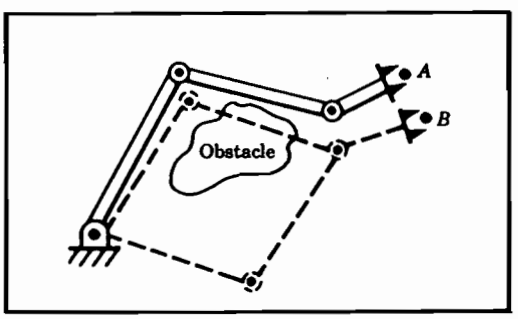

FIGURE 4.3 One of the two possible solutions to reach point B causes a collision.

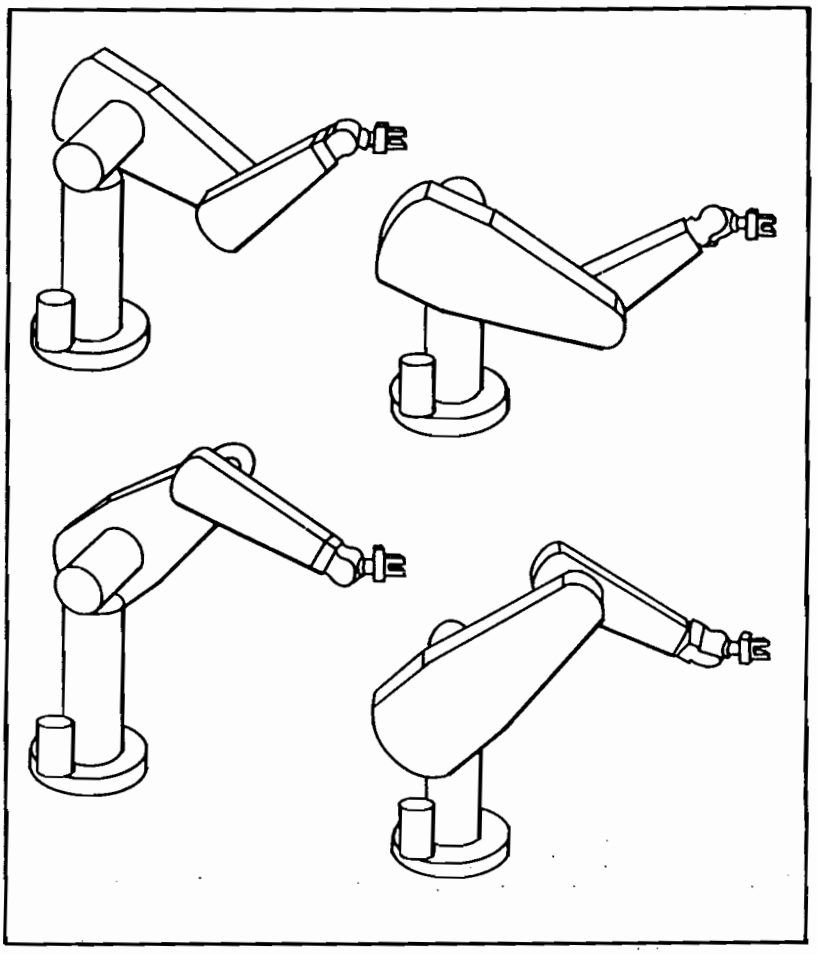

FIGURE 4.4 Four solutions of the PUMA 560.

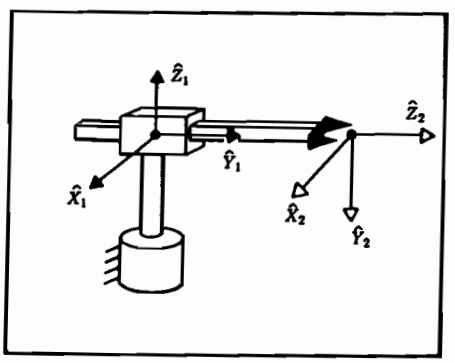

FIGURE 4.6 A polar two-link manipulator.

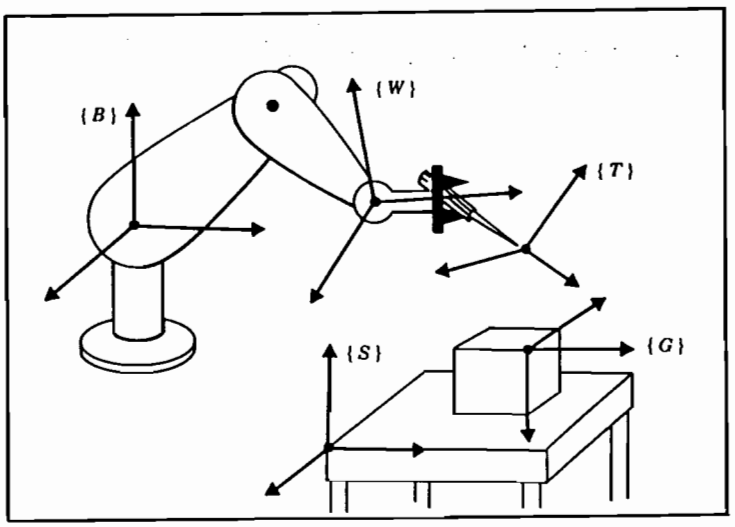

FIGURE 4.11 Location of the "standard" frames.

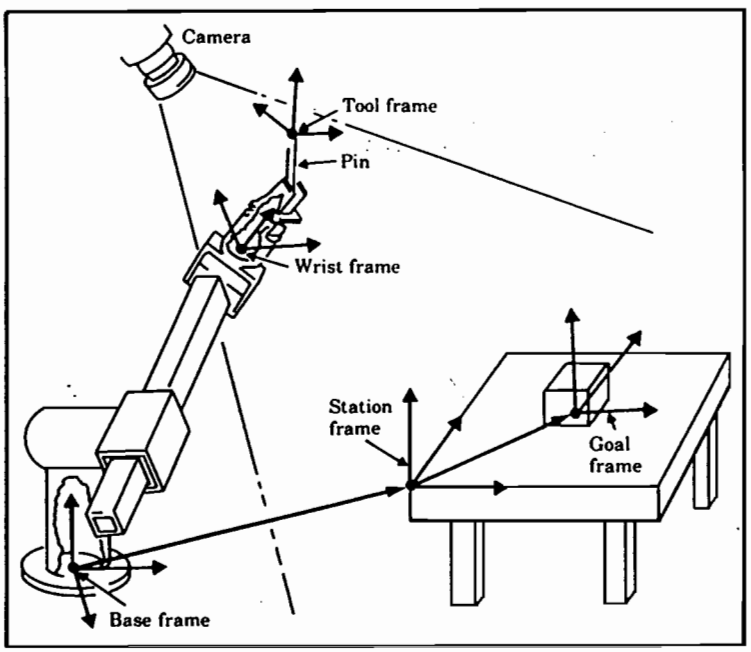

FIGURE 4.12 Example workstation.

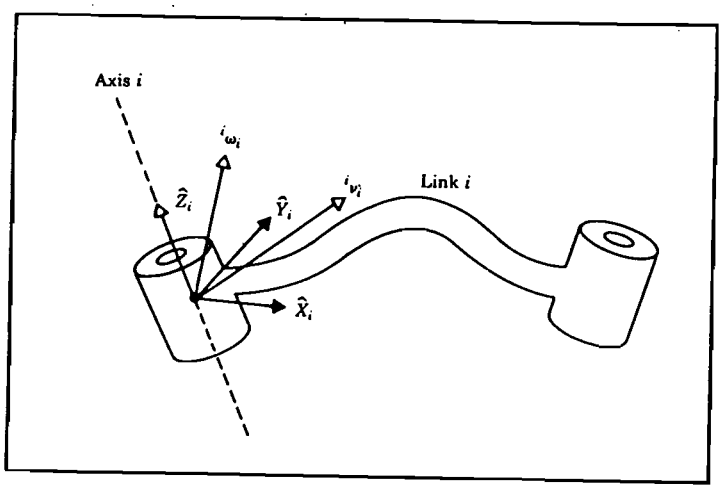

FIGURE 5.6 The velocity of link i is given by vectors v_i and ω_i which may be written in any frame, even frame $\{i\}$.

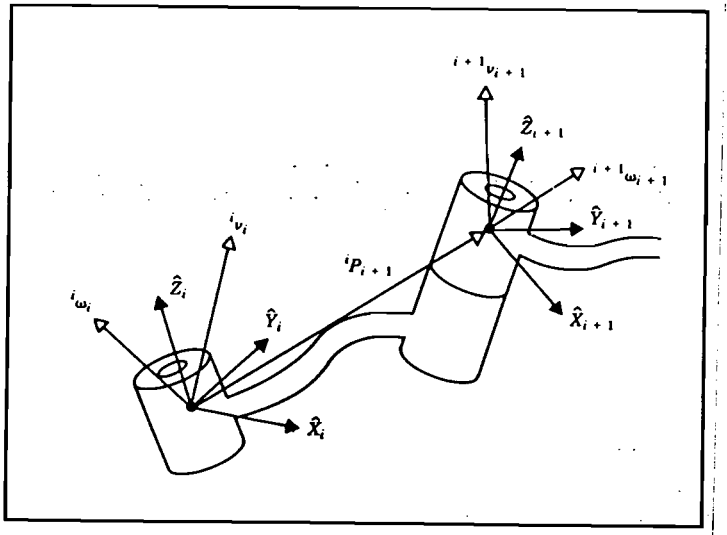

FIGURE 5.7 Velocity vectors of neighboring links.

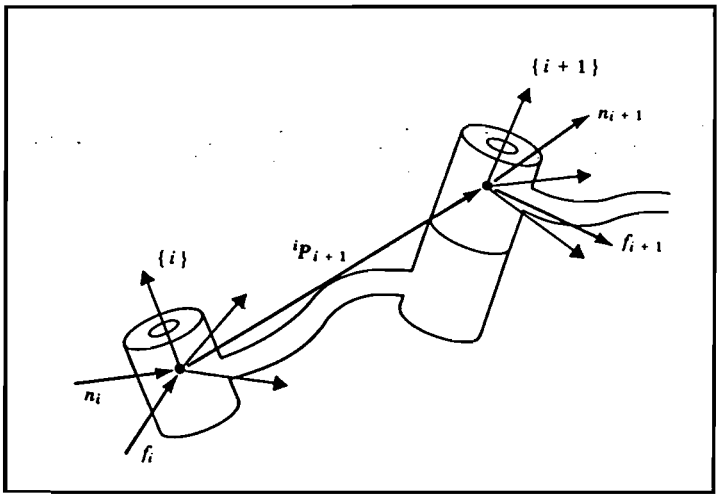

FIGURE 5.11 Static force-moment balance for a single link.

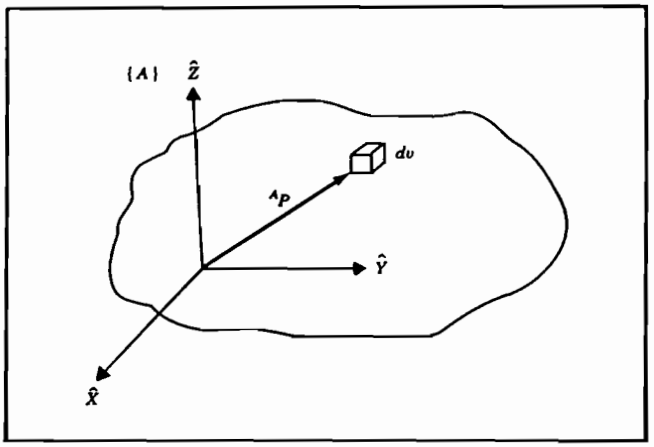

FIGURE 6.1 The inertia tensor of an object describes the object's mass distribution. Here a vector ^{A}P locates the differential volume element, dv.

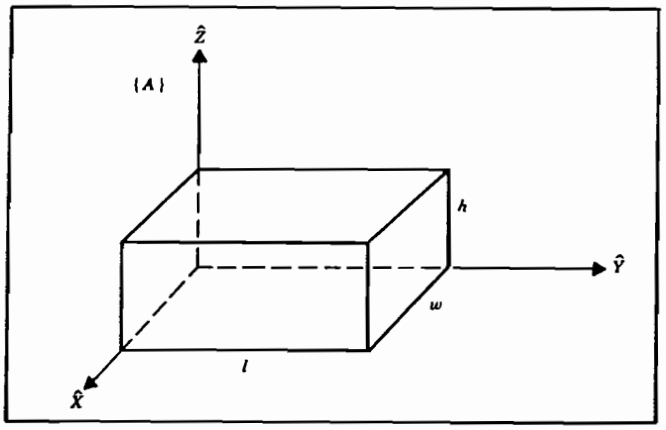

FIGURE 6.2 A body of uniform density.

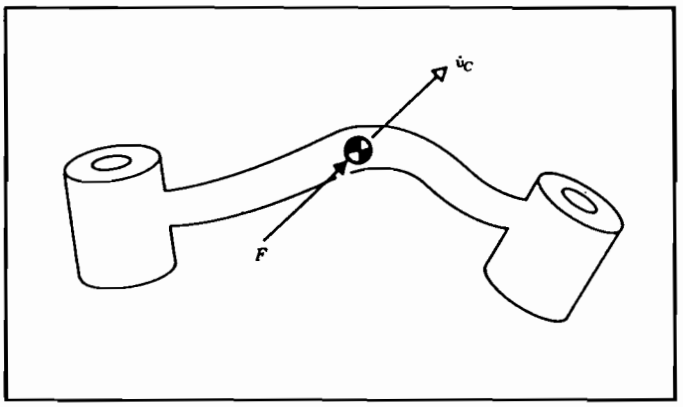

FIGURE 6.3 A force F acting at the center of mass of a body causes the body to accelerate at $\dot{v}_C$.

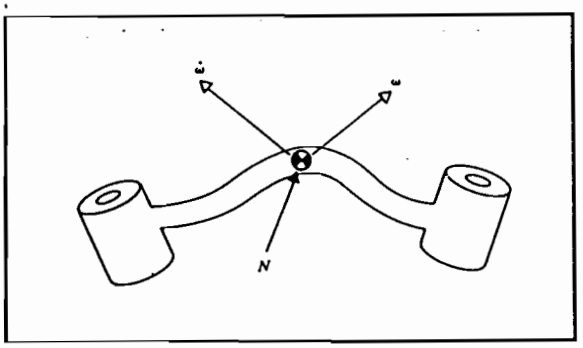

FIGURE 6.4 A moment N is acting on a body, and the body is rotating with velocity ω and accelerating at $\dot{\omega}$.

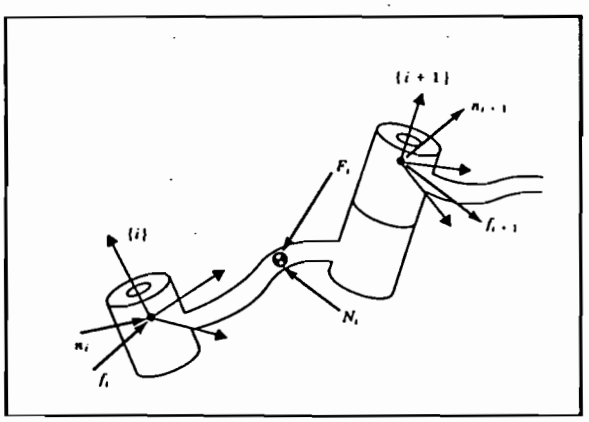

FIGURE 6.5 The force balance, including inertial forces, for a single manipulator link.

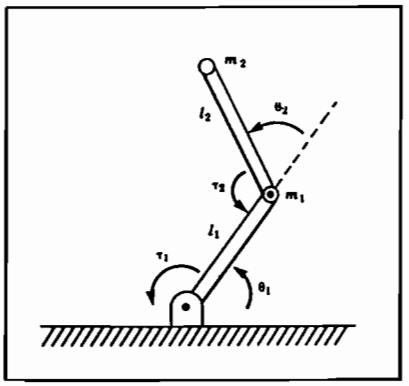

FIGURE 6.6 Two-link with point masses at distal end of links.

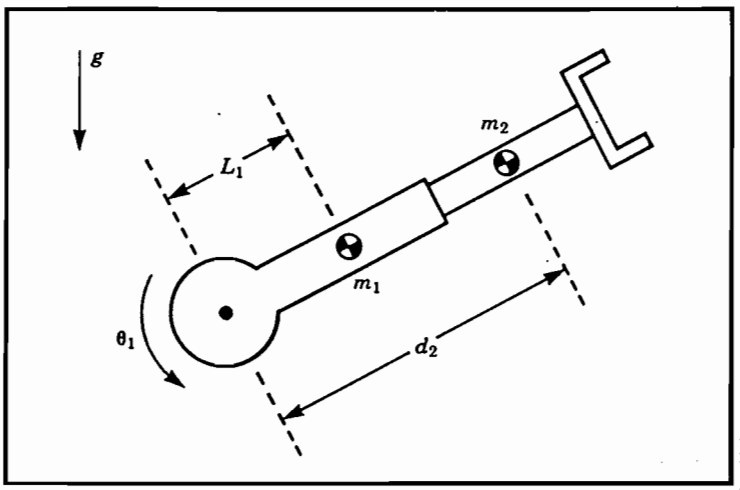

FIGURE 6.7 The RP manipulator of Example 6.5

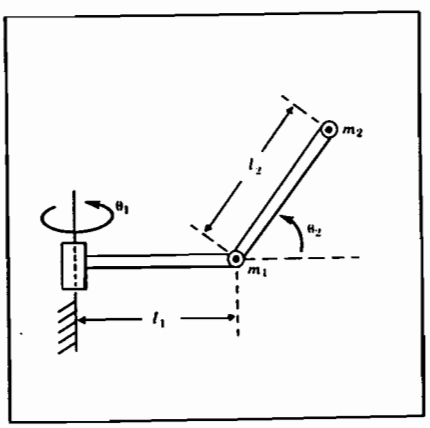

FIGURE 6.8 Two-link with point masses at distal end of links.

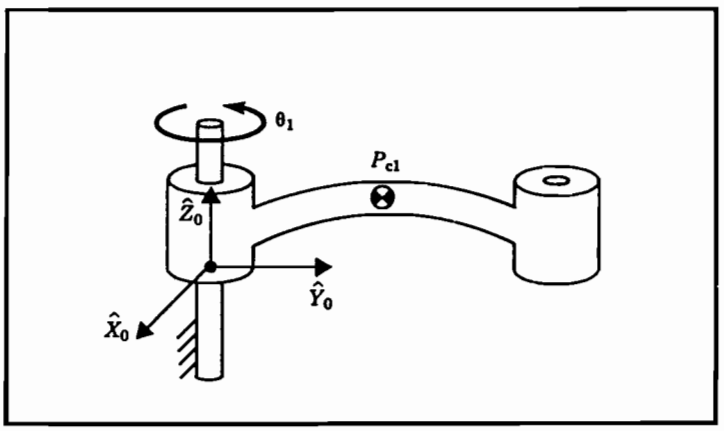

FIGURE 6.9 One-link "manipulator" of Exercise 6.12.

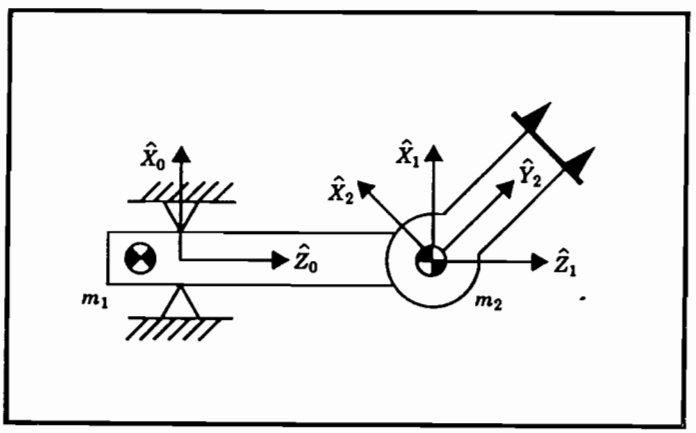

FIGURE 6.10 PR manipulator of Exercise 6.16.

Note that

$$\dot{\theta}_{i+1}\,{}^{i+1}\hat{Z}_{i+1} = {}^{i+1}\begin{bmatrix} 0 \\ 0 \\ \dot{\theta}_{i+1} \end{bmatrix}. \tag{5.44}$$

We have made use of the rotation matrix relating frames $\{i\}$ and $\{i+1\}$ in order to represent the added rotational component due to motion at the joint in frame $\{i\}$. The rotation matrix rotates the axis of rotation of joint $i+1$ into its description in frame $\{i\}$ so that the two components of angular velocity may be added.

By premultiplying both sides of (5.43) by ${}^{i+1}_{i}R$ we can find the description of the angular velocity of link $i+1$ with respect to frame $\{i+1\}$:

revolute

$$\boxed{{}^{i+1}\omega_{i+1} = {}^{i+1}_{i}R\,{}^{i}\omega_i + \dot{\theta}_{i+1}\,{}^{i+1}\hat{Z}_{i+1}.} \tag{5.45}$$

The linear velocity of the origin of frame $\{i+1\}$ is the same as that of the origin of frame $\{i\}$ plus a new component caused by rotational velocity of link i. This is exactly the situation described by (5.13), with one term vanishing because ${}^{i}P_{i+1}$ is constant in frame $\{i\}$. Therefore we have

$${}^{i}v_{i+1} = {}^{i}v_i + {}^{i}\omega_i \times {}^{i}P_{i+1}. \tag{5.46}$$

Premultiplying both sides by ${}^{i+1}_{i}R$, we compute

revolute

$$\boxed{{}^{i+1}v_{i+1} = {}^{i+1}_{i}R({}^{i}v_i + {}^{i}\omega_i \times {}^{i}P_{i+1}).} \tag{5.47}$$

Equations (5.45) and (5.47) are perhaps the most important results of this chapter. The equivalent relationships for the case that joint $i+1$ is prismatic are

prismatisk

$$\boxed{\begin{aligned} {}^{i+1}\omega_{i+1} &= {}^{i+1}_{i}R\,{}^{i}\omega_i, \\ {}^{i+1}v_{i+1} &= {}^{i+1}_{i}R({}^{i}v_i + {}^{i}\omega_i \times {}^{i}P_{i+1}) + \dot{d}_{i+1}\,{}^{i+1}\hat{Z}_{i+1}. \end{aligned}} \tag{5.48}$$

Applying these equations successively from link to link, we can compute ${}^{N}\omega_N$ and ${}^{N}v_N$, the rotational and linear velocities of the last link. Note that the resulting velocities are expressed in terms of frame $\{N\}$. This turns out to be useful, as we will see later. If the velocities are desired in terms of the base coordinate system, they can be rotated into base coordinates by multiplication with ${}^{0}_{N}R$.

EKSEMPLER

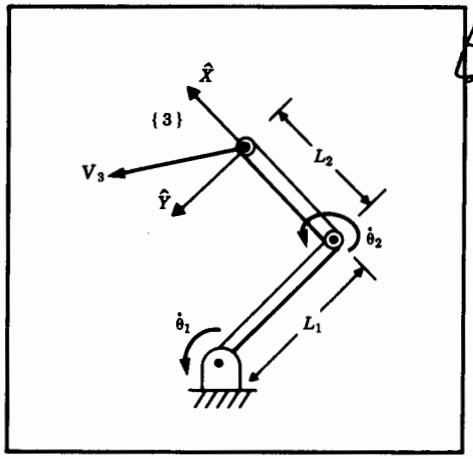

FIGURE 5.8 A two-link manipulator.

EXAMPLE 5.3

A two-link manipulator with rotational joints is shown in Fig. 5.8. Calculate the velocity of the tip of the arm as a function of joint rates. Give the answer in two forms—in terms of frame {3} and also in terms of frame {0}.

Frame {3} has been attached at the end of the manipulator as shown in Fig. 5.9, and we wish to find the velocity of the origin of this frame expressed in frame {3}. As a second part of the problem, we will express these velocities in frame {0} as well. We will start by attaching frames to the links as we have done before (see Fig. 5.9)

We will use Eqs. (5.45) and (5.47) to compute the velocity of the origin of each frame starting from the base frame {0}, which has zero velocity. Since (5.45) and (5.47) will make use of the link transformations, we compute them as:

$${}^{0}_{1}T = \begin{bmatrix} c_1 & -s_1 & 0 & 0 \\ s_1 & c_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$${}^{1}_{2}T = \begin{bmatrix} c_2 & -s_2 & 0 & l_1 \\ s_2 & c_2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{5.49}$$

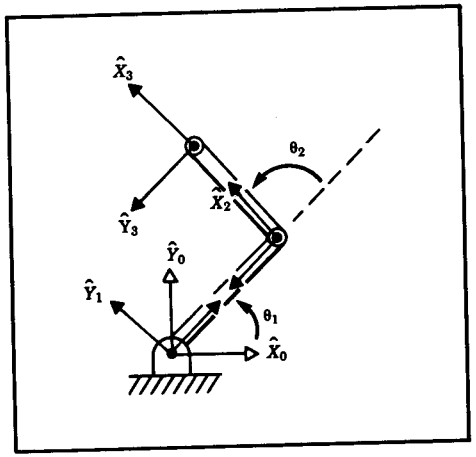

FIGURE 5.9 Frame assignments for the two-link manipulator.

$$ {}^2_3T = \begin{bmatrix} 1 & 0 & 0 & l_2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. $$

Note that these correspond to the manipulator of Example 3.3 with joint 3 permanently fixed at zero degrees. The final transformation between frames {2} and {3} need not be cast as a standard link transformation (though it may be helpful to do so). Then using (5.45) and (5.47) sequentially from link to link, we calculate

$$ {}^1\omega_1 = \begin{bmatrix} 0 \\ 0 \\ \dot{\theta}_1 \end{bmatrix}, \tag{5.50} $$

$$ {}^1v_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{5.51} $$

$$ {}^2\omega_2 = \begin{bmatrix} 0 \\ 0 \\ \dot{\theta}_1 + \dot{\theta}_2 \end{bmatrix}, \tag{5.52} $$

$$ {}^2v_2 = \begin{bmatrix} c_2 & s_2 & 0 \\ -s_2 & c_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ l_1\dot{\theta}_1 \\ 0 \end{bmatrix} = \begin{bmatrix} l_1 s_2 \dot{\theta}_1 \\ l_1 c_2 \dot{\theta}_1 \\ 0 \end{bmatrix}, \tag{5.53} $$

$$ {}^3\omega_3 = {}^2\omega_2, \tag{5.54} $$

$$ {}^3v_3 = \begin{bmatrix} l_1 s_2 \dot{\theta}_1 \\ l_1 c_2 \dot{\theta}_1 + l_2(\dot{\theta}_1 + \dot{\theta}_2) \\ 0 \end{bmatrix}. \tag{5.55} $$

Equation (5.55) is the answer. Also the rotational velocity of frame {3} is found in Eq. (5.54).

To find these velocities with respect to the nonmoving base frame, we rotate them with the rotation matrix ${}^0_3R$, which is

$$ {}^0_3R = {}^0_1R \;\; {}^1_2R \;\; {}^2_3R = \begin{bmatrix} c_{12} & -s_{12} & 0 \\ s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.56} $$

This rotation yields

$$ {}^0v_3 = \begin{bmatrix} -l_1 s_1 \dot{\theta}_1 - l_2 s_{12}(\dot{\theta}_1 + \dot{\theta}_2) \\ l_1 c_1 \dot{\theta}_1 + l_2 c_{12}(\dot{\theta}_1 + \dot{\theta}_2) \\ 0 \end{bmatrix}. \quad \blacksquare \tag{5.57} $$

It is important to point out the two distinct uses for Eqs. (5.45) and (5.47). First, they may be used as a means of deriving analytical expressions as in Example 5.2 above. Here, we manipulate the symbolic equations until we arrive at a form such as (5.55), which will be evaluated with a computer in some application. Second, they may be used directly to compute (5.45) and (5.47) as they are written. They can easily be written as a subroutine which is then applied iteratively to compute link velocities. As such they could be used for any manipulator without the need of deriving the equations for a particular manipulator. However, the computation then yields a numeric result with the structure of the equations hidden. We are often interested in the structure of an analytic result such as (5.55). Also, if we bother to do the work (that is, (5.50) through (5.57)), we generally will find that there are fewer computations left for the computer to perform in the final application.

5.7 Jacobians

The Jacobian is a multidimensional form of the derivative. Suppose, for example, that we have six functions, each of which is a function of six

To find the singular points of a mechanism we must examine the determinant of its Jacobian. Where the determinant is equal to zero, the Jacobian has lost full rank, and is singular.

$$DET\,[J(\Theta)] = \begin{vmatrix} l_1 s_2 & 0 \\ l_1 c_2 + l_2 & l_2 \end{vmatrix} = l_1 l_2 s_2 = 0. \tag{5.73}$$

Clearly, a singularity of the mechanism exists when θ_2 is 0 or 180 degrees. Physically, when $\theta_2 = 0$ the arm is stretched straight out. In this configuration, motion of the end-effector is possible only along one Cartesian direction (the one perpendicular to the arm). Therefore, the mechanism has lost one degree of freedom. Likewise when $\theta_2 = 180$ the arm is folded completely back on itself, and motion of the hand is again only possible in one Cartesian direction instead of two. We will class these singularities as workspace boundary singularities because they exist at the edge of the manipulator's workspace. Note that the Jacobian written with respect to frame {0}, or any other frame, would have yielded the same result. ■

The danger in applying (5.72) in a robot control system is that at a singular point, the inverse Jacobian blows up! This results in joint rates approaching infinity as the singularity is approached.

EXAMPLE 5.5

Consider the two-link robot from Example 5.2 moving its end-effector along the $\hat{X}$ axis at 1.0 m/s as in Fig. 5.10. Show that joint rates are reasonable when far from a singularity, but as a singularity is approached at $\theta_2 = 0$, joint rates tend to infinity.

We start by calculating the inverse of the Jacobian written in {0}:

$${}^0J^{-1}(\Theta) = \frac{1}{l_1 l_2 s_2} \begin{bmatrix} l_2 c_{12} & l_2 s_{12} \\ -l_1 c_1 - l_2 c_{12} & -l_1 s_1 - l_2 s_{12} \end{bmatrix}. \tag{5.74}$$

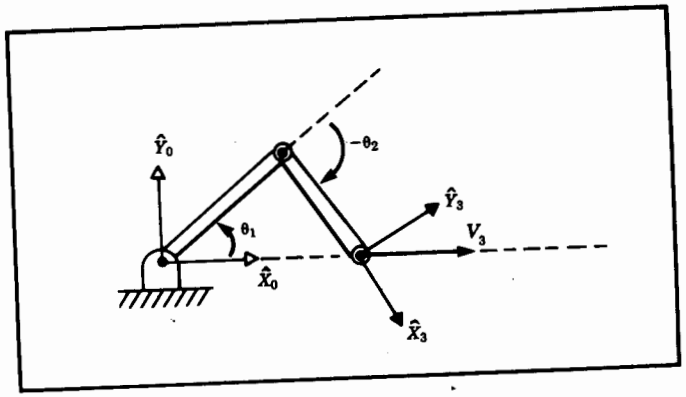

FIGURE 5.10 A two-link manipulator moving its tip at a constant linear velocity.

Then using Eq. (5.74) for a velocity of 1 m/s in the $\hat{X}$ direction we can calculate joint rates as a function of manipulator configuration:

$$\dot{\theta}_1 = \frac{c_{12}}{l_1 s_2},$$
$$\dot{\theta}_2 = -\frac{c_1}{l_2 s_2} - \frac{c_{12}}{l_1 s_2}. \tag{5.75}$$

Clearly, as the arm stretches out toward $\theta_2 = 0$ both joint rates go to infinity. ■

EXAMPLE 5.6

For the PUMA 560 manipulator, give two examples of singularities which can occur.

There is singularity when θ_3 is near −90.0 degrees. Calculation of the exact value of θ_3 is left as an exercise (see Exercise 5.14). In this situation, links 2 and 3 are "stretched out" just like the singular location of the two-link manipulator in Example 5.3. This is classed as a workspace boundary singularity.

Whenever $\theta_5 = 0.0$ degrees the manipulator is in a singular configuration. In this configuration joint axes 4 and 6 line up—both of their actions would result in the same end-effector motion, so it is as if a degree of freedom has been lost. Because this can occur interior to the workspace envelope, we will class it as a workspace interior singularity. ■

5.9 Static forces in manipulators

The chainlike nature of a manipulator leads us quite naturally to consider how forces and moments "propagate" from one link to the next. Typically the robot is pushing on something in the environment with the chain's free end (the end-effector) or is perhaps supporting a load at the hand. We wish to solve for the joint torques which must be acting to keep the system in static equilibrium.

Problem

In considering static forces in a manipulator we first lock all the joints so that the manipulator becomes a structure. We then consider each link in this structure and write a force-moment balance relationship in terms of the link frames. Finally, we compute what static torque must be acting about the joint axis in order for the manipulator to be in static equilibrium. In this way, we solve for the set of joint torques needed to support a static load acting at the end-effector.

We define special symbols for the force and torque exerted by a neighbor link:

f_i = force exerted on link i by link $i-1$,

n_i = torque exerted on link i by link $i-1$.

We will use our usual convention for assigning frames to links. Figure 5.11 shows the static forces and moments acting on link i. Summing the forces and setting them equal to zero we have

$$ {}^{i}f_i - {}^{i}f_{i+1} = 0. \tag{5.76} $$

Summing torques about the origin of frame $\{i\}$ we have

$$ {}^{i}n_i - {}^{i}n_{i+1} - {}^{i}P_{i+1} \times {}^{i}f_{i+1} = 0. \tag{5.77} $$

If we start with a description of the force and moment applied by the hand, we can calculate the force and moment applied by each link working from the last link down to the base, link 0. To do this, we formulate the force-moment expressions (5.76) and (5.77) such that they specify iterations from higher numbered links to lower numbered links. The result may be written:

$$ {}^{i}f_i = {}^{i}f_{i+1}, \tag{5.78} $$

$$ {}^{i}n_i = {}^{i}n_{i+1} + {}^{i}P_{i+1} \times {}^{i}f_{i+1}. \tag{5.79} $$

In order to write these equations in terms of only forces and moments defined within their own link frames, we transform with the rotation

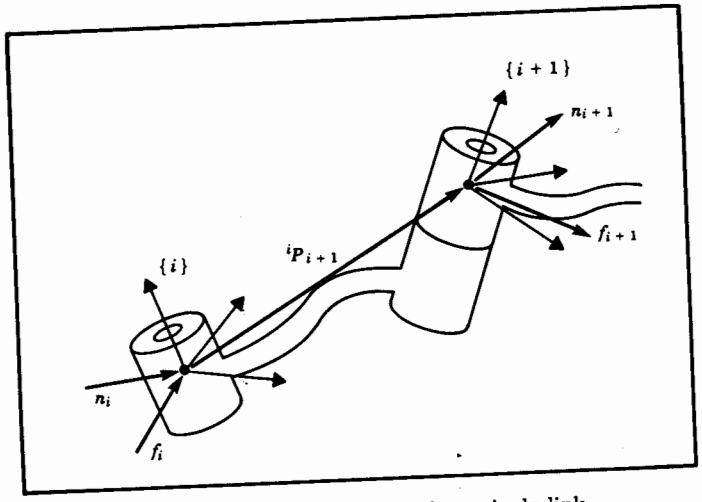

FIGURE 5.11 Static force-moment balance for a single link.

matrix describing frame $\{i+1\}$ relative to frame $\{i\}$. This leads to our most important result for static force "propagation" from link to link:

$$ {}^{i}f_i = {}^{i}_{i+1}R\, {}^{i+1}f_{i+1}, \tag{5.80} $$

$$ {}^{i}n_i = {}^{i}_{i+1}R\, {}^{i+1}n_{i+1} + {}^{i}P_{i+1} \times {}^{i}f_i. \tag{5.81} $$

Finally, the important question arises: What torques are needed at the joints in order to balance the reaction forces and moments acting on the links? All components of the force and moment vectors are resisted by the structure of the mechanism itself, except for the torque about the joint axis. Therefore, to find the joint torque required to maintain the static equilibrium, the dot product of the joint axis vector with the moment vector acting on the link is computed:

$$ \tau_i = {}^{i}n_i^T\, {}^{i}\hat{Z}_i. \tag{5.82} $$

In the case that joint i is prismatic, we compute the joint actuator force as

$$ \tau_i = {}^{i}f_i^T\, {}^{i}\hat{Z}_i. \tag{5.83} $$

Note that we are using the symbol τ even for a linear joint force.

As a matter of convention we generally define the positive direction of joint torque as the direction which would tend to move the joint in the direction of increasing joint angle.

Equations (5.80) through (5.83) give us a means to compute the joint torques needed to apply any force or moment with the end-effector of a manipulator in the static case.

EXAMPLE 5.7

The two-link manipulator of Example 5.2 is applying a force vector ${}^{3}F$ with its end-effector (consider this force to be acting at the origin of $\{3\}$). Find the required joint torques as a function of configuration and of the applied force. See Fig. 5.12.

We apply Eqs. (5.80) through (5.82) starting from the last link and going toward the base of the robot:

$$ {}^{2}f_2 = \begin{bmatrix} f_x \\ f_y \\ 0 \end{bmatrix}, \tag{5.84} $$

$$ {}^{2}n_2 = l_2\hat{X}_2 \times \begin{bmatrix} f_x \\ f_y \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ l_2 f_y \end{bmatrix}, \tag{5.85} $$

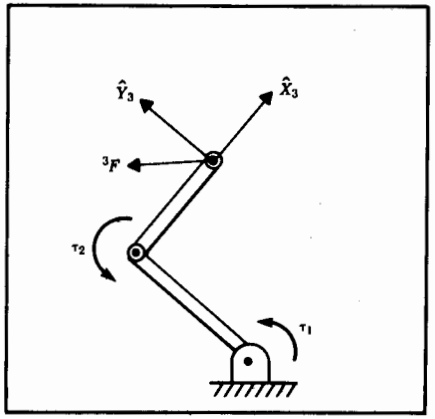

FIGURE 5.12 A two-link manipulator applying a force at its tip.

$$^1f_1 = \begin{bmatrix} c_2 & -s_2 & 0 \\ s_2 & c_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_x \\ f_y \\ 0 \end{bmatrix} = \begin{bmatrix} c_2 f_x - s_2 f_y \\ s_2 f_x + c_2 f_y \\ 0 \end{bmatrix}, \tag{5.86}$$

$$^1n_1 = \begin{bmatrix} 0 \\ 0 \\ l_2 f_y \end{bmatrix} + l_1 \hat{X}_1 \times {}^1f_1 = \begin{bmatrix} 0 \\ 0 \\ l_1 s_2 f_x + l_1 c_2 f_y + l_2 f_y \end{bmatrix}. \tag{5.87}$$

Therefore we have

$$\tau_1 = l_1 s_2 f_x + (l_2 + l_1 c_2) f_y, \tag{5.88}$$

$$\tau_2 = l_2 f_y. \tag{5.89}$$

This relationship may be written as a matrix operator as

$$\tau = \begin{bmatrix} l_1 s_2 & l_2 + l_1 c_2 \\ 0 & l_2 \end{bmatrix} \begin{bmatrix} f_x \\ f_y \end{bmatrix}. \quad \blacksquare \tag{5.90}$$

It is not a coincidence that this matrix is the transpose of the Jacobian that we found in (5.66)!

5.10 Jacobians in the force domain

We have found joint torques that will exactly balance forces at the hand in the static situation. When forces act on a mechanism, work (in the technical sense) is done if the mechanism moves through a displacement. Work is defined as a force acting through a distance and is a scalar with units of energy. The principle of **virtual work** allows us to make certain statements about the static case by allowing the amount of this displacement to go to an infinitesimal. Since work has units of energy it must be the same measured in any set of generalized coordinates. Specifically, we can equate the work done in Cartesian terms with the work done in joint space terms. In the multidimensional case, work is the dot product of a vector force or torque and a vector displacement. Thus we have

$$\mathcal{F} \cdot \delta\mathcal{X} = \tau \cdot \delta\Theta, \tag{5.91}$$

where $\mathcal{F}$ is a 6×1 Cartesian force-moment vector acting at the end-effector, $\delta\mathcal{X}$ is a 6×1 infinitesimal Cartesian displacement of the end-effector, τ is a 6×1 vector of torques at the joints, and $\delta\Theta$ is a 6×1 vector of infinitesimal joint displacements. Expression (5.91) can also be written

$$\mathcal{F}^T \delta\mathcal{X} = \tau^T \delta\Theta. \tag{5.92}$$

The definition of the Jacobian is

$$\delta\mathcal{X} = J\delta\Theta, \tag{5.93}$$

and so we may write

$$\mathcal{F}^T J \delta\Theta = \tau^T \delta\Theta, \tag{5.94}$$

which must hold for all $\delta\Theta$, and so we have

$$\mathcal{F}^T J = \tau^T. \tag{5.95}$$

Transposing both sides yields the result

$$\boxed{\tau = J^T \mathcal{F}.} \tag{5.96}$$

Equation (5.96) verifies in general what we saw in the particular case of the two-link manipulator in Example 5.6: The Jacobian transpose maps Cartesian forces acting at the hand into equivalent joint torques. When

The iterative Newton-Euler dynamics algorithm

The complete algorithm for computing joint torques from the motion of the joints is composed of two parts. First, link velocities and accelerations are iteratively computed from link 1 out to link n and the Newton-Euler equations are applied to each link. Second, forces and torques of interaction and joint actuator torques are computed recursively from link n back to link 1. The equations are summarized below for the case of all joints rotational.

Outward iterations: $i : 0 \rightarrow 5$

$$ {}^{i+1}\omega_{i+1} = {}^{i+1}_{i}R\,{}^{i}\omega_i + \dot{\theta}_{i+1}\,{}^{i+1}\hat{Z}_{i+1}, \tag{6.45} $$

$$ {}^{i+1}\dot{\omega}_{i+1} = {}^{i+1}_{i}R\,{}^{i}\dot{\omega}_i + {}^{i+1}_{i}R\,{}^{i}\omega_i \times \dot{\theta}_{i+1}\,{}^{i+1}\hat{Z}_{i+1} + \ddot{\theta}_{i+1}\,{}^{i+1}\hat{Z}_{i+1}, \tag{6.46} $$

$$ {}^{i+1}\dot{v}_{i+1} = {}^{i+1}_{i}R\left({}^{i}\dot{\omega}_i \times {}^{i}P_{i+1} + {}^{i}\omega_i \times \left({}^{i}\omega_i \times {}^{i}P_{i+1}\right) + {}^{i}\dot{v}_i\right), \tag{6.47} $$

$$ \begin{aligned} {}^{i+1}\dot{v}_{C_{i+1}} = {}& {}^{i+1}\dot{\omega}_{i+1} \times {}^{i+1}P_{C_{i+1}} \\ & + {}^{i+1}\omega_{i+1} \times \left({}^{i+1}\omega_{i+1} \times {}^{i+1}P_{C_{i+1}}\right) + {}^{i+1}\dot{v}_{i+1}, \end{aligned} \tag{6.48} $$

$$ {}^{i+1}F_{i+1} = m_{i+1}\,{}^{i+1}\dot{v}_{C_{i+1}}, \tag{6.49} $$

$$ {}^{i+1}N_{i+1} = {}^{C_{i+1}}I_{i+1}\,{}^{i+1}\dot{\omega}_{i+1} + {}^{i+1}\omega_{i+1} \times {}^{C_{i+1}}I_{i+1}\,{}^{i+1}\omega_{i+1}. \tag{6.50} $$

Inward iterations: $i : 6 \rightarrow 1$

$$ {}^{i}f_i = {}^{i}_{i+1}R\,{}^{i+1}f_{i+1} + {}^{i}F_i, \tag{6.51} $$

$$ \begin{aligned} {}^{i}n_i = {}& {}^{i}N_i + {}^{i}_{i+1}R\,{}^{i+1}n_{i+1} + {}^{i}P_{C_i} \times {}^{i}F_i \\ & + {}^{i}P_{i+1} \times {}^{i}_{i+1}R\,{}^{i+1}f_{i+1}, \end{aligned} \tag{6.52} $$

$$ \tau_i = {}^{i}n_i^T\,{}^{i}\hat{Z}_i. \tag{6.53} $$

Inclusion of gravity forces in the dynamics algorithm

The effect of gravity loading on the links can be included quite simply by setting ${}^{0}\dot{v}_0 = -G$, where G is the gravity vector. This is equivalent to saying that the base of the robot is accelerating upward with 1 G acceleration. This fictitious upward acceleration causes exactly the same effect on the links as gravity would. So, with no extra computational expense, the gravity effect is calculated.

6.6 Iterative vs. closed form

Equations (6.45) through (6.53) give a computational scheme whereby given the joint positions, velocities, and accelerations, we can compute the required joint torques. As with our development of equations to compute the Jacobian in Chapter 5, these relations can be used in two ways: as a numerical computational algorithm, or as an algorithm used analytically to develop symbolic equations.

Use of the equations as a numerical computational algorithm is attractive because the equations apply to any robot. Once the inertia tensors, link masses, P_{C_i} vectors, and ${}^{i+1}_{i}R$ matrices are specified for a particular manipulator, the equations may be applied directly to compute the joint torques corresponding to any motion.

However, we often are interested in obtaining better insight to the structure of the equations. For example, what is the form of the gravity terms? How does the magnitude of the gravity effects compare with the magnitude of the inertial effects? To investigate these and other questions, it is often useful to write **closed form** dynamic equations. These closed form equations can be derived by applying the recursive Newton-Euler equations symbolically to Θ, $\dot{\Theta}$, and $\ddot{\Theta}$. This is analogous to what we did in Chapter 5 to derive the symbolic form of the Jacobian.

6.7 An example of closed form dynamic equations

Here we compute the closed form dynamic equations for the two-link planar manipulator shown in Fig. 6.6. For simplicity, we assume that the mass distribution is extremely simple: All mass exists as a point mass at the distal end of each link. These masses are m_1 and m_2.

First we determine the value of the various quantities which will appear in the recursive Newton-Euler equations. The vectors which locate the center of mass for each link are

$$ {}^{1}P_{C_1} = l_1\hat{X}_1, $$

$$ {}^{2}P_{C_2} = l_2\hat{X}_2. $$

Because of the point mass assumption, the inertia tensor written at the center of mass for each link is the zero matrix:

$$ {}^{C_1}I_1 = 0, $$

$$ {}^{C_2}I_2 = 0. $$

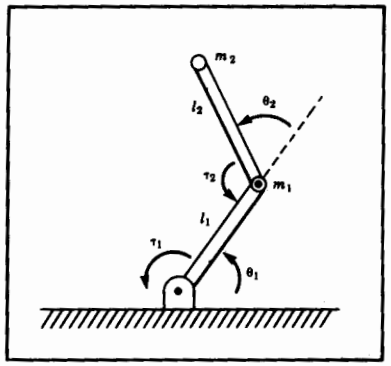

FIGURE 6.6 Two-link with point masses at distal end of links.

There are no forces acting on the end-effector, and so we have

$$f_3 = 0,$$

$$n_3 = 0.$$

The base of the robot is not rotating, and hence we have

$$\omega_0 = 0,$$

$$\dot{\omega}_0 = 0.$$

To include gravity forces we will use

$${}^0\dot{v}_0 = g\hat{Y}_0.$$

The rotation between successive link frames is given by

$${}^i_{i+1}R = \begin{bmatrix} c_{i+1} & -s_{i+1} & 0.0 \\ s_{i+1} & c_{i+1} & 0.0 \\ 0.0 & 0.0 & 1.0 \end{bmatrix},$$

$${}^{i+1}_i R = \begin{bmatrix} c_{i+1} & s_{i+1} & 0.0 \\ -s_{i+1} & c_{i+1} & 0.0 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}.$$

We now apply equations (6.45) through (6.53).

The outward iterations for link 1 are as follows:

$${}^1\omega_1 = \dot{\theta}_1 {}^1\hat{Z}_1 = \begin{bmatrix} 0 \\ 0 \\ \dot{\theta}_1 \end{bmatrix},$$

$${}^1\dot{\omega}_1 = \ddot{\theta}_1 {}^1\hat{Z}_1 = \begin{bmatrix} 0 \\ 0 \\ \ddot{\theta}_1 \end{bmatrix},$$

$${}^1\dot{v}_1 = \begin{bmatrix} c_1 & s_1 & 0 \\ -s_1 & c_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ g \\ 0 \end{bmatrix} = \begin{bmatrix} gs_1 \\ gc_1 \\ 0 \end{bmatrix},$$

$${}^1\dot{v}_{C_1} = \begin{bmatrix} 0 \\ l_1\ddot{\theta}_1 \\ 0 \end{bmatrix} + \begin{bmatrix} -l_1\dot{\theta}_1^2 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} gs_1 \\ gc_1 \\ 0 \end{bmatrix} = \begin{bmatrix} -l_1\dot{\theta}_1^2 + gs_1 \\ l_1\ddot{\theta}_1 + gc_1 \\ 0 \end{bmatrix},$$

$${}^1F_1 = \begin{bmatrix} -m_1 l_1 \dot{\theta}_1^2 + m_1 g s_1 \\ m_1 l_1 \ddot{\theta}_1 + m_1 g c_1 \\ 0 \end{bmatrix},$$

$${}^1N_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

(6.54a–f)

The outward iterations for link 2 are as follows:

$${}^2\omega_2 = \begin{bmatrix} 0 \\ 0 \\ \dot{\theta}_1 + \dot{\theta}_2 \end{bmatrix},$$

$${}^2\dot{\omega}_2 = \begin{bmatrix} 0 \\ 0 \\ \ddot{\theta}_1 + \ddot{\theta}_2 \end{bmatrix},$$

$${}^2\dot{v}_2 = \begin{bmatrix} c_2 & s_2 & 0 \\ -s_2 & c_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -l_1\dot{\theta}_1^2 + gs_1 \\ l_1\ddot{\theta}_1 + gc_1 \\ 0 \end{bmatrix} = \begin{bmatrix} l_1\ddot{\theta}_1 s_2 - l_1\dot{\theta}_1^2 c_2 + gs_{12} \\ l_1\ddot{\theta}_1 c_2 + l_1\dot{\theta}_1^2 s_2 + gc_{12} \\ 0 \end{bmatrix},$$

$${}^2\dot{v}_{C_2} = \begin{bmatrix} 0 \\ l_2(\ddot{\theta}_1 + \ddot{\theta}_2) \\ 0 \end{bmatrix} + \begin{bmatrix} -l_2(\dot{\theta}_1 + \dot{\theta}_2)^2 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} l_1\ddot{\theta}_1 s_2 - l_1\dot{\theta}_1^2 c_2 + gs_{12} \\ l_1\ddot{\theta}_1 c_2 + l_1\dot{\theta}_1^2 s_2 + gc_{12} \\ 0 \end{bmatrix},$$

(6.55a–d)

$$ {}^2F_2 = \begin{bmatrix} m_2 l_1 \ddot{\theta}_1 s_2 - m_2 l_1 \dot{\theta}_1^2 c_2 + m_2 g s_{12} - m_2 l_2 (\dot{\theta}_1 + \dot{\theta}_2)^2 \\ m_2 l_1 \ddot{\theta}_1 c_2 + m_2 l_1 \dot{\theta}_1^2 s_2 + m_2 g c_{12} + m_2 l_2 (\ddot{\theta}_1 + \ddot{\theta}_2) \\ 0 \end{bmatrix}, \qquad (6.55e\text{–}f) $$

$$ {}^2N_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. $$

Inward iterations for link 2 are as follows:

$$ {}^2f_2 = {}^2F_2 $$

$$ {}^2n_2 = \begin{bmatrix} 0 \\ 0 \\ m_2 l_1 l_2 c_2 \ddot{\theta}_1 + m_2 l_1 l_2 s_2 \dot{\theta}_1^2 + m_2 l_2 g c_{12} + m_2 l_2^2 (\ddot{\theta}_1 + \ddot{\theta}_2) \end{bmatrix} \qquad (6.56a\text{–}b) $$

Inward iterations for link 1 are as follows:

$$ {}^1f_1 = \begin{bmatrix} c_2 & -s_2 & 0 \\ s_2 & c_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} m_2 l_1 s_2 \ddot{\theta}_1 - m_2 l_1 c_2 \dot{\theta}_1^2 + m_2 g s_{12} - m_2 l_2 (\dot{\theta}_1 + \dot{\theta}_2)^2 \\ m_2 l_1 c_2 \ddot{\theta}_1 + m_2 l_1 s_2 \dot{\theta}_1^2 + m_2 g c_{12} + m_2 l_2 (\ddot{\theta}_1 + \ddot{\theta}_2) \\ 0 \end{bmatrix} + \begin{bmatrix} -m_1 l_1 \dot{\theta}_1^2 + m_1 g s_1 \\ m_1 l_1 \ddot{\theta}_1 + m_1 g c_1 \\ 0 \end{bmatrix}, $$

$$ {}^1n_1 = \begin{bmatrix} 0 \\ 0 \\ m_2 l_1 l_2 c_2 \ddot{\theta}_1 + m_2 l_1 l_2 s_2 \dot{\theta}_1^2 + m_2 l_2 g c_{12} + m_2 l_2^2 (\ddot{\theta}_1 + \ddot{\theta}_2) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ m_1 l_1^2 \ddot{\theta}_1 + m_1 l_1 g c_1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ m_2 l_1^2 \ddot{\theta}_1 - m_2 l_1 l_2 s_2 (\dot{\theta}_1 + \dot{\theta}_2)^2 + m_2 l_1 g s_2 s_{12} \\ + m_2 l_1 l_2 c_2 (\ddot{\theta}_1 + \ddot{\theta}_2) + m_2 l_1 g c_2 c_{12} \end{bmatrix}. \qquad (6.57a\text{–}b) $$

Extracting the $\hat{Z}$ components of the ${}^in_i$, we find the joint torques:

$$ \tau_1 = m_2 l_2^2 (\ddot{\theta}_1 + \ddot{\theta}_2) + m_2 l_1 l_2 c_2 (2\ddot{\theta}_1 + \ddot{\theta}_2) + (m_1 + m_2) l_1^2 \ddot{\theta}_1 - m_2 l_1 l_2 s_2 \dot{\theta}_2^2 - 2 m_2 l_1 l_2 s_2 \dot{\theta}_1 \dot{\theta}_2 + m_2 l_2 g c_{12} + (m_1 + m_2) l_1 g c_1, $$

$$ \tau_2 = m_2 l_1 l_2 c_2 \ddot{\theta}_1 + m_2 l_1 l_2 s_2 \dot{\theta}_1^2 + m_2 l_2 g c_{12} + m_2 l_2^2 (\ddot{\theta}_1 + \ddot{\theta}_2). \qquad (6.58a\text{–}b) $$

Equations (6.58) give expressions for the torque at the actuators as a function of joint position, velocity, and acceleration. Note that these rather complex functions arose from one of the simplest manipulators imaginable. Obviously, the closed form equations for a manipulator with six degrees of freedom are quite complex.

6.8 The structure of the manipulator dynamic equations

It is often convenient to express the dynamic equations of a manipulator in a single equation which hides some of the details but shows some of the structure of the equations.

The state space equation

When the Newton-Euler equations are evaluated symbolically for any manipulator, they yield a dynamic equation which can be written in the form

$$ \tau = M(\Theta)\ddot{\Theta} + V(\Theta, \dot{\Theta}) + G(\Theta), \qquad (6.59) $$

where $M(\Theta)$ is the $n \times n$ **mass matrix** of the manipulator, $V(\Theta, \dot{\Theta})$ is an $n \times 1$ vector of centrifugal and Coriolis terms, and $G(\Theta)$ is an $n \times 1$ vector of gravity terms. We use the term **state space equation** because the term $V(\Theta, \dot{\Theta})$, appearing in (6.59) has both position and velocity dependence [3].

Each element of $M(\Theta)$ and $G(\Theta)$ is a complex function which depends on Θ, the position of all the joints of the manipulator. Each element of $V(\Theta, \dot{\Theta})$ is a complex function of both Θ and $\dot{\Theta}$.

We may separate the various types of terms appearing in the dynamic equations and form the mass matrix of the manipulator, the centrifugal and Coriolis vector, and the gravity vector.

EXAMPLE 6.3

Give $M(\Theta)$, $V(\Theta, \dot{\Theta})$, and $G(\Theta)$ for the manipulator of Section 6.7.

Equation (6.59) defines the manipulator mass matrix, $M(\Theta)$: it is composed of all those terms which multiply $\ddot{\Theta}$, and is a function of Θ. Therefore we have

$$ M(\Theta) = \begin{bmatrix} l_2^2 m_2 + 2 l_1 l_2 m_2 c_2 + l_1^2 (m_1 + m_2) & l_2^2 m_2 + l_1 l_2 m_2 c_2 \\ l_2^2 m_2 + l_1 l_2 m_2 c_2 & l_2^2 m_2 \end{bmatrix}. \qquad (6.60) $$

Any manipulator mass matrix is symmetric and positive definite, and is, therefore, always invertible.

The velocity term, $V(\Theta, \dot{\Theta})$, contains all those terms which have any dependence on joint velocity. Therefore we have

$$V(\Theta, \dot{\Theta}) = \begin{bmatrix} -m_2 l_1 l_2 s_2 \dot{\theta}_2^2 - 2m_2 l_1 l_2 s_2 \dot{\theta}_1 \dot{\theta}_2 \\ m_2 l_1 l_2 s_2 \dot{\theta}_1^2 \end{bmatrix}. \quad (6.61)$$

A term like $-m_2 l_1 l_2 s_2 \dot{\theta}_2^2$ is caused by a **centrifugal force**, and is recognized as such because it depends on the square of a joint velocity. A term such as $-2m_2 l_1 l_2 s_2 \dot{\theta}_1 \dot{\theta}_2$ is caused by a **Coriolis force** and will always contain the product of two different joint velocities.

The gravity term, $G(\Theta)$, contains all those terms in which the gravitational constant, g, appears. Therefore we have

$$G(\Theta) = \begin{bmatrix} m_2 l_2 g c_{12} + (m_1 + m_2) l_1 g c_1 \\ m_2 l_2 g c_{12} \end{bmatrix}. \quad (6.62)$$

Note that the gravity term depends only on Θ, and not on its derivatives. ■

The configuration space equation

By writing the velocity dependent term, $V(\Theta, \dot{\Theta})$, in a different form, we can write the dynamic equations as

$$\tau = M(\Theta)\ddot{\Theta} + B(\Theta)\left[\dot{\Theta}\dot{\Theta}\right] + C(\Theta)\left[\dot{\Theta}^2\right] + G(\Theta), \quad (6.63)$$

where $B(\Theta)$ is a matrix of dimensions $n \times n(n-1)/2$ of Coriolis coefficients, $\left[\dot{\Theta}\dot{\Theta}\right]$ is an $n(n-1)/2 \times 1$ vector of joint velocity products given by

$$\left[\dot{\Theta}\dot{\Theta}\right] = \left[\dot{\theta}_1\dot{\theta}_2\ \dot{\theta}_1\dot{\theta}_3\ \cdots\ \dot{\theta}_{n-1}\dot{\theta}_n\right]^T, \quad (6.64)$$

$C(\Theta)$ is an $n \times n$ matrix of centrifugal coefficients, and $\left[\dot{\Theta}^2\right]$ is an $n \times 1$ vector given by

$$\left[\dot{\theta}_1^2\ \dot{\theta}_2^2\ \cdots\ \dot{\theta}_n^2\right]^T. \quad (6.65)$$

We will call (6.63) the **configuration space equation** since the matrices are functions only of manipulator position [3].

In this form of the dynamic equations, the complexity of the computation is seen to be in the form of computing various parameters which are a function of only the manipulator position, Θ. This is important in applications (such as computer control of a manipulator) in which the dynamic equations must be updated as the manipulator moves. (Equation (6.63) gives a form in which parameters which are only a function of joint position, and can be updated at a rate related to how fast the manipulator is changing configuration.) We will consider this form again with regard to the problem of manipulator control in Chapter 10.

EXAMPLE 6.4

Give $B(\Theta)$ and $C(\Theta)$ (from (6.63)) for the manipulator of Section 6.7.

For this simple two-link manipulator, we have

$$\left[\dot{\Theta}\dot{\Theta}\right] = [\dot{\theta}_1\dot{\theta}_2], \qquad \left[\dot{\Theta}^2\right] = \begin{bmatrix} \dot{\theta}_1^2 \\ \dot{\theta}_2^2 \end{bmatrix}. \quad (6.66)$$

So we see that

$$B(\Theta) = \begin{bmatrix} -2m_2 l_1 l_2 s_2 \\ 0 \end{bmatrix} \quad (6.67)$$

and

$$C(\Theta) = \begin{bmatrix} 0 & -m_2 l_1 l_2 s_2 \\ m_2 l_1 l_2 s_2 & 0 \end{bmatrix}. \quad (6.68)$$

■

6.9 Lagrangian formulation of manipulator dynamics

The Newton-Euler approach is based on the elementary dynamic formulas (6.29) and (6.30), and on an analysis of forces and moments of constraint acting between the links. As an alternative to the Newton-Euler method, in this section we briefly introduce the **Lagrangian dynamic formulation**. Whereas the Newton-Euler formulation might be said to be a "force balance" approach to dynamics, the Lagrangian formulation is an "energy-based" approach to dynamics. Of course, for the same manipulator, both will give the same equations of motion. Our statement of Lagrangian dynamics will be brief and somewhat specialized to the case of a serial chain mechanical manipulator with rigid links. For a more complete and general reference, see [4].

We start by developing an expression for the kinetic energy of a manipulator. The kinetic energy of the ith link, k_i, can be expressed as

$$k_i = \frac{1}{2} m_i v_{C_i}^T v_{C_i} + \frac{1}{2} {}^i\omega_i^T\, {}^{C_i}I_i\, {}^i\omega_i, \quad (6.69)$$

where the first term is kinetic energy due to linear velocity of the link's center of mass, and the second term is kinetic energy due to angular velocity of the link. The total kinetic energy of the manipulator is the sum of the kinetic energy in the individual links; that is,

$$k = \sum_{i=1}^{n} k_i. \quad (6.70)$$

We next develop a relationship between joint space and Cartesian acceleration, starting with the definition of the Jacobian,

$$\dot{\mathcal{X}} = J\dot{\Theta}, \tag{6.95}$$

and differentiating to obtain

$$\ddot{\mathcal{X}} = \dot{J}\dot{\Theta} + J\ddot{\Theta}. \tag{6.96}$$

Solving (6.96) for joint space acceleration leads to

$$\ddot{\Theta} = J^{-1}\ddot{\mathcal{X}} - J^{-1}\dot{J}\dot{\Theta}. \tag{6.97}$$

Substituting (6.97) into (6.94) we have

$$\mathcal{F} = J^{-T}M(\Theta)J^{-1}\ddot{\mathcal{X}} - J^{-T}M(\Theta)J^{-1}\dot{J}\dot{\Theta} + J^{-T}V(\Theta,\dot{\Theta}) + J^{-T}G(\Theta), \tag{6.98}$$

from which we derive the expressions for the terms in the Cartesian dynamics as

$$\begin{aligned} M_x(\Theta) &= J^{-T}(\Theta)\, M(\Theta)\, J^{-1}(\Theta), \\ V_x(\Theta,\dot{\Theta}) &= J^{-T}(\Theta)\, \left(V(\Theta,\dot{\Theta}) - M(\Theta)\, J^{-1}(\Theta)\, \dot{J}(\Theta)\, \dot{\Theta}\right), \\ G_x(\Theta) &= J^{-T}(\Theta)\, G(\Theta). \end{aligned} \tag{6.99}$$

The Jacobian appearing in equations (6.99) is written in the same frame as $\mathcal{F}$ and $\mathcal{X}$ in (6.91) though the choice of this frame is arbitrary.* Note that when the manipulator approaches a singularity, certain quantities in the Cartesian dynamics become infinite.

EXAMPLE 6.6

Derive the Cartesian space form of the dynamics for the two-link planar arm of Section 6.7. Write the dynamics in terms of a frame attached to the end of the second link.

For this manipulator we have already obtained the dynamics (in Section 6.7), and the Jacobian (equation (5.66)), which we restate here:

$$J(\Theta) = \begin{bmatrix} l_1 s_2 & 0 \\ l_1 c_2 + l_2 & l_2 \end{bmatrix}. \tag{6.100}$$

First compute the inverse Jacobian:

$$J^{-1}(\Theta) = \frac{1}{l_1 l_2 s_2}\begin{bmatrix} l_2 & 0 \\ -l_1 c_2 - l_2 & l_1 s_2 \end{bmatrix} \tag{6.101}$$

* Certain choices may facilitate computation.

and the time derivative of the Jacobian:

$$\dot{J}(\Theta) = \begin{bmatrix} l_1 c_2 \dot{\theta}_2 & 0 \\ -l_1 s_2 \dot{\theta}_2 & 0 \end{bmatrix}. \tag{6.102}$$

Using (6.99) and the results of Section 6.7 we obtain

$$\begin{aligned} M_x(\Theta) &= \begin{bmatrix} m_2 + \frac{m_1}{s_2^2} & 0 \\ 0 & m_2 \end{bmatrix}, \\ V_x(\Theta,\dot{\Theta}) &= \begin{bmatrix} -(m_2 l_1 c_2 + m_2 l_2)\dot{\theta}_1^2 - m_2 l_2 \dot{\theta}_2^2 - (2m_2 l_2 + m_2 l_1 c_2 + m_1 l_1 \frac{c_2}{s_2^2})\dot{\theta}_1\dot{\theta}_2 \\ m_2 l_1 s_2 \dot{\theta}_1^2 + l_1 m_2 s_2 \dot{\theta}_1 \dot{\theta}_2 \end{bmatrix}, \\ G_x(\Theta) &= \begin{bmatrix} m_1 g \frac{c_1}{s_2} + m_2 g s_{12} \\ m_2 g c_{12} \end{bmatrix}. \end{aligned} \tag{6.103}$$

When $s_2 = 0$ the manipulator is in a singular position and some of the dynamic terms go to infinity. For example, when $\theta_2 = 0$ (arm stretched straight out), the effective Cartesian mass of the end-effector becomes infinite in the $\hat{X}_2$ direction of the link 2 tip frame, as expected. In general, at a singular configuration there is a certain direction, the *singular direction* in which motion is impossible, but general motion in the subspace "orthogonal" to this direction is possible [8]. ■

The Cartesian configuration space torque equation

Combining (6.91) and (6.92) we can write equivalent joint torques with the dynamics expressed in Cartesian space:

$$\tau = J^T(\Theta)\left(M_x(\Theta)\ddot{\mathcal{X}} + V_x(\Theta,\dot{\Theta}) + G_x(\Theta)\right). \tag{6.104}$$

We will find it usefull= to write this equation in the form

$$\tau = J^T(\Theta)M_x(\Theta)\ddot{\mathcal{X}} + B_x(\Theta)\left[\dot{\Theta}\dot{\Theta}\right] + C_x(\Theta)\left[\dot{\Theta}^2\right] + G(\Theta), \tag{6.105}$$

where $B_x(\Theta)$ is a matrix of dimensions $n \times n(n-1)/2$ of Coriolis coefficients, $\left[\dot{\Theta}\dot{\Theta}\right]$ is an $n(n-1)/2 \times 1$ vector of joint velocity products given by

$$\left[\dot{\Theta}\dot{\Theta}\right] = \left[\dot{\theta}_1\dot{\theta}_2\ \dot{\theta}_1\dot{\theta}_3\ \cdots\ \dot{\theta}_{n-1}\dot{\theta}_n\right]^T, \tag{6.106}$$

$C_x(\Theta)$ is an $n \times n$ matrix of centrifugal coefficients, and $\left[\dot{\Theta}^2\right]$ is an $n \times 1$ vector given by

$$\left[\dot{\theta}_1^2\ \dot{\theta}_2^2\ \cdots\ \dot{\theta}_n^2\right]^T. \tag{6.107}$$

Note that in (6.105), $G(\Theta)$ is the same as in the joint space equation, but in general $B_x(\Theta) \neq B(\Theta)$ and $C_x(\Theta) \neq C(\Theta)$.

EXAMPLE 6.7

Determine $B_x(\Theta)$ and $C_x(\Theta)$ (from (6.105)) for the manipulator of Section 6.7.

If we form the product $J^T(\Theta)V_x(\Theta,\dot{\Theta})$ we find that

$$B_x(\Theta) = \begin{bmatrix} m_1 l_1^2 \frac{c_2}{s_2} - m_2 l_1 l_2 s_2 \\ m_2 l_1 l_2 s_2 \end{bmatrix} \tag{6.108}$$

and

$$C_x(\Theta) = \begin{bmatrix} 0 & -m_2 l_1 l_2 s_2 \\ m_2 l_1 l_2 s_2 & 0 \end{bmatrix}. \quad \blacksquare \tag{6.109}$$

6.11 Inclusion of nonrigid body effects

It is important to realize that the dynamic equations we have derived *do not encompass all the effects acting on a manipulator*. They include just those forces which arise from rigid body mechanics. The most important source of forces that are *not* included is friction. All mechanisms are, of course, affected by frictional forces. In present day manipulators in which significant gearing is typical, the forces due to friction can actually be quite large—perhaps equaling 25% of the torque required to move the manipulator in typical situations.

In order to make dynamic equations reflect the reality of the physical device, it is important to model (at least approximately) these forces of friction. A very simple model for friction is **viscous friction**, in which the torque due to friction is proportional to the velocity of joint motion. Thus we have

$$\tau_{friction} = v\dot{\theta}, \tag{6.110}$$

where v is a viscous friction constant. Another possible simple model for friction, **Coulomb friction**, is sometimes used. Coulomb friction is constant except for a sign dependence on the joint velocity:

$$\tau_{friction} = c\ sgn(\dot{\theta}), \tag{6.111}$$

where c is a Coulomb friction constant. The value of c is often taken at one value when $\dot{\theta} = 0$, the static coefficient, and at a lower value, the dynamic coefficient, when $\dot{\theta} \neq 0$. Whether a joint of a particular manipulator exhibits viscous or Coulomb friction is a complicated issue of lubrication and other effects. A reasonable model is to include both, since both effects are likely:

$$\tau_{friction} = c\ sgn(\dot{\theta}) + v\dot{\theta}. \tag{6.112}$$

It turns out that in many manipulator joints, friction also displays a dependence on the joint position. A major cause of this effect might be gears which are not perfectly round—their eccentricity would cause friction to change according to joint position. So a fairly complex friction model would have the form

$$\tau_{friction} = f(\theta,\dot{\theta}). \tag{6.113}$$

These friction models are then added to the other dynamic terms derived from the rigid body model, yielding the more complete model

$$\tau = M(\Theta)\ddot{\Theta} + V(\Theta,\dot{\Theta}) + G(\Theta) + F(\Theta,\dot{\Theta}) \tag{6.114}$$

There are other effects which are also neglected in this model. For example, the assumption of rigid body links means that we have failed to include bending effects (which give rise to resonances) in our equations of motion. However, these effects are extremely difficult to model, and are beyond the scope of this book (see [9,10]).

6.12 Dynamic simulation

To simulate the motion of a manipulator we must make use of a model of the dynamics, such as we have just developed. Given the dynamics written in closed form as in (6.59), simulation requires solving the dynamic equation for acceleration:

$$\ddot{\Theta} = M^{-1}(\Theta)\left[\tau - V(\Theta,\dot{\Theta}) - G(\Theta) - F(\Theta,\dot{\Theta})\right]. \tag{6.115}$$

We may then apply any of several known **numerical integration** techniques to integrate the acceleration to compute future positions and velocities.

Given initial conditions on the motion of the manipulator, usually in the form:

$$\begin{aligned} \Theta(0) &= \Theta_0, \\ \dot{\Theta}(0) &= 0, \end{aligned} \tag{6.116}$$

we numerically integrate (6.115) forward in time by steps of size Δt. There are many methods of performing numerical integration [11]. Here we introduce the simplest integration scheme, called *Euler integration*, which is accomplished as follows: Starting with $t = 0$, iteratively compute

$$\begin{aligned} \dot{\Theta}(t+\Delta t) &= \dot{\Theta}(t) + \ddot{\Theta}(t)\Delta t, \\ \Theta(t+\Delta t) &= \Theta(t) + \dot{\Theta}(t)\Delta t + \frac{1}{2}\ddot{\Theta}(t)\Delta t^2, \end{aligned} \tag{6.117}$$

from Equation (13-70) so that

$$u = \tilde{B}_\alpha^{-1}(x, \{\tilde{y}_i^{(j)} \mid 1 \le i \le \alpha - 1, i \le j \le \alpha - 1\}) \times (\tilde{Y}_\alpha - \tilde{A}_\alpha(x, \{\tilde{y}_i^{(j)} \mid 1 \le i \le \alpha - 1, i \le j \le \alpha\})) \stackrel{\text{def}}{=} \phi(x, \{\tilde{y}_i^{(j)} \mid 1 \le i \le \alpha, i \le j \le \alpha\})) \tag{13-71}$$

is well defined in a neighborhood of $(x_0, \{\tilde{y}_{i0}^{(j)} (1 \le i \le \alpha - 1, i \le j \le \alpha - 1\}))$. Let γ_i be the lowest time derivative and δ_i the highest time derivative of y_i appearing in Equation (13-71). Then it can be shown easily that ϕ is of the form

$$\phi(x, \{y_i^{(j)} \mid 1 \le i \le m, \gamma_i \le j \le \delta_i\}) = \phi_1(x, \{y_i^{(j)} \mid 1 \le i \le m, \gamma_i \le j \le \delta_i - 1\} + \sum_{k=1}^{m} \phi_{2k}(x, \{y_i^{(j)} \mid 1 \le i \le m, \gamma_i \le j \le \delta_i - 1\}) y_i^{(\delta_i)}$$

Then a desired dynamic feedback control law can be synthesized following the procedure of Remark 13-6.

13.4 NONLINEAR AUTOPILOT DESIGN

In this section, we will illustrate the application of the inversion-based controllers to the flight control systems. Specifically, the six-degree-of-freedom model for a flight vehicle will be presented in Section 13.4.1. The autopilots for the flight vehicle will be designed using the input–output feedback linearization and sliding mode control approaches in Sections 13.4.2 and 13.4.3, respectively. Simulations showing the performance of the autopilots for the HAVE DASH II bank-to-turn missile are also included.

13.4.1 Six-degree-of-freedom Flight Vehicle Model

The motion of a rigid-body flight vehicle is governed by its translational motion and rotational motion. Let (X, Y, Z) denote a coordinate frame fixed in the vehicle body with the origin located at the vehicle's center of gravity. Let the axis, X, point forward out of the nose of the vehicle, the axis, Z, point down, and then the axis Y be chosen to form a right-hand orthogonal triad. Under these assumptions, the rotational equations of motion are

$$\left.\begin{aligned} L &= I_{xx}\dot{P} + (I_{zz} - I_{yy})QR \\ M &= I_{yy}\dot{Q} + (I_{xx} - I_{zz})PR \\ N &= I_{zz}\dot{R} + (I_{yy} - I_{xx})PQ \end{aligned}\right\} \text{Eulerlikn.} \tag{13-72}$$

where P, Q, and R are the angular rates of the vehicle in the axes (X, Y, Z) of the body fixed frame; L, M, and N are rolling, pitching, and yawing moments with respect to the body fixed frame; and I_{xx}, I_{yy}, and I_{zz} are moments of inertia about the same frame; all cross-coupling terms are neglected.

The translational equations of motion are

$$\left.\begin{aligned} F_x + T_x + g_x &= m[\dot{U} + QW - RV] \\ F_y + g_y &= m[\dot{V} + RU - PW] \\ F_z + g_z &= m[\dot{W} + PV - QU] \end{aligned}\right\} \quad (13\text{-}73)$$

where U, V, and W are three components of the vehicle velocity along the (X, Y, Z) axes of the body fixed frame; F_x, F_y, and F_z are three components of the aerodynamic force; g_x, g_y, and g_z are three components of the gravitational force, all along the (X, Y, Z) axes of the body fixed frame; T_x is the thrust along the x axis; and m is the mass of the vehicle.

For the autopilot design, it is more convenient to express the translational motion in terms of velocity V_m, angle of attack α, and sideslip angle β. They are related to U, V, and W as follows:

$$V_m = \sqrt{V^2 + U^2 + W^2}$$

$$\alpha = \tan^{-1}\left(\frac{W}{U}\right)$$

$$\beta = \sin^{-1}\left(\frac{V}{V_m}\right)$$

In terms of V_m, α, and β, the translational motion equation (13-73) can be written as

$$\begin{aligned} \dot{V}_m &= \frac{1}{m}\{\cos\alpha\cos\beta(F_x + g_x + T_x) + \sin\beta(F_y + g_y) \\ &\quad + \sin\alpha\cos\beta(F_z + g_z)\} \\ \dot{\alpha} &= Q + \frac{1}{V_m\cos\beta}\{-PV_m\cos\alpha\sin\beta - RV_m\sin\alpha\sin\beta \\ &\quad - \sin\alpha(F_x + g_x + T_x) + \cos\alpha(F_z + g_z)\} \\ \dot{\beta} &= P\sin\alpha - R\cos\alpha + \frac{1}{mV_m}\{-\cos\alpha\sin\beta(F_x + g_x + T_x) \\ &\quad + \cos\beta(F_y + g_y) - \sin\alpha\sin\beta(F_z + g_z)\} \end{aligned} \quad (13\text{-}74)$$

The three components of the gravitational force in the body frame depend on the vehicle's attitude relative to an inertial frame. Let (X_I, Y_I, Z_I) denote the inertial frame with its origin fixed at the center of the earth. Let the axis, X_I, be chosen to point north, the axis Y_I then points east, and the axis, Z_I, points down. Assume that the orientation sequence of the body fixed frame with respect to the inertial frame is in the order of yaw angle Ψ, pitch angle Θ, and roll angle Φ. Also assume that the gravity acts at the center of the vehicle. Then we have gravity of

$$\begin{aligned} g_x &= -mg\sin\Theta \\ g_y &= mg\cos\Theta\sin\Phi \\ g_z &= mg\cos\Theta\cos\Phi \end{aligned}$$

The angular rates in the body frame P, Q, and R are related to the vehicle attitude through

$$\begin{aligned} P &= \dot{\Phi} - \dot{\Psi}\sin\Theta \\ Q &= \dot{\Theta}\cos\Phi + \dot{\Psi}\cos\Theta\sin\Phi \\ R &= -\dot{\Theta}\sin\Phi + \dot{\Psi}\cos\Theta\cos\Phi \end{aligned} \tag{13-75}$$

The forces and moments are given by

$$\begin{aligned} F_x &= k_F\rho V_m^2 C_x \\ F_y &= k_F\rho V_m^2 C_y \\ F_z &= k_F\rho V_m^2 C_z \\ L &= k_M\rho V_m^2 C_l \\ M &= k_M\rho V_m^2 C_m \\ N &= k_M\rho V_m^2 C_n \end{aligned}$$

where ρ is the atmospheric density, k_F and k_M are constants determined by the vehicle geometry, and C_x, C_y, C_z, C_l, C_m, and C_n are aerodynamic coefficients that are given in terms of an aerodynamic table and take the following functional form:

$$\begin{aligned} C_x &= C_{x0}(\alpha,\beta,M_m) + C_{x\delta_e}(\alpha,\delta_e,M_m) + C_{x\delta_a}(\alpha,\delta_a,M_m) + C_{x\delta_r}(\alpha,\delta_r,M_m) \\ C_y &= C_{y0}(\alpha,\beta,M_m) + C_{y\delta_e}(\alpha,\delta_e,M_m) + C_{y\delta_a}(\alpha,\delta_a,M_m) + C_{y\delta_r}(\alpha,\delta_r,M_m) \\ C_z &= C_{z0}(\alpha,\beta,M_m) + C_{z\delta_e}(\alpha,\delta_e,M_m) + C_{z\delta_a}(\alpha,\delta_a,M_m) + C_{z\delta_r}(\alpha,\delta_r,M_m) \\ C_l &= C_{l0}(\alpha,\beta,M_m) + C_{l\delta_e}(\alpha,\delta_e,M_m) + C_{l\delta_a}(\alpha,\delta_a,M_m) + C_{l\delta_r}(\alpha,\delta_r,M_m) \\ C_m &= C_{m0}(\alpha,\beta,M_m) + C_{m\delta_e}(\alpha,\delta_e,M_m) + C_{m\delta_a}(\alpha,\delta_a,M_m) + C_{m\delta_r}(\alpha,\delta_r,M_m) \\ C_n &= C_{n0}(\alpha,\beta,M_m) + C_{n\delta_e}(\alpha,\delta_e,M_m) + C_{n\delta_a}(\alpha,\delta_a,M_m) + C_{n\delta_r}(\alpha,\delta_r,M_m) \end{aligned}$$

Here M_m denotes the Mach number given by $M_m = V_m/c$, where c is the speed of sound and δ_e, δ_a, δ_r are the effective pitch, roll, and yaw fin control deflections, respectively.

Equations (13-72), (13-74), and (13-75) constitute the six-degree-of-freedom flight vehicle dynamic model. After some simple manipulations, the model can be put into the following standard form:

$$\begin{aligned} \dot{P} &= -\frac{I_{zz} - I_{yy}}{I_{xx}}QR + \frac{L}{I_{xx}} \\ \dot{Q} &= -\frac{I_{xx} - I_{zz}}{I_{yy}}PR + \frac{M}{I_{yy}} \\ \dot{R} &= -\frac{I_{yy} - I_{xx}}{I_{zz}}PQ + \frac{N}{I_{zz}} \\ \dot{\Psi} &= \sin\Phi\sec\Theta Q + \cos\Phi\sec\Theta R \end{aligned}$$

$$\dot{\Theta} = \cos\Phi Q - \sin\Phi R$$

$$\dot{\Phi} = P + \sin\Phi\tan\Theta Q + \cos\Phi\tan\Theta R$$

$$\begin{aligned}\dot{V}_m = \frac{1}{m}\{&\cos\alpha\cos\beta(F_x + g_x + T_x) \\ &+ \sin\beta(F_y + g_y) \\ &+ \sin\alpha\cos\beta(F_z + g_z)\}\end{aligned} \tag{13-76}$$

$$\begin{aligned}\dot{\alpha} = {} & Q - (P\cos\alpha + R\sin\alpha)\tan\beta \\ & + \frac{1}{mV_m\cos\beta}\{-\sin\alpha(F_x + g_x + T_x) + \cos\alpha(F_z + g_z)\}\end{aligned}$$

$$\begin{aligned}\dot{\beta} = {} & P\sin\alpha - R\cos\alpha \\ & + \frac{1}{mV_m}\{-\cos\alpha\sin\beta(F_x + g_x + T_x) + \cos\beta(F_y + g_y) \\ & - \sin\alpha\sin\beta(F_z + g_z)\}\end{aligned}$$

The basic requirement for the autopilot is a rapid and precise command following ability while satisfying certain physical constraints. It is well known that, for most practical designs, maneuvers are realized by developing the appropriate angle of attack. For BTT missiles, not only must the airframe rotate to develop appropriate angle of attack, but it must be rolled appropriately to satisfy sideslip constraints. The most difficult maneuver for a given response requirement is the rapid roll in conjunction with an angle-of-attack change [Arrow, 1985]. Reflecting these observations, we choose the performance output to be the following:

$$y_1 = P\cos\alpha + R\sin\alpha \overset{\text{def}}{=} P_s$$

$$y_2 = \alpha$$

$$y_3 = \beta$$

where P_s represents the roll rate with respect to the stability axis. We require that the closed-loop system have rapid and precise command following the ability of P_s and α. Also we require that the induced sideslip angle β be less than one degree. Let

$$x = \begin{bmatrix} P \\ Q \\ R \\ V_m \\ \alpha \\ \beta \\ \Psi \\ \Phi \\ \Theta \end{bmatrix}, \quad u = \begin{bmatrix} \delta_e \\ \delta_a \\ \delta_r \end{bmatrix}, \quad y = \begin{bmatrix} \Phi \\ \alpha \\ \beta \end{bmatrix}$$

Then the vehicle dynamical model can be compactly written as follows:

$$\begin{aligned} \dot{x} &= f(x, u) \\ y &= h(x) \end{aligned} \tag{13-77}$$

13.4.2 Input–Output Feedback Linearization Approach

To apply the feedback linearization approach, we need to approximate the aerocoefficients, which are given in the form of a table, by well-behaved affine functions as follows:

$$\begin{aligned} C_F &\stackrel{\text{def}}{=} \begin{bmatrix} C_x \\ C_y \\ C_z \end{bmatrix} = \begin{bmatrix} C_{x0} \\ C_{y0} \\ C_{z0} \end{bmatrix} + \begin{bmatrix} C_{xe} & C_{xa} & C_{xr} \\ C_{ye} & C_{ya} & C_{yr} \\ C_{ze} & C_{za} & C_{zr} \end{bmatrix} \begin{bmatrix} \delta_e \\ \delta_a \\ \delta_r \end{bmatrix} \\ &\stackrel{\text{def}}{=} C_{F0} + C_{Fu}u \end{aligned} \tag{13-78}$$

$$\begin{aligned} C_M &\stackrel{\text{def}}{=} \begin{bmatrix} C_l \\ C_m \\ C_n \end{bmatrix} = \begin{bmatrix} C_{l0} \\ C_{m0} \\ C_{n0} \end{bmatrix} + \begin{bmatrix} C_{le} & C_{la} & C_{lr} \\ C_{me} & C_{ma} & C_{mr} \\ C_{ne} & C_{na} & C_{nr} \end{bmatrix} \begin{bmatrix} \delta_e \\ \delta_a \\ \delta_r \end{bmatrix} \\ &\stackrel{\text{def}}{=} C_{M0} + C_{Mu}u \end{aligned} \tag{13-79}$$

Then the force F_x, F_y, F_z and the moment L, M, N are given by

$$\begin{bmatrix} F_x \\ F_y \\ F_z \end{bmatrix} = k_F \rho V_m^2 (C_{F0} + C_{Fu}u), \qquad \begin{bmatrix} L \\ M \\ N \end{bmatrix} = k_M \rho V_m^2 (C_{M0} + C_{Mu}u)$$

The entries of the matrices C_{F0}, C_{Fu}, C_{M0}, and C_{Mu} are generally functions of α, β, and Mach number. It is clear that good approximation of the aerocoefficients will lead to a better characterization of the aerodynamics. For a HAVE DASH II missile, based on some analysis on the aerocoefficient table, C_{F0}, C_{Fu}, C_{M0}, and C_{Mu} are approximated by

$$\begin{aligned} C_{F0} &= \begin{bmatrix} -0.57 + 0.0083\alpha \\ -0.21\beta \\ C_{z0}(\alpha, M_m) \end{bmatrix}, \qquad C_{Fu} = \begin{bmatrix} 0.004 & 0 & 0 \\ 0 & 0 & 0.08 \\ -0.09 & 0 & 0 \end{bmatrix} \\ C_{M0} &= \begin{bmatrix} 0.116\beta \\ C_{m0}(\alpha, M_m) \\ 0.08\beta \end{bmatrix}, \qquad C_{Mu} = \begin{bmatrix} 0 & -0.127 & 0 \\ -0.675 & 0 & 0 \\ 0 & 0 & -0.584 \end{bmatrix} \end{aligned} \tag{13-80}$$

where $C_{z0}(\alpha, M_m)$, and $C_{m0}(\alpha, M_m)$ are given as follows

$$\begin{aligned} C_{z0}(\alpha, M_n) &= C_{z1}(\alpha) + C_{z2}(\alpha)M_m \\ C_{m0}(\alpha, M_m) &= C_{m1}(\alpha) + C_{m2}(\alpha)M_m \end{aligned}$$

5 STACEY, W. M., Jnr.: 'Space-time nuclear reactor kinetics', Nuclear Science and Engineering Monograph 5 (Academic Press, 1969)
6 OWENS, D. H.: 'Multivariable control analysis of distributed parameter nuclear reactors', Ph.D. Thesis, Imperial College, 1973
7 KAPLAN, S.: 'The property of finality and the analysis of problems in reactor space-time kinetics by various modal expansions', *Nucl. Sci. Eng.*, 1961, **9**, pp. 357–361
8 WIBERG, D. M.: 'Optimal control of nuclear reactor systems', Advances in control systems 5 (Academic Press, 1967)
9 SUMNER, H. M.: 'ZIP Mk 2: A FORTRAN code for calculating the eigenvalues of large sets of linear equations', AEEW-R 543, 1969, HMSO
10 OWENS, D. H.: 'XENFER: A FORTRAN code for the calculation of spatial transfer functions of thermal reactors subject to xenon poison effects", AEEW-R 816, 1973, HMSO
11 HALMOS, P. R.: 'Finite-dimensional vector spaces' (Van Nostrand, 1958)
12 OWENS, D. H.: 'Feedback and multivariable systems', Control Engineering Series 7 (Peter Peregrinus, 1978)
13 OWENS, D. H.: 'Calculation of nuclear reactor spatial transfer functions using the computer programme XENFER', AEEW-R 817, 1973, HMSO
14 OWENS, D. H.: 'Multivariable-control-systems design concepts in the failure analysis of a class of nuclear reactor spatial control systems', *Proc. IEE*, 1973, **120**(1), pp. 119–125
15 OWENS, D. H.: 'Dyadic approximation method for multivariable control-systems design with a nuclear reactor application', ibid, pp. 801–809
16 BENNET, D. J.: 'The elements of nuclear power' (Longmans, 1972)

Chapter 5

Aerospace systems

J. M. Lipscombe

List of principal symbols

A	state matrix
a_x, a_y, a_z	co-ordinates of centre of gravity (Section 5.3)
a_x, a_y, a_z	accelerations in platform axes (Section 5.7)
a'_x, a'_y, a'_z	corrected accelerations (Section 5.7)
B	input matrix
b_x, b_y, b_z	accelerometer co-ordinates
C	output matrix
C^*	aircraft handling parameter (Section 5.6.3)
c_x, c_y, c_z	engine thrust moment arms (eqn. 5.9)
D	direct coupling matrix
E, E_x, E_y, E_z	engine thrust and its components
e	specific engine thrust, E/m
e_0, e_1, e_2, e_3	attitude parameters (Section 5.7.4)
f_L	line of sight acceleration
f_T	target acceleration
f_m	missile lateral acceleration (latax)
f_x, f_y, f_z	components of vehicle acceleration
$\lvert f_{max} \rvert$	maximum latax
G	error covariance matrix (Section 5.7.5)
g	acceleration due to gravity
g_x, g_y, g_z	components of g (Section 5.7)
H	nominal height (Section 5.6.1)
h	height
h_a	height measured by air data system
$\hat{h}$	best estimate of height

I	unit matrix
I_x, I_y, I_z	moments of inertia about axes through c.g.
K	state feedback matrix
K	constant in human pilot model (Section 5.6.4)
k	proportional navigation gain (Section 5.5.4)
k	constant defined by eqn. 5.63 (Section 5.6.1)
k_a	accelerometer feedback gain
k_r	rate gyroscope feedback gain
k_1, k_2	components of K
L	moment about Ox (Fig. 5.1)
L_a	moment about Ox due to aerodynamic effects
L_e	moment about Ox due to engine thrust
L_g	moment about Ox due to gravitation
L_p	$\partial L/\partial p$
L_r	$\partial L/\partial r$
L_v	$\partial L/\partial v$
L_ζ	$\partial L/\partial \zeta$ } aerodynamic derivatives for rolling moment
L_ξ	$\partial L/\partial \xi$
$L_{\dot p}$	$\partial L/\partial \dot p$
$L_{\dot r}$	$\partial L/\partial \dot r$
l_p	$\doteq L_p/I_x$
l_r	$\doteq L_r/I_x$
l_v	$\doteq L_v/I_x$ } for exact values see Section 5.6.2
l_ζ	$\doteq L_\zeta/I_x$
l_ξ	$\doteq L_\xi/I_x$
M	moment about Oy (Fig. 5.1)
M	mach number (Fig. 5.3)
M_a	moment about Oy due to aerodynamic effects
M_e	moment about Oy due to engine thrust
M_g	moment about Oy due to gravitation
M_q	$\partial M/\partial q$
M_u	$\partial M/\partial u$
M_w	$\partial M/\partial w$ } aerodynaimic derivatives for pitching moment
M_η	$\partial M/\partial \eta$
$M_{\dot q}$	$\partial M/\partial \dot q$
$M_{\dot w}$	$\partial M/\partial \dot w$
m	vehicle mass
m_e	$\doteq mc_x/I_y$
m_q	$\doteq M_q/I_y$
m_u	$\doteq M_u/I_y$ } for exact values see Section 5.6.1
m_w	$\doteq M_w/I_y$
N	moment about Oz (Fig. 5.1)
N_a	moment about Oz due to aerodynamic effects
N_e	moment about Oz due to engine thrust
N_g	moment about Oz due to gravitation

N_p	$\partial N/\partial p$
N_r	$\partial N/\partial r$
N_v	$\partial N/\partial v$
N_ζ	$\partial N/\partial \zeta$ } aerodynamic derivatives for yawing moment
N_ξ	$\partial N/\partial \xi$
$N_{\dot p}$	$\partial N/\partial \dot p$
$N_{\dot r}$	$\partial N/\partial \dot r$
n_p	$\doteq N_p/I_z$
n_r	$\doteq N_r/I_z$
n_v	$\doteq N_v/I_z$ } for exact values see Section 5.6.2
n_ζ	$\doteq N_\zeta/I_z$
n_ξ	$\doteq N_\xi/I_z$
P	angular velocity about Ox (Fig. 5.1)
P	Kalman filter gain matrix (Section 5.7.5)
$P(s)$	transfer function model of human pilot (Section 5.6.4)
p	perturbation angular velocity about Ox
p_1, p_2	elements of Kalman filter gain matrix, P
Q	angular velocity about Oy (Fig. 5.1)
$Q(s)$	transfer function of pilot plus plant (Fig. 5.11)
q	perturbation angular velocity about Oy
q	$q\delta(t) = \epsilon\alpha^2$ (Section 5.7.5)
R	angular velocity about Oz (Fig. 5.1)
R	radius of the earth (Section 5.7)
R_m	missile range from tracker (Fig. 5.5)
R_r	distance of missile from target (Fig. 5.7)
r	perturbation angular velocity about Oz
r	$r\dot\delta(t) = \epsilon\beta^2$ (Section 5.7.5)
s	Laplace transform variable
T	time of missile-target impact
t	time variable
U	velocity along Ox (Fig. 5.1)
U_G	gust velocity along Ox
U_T	target velocity
U_m	missile velocity
U_r	relative velocity of missile and target
U_0	velocity along O_0x_0
U_{co}	cross-over velocity (Section 6.6.3)
U_{0G}	gust velocity along O_0x_0
u	perturbation velocity along Ox
$\mathbf{u}$	input vector
V	velocity along Oy (Fig. 5.1)
V_G	gust velocity along Oy
V_T	$(U^2 + V^2 + W^2)^{1/2}$
V_x, V_y, V_z	velocities with respect to a tangent plane rotating with the earth (Section 5.7.1)

V_0 velocity along $O_0 y_0$

V_{xe}, V_{ye}, V_{ze} velocity of the tangent plane due to the earth's rotation (Section 5.7.1)

V_{0G} gust velocity along $O_0 y_0$

$\dot{V}_{zi}$ inertial platform measurement of vertical acceleration

v perturbation velocity along Oy

W velocity along Oz (Fig. 5.1)

W_G gust velocity along Oz

W_0 velocity along $O_0 z_0$

W_{0G} gust velocity along $O_0 z_0$

w perturbation velocity along Oz

X force along Ox (Fig. 5.1)

X_a force along Ox due to aerodynamic effects

X_e force along Ox due to engine thrust

X_g force along Ox due to gravitation

X_u $\partial X/\partial u$
X_w $\partial X/\partial w$
X_η $\partial X/\partial \eta$
} aerodynamic derivatives for longitudinal force

x_u $\doteq X_u/m$
x_w $\doteq X_w/m$
x_η $\doteq X_\eta/m$
} for exact values see Section 5.6.1

x_1, x_2 components of x

x state vector

$\hat{x}_1$, $\hat{x}_2$ components of $\hat{x}$

$\hat{x}$ best estimate of x

Y force along Oy (Fig. 5.1)

Y_a force along Oy due to aerodynamic effects

Y_e force along Oy due to engine thrust

Y_g force along Oy due to gravitation

Y_r $\partial Y/\partial r$
Y_v $\partial Y/\partial v$
Y_ζ $\partial Y/\partial \zeta$
} aerodynamic derivatives for lateral force

y_v $\doteq Y_v/m$
y_ζ $\doteq Y_\zeta/m$
} for exact values see Section 5.6.2

y output vector

Z force along Oz (Fig. 5.1)

Z_a force along Oz due to aerodynamic effects

Z_e force along Oz due to engine thrust

Z_g force along Oz due to gravitation

Z_q $\partial Z/\partial q$
Z_u $\partial Z/\partial u$
Z_w $\partial Z/\partial w$
Z_η $\partial Z/\partial \eta$
$Z_{\dot{q}}$ $\partial Z/\partial \dot{q}$
$Z_{\dot{w}}$ $\partial Z/\partial \dot{w}$
} aerodynamic derivatives for normal force

z_e $\doteq \epsilon$
z_q $\doteq Z_q/m$
z_u $\doteq Z_u/m$
z_w $\doteq Z_w/m$
z_η $\doteq Z_\eta/m$
} for exact values see Section 5.6.1

A angle of attack

α perturbation angle of attack

α acceleration constraint constant (Section 5.5.4)

α phase advance constant (Section 5.6.4)

α inertial platform error (Section 5.7.5)

B angle of sideslip

β perturbation angle of slideslip

β air data computer error (Section 5.7.5)

γ nominal value of ϕ_m

$\delta(t)$ unit impulse function

Δy, $\Delta\theta$, $\Delta\psi$ defined in Fig. 5.5 (Section 5.5.2)

Δf_x, Δh, Δu
ΔV_y, Δx, $\Delta\beta$
$\Delta\theta$, $\Delta\omega_x$
} errors associated with inertial navigation unit (Section 5.7)

δt, δy, $\delta\theta$
$\delta\phi_m$
} defined in Fig. 5.9 (Section 5.5.4)

ϵ E_z/E

ζ rudder deflection

ζ_c rudder command from guidance system (Section 5.5)

η elevator deflection

Θ pitch angle

θ perturbation pitch angle

θ angle of line of sight (Section 5.5)

Λ longitude

λ latitude

μ damping ratio

μ_d damping ratio of dutch roll mode

μ_p damping ratio of phugoid mode

μ_q damping ration of short period mode

ξ aileron deflection

ρ air density

τ_r time constant of roll subsidence mode

τ_s time constant of spiral mode

τ_a, τ_d, τ_e time constants of pilot behaviour (Section 5.6.4)

Φ roll angle

ϕ perturbation roll angle

ϕ_T target heading angle (Section 5.5.4)

ϕ_m missile heading angle (Section 5.5.4)

Ψ azimuth angle

ψ perturbation azimuth angle

Ω	rotational speed of the earth
ω_e	bandwidth associated with human pilot model (Section 6.4)
ω_d	natural frequency of dutch roll mode
ω_n	natural frequency
ω_p	natural frequency of phugoid mode
ω_q	natural frequency of short period mode
$\omega_x, \omega_y, \omega_z$	inertial platform rotation rates
ω_{xe}	drift error in x axis integrating rate gyroscope
Ω_{xe}	a constant value of ω_{xe}

Mathematical notation

$\dot{x}$	derivative of x
x	column matrix col $[x_1, x_2, x_3, \ldots, x_m]$
$\bar{x}$	Laplace transform of x
∂E	perturbation of E
A^T	the transpose of A
ϵw	expectation of w
det A	determinant of A
$\Rightarrow$	implies

Abbreviations

c.g.	centre of gravity
DCM	direction cosine matrix
DLC	direct lift control
DSFC	direct side force control
INU	inertial navigation unit
IRG	integrating rate gyroscope
latax	lateral acceleration at missile c.g.
LOS	line of sight
SAS	stability augmentation system
SL	sight line

5.1 Introduction

Traditionally the modelling of aerospace systems has fallen into two parts

(*a*) three-dimensional geometry and kinematics
(*b*) the dynamics of a rigid body and the fluid (air) through which it moves.

Developments through the past few decades have forced the attention of aerospace engineers toward other models such as

(*c*) structural elasticity (structural mode controllers) and the interaction with the fluid (flutter suppression controllers, gust alleviation systems, etc.)
(*d*) models of the propulsion system

(*e*) dynamics of the hydraulic or other actuators which power the aerodynamic control surfaces
(*f*) the behaviour of the human pilot in response to natural and artificial cues
(*g*) the navigation system (this largely applies to aircraft and covers an enormous range of disciplines from the errors in inertial navigation units to the radio wave propagation problems of instrument landing systems)
(*h*) guidance problems (this is roughly the missile equivalent of (*g*), and includes models as diverse as laser beam riding for antitank weapons and minimum-fuel rendezvous trajectories for space vehicles).

The above models are only a selection of those currently used by aerospace engineers, and ignore the control system itself; typically (in an aircraft) this consists of a large number of redundant sensors and separated digital computing systems which not only have complex internal dynamics, but give rise to difficult and interesting modelling problems in terms of reliability, maintainability, cost of ownership and overall system integrity. This chapter concentrates on models (*a*) and (*b*), but inevitably the characteristics of some of the other models intrude.

The layout of the chapter is as follows. The basic equations of vehicle geometry and dynamics are presented in Sections 5.2 and 5.3, respectively. They are reduced to a set of linear perturbation equations in Section 5.4. Three examples of their use in modelling various aerospace systems are given in the remaining three sections. The lateral equations of motion of a line-of-sight guided missile are given in Section 5.5, the equations of motion of a conventional aircraft in Section 5.6, and the behaviour of an inertial navigation unit in Section 5.7.

5.2 Geometry of aerospace vehicles

This model and its various simplifications describe the position and attitude of the vehicle with respect to a set of reference axes. For the purposes of evaluating the aircraft performance, the reference axes are stationary with respect to the fixed stars (i.e. are 'inertial axes'), but when navigation systems are modelled a third set describing navigation co-ordinates on a spherical rotating earth must be invoked (Section 5.7.1).

5.2.1 *Definitions*

Body axes: A right-handed set of orthogonal axes fixed to the aircraft as shown in Fig. 5.1

Ox points 'forward' (*longitudinal* axis)
Oy points to starboard (*lateral* axis)
Oz points 'down' (*normal* axis)

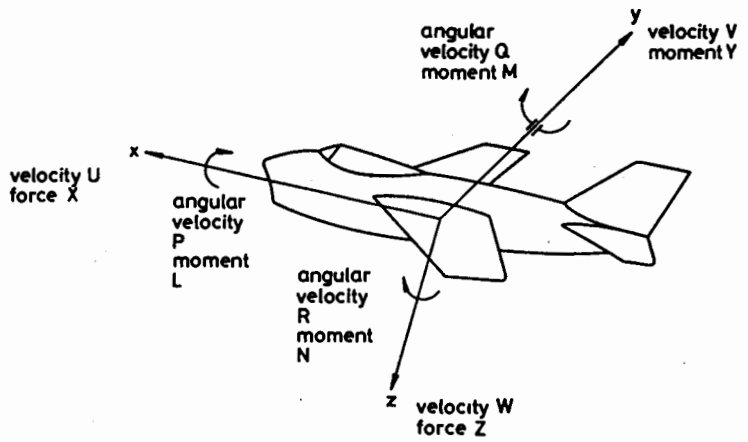

Fig. 5.1 *Diagram of body axes and velocities*

Reference axes: A right-handed set of orthogonal axes in the tangent plane to the earth's surface

O_0x_0 points east (or sometimes north)

O_0y_0 points south (if O_0x_0 points north, O_0y_0 must point east)

O_0z_0 points towards the centre of the earth.

Body velocities: The aircraft has velocities

U along Ox

V along Oy

W along Oz

and angular velocities

P about Ox (*roll* rate)

Q about Oy (*pitch* rate)

R about Oz (*yaw* rate)

The angular velocities are right handed, i.e. clockwise when looking in the direction of the appropriate axis.

Velocity with respect to the reference axes: Let the vehicle have velocity

U_0 along O_0x_0

V_0 along O_0y_0

W_0 along O_0z_0

Note that $W_0 = -dh/dt$ where h is the *height* of the aircraft.

Aircraft attitude: The orientation of $Oxyz$ with respect to $O_0x_0y_0z_0$ is defined by the following sequence of rotations

(i) let an intermediate axis set $O_ix_iy_iz_i$ be coincident with $O_0x_0y_0z_0$.

(ii) rotate $O_ix_iy_iz_i$ about the axis Oz_i by an angle Ψ (*azimuth* angle)

(iii) rotate $O_ix_iy_iz_i$ about the axis Oy_i by an angle Θ (*pitch* angle)

(iv) rotate $O_ix_iy_iz_i$ about the axis Ox_i by an angle Φ (*roll* angle)

(v) these rotations have been made in such a way that $O_ix_iy_iz_i$ is now coincident with $Oxyz$.

Then Ψ, Θ, Φ define the attitude of $Oxyz$ with respect to $O_0x_0y_0z_0$ and are called the *Euler* angles of the aircraft. They are positive if the rotations are right handed.

5.2.2 *Relationships between body and reference axes*

It can be shown[1] that

$$\begin{bmatrix} U_0 \\ V_0 \\ W_0 \end{bmatrix} = \begin{bmatrix} c\Theta c\Psi & s\Phi s\Theta c\Psi - c\Phi s\Psi & c\Phi s\Theta c\Psi + s\Phi s\Psi \\ c\Theta s\Psi & s\Phi s\Theta s\Psi + c\Phi c\Psi & c\Phi s\Theta s\Psi - s\Phi s\Psi \\ -s\Theta & s\Phi c\Theta & s\Phi s\Theta \end{bmatrix} \begin{bmatrix} U \\ V \\ W \end{bmatrix} \tag{5.1}$$

where c stands for cosine and s for sine

Clearly the 3×3 matrix in eqn. (5.1) is a *direction* cosine *matrix* (DCM) written in terms of the Euler angles instead of the more usual form (in applied mathematics texts) which uses the direction cosines between $Oxyz$ and $O_0x_0y_0z_0$.

The attitude of the aircraft changes in response to its roll, pitch and yaw velocities. It is simple to show[1] that the relationship is

$$\begin{bmatrix} \dot{\Phi} \\ \dot{\Theta} \\ \dot{\Psi} \end{bmatrix} = \begin{bmatrix} 1 & \sin\Phi\tan\Theta & \cos\Phi\tan\Theta \\ 0 & \cos\Phi & -\sin\Phi \\ 0 & \sin\Phi\sec\Theta & \cos\Phi\sec\Theta \end{bmatrix} \begin{bmatrix} P \\ Q \\ R \end{bmatrix} \tag{5.2}$$

Equation (5.2) indicates one of the main disadvantages of representing attitude by Euler angles: the equation breaks down at $\Theta = \pm 90°$, and therefore the description is of limited use for simulation purposes, (see eqns. (5.88), (5.89) and (5.90) for a suitable alternative). However, the Euler angles are convenient for analytical work because they are easy to visualize. For example, $[\Phi\Theta\Psi] = [5°\ 30°\ 45°]$ immediately conveys the information that the aircraft is climbing at 30°, rolled slightly to starboard and on a heading 45° to the reference direction O_0x_0.

The inverse of eqn. (5.2) is

$$\begin{bmatrix} P \\ Q \\ R \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\sin\Theta \\ 0 & \cos\Phi & \sin\Phi\cos\Theta \\ 0 & -\sin\Phi & \cos\Phi\cos\Theta \end{bmatrix} \begin{bmatrix} \dot{\Phi} \\ \dot{\Theta} \\ \dot{\Psi} \end{bmatrix} \tag{5.3}$$

Notice that P, Q Φ and Θ have been named in a rather unfortunate way because the pitch and roll angles (Θ and Φ) are not in general the integral of the pitch and roll rates (Q and P).

Figure 5.2 shows how eqns. (5.1) and (5.2) transform body velocities into attitude and velocity with respect to the reference axes.

5.2.3 *Angles of attack and sideslip*

Before leaving geometrical ideas, it is possible to transform U, V and W into total aircraft velocity, and the angles of attack and sideslip, which are perhaps easier to visualize.

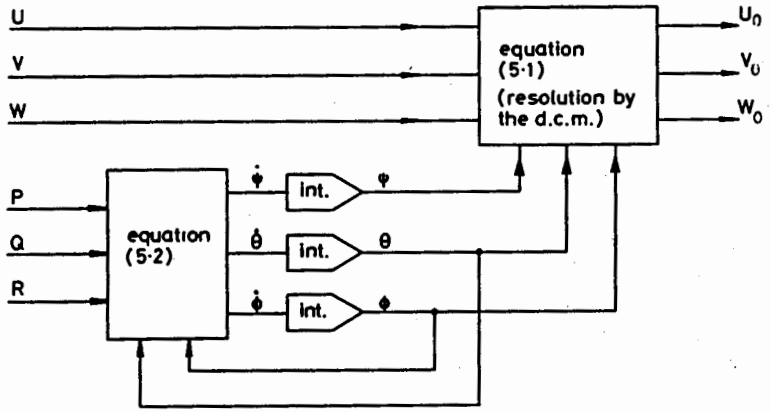

Fig. 5.2 *Rigid body geometry*

Define the *total velocity* of the aircraft

$$V_T = (U^2 + V^2 + W^2)^{1/2}$$

Then the *angle of attack* is defined by

$$\sin \mathrm{A} = \frac{W}{(U^2 + W^2)^{1/2}}$$

and the *angle of sideslip* is defined by

$$\sin \mathrm{B} = \frac{V}{V_T}$$

Thus

$$\begin{aligned} U &= V_T \cos \mathrm{A} \cos \mathrm{B} \\ V &= V_T \sin \mathrm{B} \\ W &= V_T \sin \mathrm{A} \cos \mathrm{B} \end{aligned} \tag{5.4}$$

5.3 Dynamics

It is conventional[1, 2] to start with the so-called *Euler equations* of a rigid body moving in a vacuum under the influence of external forces. In fact these equations are in error because they do not properly account for the mass of the displaced fluid. However, for aircraft and missiles this effect is quite small and can be approximated by adding appropriate external forces to the right-hand side of the equations. The models of submarines or airships may sometimes be formulated with more complicated dynamics.

5.3.1 Euler equations

$$\begin{aligned} &[\dot{U} + WQ - RV - a_x(R^2 + Q^2) + a_y(PQ - \dot{R}) + a_z(PR + \dot{Q})]m \\ &\quad = X \\ &[\dot{V} + UR - PW + a_x(PQ + \dot{R}) - a_y(P^2 + R^2) + a_z(RQ - \dot{P})]m \\ &\quad = Y \\ &[\dot{W} + VP - QU + a_x(RP - \dot{Q}) + a_y(RQ + \dot{P}) - a_z(Q^2 + P^2)]m \\ &\quad = Z \\ &I_x\dot{P} + (I_z - I_y)RQ - I_{yz}(Q^2 - R^2) - I_{xz}(\dot{R} + PQ) - I_{xy}(\dot{Q} - PR) \\ &\quad = L + Ya_z - Za_y \\ &I_y\dot{Q} + (I_x - I_z)PR - I_{yz}(\dot{R} - PQ) - I_{xz}(R^2 - P^2) - I_{xy}(\dot{P} + QR) \\ &\quad = M + Za_x - Xa_z \\ &I_z\dot{R} + (I_y - I_x)QP - I_{yz}(\dot{Q} + RP) - I_{xz}(\dot{P} - RQ) - I_{xy}(P^2 - Q^2) \\ &\quad = N + Xa_y - Ya_x \end{aligned} \tag{5.5}$$

where m is the mass of the aircraft, I_x, I_y, I_z, I_{xy}, I_{yz}, I_{xz} are the moments of inertia about axes through the centre of gravity (c.g.) but parallel to Ox, Oy and Oz, X, Y and Z are the forces along Ox, Oy and Oz, L, M and N are the moments about Ox Oy and Oz (in a right-handed sense) and a_x, a_y and a_z are the co-ordinates of the aircraft centre of gravity with respect to Oxyz.

For these equations to be valid U, V, W, P, Q and R must be absolute velocities. Application of eqns. (5.1) and (5.2) therefore result in U_0, V_0, W_0, Φ, Θ and Ψ which are with respect to inertial axes, i.e. a fixed flat earth. However, for aircraft or missile performance modelling this is quite satisfactory. The navigational model is a different matter, and this will be discussed in Section 5.7.

Equation (5.5) simplifies if the origin of the body axes can be placed at the c.g. ($a_x = a_y = a_z = 0$), and the assumption made that the axes are principal ones ($I_{xy} = I_{yz} = I_{zx} = 0$). The Euler equations then reduce to

$$\begin{aligned} m(\dot{U} + WQ - RV) &= X \\ m(\dot{V} + UR - PW) &= Y \\ m(\dot{W} + VP - QU) &= Z \\ I_x\dot{P} + (I_z - I_y)RQ &= L \\ I_y\dot{Q} + (I_x - I_z)PR &= M \\ I_z\dot{R} + (I_y - I_x)QP &= N \end{aligned} \tag{5.6}$$

5.3.2 *Aircraft acceleration*

In terms of the body velocities (U, V, W, P, Q, R), the acceleration experienced in the aircraft at a point with co-ordinates ($b_x b_y b_z$) with respect to $Oxyz$ is given by[2]

$$\begin{aligned} f_x &= \dot{U} + WQ - RV - b_x(R^2 + Q^2) + b_y(PQ - \dot{R}) \\ &\quad + b_z(PR + \dot{Q}) + g \sin \Theta \\ f_y &= \dot{V} + UR - PW + b_x(PQ + \dot{R}) - b_y(P^2 + R^2) \\ &\quad + b_z(RQ - \dot{P}) - g \cos \Theta \sin \Phi \\ f_z &= \dot{W} + VP - QU + b_x(RP - \dot{Q}) + b_y(RQ + \dot{P}) \\ &\quad - b_z(Q^2 + P^2) - g \cos \Theta \cos \Phi \end{aligned} \tag{5.7}$$

where f_x, f_y and f_z are the accelerations along Ox, Oy and Ox, respectively. The reader will find it instructive to compare eqn. (5.7) with the first three equations of (5.5), and put a physical meaning to each term.

5.3.3 *Gravitional forces*

Equation (5.6) (or in awkward cases (5.5)) transforms a set of forces and moments acting on the aircraft into body velocities. These can then be transformed by eqns. (5.1) and (5.2) into vehicle motion with respect to the earth. Clearly the next task is to build up a model to generate the forces X, Y and Z and the moments L, M and N.

First consider the forces and moments which are due to gravitational attraction. Simple geometry shows that

$$\begin{bmatrix} X_g \\ Y_g \\ Z_g \\ L_g \\ M_g \\ N_g \end{bmatrix} = \begin{bmatrix} -\sin \Theta \\ \cos \Theta \sin \Phi \\ \cos \Theta \cos \Phi \\ a_y \cos \Theta \cos \Phi - a_z \cos \Theta \sin \Phi \\ -a_z \sin \Theta - a_x \cos \Theta \cos \Phi \\ a_y \sin \Theta + a_x \cos \Theta \sin \Phi \end{bmatrix} mg \tag{5.8}$$

where the subscript g means 'due to gravitation'.

If the origin of the body axes is placed at the c.g., $a_x = a_y = a_z = 0$ and the three gravitational moments become zero.

5.3.4 *Aerodynamic forces*

It is assumed that each force is a function of

(i) air density, ρ
(ii) relative motion of the vehicle through the air, and
(iii) the vehicle configuration.

Consider each of these in turn. Clearly the density of the air (which is mainly a function of height, h) influences the magnitude of any aerodynamically generated force. The variable h forms the ordinate of the flight envelope (Fig. 5.3) which is described in Section 5.4.1.

The relative motion of the vehicle through the air is not only a function of the body velocities, but also a function of any movement of the air relative to the reference axes. Let these *gust velocities* be U_{OG}, V_{OG} and W_{OG}. Then transformation through the direction cosine matrix (DCM) by using the inverse of eqn. (5.1) gives the gust velocities resolved along the aircraft body axes, U_G, V_G and W_G.

A reasonable assumption might be that each aerodynamic force or moment is a function of (U, V, W, P, Q, R, $\dot{U}$, $\dot{V}$, $\dot{W}$, $\dot{P}$, $\dot{Q}$, $\dot{R}$, U_G, V_G, W_G, $\dot{U}_G$, $\dot{V}_G$, $\dot{W}_G$, ρ). One of the main reasons for reformulating the equations in terms of perturbation velocities (Section 5.4) is to reduce this function to one which can be handled analytically. Even simulation studies usually use simplified semi-analytical approximations.

The word configuration is usually used to denote a 'steady-state' aircraft shape, i.e. whether it has the undercarriage extended or not, or the position of the landing flaps. In this section, however, the meaning is extended to include control surface deflections. There are conventionally three aerodynamic controls on an aircraft, each designed to produce a moment about one of the three axes. The *ailerons* deflect by an angle ξ to produce a moment about Ox; the *elevator* deflects by an angle η to produce a moment about Oy; and the *rudder* deflects by an angle ζ to produce a moment about Oz. In each case a *positive* deflection of the control surface produces a *negative* moment about the appropriate axis.

In traditional aircraft the ailerons, elevator and rudder are hinged flaps placed at the wing tips and tail. On most modern aircraft the ailerons are supplemented by additional aerodynamic surfaces on the wing. Indeed on

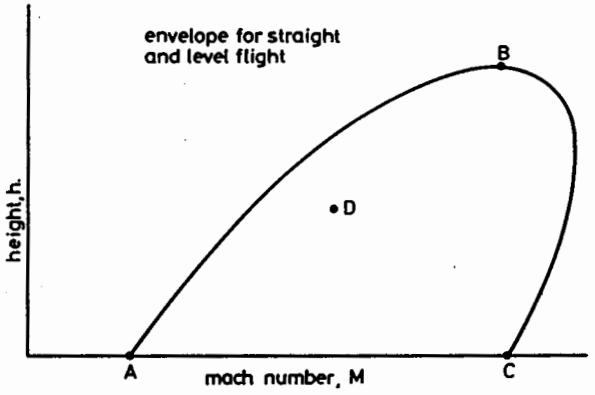

Fig. 5.3 *Sketch of a flight envelope*

some high performance aircraft every square inch of the aerodynamic surfaces seem movable for some control purpose. A selection of the control functions are:

(i) additional surfaces on the wing to alter its lift without creating a moment (direct lift control: DLC)
(ii) additional surfaces (which may work in conjunction with the rudder) to produce a lateral force without creating a moment (direct side force control: DSFC). These five aerodynamic controls, ξ, η, ζ, DLC and DSFC together with engine thrust and aerodynamic air-brakes to produce longitudinal forces give six control forces and moments to control the six degree of freedom dynamics expressed in equation (5.5).
(iii) Additional surfaces to resolve structural mode and aeroelasticity (flutter) problems. These surfaces may be necessary because in control theory language, the modes may be uncontrollable with respect to the main control surfaces. Additional instrumentation is often required to make them observable.
(iv) Variable sweep-back or flap deflection as an automatic function of aircraft manoeuvre. This extends the flight envelope without adding to the pilot's workload, and can improve the combat manoeuvrability.

5.3.5 *Propulsion forces*

An aircraft engine is a complicated machine which is as challenging to model as the aircraft itself. Fortunately, when studying the behaviour of an aircraft, drastic simplifications to the engine model may usually be made without compromising the results.

Let the engine have thrust E with components $E_x E_y$ and E_z along the axes Ox, Oy and Oz, and the perpendicular from the thrust line (i.e. the direction of E) to the axis origin have components $c_x c_y$ and c_z.
Then

$$\begin{aligned} X_e &= E_x & L_e &= c_y E_z - c_z E_y \\ Y_e &= E_y & M_e &= c_z E_x - c_x E_z \\ Z_e &= E_z & N_e &= c_x E_y - c_y E_x \end{aligned} \tag{5.9}$$

where the subscript e means 'due to engine thrust'.

5.3.6 *Summary of Section 5.3*

At this point, a respectable model of the aircraft motion has been constructed. Inputs in the form of gust velocities, aerodynamic configuration changes and engine commands modify the forces and moments acting on the aircraft. These, via eqn. (5.5) or (5.6), produce body velocities which can be translated into movement with respect to fixed axes by eqns. (5.1) and (5.2). A feedback loop exists from the body velocities via the aerodynamic functions (which have not yet been specified in a detailed form) to create aerodynamic forces,

X_a, Y_a, Z_a, L_a, M_a and N_a. Another feedback loop is created from the aircraft attitude Φ, Θ, Ψ via the gravitational forces by eqn. (5.8). The outputs of the model are, of course, the variables which are of interest for performance estimation, or which are easy to measure. Examples are angle of attack (definitions preceding eqn. (5.4)) and aircraft acceleration [eqn. (5.7)], respectively. It is an instructive exercise for the reader to complete Fig. 5.2 by adding the extra blocks and feedback loops.

The model is almost suitable for simulation. An analytical/look-up table description of the aerodynamic functions must be created for the particular aircraft under simulation, and perhaps some structural dynamics, actuator lags and engine characteristics added. (Where would these blocks appear in the extended version of Fig. 5.2?)

It is conventional to use the inverse of eqn. (5.4) to replace the variables U, V and W with V_T, A and B in many flight simulations for accuracy and bandwidth reasons.[3] This is a special case of the general mathematical modelling problem of transforming the state variables into the best form for processing through the computer. In the linear time-invariant case, for example, it often means putting the state matrix into companion form.

5.4 Perturbation equations

For analytical studies the above model is too complex. In this section perturbation methods are used to reduce the equations to linear time-invariant form suitable for control system design studies.

5.4.1 *Flight envelope*

A characteristic which aircraft and missiles share is that the total velocity V_T is predominantly composed of the forward velocity U. The aerodynamic characteristics of the aircraft vary greatly as the speed is changed. It has already been noted that the aerodynamics are a function of air density, ρ, and these two ideas lead to the concept of a *flight envelope* for a particular vehicle. Figure 5.3 shows a sketch of a flight envelope. The vertical scale is height h (representing ρ), and the horizontal scale is Mach number M (which is V_T divided by the ambient speed of sound). The interior of the envelope is the operational region of the aircraft. If it goes outside curve AB the wings are unable to generate enough lift to support the aircraft weight. Structural, thermal or power limitations prevent the aircraft crossing the boundary BC. The first step in forming the perturbation equations is to choose a point within the flight envelope around which to linearize. A full-scale study looks at many points and control laws are devised for each point. These are mechanized on the aircraft in such a way that the flight control system interpolates between them as the air data computer calculates where the aircraft is positioned in its envelope. In the next section the model will be formulated for just one point, D say.

5.4.2 Stability axes

If the body axes are oriented in the vehicle so that under nominal conditions U is the forward speed and $V = W = 0$, then the axes are called *stability* axes for that particular flight condition. Naturally the axes move with respect to the body as the aircraft changes its position in the flight envelope, but at a particular nominal operating condition the axes are fixed with respect to the aircraft. Assume steady horizontal flight so that $P = Q = R = \Phi = \Theta = \Psi = 0$. Small departures from the nominal condition give rise to perturbation velocities along the body-fixed stability axes of $[u, v, w]$ and angular velocities about them of $[p, q, r]$. The Euler angles are perturbed to $[\phi, \theta, \psi]$.

Neglecting second-order terms, eqn. (5.1) becomes

$$\begin{bmatrix} U_0 \\ V_0 \\ W_0 \end{bmatrix} = \begin{bmatrix} U \\ U\psi + v \\ -U\theta + w \end{bmatrix} \tag{5.10}$$

The last row is more usually written

$$\dot{h} = U\theta - w \tag{5.11}$$

Equation (5.2) becomes

$$\begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} p \\ q \\ r \end{bmatrix} \tag{5.12}$$

The three definitions of Section 5.2.3 reduce to

$$U = V_T$$

$$\alpha = \frac{w}{U} \quad \text{and} \quad \beta = \frac{v}{U} \tag{5.13}$$

where α and β are perturbation angies of attack and sideslip, respectively.

Stability axes are naturally placed at the centre of gravity, reducing eqns. (5.5) to

$$\begin{aligned} m\dot{u} &= \partial X \\ m(\dot{v} + Ur) &= \partial Y \\ m(\dot{w} - qU) &= \partial Z \\ I_x\dot{p} - I_{xz}\dot{r} - I_{xy}\dot{q} &= \partial L \\ I_y\dot{q} - I_{yz}\dot{r} - I_{xy}\dot{p} &= \partial M \\ I_z\dot{r} - I_{yz}\dot{q} - I_{xz}\dot{p} &= \partial N \end{aligned} \tag{5.14}$$

where ∂ stands for 'the perturbation in'.

Sometimes the stability axes are far from being axes of symmetry and I_{xz} is not negligible, usually under low-speed conditions. Readers who have seen Concorde land, and recall that Ox is along the total unperturbed velocity vector, will appreciate the situation. Often, however, the cross products of

inertia are relatively small, and eqn. (5.14) can be rewritten

$$\begin{aligned} m\dot{u} &= \partial X \\ m(\dot{v} + Ur) &= \partial Y \\ M(\dot{w} - qU) &= \partial Z \\ I_x\dot{p} &= \partial L \\ I_y\dot{q} &= \partial M \\ I_z\dot{r} &= \partial N \end{aligned} \tag{5.15}$$

Equation (5.7) becomes

$$\begin{aligned} f_x &= \dot{u} - b_y\dot{r} + b_z\dot{q} + g\theta \\ f_y &= \dot{v} + Ur + b_x\dot{r} - b_z\dot{p} - g\phi \\ f_z &= \dot{w} - Uq - b_x\dot{q} + b_y\dot{p} \end{aligned} \tag{5.16}$$

where f_x, f_y and f_z are now perturbation accelerations. Equation (5.8) becomes (with $a_x = a_y = a_z = 0$)

$$\begin{aligned} \partial X_g &= -mg\theta \\ \partial Y_g &= mg\phi \\ \partial Z_g &= \partial L_g = \partial M_g = \partial N_g = 0 \end{aligned} \tag{5.17}$$

Equation (5.9) becomes

$$\begin{aligned} \partial X_e &= \partial E \\ \partial Z_e &= \epsilon\, \partial E \\ \partial M_e &= c_z\, \partial E \\ \partial Y_e &= \partial L_e = \partial N_e = 0 \end{aligned} \tag{5.18}$$

where $\epsilon = E_z/E$, the thrust line inclination, and E_y, c_y and $E_z c_x$ are assumed negligible.

5.4.3 Aerodynamic derivatives

The aerodynamic force and moment functions mentioned in Section 5.3.4 can be expanded about their nominal value in a multidimensional Taylor series. For example, for a particular value of U and ρ (i.e. a particular point in the flight envelope)

$$\partial X_a = \frac{\partial X}{\partial u}.u + \frac{\partial X}{\partial v}.v + \cdots \frac{\partial X}{\partial \dot{q}}\dot{q} + \frac{\partial X}{\partial \dot{r}}\dot{r} + \frac{\partial X}{\partial \xi}\xi + \cdots + \text{higher order terms} \tag{5.19}$$

Note that configuration changes are included in the expansions [an aileron term $(\partial X/\partial \xi)\xi$ is shown]. The aileron, elevator and rudder deflections ξ, η and ζ must be regarded as perturbation deflections from the nominal values for that point in the flight envelope.

Those terms in expansion (5.19), which experiment or intuition indicate are negligible, are rejected and the remainder incorporated in the mathematical model. Under some extreme flight conditions certain second-order terms may be non-negligible, but this is unusual. For example, a typical aircraft expansion might be

$$\partial X_a = \frac{\partial X}{\partial u}u + \frac{\partial X}{\partial w}w + \frac{\partial X}{\partial \eta}\eta$$

It is conventional to write this

$$\partial X_a = X_u u + X_w w + X_\eta \eta$$

where X_u, X_w and X_η are called *aerodynamic derivatives.*

5.5 Model of the Lateral Equations of Motion of a Guided Missile

To illustrate the use of the equations given in Section 5.4, consider the model of lateral motion of the cruciform missile illustrated in Fig. 5.4. Control is via the deflection of the four rear fins, and it is assumed that the missile is stabilized in roll so that $\Phi = 0$. Further, assume that the pitch motion and yaw motion can be analysed separately because the symmetry of the missile eliminates aerodynamic coupling. (In fact this assumption breaks down at high angles of attack.) Geometrical coupling [eqn. (5.3)] is also ignored.

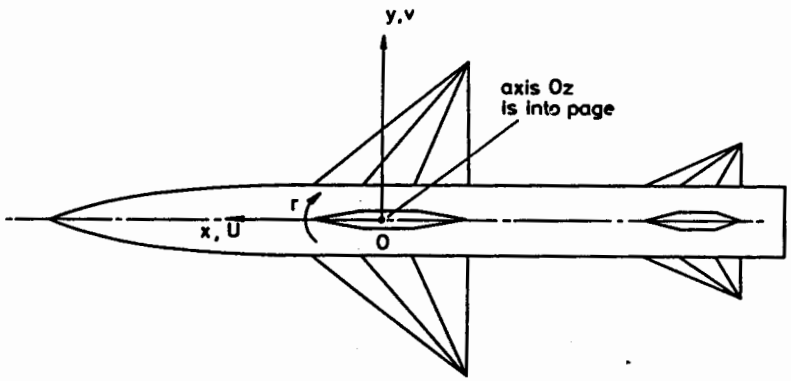

Fig. 5.4 *Diagram of a Cruciform Missile*

5.5.1 *Lateral dynamics of the Missile*

Consider each of the constituents of eqn. (5.15) in turn. The first is $m\dot{u} = \partial X$. If it is assumed that the missile performance is insensitive to small changes in forward speed, then the perturbation velocity u may be taken as zero. The first equation is therefore irrelevant and can be discarded.

The second equation, $m(\dot{v} + Ur) = \partial Y$, is the lateral force balance and is obviously of importance.

The third equation, $m(\dot{w} - qU) = \partial Z$, is the normal force balance and, by the initial assumptions, does not affect the lateral behaviour.

The fourth equation, $I_x \dot{p} = \partial L$, is the moment balance about the Ox axis. Because the missile is roll stabilized, the control system automatically adjusts ∂L by differential movement of the tail fins to keep $\Phi = 0$. Assume that the operation of this roll control loop does not influence the lateral motion of the missile.

The fifth equation is $I_y \dot{q} = \partial M$ and describes the pitching motion of the vehicle. By the initial assumptions, this is of no interest.

Finally the sixth equation, $I_z \dot{r} = \partial N$, which describes the moment balance about the Oz axis, is clearly of importance.

Thus the significant equations are

$$\begin{aligned} m(\dot{v} + Ur) &= \partial Y \\ I_z \dot{r} &= \partial N \end{aligned} \tag{5.20}$$

The next stage in modelling the dynamics is to obtain expressions for ∂Y and ∂N. Examination of eqns. (5.17) and (5.18) reveals that $\partial N_\theta = \partial Y_\epsilon = \partial N_\epsilon = 0$ and $\partial Y_\theta = mg\phi$. Since $\phi = 0$ (roll stabilization), $\partial Y_\theta = 0$. Thus the only forces and moments are aerodynamic. Experience shows that the important aerodynamic derivatives are

Y_v Lateral force due to lateral velocity. Clearly this is always negative.
Y_r Lateral force due to yaw rate.
N_v Yaw moment due to lateral velocity. If the centre at which the aerodynamic forces act is behind the centre of gravity, then the missile 'turns into the wind' and N_v is positive. This is called *weathercock* or *static stability.* In principle N_v may be made positive by suitably positioning the wings and tail. However, both the centre of gravity and aerodynamic centre are subject to relatively large excursions during the missile flight.
N_r Yaw moment due to yaw rate. Clearly this is negative.
Y_ζ Lateral force due to rudder (rear fin) displacement.
N_ζ Yaw moment due to rudder displacement. The definition of positive rudder displacement is one which causes a negative moment. Thus N_ζ is always negative. The side force at the rear fin which is needed to produce this negative moment is in the direction of Oy, and therefore Y_ζ is positive.

Substituting these derivatives into eqn. (5.20) gives

$$m\dot{v} = Y_v v + (Y_r - mU)r + Y_\zeta \zeta$$
$$I_z \dot{r} = N_v v + N_r r + N_\zeta \zeta \tag{5.21}$$

Because of the missile's high speed $|mU| \gg |Y_r|$. Thus Y_r may be neglected. (The same is not true for slow vehicles: for example most torpedos rely on Y_r to provide uncontrolled dynamic stability.)

Let

$$\frac{Y_v}{m} = y_v \qquad \frac{Y_\zeta}{m} = y_\zeta$$

$$\frac{N_v}{I_z} = n_v \qquad \frac{N_r}{I_z} = n_r \qquad \text{and} \qquad \frac{N_\zeta}{I_z} = n_\zeta$$

Equation (5.21) can then be written in the state form of $\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}$:

$$\begin{bmatrix}\dot{v}\\ \dot{r}\end{bmatrix} = \begin{bmatrix}y_v & -U\\ n_v & n_r\end{bmatrix}\begin{bmatrix}v\\ r\end{bmatrix} + \begin{bmatrix}y_\zeta\\ n_\zeta\end{bmatrix}\zeta \tag{5.22}$$

For readers familiar with system theory, the characteristic polynomial is seen to be $\det(sI - A) = s^2 - (y_v + n_r)s + (y_v n_r + n_v U)$. For the missile to be asymptotically stable (called in missile literature *dynamically stable*) all the zeros must lie in the open left half s-plane, i.e. $-(y_v + n_r) > 0$ and $y_v n_r + n_v U > 0$. The first of these inequalities is always true because y_v and n_r are both negative. Because U is large the practical implication of the second inequality is that n_v must be greater than zero to give dynamic stability. This, of course is the same as the criterion for static stability mentioned previously. As an exercise the reader may care to prove that for a vehicle such as a torpedo where y_r and $y_v n_r$ are not negligible, the condition for dynamic stability is $1 - [n_v(y_r - U)/y_v n_r] > 0$, which is the so-called G-criterion.

5.5.2 *Stability augmentation*

If natural frequency or damping ratio of the missile characteristic polynomial are not satisfactory they may be changed by a combination of accelerometer and rate gyroscope feedback. Let the gyroscope feedback gain be k_r and the accelerometer feedback gain be k_a, so that

$$\zeta = k_r r + k_a f_y + \zeta_c \tag{5.23}$$

where ζ_c is the command from the missile guidance system. From eqn. (5.16), (assuming $\dot{p}$ and ϕ are zero)

$$f_y = \dot{v} + Ur + b_x \dot{r} \tag{5.24}$$

Substituting for $\dot{v}$ and $\dot{r}$ from eqn. (5.22) gives

$$f_y = (y_v + b_x n_v)v + b_x n_r r + (y_\zeta + b_x n_\zeta)\zeta \tag{5.25}$$

Equation (5.25) shows that if f_y is used to drive the rudder via eqn. (5.23), then measured rudder position scaled by $k_a(y_\zeta + b_x n_\zeta)$ is part of the drive signal. This type of feedback is known as an *algebraic loop* and if the gain is too high may lead to high-frequency stability problems. To minimize this the accelerometer is usually placed in front of the missile centre of gravity near the position

$$b_x = -y_\zeta / n_\zeta$$

Equation (5.25) then reduces to

$$f_y = (y_v - y_\zeta n_v/n_\zeta)v - (y_\zeta n_r/n_\zeta)r \tag{5.26}$$

Substituting eqn. (5.26) and (5.23) into (5.22) gives

$$\begin{bmatrix}\dot{v}\\ \dot{r}\end{bmatrix} = \begin{bmatrix}y_v + k_1 y_\zeta & -U + k_2 y_\zeta\\ n_v + k_1 n_\zeta & n_r + k_2 n_\zeta\end{bmatrix}\begin{bmatrix}v\\ r\end{bmatrix} + \begin{bmatrix}y_\zeta\\ n_\zeta\end{bmatrix}\zeta \tag{5.27}$$

where

$$k_1 = k_a(y_v - y_\zeta n_v/n_\zeta)$$
$$k_2 = k_r - k_a y_\zeta n_r/n_\zeta.$$

The matrix $K = [k_1 \;\; k_2]$ is called the *state feedback matrix*.

If the characteristic polynomial of state equation (5.27) is evaluated, and small terms ignored, the result is

$$s^2 - (y_v + n_r + y_\zeta k_1 + n_\zeta k_2)s + (n_v + n_\zeta k_1)U.$$

The constant term $(n_v + n_\zeta k_1)U$ of this polynomial is denoted by ω_n^2, where ω_n is the missile natural frequency. Since k_1 is purely a function of k_a, the missile natural frequency may be adjusted by accelerometer feedback. The coefficient of s is denoted by $2\mu\omega_n$, where μ is the missile damping ratio. For any given k_a (chosen to give the desired value of ω_n), the value of k_2 can be altered by changing the rate gyro feedback gain k_r to give the required missile damping.

The values of ω_n and μ so obtained are sometimes called the *controlled* but *unguided* natural frequency and damping ratio.

To obtain the missile characteristics including the guidance system, a model of the guidance geometry must be added.

5.5.3 *The geometry of line-of-sight (LOS) guidance*

The geometry of a missile following a line of sight (LOS) toward a target is shown on Fig. 5.5. The symbols shown on the diagram are

U	velocity along the missile x-axis
v	velocity along the missile y-axis
$\Delta\psi$	angle between the missile x-axis and the LOS
Δy	length of perpendicular from LOS to missile

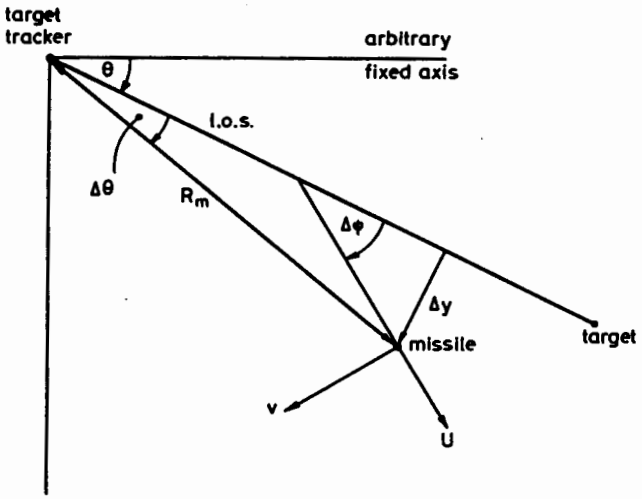

Fig. 5.5 *Geometry of line-of-sight guidance*

R_m range (distance) of missile from the ground-based tracking station
θ angle of LOS to an arbitrary fixed reference line through the tracker
$\Delta\theta$ angle between LOS and line joining tracker to missile.

From this figure,

$$\Delta\dot{y} = v\cos\Delta\psi + U\sin\Delta\psi - R_m\dot{\theta} \tag{5.28}$$

$$\Delta\psi = r - \dot{\theta} \tag{5.29}$$

Assuming that $\Delta\psi$ is small reduces eqn. (5.28) to

$$\Delta\dot{y} = v + U\,\Delta\psi - R_m\dot{\theta} \tag{5.30}$$

Equations (5.29) and (5.30) can be combined with eqn. (5.22) to give

$$\begin{bmatrix}\dot{v}\\ \dot{r}\\ \Delta\dot{\psi}\\ \Delta\dot{y}\end{bmatrix} = \begin{bmatrix}y_v & -U & 0 & 0\\ n_v & n_r & 0 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & U & 0\end{bmatrix}\begin{bmatrix}v\\ r\\ \Delta\psi\\ \Delta y\end{bmatrix} + \begin{bmatrix}y_\zeta\\ n_\zeta\\ 0\\ 0\end{bmatrix}\zeta + \begin{bmatrix}0\\ 0\\ -1\\ -R_m\end{bmatrix}\dot{\theta} \tag{5.31}$$

This is a typical 'aerospace model' in that it combines several rows of dynamics with several rows of geometry. The same pattern occurs in the aircraft equations derived in Section 5.6.

Consider the last row of eqn. (5.31)

$$\Delta\dot{y} = v + U\Delta\psi - R_m\dot{\theta}$$

Differentiating and substituting for $\Delta\dot{\psi}$ from row 3 gives

$$\Delta\ddot{y} = \dot{v} + Ur - (U + \dot{R}_m)\dot{\theta} - R_m\ddot{\theta}$$

If $\Delta\psi$ is small, then $U \doteq \dot{R}_m$. Hence

$$\Delta\ddot{y} = (\dot{v} + Ur) - (2\dot{R}_m\dot{\theta} + R_m\ddot{\theta}) \tag{5.32}$$

From eqn. (5.24), the term $\dot{v} + Ur$ is the lateral acceleration of the missile at its centre of gravity f_m, say. Elementary kinematics show that $R_m\ddot{\theta} + 2\dot{R}_m\dot{\theta}$ is the lateral acceleration of the LOS at a distance R_m from the tracker f_L, say. Thus

$$\Delta\ddot{y} = f_m - f_L \tag{5.33}$$

where

$$f_m = \dot{v} + Ur \tag{5.34}$$

$$f_L = R_m\ddot{\theta} + 2\dot{R}_m\dot{\theta} \tag{5.35}$$

Figure 5.6 shows a typical control structure arising from this model. An estimate of LOS acceleration is calculated on the ground from measured values of $\dot{\theta}$ and $\ddot{\theta}$, and this forms the basic guidance command to the missile. The displacement of the missile from the LOS is also measured and used as a 'double integral' feedback correction term. Compensation is necessary to give adequate stability to the geometrical (guidance) loop, and this may conflict with noise rejection requirements. Alternative feedback structures including optimal filtering, and full state feedback will probably occur to those readers who are control specialists.

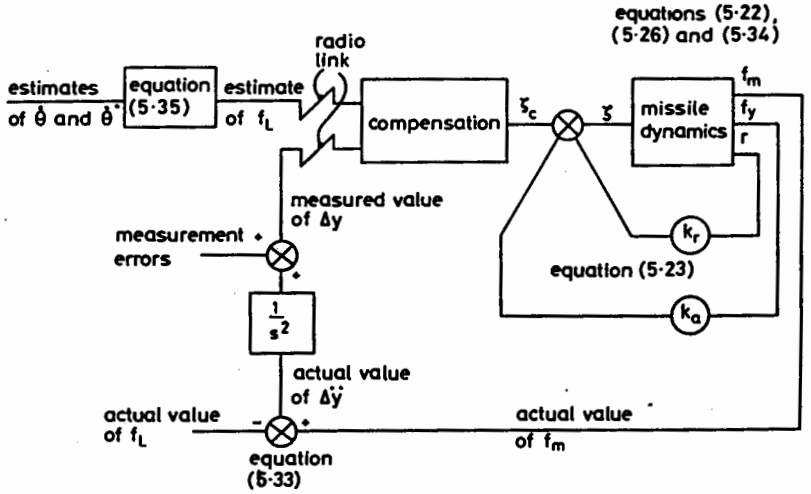

Fig. 5.6 *Missile LOS control block diagram*

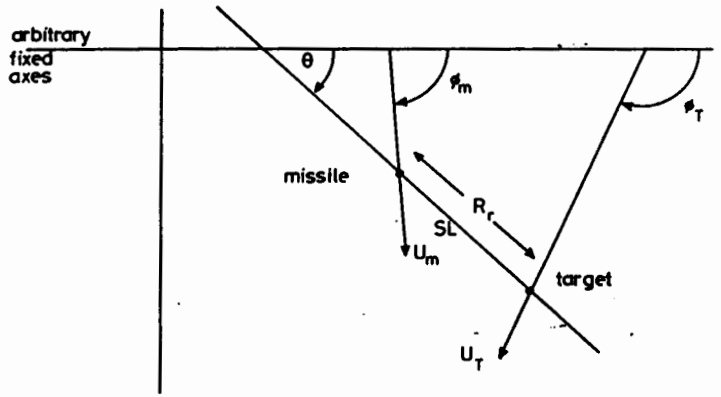

Fig. 5.7 *Homing geometry*

5.5.4 Geometry of terminal homing

Figure 5.7 shows the geometry of a missile homing onto a target. The sight line (SL) which joins the two differs from the LOS described in Section 5.5.3 because the 'end' is not fixed to the ground by a target tracker.

It is conventional when modelling the homing performance to replace the state variables v and ψ by

$$U_m = (U^2 + v^2)^{1/2}$$

$$\phi_m = \psi + \arctan\frac{v}{U}$$

In simulation studies this added complexity may be unnecessary. However, for analysis, if the missile dynamics are fast compared with the rates of change of geometry, then it may be assumed that $\dot{\phi}_m$ is controlled directly. The missile thus becomes equivalent to a particle with lateral acceleration

$$f_m = U_m \dot{\phi}_m \tag{5.36}$$

travelling at speed U_m.

Similarly the target has speed U_T directed at angle ϕ_T and has lateral acceleration f_T

$$f_T = U_T \dot{\phi}_T \tag{5.37}$$

The mathematical model is obtained by resolving along the SL:

$$\dot{R}_r = U_T \cos(\phi_T - \theta) - U_m \cos(\phi_m - \theta) \tag{5.38}$$

and perpendicular to the SL:

$$R_r\dot{\theta} = U_T \sin(\phi_T - \theta) - U_m \sin(\phi_m - \theta) \tag{5.39}$$

The way in which these equations are used depends upon the *homing policy*, i.e. the way in which the missile attacks the target. For example, *pursuit homing* has the policy

$$\phi_m = \theta \tag{5.40}$$

Thus the missile always points its velocity vector at the target.

Substituting eqns. (5.40) and (5.36) into (5.39) gives

$$f_m = \frac{U_m U_T}{R_r} \sin(\phi_T - \phi_m)$$

If the maximum achievable lateral acceleration of the missile is $|f_{max}|$, then to follow a pursuit trajectory

$$R_r \geq \frac{U_m U_T}{|f_{max}|} \sin(\phi_T - \phi_m) \tag{5.41}$$

This inequality is shown graphically on Fig. 5.8, and illustrates why missiles that employ this mode of guidance usually end in a tail chase (good if the missile has an infrared heat-seeking head).

One way of maintaining moderate acceleration levels is to use *proportional navigation* guidance

$$\dot{\phi}_m = k\dot{\theta} \tag{5.42}$$

An interesting expression for the value of k can be obtained by linearizing about the trajectory which would give interception to a constant-velocity

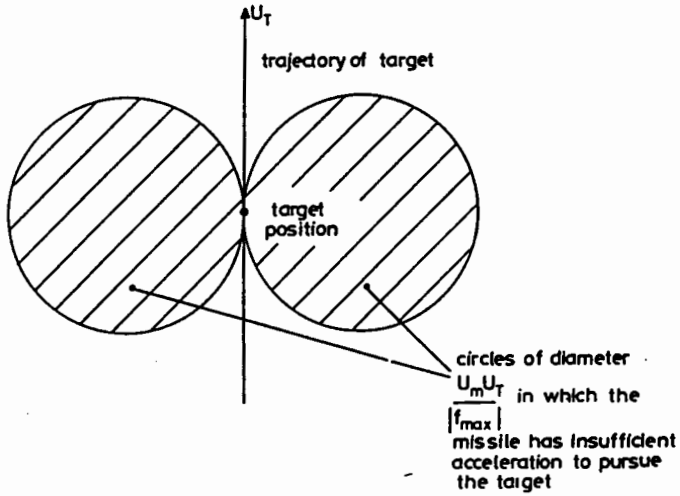

Fig. 5.8 *Pursuit homing constant acceleration contours*

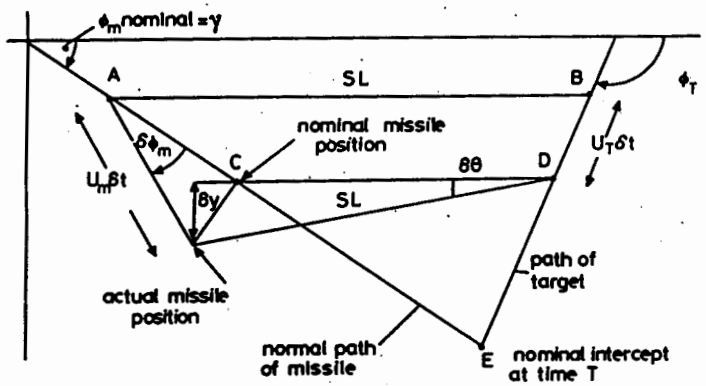

Fig. 5.9 *Proportional navigation homing geometry*

target. The similar triangles *ABE* and *CDE* of Fig. 5.9 show that if ϕ_T, U_T and U_m are constant, then ϕ_m should be constant. Further, the *SL* angle θ is constant. This is confirmed by eqn. (5.39). For $\dot{\theta} = 0$,

$$U_T \sin(\phi_T - \theta) - U_m \sin(\phi_m - \theta) = 0$$

Arbitrarily choose $\theta = 0$, and perturb the missile from the nominal trajectory $\phi_m = \gamma$ by $\delta\phi_m$. Let this perturbation cause deviations after time δt of δy and $\delta\theta$ as shown on Fig. 5.9.

Then from the figure,

$$\delta y = R_r\, \delta\theta \tag{5.43}$$

where R_r is the distance between missile and target. For small perturbations, the closing velocity of the missile and target is constant, U_r say. Thus

$$U_r = \dot{R}_r \tag{5.44}$$

Differentiating eqn. (5.43) gives

$$\delta\dot{y} = R_r\dot{\theta} - U_r\, \delta\theta \tag{5.45}$$

Equation (5.39) for $\theta = 0$ is

$$\begin{aligned} R_r\dot{\theta} &= U_T \sin\phi_T - U_m \sin(\gamma + \delta\phi_m) \\ &= U_T \sin\phi_T - U_m \sin\gamma - U_m\, \delta\phi_m \cos\gamma \end{aligned}$$

But γ has been chosen to make $U_T \sin\phi_T - U_m \sin\gamma$ zero. Thus the *SL* rotation caused by deviation $\delta\phi_m$ is

$$\dot{\theta} = -\frac{U_m \cos\gamma}{R_r}\, \delta\phi_m \tag{5.46}$$

Substitute eqn. (5.46) into (5.45) and differentiate

$$\delta\ddot{y} = -U_m \cos\gamma\, \dot{\phi}_m - U_r\dot{\theta}$$

Substituting again for $\dot{\theta}$ gives

$$\delta\ddot{y} = -U_m \cos\gamma\, (\dot{\phi}_m - \delta\phi_m R_r/U_r)$$

It is convenient to ignore the second term and assume that $\delta\ddot{y}$ can be regarded as an input to the model determined by the turn rate $\dot{\phi}_m$. Thus

$$\delta\ddot{y} = -U_m \cos\gamma\, \dot{\phi}_m \tag{5.47}$$

Now that δy, $\delta\dot{y}$ and $\delta\ddot{y}$ have been found in terms of the usual missile description (R_r, $\delta\theta$, U_m, etc.), the geometry can be expressed by the extremely simple model

$$\frac{d\delta y}{dt} = \delta\dot{y} \qquad \text{and} \qquad \frac{d\delta\dot{y}}{dt} = \delta\ddot{y}$$

In state form

$$\begin{bmatrix} \delta\dot{y} \\ \delta\ddot{y} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \delta y \\ \delta\dot{y} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \delta\ddot{y} \tag{5.48}$$

A technique which can now be applied is to find the state feedback matrix which minimizes the cost function

$$V = \delta y(T)^2 + \int_0^T \frac{\delta\ddot{y}^2}{\alpha^3}\, dt \tag{5.49}$$

The crucial term of eqn. (5.49) is $\delta y(T)$, the distance perpendicular to the *SL* by which the missile misses the target. However, the integral of $\delta\ddot{y}^2/\alpha^3$ puts a constraint on the missile's lateral acceleration whose magnitude is determined by an appropriate choice of α.

The method of minimizing functions such as V is well known to control engineers,[5] and basically consists of solving a non-linear matrix differential equation called the matrix Riccati equation. The particular form arising from (5.48) and (5.49) is unusual because it has an analytic solution, and results in the state feedback relationship

$$\delta\ddot{y} = -\frac{3\alpha^3}{3 + \alpha^3(\dot{R}_r/U_r)^3} \left[\frac{R_r}{U_r} \left(\frac{R_r}{U_r} \right)^2 \right] \begin{bmatrix} \delta y \\ \delta\dot{y} \end{bmatrix} \tag{5.50}$$

Substituting for $\delta\ddot{y}$, $\delta\dot{y}$ and δy from eqns. (5.47), (5.43) and (5.45) gives

$$\dot{\phi}_m = \frac{3U_r\alpha^3(R_r/U_r)^3}{U_m\cos\gamma(3+\alpha^3(R_r/U_r)^3)}\dot{\theta}$$

Clearly when the range is large compared with the closing speed

$$\alpha^3(R_r/U_r)^3 \gg 3 \tag{5.51}$$

(note the cube law) and

$$\dot{\phi}_m = \frac{3U_r}{U_m\cos\gamma}\dot{\theta} \tag{5.52}$$

Thus the value of k in the proportional navigation equation (5.42) is

$$\frac{3U_r}{U_m\cos\gamma}$$

If inequality (5.51) does not hold, the assumptions of the mathematical model are also suspect: for example $\delta\phi_m R_r/U_r$ should perhaps not be ignored in obtaining eqn. (5.47).

5.6 Model of the motion of an aeroplane

As a second example of the use of the equations developed in Section 5.4, consider the motion of a conventional aeroplane. The aerodynamic derivatives which are usually non-negligible are[2]

$$\begin{array}{ccccccc}
X_u & X_w & X_\eta & & & & \\
Y_v & Y_r & Y_\zeta & & & & \\
Z_u & Z_w & Z_q & Z_{\dot{w}} & Z_{\dot{q}} & Z_\eta & \\
L_v & L_p & L_r & L_{\dot{p}} & L_{\dot{r}} & L_\xi & L_\zeta \\
M_u & M_w & M_q & M_{\dot{w}} & M_{\dot{q}} & M_\eta & \\
N_v & N_p & N_r & N_{\dot{p}} & N_{\dot{r}} & N_\xi & N_\zeta
\end{array}$$

If the state vector is taken in the obvious order

$$[u\ v\ w\ p\ q\ r\ \phi\ \theta\ \psi\ h]^T,$$

an attractive pattern of zero and non-zero elements are seen when the equations are written in matrix form. A certain amount of trial and error on the mathematics, or alternatively and far better, an appreciation of the physical implications of the model, results in a block diagonal form if the state vector is reordered

$$[u\ w\ q\ \theta\ h\ \vdots\ v\ p\ r\ \phi\ \psi]^T$$

$$\left[\begin{array}{ccccc:ccccc}
m & 0 & 0 & 0 & 0 & & & & & \\
0 & m-Z_{\dot{w}} & -Z_{\dot{q}} & 0 & 0 & & & & & \\
0 & -M_{\dot{w}} & I_y-M_{\dot{q}} & 0 & 0 & & & 0 & & \\
0 & 0 & 0 & 1 & 0 & & & & & \\
0 & 0 & 0 & 0 & 1 & & & & & \\
\hdashline
 & & & & & m & 0 & 0 & 0 & 0 \\
 & & & & & 0 & I_x-L_{\dot{p}} & -L_{\dot{r}} & 0 & 0 \\
 & & 0 & & & 0 & -N_{\dot{p}} & I_z-N_{\dot{r}} & 0 & 0 \\
 & & & & & 0 & 0 & 0 & 1 & 0 \\
 & & & & & 0 & 0 & 0 & 0 & 1
\end{array}\right]\left[\begin{array}{c}\dot{u}\\\dot{w}\\\dot{q}\\\dot{\theta}\\\dot{h}\\\hdashline\dot{v}\\\dot{p}\\\dot{r}\\\dot{\phi}\\\dot{\psi}\end{array}\right]$$

$$=\left[\begin{array}{ccccc:ccccc}
X_u & X_w & 0 & -mg & 0 & & & & & \\
Z_u & Z_w & Z_q+mU & 0 & 0 & & & & & \\
M_u & M_w & M_q & 0 & 0 & & & 0 & & \\
0 & 0 & 1 & 0 & 0 & & & & & \\
0 & -1 & 0 & U & 0 & & & & & \\
\hdashline
 & & & & & Y_v & 0 & Y_r-mU & mg & 0 \\
 & & & & & L_v & L_p & L_r & 0 & 0 \\
 & & 0 & & & N_v & N_p & N_r & 0 & 0 \\
 & & & & & 0 & 1 & 0 & 0 & 0 \\
 & & & & & 0 & 0 & 1 & 0 & 0
\end{array}\right]\left[\begin{array}{c}u\\w\\q\\\theta\\h\\\hdashline v\\p\\r\\\phi\\\psi\end{array}\right]$$

$$+\left[\begin{array}{cc:cc}
X_\eta & 1 & & \\
Z_\eta & \epsilon & & \\
M_\eta & c_z & 0 & \\
0 & 0 & & \\
0 & 0 & & \\
\hdashline
 & & 0 & Y_\zeta \\
 & & L_\xi & L_\zeta \\
0 & & N_\xi & N_\zeta \\
 & & 0 & 0 \\
 & & 0 & 0
\end{array}\right]\left[\begin{array}{c}\eta\\\partial E\\\hdashline\xi\\\zeta\end{array}\right] \tag{5.53}$$

The equation comes from the following sources:

row 1	longitudinal force balance	1st eqn. of (5.15)
row 2	normal force balance	3rd eqn. of (5.15)
row 3	pitch moment balance	5th eqn. of (5.15)
row 4	pitch rate to pitch angle	2nd row of (5.12)
row 5	body velocity to reference velocity	eqn. (5.11)
row 6	lateral force balance	2nd eqn. of (5.15)

row 7	roll moment balance	4th eqn. of (5.15)
row 8	yaw moment balance	6th eqn. of (5.15)
row 9	roll rate to roll angle	1st row eqn. of (5.12)
row 10	yaw rate to yaw angle	3rd row eqn. of (5.12)

The gravitational forces and moments occurring in rows, 1, 2, 3, 6, 7 and 8 are specified by eqn. (5.17), the propulsion forces and moments by (5.18) and the aerodynamic forces and moments by the assumed non-zero derivatives listed above.

For physical reasons, which it is hoped are obvious to the reader, the first five rows are called the *longitudinal* equations of motion and the second five rows are called the *lateral* equations of motion. The block diagonal form shows that the motions are independent.

5.6.1 *Longitudinal equations of motion*

It can be seen from eqn. (5.53) that the longitudinal model is described by three dynamic equations and two geometrical equations. Together they give rise to three modes* of motion. Put the equations in to state form

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u} \tag{5.54}$$

by premultiplying by the inverse of the 'mass' matrix

$$\begin{bmatrix} m & 0 & 0 & 0 & 0 \\ 0 & m - Z_{\dot{w}} & -Z_{\dot{q}} & 0 & 0 \\ 0 & -M_{\dot{w}} & I_y - M_{\dot{q}} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The inverse always exists because the Gershgorin discs do not allow any zero eigenvalues (i.e. $|m - Z_{\dot{w}}| > |Z_{\dot{q}}|$ and $|I_y - M_{\dot{q}}| > |M_{\dot{w}}|$). The longitudinal equation of motion becomes

$$\begin{bmatrix} \dot{u} \\ \dot{w} \\ \dot{q} \\ \dot{\theta} \\ \dot{h} \end{bmatrix} = \begin{bmatrix} x_u & x_w & 0 & -g & 0 \\ z_u & z_w & z_q + U & 0 & 0 \\ m_u & m_w & m_q & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & U & 0 \end{bmatrix} \begin{bmatrix} u \\ w \\ q \\ \theta \\ h \end{bmatrix} + \begin{bmatrix} x_\eta & 1 \\ z_\eta & z_e \\ m_\eta & m_e \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \eta \\ e \end{bmatrix} \tag{5.55}$$

The lower-case letters for the derivatives indicate that to a good degree of accuracy (and in most cases exactly), $x_u = X_u/m$, $m_w = M_w/I_y$, etc. Note that the perturbation engine thrust ∂E has been replaced by the specific perturbation thrust $e = \partial E/m$. Naturally e is in units of acceleration, so element (1, 2) of the input matrix turns out to be 1. If the characteristic polynomial

* Readers familiar with system theory will find that the name 'mode' is used in a similar but not quite identical sense to $\mu e^{\lambda t}$.

det $(sI - A)$ is evaluated for eqn. (5.55), it is usually found to be of the form $s(s^2 + 2\mu_q\omega_q s + \omega_q^2)(s^2 + 2\mu_p\omega_p s + \omega_p^2)$ corresponding to the *height integration* mode, *short period* (quick) mode and *phugoid* mode of motion, respectively.

The height integration mode. The last column of the A-matrix of eqn. (5.55) is zero. Thus h does not affect any other state variable, and the last row of the matrix therefore represents pure integration. This accounts for the s term in det $(sI - A)$. Clearly the last row and column may be removed without affecting the behaviour of the rest of the system, and eqn. (5.55) reduces to

$$\begin{bmatrix} \dot{u} \\ \dot{w} \\ \dot{q} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} x_u & x_w & 0 & -g \\ z_u & z_w & z_q + U & 0 \\ m_u & m_w & m_q & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ w \\ q \\ \theta \end{bmatrix} + \begin{bmatrix} x_\eta & 1 \\ z_\eta & z_e \\ m_\eta & m_e \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \eta \\ e \end{bmatrix} \tag{5.56}$$

The short period mode. In a classical aircraft the forward speed changes very slowly compared with pitch motions, and it is reasonable to assume that these fast transients occur at effectively constant speed, say $u = 0$. (Recall that u is the perturbation in forward speed about the steady value U.)

The first row of (5.56) may therefore be neglected when investigating the fast pitch motion

$$\begin{bmatrix} \dot{w} \\ \dot{q} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} z_w & z_q + U & 0 \\ m_w & m_q & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ w \\ \theta \end{bmatrix} + \begin{bmatrix} z_\eta & z_e \\ m_\eta & m_e \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \eta \\ e \end{bmatrix} \tag{5.57}$$

The pitch angle now has no effect on the motion because the last column of eqn. (5.57) is zero. The equations describing the fast pitch transients therefore reduce to

$$\begin{bmatrix} \dot{w} \\ \dot{q} \end{bmatrix} = \begin{bmatrix} z_w & z_q + U \\ m_w & m_q \end{bmatrix} \begin{bmatrix} w \\ q \end{bmatrix} + \begin{bmatrix} z_\eta & z_e \\ m_\eta & m_e \end{bmatrix} \begin{bmatrix} \eta \\ e \end{bmatrix} \tag{5.58}$$

It is usually found on a classical aircraft that $|z_q| \ll U$. Also the engine thrust is almost along the Ox axis making z_e and m_e small. Equation (5.58) further reduces to

$$\begin{bmatrix} \dot{w} \\ \dot{q} \end{bmatrix} = \begin{bmatrix} z_w & U \\ m_w & m_q \end{bmatrix} \begin{bmatrix} w \\ q \end{bmatrix} + \begin{bmatrix} z_\eta \\ m_\eta \end{bmatrix} \eta \tag{5.59}$$

The characteristic polynomial det $(sI - A)$ of this state equation is

$$s^2 + (-z_w - m_q)s + (z_w m_q - m_w U) \tag{5.60}$$

It is usually found that this is a very good approximation of the factor $(s^2 + 2\mu_q\omega_q s + \omega_q^2)$ appearing in the full characteristic polynomial of (5.55).

Readers may like to compare eqn. (5.59) with eqn. (5.22) which described the lateral behaviour of a guided missile. Allowing for the sign change in U

which is due to v and r having a different mutual geometrical relationship to that of w and q (corresponding sign changes occur in the aerodynamic derivatives), the equations are identical and describe essentially the same behaviour. Using the same arguments as the missile case, the practical implication of polynomial (5.60) is that the short-period mode is stable if $m_w < 0$.

Readers should study Fig. 5.1 to convince themselves that $m_w < 0$ and $n_v > 0$ both imply weathercock stability (a vehicle suspended at its c.g. will 'turn into wind') and it is merely geometry which causes the sign reversal.

The Phugoid mode. Because the short-period mode is fast, it may be assumed that its differential equations are instantaneously satisfied compared with the slower motions. The zero-input ($\eta = e = 0$) version of eqn. (5.56) therefore becomes

$$\begin{bmatrix} \dot{u} \\ 0 \\ 0 \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} x_u & x_w & 0 & -g \\ z_u & z_w & U & 0 \\ m_u & m_w & m_q & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ w \\ q \\ \theta \end{bmatrix} \tag{5.61}$$

Solving the algebraic second and third rows to obtain w and q in terms of u, and substituting into the first and fourth rows gives

$$\begin{bmatrix} \dot{u} \\ \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} x_u - x_w \dfrac{m_q z_u - m_u U}{m_q z_w - m_w U} & -g \\ \\ -\dfrac{z_u m_w - m_u z_w}{U m_w - m_q z_w} & 0 \end{bmatrix} \begin{bmatrix} u \\ \\ \theta \end{bmatrix} \tag{5.62}$$

The characteristic polynomial is

$$s^2 + \left[x_u - x_w\left(\frac{m_q z_u - m_u U}{m_q z_w - m_w U}\right)\right]s - g\left(\frac{z_u m_w - m_u z_w}{U m_w - m_q z_w}\right)$$

It is usually found that this is a reasonable approximation of the factor $(s^2 + 2\mu_p \omega_p s + \omega_p^2)$ appearing in the full characteristic polynomial of (5.55).

Physical interpretation of the modes. The interpretation of height integration is rather obvious; if the aircraft is climbing there is no feedback to prevent it, so height continues to increase.

The short-period motion is due to the 'arrow stability' of the aircraft: the distance of the aerodynamic centre from the c.g. and the magnitude of the various aerodynamic damping terms.

The phugoid mode has a more subtle but rather interesting interpretation. Assume (and this is true for a typical subsonic classical aircraft) that $|z_w m_u| \gg |m_w z_u|$ and $|U m_w| \gg |m_q z_w|$. Thus $\omega_p \doteq \sqrt{-g z_u / U}$. The negative sign causes no concern because z_u is negative. Now under equilibrium conditions the total lift is proportional to U^2, and in the opposite direction of Oz.

$$Z = -kU^2 = -mg \tag{5.63}$$

and z_u can be calculated from

$$z_u = \frac{1}{m}\frac{\partial Z}{\partial u}$$

Now $\partial u = \partial U$, so

$$z_u = \frac{1}{m}\frac{\partial}{\partial U}(-kU^2) = -\frac{2kU}{m}$$

Substituting from (5.63) for k gives

$$z_u = -\frac{2g}{U}$$

which can be substituted into the simplified expression for ω_p to give

$$\omega_p \doteq \sqrt{2}\,\frac{g}{U} \tag{5.64}$$

Now consider a very simple model of a constant energy vehicle

$$\tfrac{1}{2}mU^2 + mgH = \text{constant}$$

supported by a force proportional to speed squared

$$m\frac{d^2H}{dt^2} = kU^2 - mg$$

Perturbations about these two equations give

$$Uu + gh = 0 \qquad \text{and} \qquad m\frac{d^2h}{dt^2} = 2kUu$$

Eliminating u between them, and substituting for k from eqn. (5.63) gives

$$\ddot{h} = -2\left(\frac{g}{U}\right)^2 h$$

Thus this very simple model gives an undamped height oscillation of the same frequency as eqn. (5.64). This analysis therefore indicates that the phugoid is basically an oscillation reflecting an interchange of potential and kinetic energy, although this is heavily disguised in eqn. (5.62).

5.6.2 Lateral equations of motion

The last five rows of eqn. (5.53) describe the lateral motion of a classical aircraft. They can be put into state form by a similar procedure to that employed in Section 5.6.1.

$$\begin{bmatrix}\dot{v}\\ \dot{p}\\ \dot{r}\\ \dot{\phi}\\ \dot{\psi}\end{bmatrix} = \begin{bmatrix} y_v & 0 & y_r - U & g & 0\\ l_v & l_p & l_r & 0 & 0\\ n_v & n_p & n_r & 0 & 0\\ 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\end{bmatrix}\begin{bmatrix}v\\ p\\ r\\ \phi\\ \psi\end{bmatrix} + \begin{bmatrix} 0 & y_\zeta\\ l_\xi & l_\zeta\\ n_\xi & n_\zeta\\ 0 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\xi\\ \zeta\end{bmatrix} \tag{5.65}$$

Like the longitudinal case, the first three rows are dynamics and the last two geometry. The lower-case letters indicate multiplication by the inverse of the 'mass' matrix so that to a high degree of accuracy $y_v = Y_v/m$, $l_v = L_v/I_x$, etc. If the characteristic polynomial det $(sI - A)$ is evaluated it will usually be of the form

$$s\left(s + \frac{1}{\tau_r}\right)(s^2 + 2\mu_d\omega_d s + \omega_d^2)\left(s + \frac{1}{\tau_s}\right) \tag{5.66}$$

which correspond to the *heading integration* mode, the *roll subsidence* mode, the *dutch roll* and the *spiral* mode, respectively.

Heading integration mode. Like the height integration mode, this arises because the last column of the state matrix is zero. Equation (5.65) can therefore be reduced to

$$\begin{bmatrix}\dot{v}\\ \dot{p}\\ \dot{r}\\ \dot{\phi}\end{bmatrix} = \begin{bmatrix} y_v & 0 & y_r - U & g\\ l_v & l_p & l_r & 0\\ n_v & n_p & n_r & 0\\ 0 & 1 & 0 & 0\end{bmatrix}\begin{bmatrix}v\\ p\\ r\\ \phi\end{bmatrix} + \begin{bmatrix} 0 & y_\zeta\\ l_\xi & l_\zeta\\ n_\xi & n_\zeta\\ 0 & 0\end{bmatrix}\begin{bmatrix}\xi\\ \zeta\end{bmatrix} \tag{5.67}$$

Roll subsidence mode. For the model of a classical aircraft, τ_r in polynomial (5.66) is small compared with τ_s or with $1/\omega_d$. It is caused primarily by the high damping of the wings in roll, which can be described by

$$\dot{p} = l_p p$$

Usually the approximation $\tau_r = -1/l_p$ is quite accurate.

Dutch roll mode. Because the roll subsidence decays so rapidly, the zero-input version of eqn. (5.67) can be written

$$\begin{bmatrix}\dot{v}\\ 0\\ \dot{r}\\ \dot{\phi}\end{bmatrix} = \begin{bmatrix} y_v & 0 & y_r - U & g\\ l_v & l_p & l_r & 0\\ n_v & n_p & n_r & 0\\ 0 & 1 & 0 & 0\end{bmatrix}\begin{bmatrix}v\\ p\\ r\\ \phi\end{bmatrix}$$

The second row can be solved to give

$$p = -\frac{1}{l_p}(l_v v + l_r r) \tag{5.68}$$

Ignoring the effect of gravitation ϕ has no effect on the motion, so that the last row may be discarded. Assuming that $|y_r| \ll U$ the equation reduces to

$$\begin{bmatrix}\dot{v}\\ \dot{r}\end{bmatrix} = \begin{bmatrix} y_v & -U\\ n_v - \dfrac{n_p l_v}{l_p} & n_r - \dfrac{n_p l_r}{l_p}\end{bmatrix}\begin{bmatrix}v\\ r\end{bmatrix} \tag{5.69}$$

The characteristic polynomial det $(sI - A)$ is

$$s^2 + \left(-y_v - n_r + \frac{n_p l_r}{l_p}\right)s + \left(y_v n_r + n_v U - \frac{y_v n_p l_r}{l_p} - \frac{n_p l_v U}{l_p}\right)$$

This polynomial is usually a close approximation to the factor $(s^2 + 2\mu_d\omega_d s + \omega_d^2)$ in the characteristic polynomial of eqn. (5.65).

Comparing eqn. (5.69) with (5.22) and (5.59), it is clear that the Dutch roll has similar origins to the yawing motion of the missile, or the short-period pitch oscillations of the aircraft. However, it is modified by the accompanying rolling motion which is necessary to satisfy eqn. (5.68).

Spiral mode. The final time constant in the characteristic polynomial τ_s is usually very slow indeed.

It can be approximated by solving the first three rows of the equation

$$\begin{bmatrix}0\\ 0\\ 0\\ \dot{\phi}\end{bmatrix} = \begin{bmatrix} y_v & 0 & -U & g\\ l_v & l_p & l_r & 0\\ n_v & n_p & n_r & 0\\ 0 & 1 & 0 & 0\end{bmatrix}\begin{bmatrix}\dot{v}\\ p\\ r\\ \phi\end{bmatrix}$$

to eliminate v and r. Putting $p = \dot{\phi}$ (last row) this gives

$$\dot{\phi} = g\left[\frac{y_v(n_r l_p - l_r n_p)}{n_r l_v - l_r n_v} - \frac{U(n_v l_p - l_v n_p)}{n_v l_r - l_v n_r}\right]^{-1}\phi$$

Usually the first term in the square brackets is small compared with the second, and a reasonable approximation to τ_s is

$$\tau_s = \frac{U}{g}\left(\frac{n_v l_p - l_v n_p}{n_v l_r - l_v n_r}\right) \tag{5.70}$$

Physical interpretation of the modes. The heading integration mode, like the height integration mode, is purely geometrical in nature and reflects the fact that there is no feedback of heading angle into the equations of motion.

The roll subsidence mode is caused by the high roll damping effect of the wing.

The Dutch roll is basically the yaw equivalent to the pitch short-period oscillation. However, the motion couples into roll, and this modifies the characteristic polynomial. In some aircraft the effect of the roll coupling on natural frequency and damping ratio is quite small. i.e.

$$\left|\frac{n_p l_v}{l_p}\right| \ll |n_v| \quad \text{and} \quad \left|\frac{n_p l_r}{l_p}\right| \ll |n_r|.$$

However, the impression gained by an observer in the aircraft may be of a predominantly rolling oscillation because of the large values of p caused via eqn. (5.68).

The spiral mode is a gravitional effect. When the aircraft rolls, gravity causes it to sideslip and yaw. The resulting forces and moments may cause ϕ to decrease giving a stable spiral mode, but on many aircraft the resulting motion is unstable. However, it is so slow that only in exceptional circumstances do pilots have difficulty in controlling it.

5.6.3 *Aircraft control-system design*

Although this chapter is concerned with modelling rather than control, a few remarks about the use of the model obtained in Sections 5.6.2 and 5.6.3 are in order. Traditionally the control system has been designed in two stages: the inner loop and the outer loop. The function of the inner loop is to make the aircraft easy and pleasant to fly, and is often called a *stability augmentation system* (SAS). The outer loop is to replace the pilot for certain simple flight manoeuvres such as maintaining height, maintaining speed, turning onto a specified heading, climbing at a specified rate etc. This system is often called an *autopilot*. Outside both of these loops is the *navigation loop*. Commands derived from the inertial navigation system or radio navigation aids can be fed into the autopilot to demand heading changes. A block diagram of these functions is shown on Fig. 5.10.

In more recent aircraft there may be additional functions such as structural mode control, gust alleviation, automatic landing, remote weapon control, etc.

All of these areas provide interesting and challenging design problems, not least in providing hardware and software of sufficiently high integrity. Current discussion centres around the requirement that the probability of a catastrophic accident due to control system failure should be less than 10^{-9} per hour for passenger aircraft. (This is about ten thousand times better than the probability of being killed or seriously injured in a car accident in Britain). Some idea of the difficulty in proving that a design meets this specification may be gained by equating 10^9 hours with one hundred thousand years.

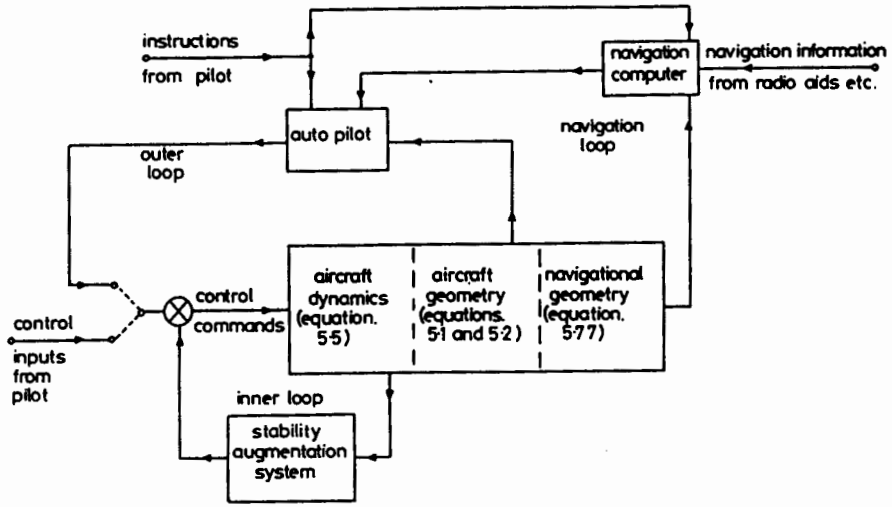

Fig. 5.10 *Block diagram of traditional aircraft control loop*

To return to the more traditional aspects of control system design, the SAS is a particularly interesting case because of the difficulty in specifying whether an aircraft handles well or badly. Reference 5 must be one of the most comprehensive and detailed specifications of a particular classs of control system performance that has ever been written, and is applied very widely as a design criterion. For aircraft whose dynamics are rather different to those of a classical aircraft, a very simple idea has been used to assess longitudinal handling, called the *C* criterion*[6].

It assumes that at low speeds a pilot senses longitudinal response visually via the pitch rate q. At high speeds he senses it by acceleration f_z through the 'seat of his pants'. A speed, U_{co} is chosen at which the two effects are assumed to be sensed equally by the average pilot. From eqn. (5.16), under steady pitch rate conditions $f_z = -Uq$. Therefore the expression $f_z + U_{co}q$ weights the terms equally at $U = U_{co}$. It gives more weight to acceleration when $U > U_{co}$, and more weight to pitch rate when $U < U_{co}$. The parameter C^* is defined by (the bar indicates Laplace transformation)

$$C^* = \frac{1}{g}\left(\frac{\bar{f}_z}{\bar{\eta}} + U_{co}\frac{\bar{q}}{\bar{\eta}}\right)$$

and either the Bode plot of $C^*(j\omega)$ or the step response of $C^*(t)$ must lie within certain regions. On some aircraft the control laws are continuously modified in flight by means of a so-called C^* box to keep the criterion satisfied.

5.6.4 *Model of the human pilot*

It is clear that a mathematical model of the human pilot is complementary to the specifications of handling quality. There are many variations possible according to the circumstances,[7] but a widely accepted model is that of the pilot being represented by the transfer function

$$P(s) = K\frac{e^{-\tau_d s}}{\tau_e s + 1} \cdot \frac{\alpha\tau_a s + 1}{\tau_a s + 1}$$

The values of time delay τ_d and the time constant τ_e depend upon the pilot's fatigue level, workload, physical health etc., and both lie between 1/20 and 1/4 seconds.

The parameters K, τ_a and α are instinctively chosen by the pilot to make the transfer function $Q(s)$ shown on Fig. 5.11 approximate to $\omega_c e^{-\tau_d s}/s$. Ignoring the pure time delay τ_d, this implies a closed loop behaviour of $\omega_c/(s + \omega_c)$, so ω_c can be considered as the bandwidth which the pilot considers necessary for adequate control. The pilot task is difficult if the bandwidth requirement is too high ($1/\omega_c$ fast compared with τ_e and τ_d), or the behaviour $\omega_c e^{-\tau_d s}/s$ cannot be achieved, or can only be achieved with excessive phase advance (a high value of α). For example the manual control of a double integrator plant $1/s^2$ is difficult because the desired $P(s)$ is $\omega_c e^{-\tau_d s}.s$ showing that τ_e must be small compared with $1/\omega_c$ and the phase advance required is 90°.

One disadvantage of the usual models of pilot behaviour is that they may be difficult to apply in the design situation. A very simple criterion was developed in Reference 8. It is assumed that the pilot has transfer function $K/0{\cdot}25s + 1$. The designer attempts to choose a value of K which gives the dominant system poles a damping ratio of 0·7. If this is possible, and other factors (such as the stability of the remaining poles, system bandwidth, steady error accuracy etc.) are within specification, then the pilot will rate the handling as 'easy'.

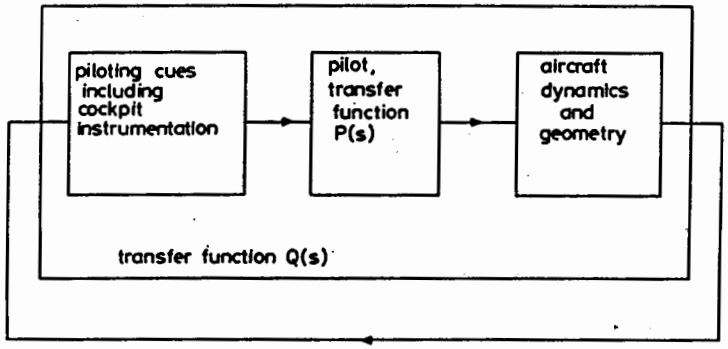

Fig. 5.11 *Human Pilot Block Diagram*

5.7 Model of an inertial navigation unit

A mathematical model of a gimballed inertial navigation unit (INU) will serve as a third example of the use of the equations developed in earlier sections. Traditionally the behaviour of INU's have been described in terms of the equations of vector mechanics. Rather than rewrite Section 5.3 in this language, the matrix description is retained. For the purposes of this chapter this has the advantage of continuity of style, and there is a further advantage which is mentioned in Section 5.7.3. The disadvantage is that the standard texts on IN may look unfamiliar.

The principle of operation of an INU may be understood from Fig. 5.12(*a*), which shows an accelerometer mounted on a railway truck. The displacement

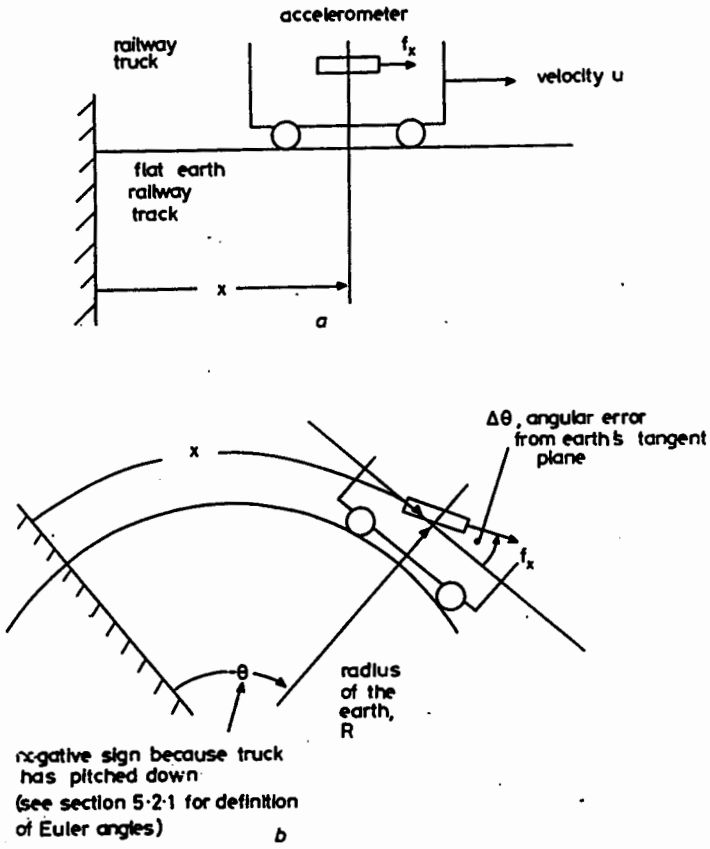

Fig. 5.12 *(a) Flat earth navigational system*
(b) Round earth navigational system

of the truck along the track could perhaps be deduced from double integration of the accelerometer output. The state equation is

$$\begin{bmatrix}\dot{x}\\ \dot{u}\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}x\\ u\end{bmatrix} + \begin{bmatrix}0\\ 1\end{bmatrix} f_x$$

Clearly the minimum polynomial of the state matrix is s^2. There is thus a repeated imaginary zero and the measurement of x is unstable. It is remarkable that the equivalent procedure for the vastly more complicated situation of a spherical rotating earth can give stable (although not asymptotically stable) navigation. The trick is to pivot the accelerometer on gimbals so that gravitation can stabilize the accelerometer errors.

To illustrate this consider the 'round stationary earth' railway track shown in Fig. 5.12(*b*). The distance x is now given by

$$x = -R\theta$$

so that distance may be determined by the angle through which the accelerometer axis has rotated. To control angular motion assume that a device called an *integrating rate gyroscope* (IRG) is mounted on the gimballed platform alongside the accelerometer. The IRG is coupled up to a motor on the gimbal system in such a way that a command to the IRG rotates the platform at any desired angular rate $\dot{\theta}$.

Suppose that the system has produced an accelerometer error Δf_x which in turn has caused a velocity error ΔU and an angular misalignment error $\Delta\theta$. Now to keep the accelerometer in the tangent plane of the earth, the IRG should be rotating the platform at a rate

$$\dot{\theta} = -\frac{U}{R} \tag{5.71}$$

However, the velocity error ΔU is causing rate of change of misalignment error

$$\Delta\dot{\theta} = -\frac{\Delta U}{R} \tag{5.72}$$

Additionally, the velocity error is being integrated to produce the displacement error.

$$\Delta\dot{x} = \Delta U \tag{5.73}$$

Because the accelerometer is tilted at $\Delta\theta$ to the tangent plane of the earth it senses not only the true acceleration $\dot{U}$ and its error Δf_x, but also a component of the gravitational field, $g \sin \Delta\theta \doteq g\,\Delta\theta$. It is this total accelerometer output which is integrated to compute velocity

$$U + \Delta U = \int (\dot{U} + \Delta f_x + g\,\Delta\theta)\,dt$$

Differentiating both sides gives

$$\Delta\dot{U} = \Delta f_x + g\,\Delta\theta \tag{5.74}$$

Writing eqns. (5.72), (5.73) and (5.74) in state form gives

$$\begin{bmatrix}\Delta\dot{x}\\ \Delta\dot{U}\\ \Delta\dot{\theta}\end{bmatrix} = \begin{bmatrix}0 & 1 & 0\\ 0 & 0 & g\\ 0 & -\frac{1}{R} & 0\end{bmatrix}\begin{bmatrix}\Delta x\\ \Delta u\\ \Delta\theta\end{bmatrix} + \begin{bmatrix}0\\ 1\\ 0\end{bmatrix}\Delta f_x \tag{5.75}$$

The minimum polynomial of the state matrix is

$$s\left(s^2 + \frac{g}{R}\right) \tag{5.76}$$

which implies stability, although not asymptotic stability. Notice that the limit cycle has a frequency $\sqrt{g/R}$. The period of this is about 84 minutes and is called the Schuler period. The transfer function matrix, $(sI - A)^{-1}B$ is given by

$$\begin{bmatrix}\overline{\Delta x}\\ \overline{\Delta U}\\ \overline{\Delta\theta}\end{bmatrix} = \begin{bmatrix}\frac{1}{s^2 + g/R}\\ \frac{s}{s^2 + g/R}\\ \frac{1}{R(s^2 + g/R)}\end{bmatrix}\overline{\Delta f_x}$$

This shows that the zero of polynomial (5.76) at $s = 0$ is uncontrollable with respect to Δf_x, and hence the error Δx caused by a step in Δf_x is bounded and purely oscillatory with a period of 84 minutes.

The interested reader may like to anticipate a result presented in Section 5.7.3 and show that this happy circumstance does not occur if the error lies in the IRG instead of the accelerometer. [Replace eqn. (5.72) by

$$\Delta\dot{\theta} + \dot{\theta}_e = -\frac{\Delta U}{R}$$

where $\dot{\theta}_e$ is the IRG error.]

5.7.1 Spherical earth geometry

In Section 2.1, reference axes O_0x_0, O_0y_0, O_0z_0 were defined in the tangent plane of the earth. It is now convenient to regard then as attached to the platform which, of course, should support accelerometers in this plane. In the past the orientation of $O_0x_0y_0z_0$ adopted in various publications has been rather arbitrary, but for the reasons given at the end of Section 5.7.4 the orientation of O_0x_0 east, O_0y_0 south and O_0z_0 down (ESD axes) are chosen for this chapter.

Let the platform have velocities V_x, V_y, V_z along these axes. Then the simple geometry of Fig. 5.13 shows that

$$\dot{\lambda} = -V_y/(R+h)$$
$$\dot{\Lambda} = [(V_x/(R+h)]\sec\lambda \qquad (5.77)$$
$$\dot{h} = -V_z$$

where λ is latitude, Λ is longitude and h is the height of O_0 above the surface of the earth.

Note that V_x, V_y, V_z are velocities with respect to a tangent plane rotating with the earth, and at a distance $(R+h)$ from the centre of the earth. If $V_x = V_y = V_z = 0$, the platform would remain at a height h above a particular geographical point on the earth's surface.

Equations (5.77) represent the transformation of V_x, V_y, V_z into 'conventional' navigation co-ordinates. For special purposes other transformations may be used—for example grid co-ordinates of a given map projection.

5.7.2 Platform gimbal geometry and precession rates

Let the aircraft be at rest on the surface of the earth, and $Oxyz$ lined up with $O_0x_0y_0z_0$ (the ESD axes). Then a brief look at the definitions of Euler angles given in Section 5.2.1 will convince the reader that during the subsequent flight the gimbal system will measure Φ, Θ and Ψ directly, if the outer pivot is aligned along O_0x_0, the intermediate pivot is aligned along O_0y_0 and the inner pivot is aligned along O_0z_0. Now the purpose of the gimbal system is to keep the platform isolated from the angular manoeuvres performed by the

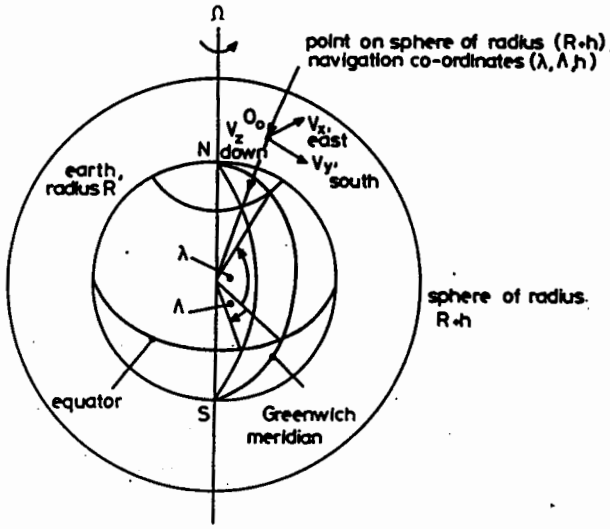

Fig. 5.13 *Spherical earth navigation geometry*

aircraft. The necessary pivot velocities are given by eqn. (5.2). It can be seen that $\dot{\Phi}$ and $\dot{\Psi} \to \infty$ as $\Theta \to \pm 90°$. This situation is known as *gimbal lock*, and reflects the mechanical situation of two pivots (outer and inner) becoming aligned onto a common axis, and one degree of kinematic freedom being consequently lost. To overcome this a fourth redundant gimbal is often slaved to the motion of the other three so that freedom is maintained at all orientations.

Let the rotation rates necessary to keep the platform aligned with ESD axes be ω_x, ω_y and ω_z about O_0x_0, O_0y_0 and O_0z_0, respectively. These are called the *precession rates* of the platform. Each is made up of two components: (i) compensation for the rotation of the earth and (ii) compensation for the movement of the platform over the earth [a more complicated version of eqn. (5.71)]. Simple geometry applied to Fig. (5.13) gives

$$\omega_x = 0 + V_y/(R+h)$$
$$\omega_y = -\Omega\cos\lambda - V_x/(R+h) \qquad (5.78)$$
$$\omega_z = -\Omega\sin\lambda - [V_x/(R+h)]\tan\lambda$$

The first term in each equation compensates for the earth's rotation, and the second for the platform's geographical velocity.

The angular velocities at the pivots to give the required values of ω_x, ω_y and ω_z can be obtained from a transformation similar (but not identical) to eqn. (5.2). A fourth gimbal makes the equations more complicated and introduces a constraint relationship. It can be seen that the value of ω_z may be large at high latitudes. For transpolar navigation Ox_0 and Oy_0 are allowed to wander from east and south, their orientation being stored in a computer. This complicates the navigation equations by an extra angular resolution, and will not be considered in this chapter.

Calculation of platform velocity. The velocities V_x, V_y and V_z are basically deduced from the three accelerometer outputs a_x, a_y and a_z. Equation (5.7) of Section 5.3.2 holds, but the variables are now in terms of platform axes, not aircraft axes. Thus

$$\begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} \to \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \qquad (5.79)$$

$$\begin{bmatrix} P \\ Q \\ R \end{bmatrix} \to \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \qquad (5.80)$$

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} \to \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} + \begin{bmatrix} V_{xe} \\ V_{ye} \\ V_{ze} \end{bmatrix} \qquad (5.81)$$

where V_{xe}, V_{ye} and V_{ze} are due to the earth's movement. For aircraft navigation the rotation of the earth around the sun is usually ignored, so

$$\begin{bmatrix} V_{xe} \\ V_{ye} \\ V_{ze} \end{bmatrix} = \begin{bmatrix} \Omega(R+h)\cos\lambda \\ 0 \\ 0 \end{bmatrix} \tag{5.82}$$

Differentiating (5.81) and (5.82) and substituting (5.77) gives

$$\begin{bmatrix} \dot{U} \\ \dot{V} \\ \dot{W} \end{bmatrix} \rightarrow \begin{bmatrix} \dot{V}_x + V_y\Omega\sin\lambda - V_z\Omega\cos\lambda \\ \dot{V}_y \\ \dot{V}_z \end{bmatrix} \tag{5.83}$$

Let the components of mass attraction along O_0x_0, O_0y_0 and O_0z_0 be g_x, g_y and g_z, respectively. Equation (5.7) may now be written

$$\begin{aligned} \dot{V}_x &= -2V_y\Omega\sin\lambda + 2V_z\Omega\cos\lambda + \frac{1}{R+h} \\ &\qquad \times (V_zV_x - V_yV_x\tan\lambda) + [g_x] + a_x \\ \dot{V}_y &= -2V_x\Omega\sin\lambda + \frac{1}{R+h}(V_x^2\tan\lambda + V_yV_z) \\ &\qquad + [\Omega^2(R+h)\cos\lambda\sin\lambda + g_y] + a_y \\ \dot{V}_z &= -2V_x\Omega\cos\lambda - \frac{1}{R+h}(V_x^2 + V_y^2) \\ &\qquad + [-\Omega^2(R+h)\cos^2\lambda + g_z] + a_z \end{aligned} \tag{5.84}$$

The terms in square brackets are due to the mass attraction and rotation of the earth, and depend upon λ, Λ and h. Their values are stored in the computer, and the raw acceleration signal compensated to give

$$\begin{aligned} a'_x &= g_x + a_x \\ a'_y &= \Omega^2(R+h)\cos\lambda\sin\lambda + g_y + a_y \\ a'_z &= -\Omega^2(R+h)\cos^2\lambda + g_z + a_z \end{aligned} \tag{5.85}$$

A block diagram for the solution of the navigation equations is shown in Fig. (5.14).

5.7.3 Accuracy of the inertial navigation unit

To assess the accuracy of the INU, a perturbation model of eqns. (5.77), (5.78), (5.84) and (5.85) are formed. For example the first equation of (5.78) becomes

$$\Delta\dot{\omega}_x = \frac{\Delta\dot{V}_y}{R+h} - \frac{V_y\,\Delta h}{(R+h)^2} \tag{5.86}$$

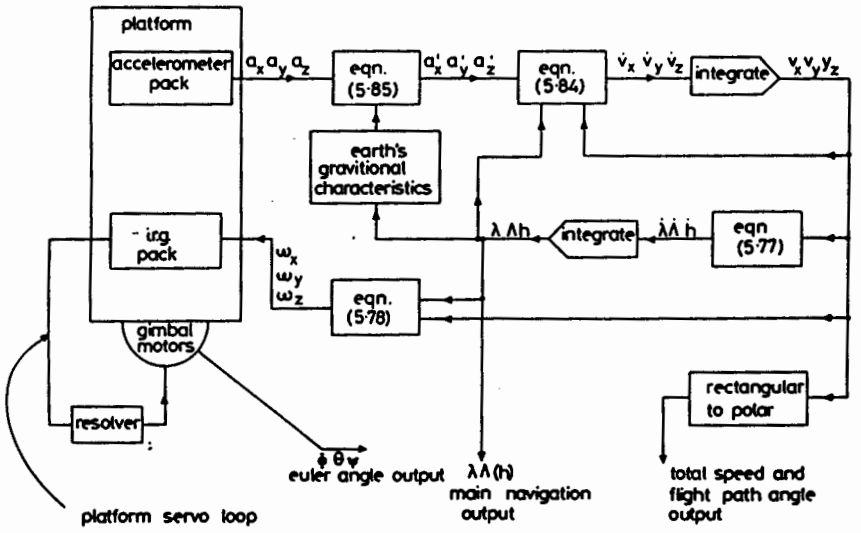

Fig. 5.14 *Block Diagram of Inertial Navigation Unit*

To these equations are added all the identifiable error sources. The resulting model then predicts the effect that each error has on navigational accuracy. Clearly the equations are rather complicated, although straightforward to produce, and there is little point in quoting them here. However, to illustrate the incredible degree of precision required for satisfactory performance, consider the effect of a drift error ω_{xe} in the O_0x_0-axis IRG. Let the INU be geographically stationary ($V_x = V_y = V_z = 0$) and at sea level ($h = 0$).

Assume that the drift error causes a velocity error in the O_0y_0 direction of ΔV_y. Then from eqn. (5.86)

$$\Delta\omega_x = \omega_{xe} + \frac{\Delta V_y}{R}$$

This false rotation rate causes a misalignment error $\Delta\beta$ to develop between the tangent plane and O_0y_0

$$\Delta\dot{\beta} = \Delta\omega_x$$

A component of the gravitational field is thus sensed by the O_0y_0 accelerometer, so that a'_y is not zero, but equal to $-g\,\Delta\beta$. Linearization of the second equation of (5.84) is therefore $\Delta\dot{V}_y = -g\,\Delta\beta$.

The error state equation is therefore

$$\begin{bmatrix} \Delta\dot{\beta} \\ \Delta\dot{V}_y \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{R} \\ -g & 0 \end{bmatrix}\begin{bmatrix} \Delta\beta \\ \Delta V_y \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix}\omega_{xe} \tag{5.87}$$

Solution of this for a constant value of drift rate $\omega_{xe} = \Omega_{xe}$ gives

$$\Delta V_y = R\left(\cos\sqrt{\frac{g}{R}}t - 1\right)\Omega_{xe}$$

This is an oscillatory error at the Schuler period with average value

$$\Delta V_y = -R\Omega_{xe}$$

This equation shows that a navigational accuracy of 1 nautical mile per hour, implies an IRG drift rate of one minute of arc per hour: 0·017°/h or one revolution in two-and-a-half years. In fact, to allow for other INU errors, the IRG's are considerably better than this, and it is remarkable that instruments with this level of performance can be made at all, let alone sufficiently light and robust for the aircraft environment.

5.7.4 *Strapdown inertial navigation units*

Over the last decade there has been increasing research into INU's which replace the mechanical gimbal system of the platform with an appropriate computer algorithm. The gyroscopes and accelerometers are now fixed directly to the airframe, and the navigator is called a *strapdown* system. A major difficulty has been the design of a gyroscope with sufficient dynamic range. A combat aircraft may have roll rates in excess of 300°/s, and the acceptable drift rate was shown in the last section to be about 0·01°/h. The dynamic range requirement is therefore a staggering value of 10^8. Acceptable gyroscopes in several forms (ring laser, tuned rotor etc.) are now available, and give rise to interesting mathematical modelling problems in their own right. However, it is the computer equivalent of the mechanical gimbal system that will be discussed in this section.

Basically, it is the direction cosine matrix (DCM) of eqn. (5.1) which distinguishes the aircraft axis set ($Oxyz$) from the tangent plane set ($O_0x_0y_0z_0$). However, as formulated in (5.1) it is a function of the Euler angles Φ, Θ and Ψ. It has already been mentioned that these suffer from the mathematical equivalent or gimbal lock [remarks after eqn. (5.2)]. A further disadvantage is the computational complexity of generating trignometrical functions.

A popular alternative is to describe attitude by a four-parameter system, sometimes called Euler symmetrical parameters, Rodrigues parameters, Caley-Klein parameters or quaternion representation.[9] Let the four parameters be e_0, e_1, e_2 and e_3. Then the equivalent of eqn. (5.2) is

$$\begin{bmatrix}\dot{e}_0\\ \dot{e}_1\\ \dot{e}_2\\ \dot{e}_3\end{bmatrix} = \frac{1}{2}\begin{bmatrix}-e_1 & -e_2 & -e_3\\ e_0 & -e_3 & e_2\\ e_3 & e_0 & -e_1\\ -e_2 & e_1 & e_0\end{bmatrix}\begin{bmatrix}P\\ Q\\ R\end{bmatrix} \qquad (5.88)$$

Since the three Euler angles are replaced by four parameters, there is obviously a redundancy equation. This turns out to be

$$e_0^2 + e_1^2 + e_2^2 + e_3^2 = 1 \qquad (5.89)$$

The DCM in terms of the parameters is

$$\begin{bmatrix}e_0^2+e_1^2-e_2^2-e_3^2 & 2(e_1e_2-e_0e_3) & 2(e_0e_2+e_1e_3)\\ 2(e_0e_3+e_1e_2) & e_0^2-e_1^2+e_2^2-e_3^2 & 2(e_1e_3-e_0e_1)\\ 2(e_1e_3-e_0e_2) & 2(e_0e_1+e_2e_3) & e_0^2-e_1^2-e_2^2+e_3^2\end{bmatrix} \qquad (5.90)$$

and can be used directly in eqn. (5.1). Apart from representing attitude without any singularities, the parameters are particularly suited to digital computation. Their value always lies in the interval $[-1\ 1]$, so fixed point arithmetic can be used, and the constraint eqn. (5.89) helps (5.88) to be integrated in an accurate and stable way.

The main disadvantage is that they do not enable attitude to be visualized by the human observer, but the Euler angles may be deduced by comparing the terms in (5.90) with those in (5.1). Thus from element (31)

$$\sin\Theta = -2(e_1e_3 - e_0e_2)$$

Dividing element (21) by element (11) gives

$$\tan\Psi = 2(e_0e_3 + e_1e_2)/(e_0^2 + e_1^2 - e_2^2 - e_3^2)$$

Dividing element (32) by element (33) gives

$$\tan\Phi = 2(e_0e_1 + e_2e_3)/(e_0^2 - e_1^2 - e_2^2 + e_3^2)$$

The equations describing the behaviour of a strapdown INU explicitly involve the description of the aircraft motion because the sensors are aligned along the aircraft axes. Obviously both models should be in the same mathematical language, which is the additional justification mentioned at the beginning of this section, for writing the IN equations in a matrix form rather than vector-mechanics form.

The earth reference axes stored in the computer should also align with the aircraft axes at some simple attitude. This is why O_0z_0 down has been chosen. The choice of O_0x_0 pointing east was made because the aircraft moves east at $\Omega R\cos\lambda$ when it is geographically stationary.

However, there is some argument for O_0x_0 to point north because the azimuth angle then lines up with compass card conventions.

5.7.5 *Mixed navigation systems*

As a final example of mathematical modelling in aerospace systems, the problem of mixing information generated in the INU with data from other sources will be considered. This usually arises because radio navigation aids such as Omega (a very low frequency phase comparison system) or doppler velocity measuring systems provide essentially the same information, but corrupted

with different errors from those characteristic of the INU. By mixing the various sources of information in a Kalman filter a considerably better performance can be obtained than from any one navigator. The INU itself often contains a Kalman filter anyway to identify and compensate for the various error terms, and to align the platform after switch-on.

Most Kalman filter implementations are discrete-time and based on error states [such as eqn. (5.87)]. However for the sake of a simple illustration, the problem of mixing the height obtained from the INU, h_I and the height signal from the air data computer (barometric) will be considered.

From eqn. (5.77), $\dot{h} = -V_z$, and from eqn. (5.84) and (5.85)

$$\dot{V}_z = -2V_x\Omega\cos\lambda - \frac{1}{R+h}(V_x^2 + V_y^2) + a_z' \tag{5.91}$$

Assume that errors in the INU values of V_x, V_y and h plus accelerometer and gravitational compensation errors combine to form a random disturbance α.

Thus (5.91) can be written

$$\ddot{h} = -\dot{V}_{zI} + \alpha \tag{5.92}$$

where $\dot{V}_{zI}$ is the value of the right-hand side of (5.91) as calculated in the INU.

Let $h = x_1$ and $\dot{h} = x_2$. Thus

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} + \begin{bmatrix}0\\ -1\end{bmatrix}\dot{V}_{zI} + \begin{bmatrix}0\\ 1\end{bmatrix}\alpha \tag{5.93}$$

The air data computer produces a height estimate

$$h_a = h + \beta \tag{5.94}$$

where β represents the error in barometric data. Let $h_a = y$, and (5.94) becomes

$$y = [1 \quad 0]\begin{bmatrix}x_1\\ x_2\end{bmatrix} + \beta \tag{5.95}$$

Equations (5.93) and (5.95) can be used as the state and output equations of a Kalman filter problem. For simplicity assume that α and β are white noise disturbances (unlikely to be valid assumptions in practice), and let

$$\epsilon\alpha^2 = q\delta \qquad \text{and} \qquad \epsilon\beta^2 = r\delta$$

The standard theory[4] shows that the steady-state Kalman filter is

$$\dot{\hat{x}} = A\hat{x} + Bu + P(y - C\hat{x} - Du) \tag{5.96}$$

where ^ means 'estimate of'.

The *filter gain matrix* P is given by $P = GC^TR^{-1}$, and G, the *error covariance matrix* defined by

$$G = \epsilon(x - \hat{x})(x - \hat{x})^T$$

can be found from the solution of the quadratic matrix equation (sometimes called the algebraic Riccati equation)

$$AG + GA^T - GC^TR^{-1}CG + Q = 0$$

In this case

$$R = r, \quad Q = \begin{bmatrix}0 & 0\\ 0 & q\end{bmatrix} \qquad A = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix} \qquad \text{and} \qquad C = [1 \quad 0]$$

Hence

$$G = \begin{bmatrix}\sqrt{2}q^{1/4}r^{3/4} & \sqrt{qr}\\ \sqrt{qr} & \sqrt{2}q^{3/4}r^{1/4}\end{bmatrix}$$

and

$$P = \begin{bmatrix}p_1\\ p_2\end{bmatrix} = \begin{bmatrix}\sqrt{2}(q/r)^{1/4}\\ (q/r)^{1/2}\end{bmatrix} \tag{5.97}$$

Equation (5.96) written in full is

$$\begin{bmatrix}\dot{\hat{h}}\\ \ddot{\hat{h}}\end{bmatrix} = \begin{bmatrix}\dot{\hat{h}}\\ 0\end{bmatrix} + \begin{bmatrix}0\\ -\dot{V}_{zI}\end{bmatrix} + \begin{bmatrix}p_1(h_a - \hat{h})\\ p_2(h_a - \hat{h})\end{bmatrix} \tag{5.98}$$

The block diagram is shown in Fig. (5.15).

The first element of the transfer function matrix of equation (5.98) is

$$\bar{\hat{h}} = \frac{(-\dot{V}_{zI}) + (p_1 s + p_2)\bar{h}_a}{s^2 + p_1 s + p_2} \tag{5.99}$$

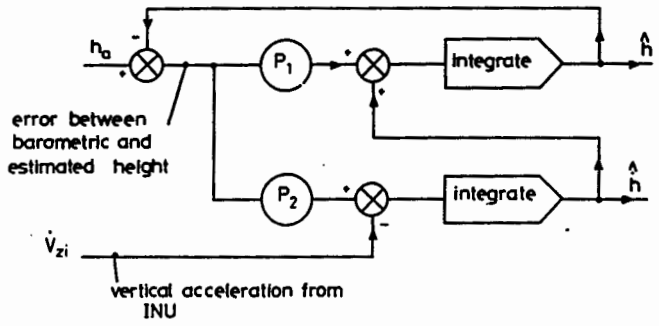

Fig. 5.15 *Block diagram of Kalman filter for h*

showing that the Kalman filter is acting as a so-called complementary filter with zero position and velocity-lag errors. As a matter of interest, substituting for p_1 and p_2 from eqn. (5.97) shows that the damping ratio of (5.99) is $1/\sqrt{2}$ for all disturbance strengths q and r.

5.8 References

1 McRUER, D., ASHKENAS, I., and GRAHAM, D.: 'Aircraft Dynamics and Automatic Control' (Princeton University Press, 1973)
2 BABISTER, A. W.: 'Aircraft Stability and Control' (Pergamon Press, 1961)
3 BEKEY, G. A., and KARPLUS, W. J.: 'Hybrid Computation' (John Wiley, 1968)
4 KWAKERNAAK, H., and SIVAN, R.: 'Linear Optimal Control Systems' (Wiley-Interscience, 1972)
5 FLYING QUALITIES OF PILOTED AIRPLANES: U.S. Government Printing Office MIL-F-8785B(ASG), 1969
6 TOBIE, H. N., ELLIOTT, E. M., and MALCOM, L. G.: 'A New Longitudinal Handling Quality Criterion', Proc. Nat. Aerosp. Elect. Conf., May 1966, Dayton, Ohio
7 McRUER, D. T. and KRENDEL, E. S.: 'Mathematical Models of Human Pilot Behaviour NATO Advisory Group for Aerospace Research and Development', AGARD-AG-188, January, 1974
8 ADAMS and HATCH: *Journal of Aircraft AIAA*, 1971, **8**, pp. 319–324
9 VANBRONKHORST, A.: 'Strapdown System Algorithms', NATO Advisory Group for Aerospace Research and Development AGARD-LS-95 pp. 3-1 to 3-22, May, 1978

Chapter 6

Marine systems

D. A. Linkens

6.1 Introduction

Marine vehicles are required to accomplish certain operating objectives in the environment of the sea. Within each vehicle there may be a number of subsystems, such as the power plant (propulsion and energy conversion), the control system (for ship steering), the cargo handling system and numerous military missions involving search-and-destroy systems. Each subsystem requires dynamic modelling and design to operate within its subenvironment 'optimally', together with satisfactory overall performance and survival of the integrated vehicle. In this chapter the modelling and control of one of the marine subsystems, being that of ship steering and manoeuvring, is considered.

Within the class of marine vehicles a number of different types can be distinguished. These include torpedoes, submarines and submersibles, military surface ships, commercial and fishing vessels, hydrofoil boats, hovercraft, towed oceanographic instruments and barges. Clearly, each of these categories have their own particular requirements and problems in dynamic modelling. For underwater vehicles, motion in three dimensions is involved giving similar considerations to those for the study of aircraft. For surface ships the motion can usually be simplified to that of one plane for horizontal manoeuvring. Thus, yaw response can be considered without coupling from heeling and pitching motion, unlike aircraft dynamics. This simplification of dynamics is accompanied by problems, however, particularly for large commercial ships. Thus, for relatively slow-moving ships performing large-angle manoeuvres the dynamics become nonlinear, and linearized mathematical models and autopilot controller design become invalid. In this chapter we shall concentrate on surface ship manoeuvring, emphasising the nonlinear aspects.

Having concentrated on the surface ship steering subsystems a number of considerations still need to be made. The operating requirements and

1
An Outline of Inertial Navigation

Navigation's Beginnings

Pioneers returning from their journeys provided travel instructions for those who wished to repeat their journeys. They wrote descriptions of their routes and made charts or maps pointing out landmarks and hazards like rivers and mountains on land, or shoals and rocks at sea. Map makers devised a global coordinate system using a grid of latitude and longitude circles, by which the position of any place on earth could be defined.

Buoys and lighthouses provided route markers for sailors close to land, but once they ventured into the featureless seas, their pathfinding became much more difficult. Magnetic compasses were known as early as the 11th century, but were not very accurate. The 14th- century navigator used the stars to find latitude, and by the 18th century, technology had provided instruments like the sextant for more accurately measuring the positions of the celestial bodies.

Finding longitude was more difficult, for the positions of the celestial bodies depend on the rotational position of the earth relative to the stars, that is, the time of day. Longitude is measured from an arbitrarily chosen zero at Greenwich, near London, England; to determine longitude, navigators need to know the time where they are and the time at Greenwich (Greenwich Mean Time, GMT). They can find local time from observations of the sun, but finding GMT means observing the satellites of Jupiter, or lunar distances [1], which are uniquely positioned at any time. But these observations require complicated calculations to yield time, so it is much better to have a clock set to GMT on board. It took the invention of the spring escapement clock in the mid-18th century to make accurate timekeeping available at sea. How the 18th-century mariner would envy today's $10 quartz watch, accurate beyond all his wildest dreams! Nowadays, GMT is broadcast by radio, worldwide.

This process of following a path on a map through predetermined latitudes and longitudes is called *navigation*, and the process of pointing the vehicle to follow a chosen path is called *guidance*. Navigation at sea uses the process of *deduced reckoning* (called "dead" reckoning), illustrated in Figure 1.1. If a ship starts from a known latitude and longitude and travels in a known direction for a known time, its position on the previously charted sea is known and can be

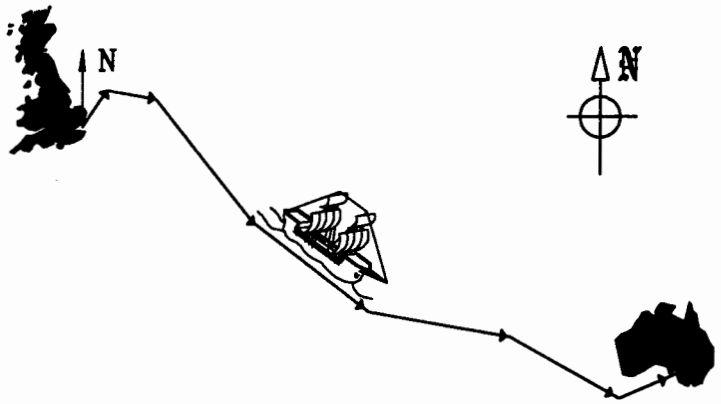

Figure 1.1. Inertial navigation is dead reckoning.

verified by sun (or other celestial body) sightings. Provided the ship's direction and speed are known accurately, the sailors will land in the correct harbor.

When aircraft began to fly over the sea, they got direction and speed from the magnetic compass and airspeed indicator. As radar evolved, the Doppler set could give more accurate ground speed. Dead reckoned courses tried to follow planned routes, corrected by celestial navigation, but airplanes fly so fast that the navigator was hard pressed to make celestial measurements quickly enough. More accurate, safer, navigation was possible when land-based radio beacons and systems like Loran and Omega became available.

By 1908, the magnetic compass had been replaced by the gyrocompass. Gyrocompasses are important marine navigation instruments, but they are not generally useful for aircraft, spacecraft and guided missiles. Because our focus is on the instruments for inertial navigators, we are not going to consider the gyrocompass in this book; they are well described in general navigation books such as Dutton's [2], by Wrigley in the *Encyclopedia Britannica* under "Gyrocompass," and by others [3,4].

Military aircraft prefer not to depend on radio beacons in time of war, for an enemy can jam them, and, conversely, aircraft can be detected by the enemy if they transmit Doppler radar speed measurement signals. They need a self-contained navigation system. If one wants to guide unmanned vehicles (such as a space probe or an intercontinental missile), one must also have a self-contained navigation system. Civil aircraft need to have reliable navigation at the lowest possible cost, and human navigators are expensive. The solution to all these needs is to automate navigation, which means that one must find a way to keep track of position and attitude continuously, in all weather conditions, in space, in the sky, and under the sea. Inertial navigation serves this purpose.

Inertial Navigation

Gyroscopes and accelerometers can provide the necessary signals for automatic navigation. Gyroscopes measure rotation, and accelerometers measure acceleration. Integrating the output from an accelerometer gives speed, and integrating speed gives distance traveled. The gyroscopes provide information on where the accelerations are directed, and therefore heading and distance are known, the essential ingredients for dead reckoning. As these instruments use the inertial properties of matter (or of light) for their operation, dead reckoning with gyros and accelerometers is called *inertial navigation.*

The first inertial navigators were used in the German V1 and V2 weapons in World War II. After the war, a group of German scientists, under Werner Von Braun, developed this technology at Redstone Arsenal, Huntsville, Alabama, for ICBMs and spacecraft, building inertial navigation systems for U.S. Army missiles.

Other U.S. groups developed IN systems, notably one under Draper at the Massachusetts Institute of Technology [5,6]. Their first aircraft inertial navigators were flown in 1949, followed in 1954 by the Navy Ship's Inertial Navigation System. The Draper Laboratory made inertial systems for Polaris submarines; one of them, the *Nautilus*, successfully navigated under the polar ice past the North Pole. In the 1960s the Apollo program took inertial guidance into space, and now inertial systems are being used in "smart" munitions. For more on the history of inertial navigation, see Wrigley's summary [7].

Maps and Reference Frames

Before we go anywhere, we need to orient ourselves in the space in which we will navigate, which we do with *maps* and *reference frames*. A map or chart is drawn to some scale so that the user can calculate distances between places. Over the distance covered by a town map we assume that the flat sheet represents a flat area of the earth, and although there may be hills and valleys in the town, the town is represented as if it were a model on a flat board. But for larger areas, the earth, being roughly a sphere, cannot be precisely represented by flat maps. Many different projections have been developed for making flat navigation maps [2], but we will leave those readers who are interested to pursue this field elsewhere.

Maps are oriented in a known direction, usually with North up the map. Navigation requires us to define frames of reference (or coordinate sets) so that we can orient ourselves in the mapped space, and in this instance "north" specifies the map's reference frame. There is a frame fixed in the "fixed stars," the *inertial frame*, that Newton postulated for his laws of inertia in 1687, and inertial navigation is navigation in this frame. This is a reference frame that is independent of the motion of the vehicle.

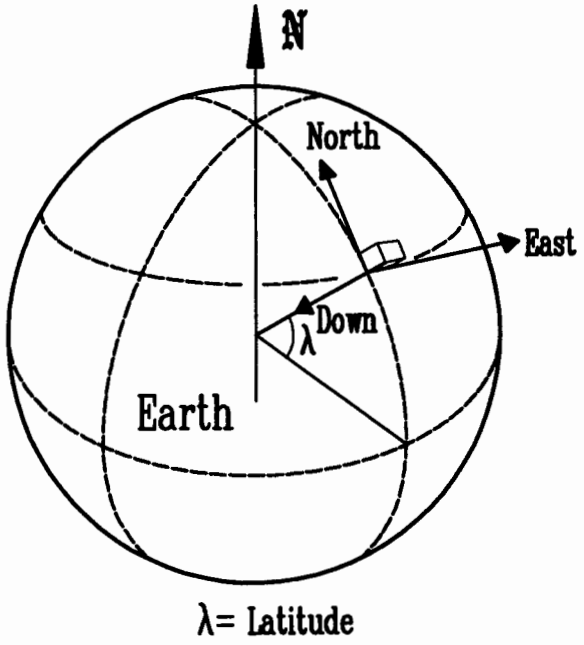

Figure 1.2. Earth-fixed axes.

On (or close to) the earth's surface it is more useful to work in a reference frame oriented to surrounding objects, a *local level* frame with North, East, and Vertically-Down as axes. In the simple (spherical homogeneous) earth local level frame, Figure 1.2, Down is a vector pointing at the center of the earth at the angle of latitude λ, North is the horizontal vector in the plane of Down and the earth's spin axis (True North), and East is the horizontal normal to that plane at the observer's site. The earth rotates once in 24 hours, so its rotation rate is 15 deg/h; the horizontal and vertical earth's rate components Ω_{eh} and Ω_{ev} in local level axes are

$$\Omega_{eh} = 15 \cos \lambda$$
$$\Omega_{ev} = 15 \sin \lambda \text{ deg/h}$$

where λ = latitude.

East is the direction in which there is no earth's rate, a fact that is used to align inertial navigation systems because it is much easier to find a null than the peak of a maximum. Gyroscopic north-finding systems find East!

Vehicles have their own axis set, Roll, Pitch, and Yaw, shown in Figure 1.3, corresponding to the conventional use of the terms.

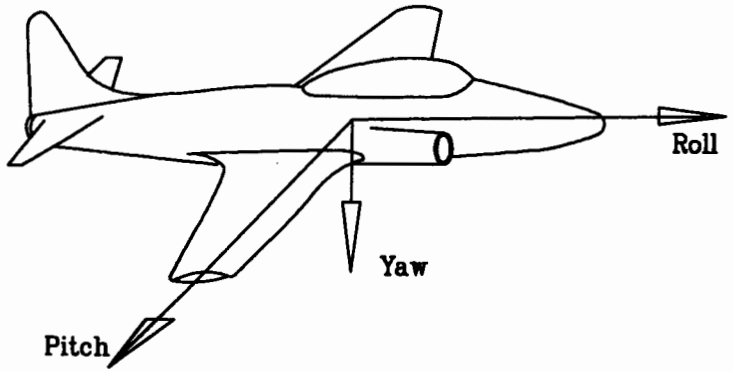

Figure 1.3. Vehicle axes set.

The Inertial Navigation Process

The inertial navigation system, INS (or unit, INU), is made from a navigation computer and a set of gyroscopes and accelerometers that measure in Newton's inertial axes, generally called *inertial sensors*. The group of inertial sensors is commonly called an *inertial measurement unit* (IMU) or an *inertial reference unit* (IRU). Once aligned to a set of reference axes (such as the "North, East, Down" set), the sensors provide distance measurements and the navigation computer carries out the continuous dead reckoning calculations.

The inertial sensors might be mounted in a set of gimbals so that they stay level and head in a fixed direction whatever the vehicle's motion. This construction is called a *navigation platform*. Alternatively, the instruments might be attached to the vehicle, in which case they measure its motion components in the vehicle axes set, and the system computes direction traveled in the reference axes by transforming the measurements from the vehicle axes to the reference axes. This is called a *strapdown* system, jargon for instruments "strapped down" to the vehicle.

To navigate inertially, we first measure the accelerations in the directions of the navigation axes, and if our instruments are not perfect, we might compensate their readings by removing biases or scale factor errors (defined in Chapter 2), perhaps known as a function of the system's measured temperature.

Second, to find the vehicle's vertical acceleration, we subtract gravity from the "Down" accelerometer output, perhaps using a gravity model to allow for the variation of gravity with latitude and longitude. The earth's rotation causes a centripetal acceleration that is greatest (about 3 mg) at the equator and zero at the poles. In addition, the earth is not a sphere; rotation has flattened it, causing

a mass concentration at the equator, so that the value of g varies with latitude. Precise navigators, like those in nuclear submarines, carry gravity gradiometers to correct for gravity vector errors; because of this unevenness in mass distribution, the vector direction is not necessarily along an earth's radius [8,9]. In space we can use Newton's Law of Gravitation to compute the local gravity, but once on a foreign planet, we lose accuracy until we can map its gravity fields.

Third, we integrate the accelerations over a known time, once to get the velocity, twice to get the distance traveled. As integration is the process of summing the outputs at frequent, known intervals, we must know the time interval accurately for it enters as a squared term in the distance computation.

Fourth, we measure the rotation rates either from gimbal motions in a stabilized platform, or directly with gyroscopes in a strapdown system. We then compensate for gyro bias and possibly scale factor errors, and determine a new heading.

Fifth, we compensate for earth's rotation if we are in local level axes, for otherwise the platform would be space stabilized and would seem to tilt in the vehicle axes set.

Finally, the combined distance and heading data give us an updated dead reckoned position to display. Then we go back to the beginning and do it all over again, until the end of the journey.

Inertial Platforms

An inertial platform uses gyros to maintain the accelerometers in a fixed attitude, i.e., the gimbaled platform serves to define the directions for the measurements of acceleration. A single-axis platform, Figure 1.4, consists of a gyro mounted so that its sensing axis or *input axis* (IA) is along the axis of the platform, which is set into the vehicle in bearings. Electrical power for the gyro comes in, and the gyro's output goes out, through slip rings. The platform can be driven around its axis by an electric torque motor. The gyro output, which indicates an unwanted platform rotation, drives the torque motor through a servo amplifier; it basically provides the torque to overcome the friction in the slip rings and bearings.

If the platform had no gyro and the vehicle rotated around the platform axis, the platform's inertia would tend to keep the platform aligned in space. But the platform would slowly accelerate due to the torque transmitted through the bearings and slip rings, and the inertial reference would be lost. By adding the gyro, we provide a means of sensing any platform rotation with respect to inertial space, down to the limit of the gyro's resolution. The gyro does not measure anything in the inertial platform; it is only used to maintain a fixed position (null operation), and so does not need to be able to measure large rotation rates.

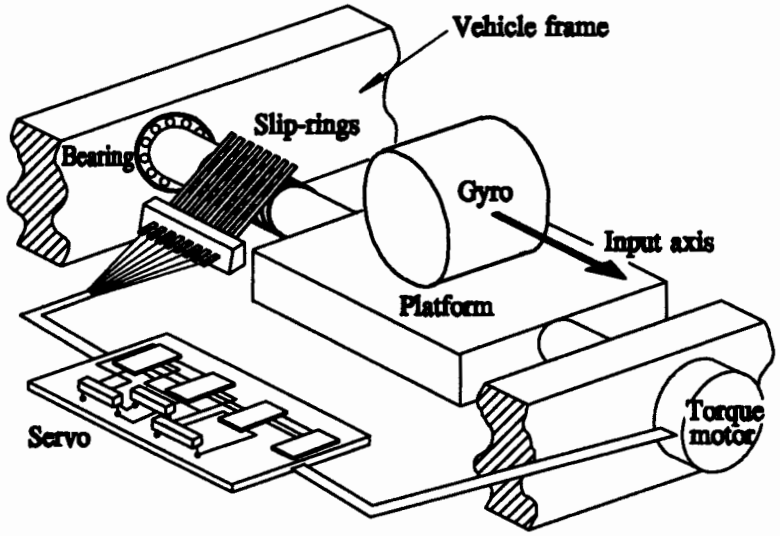

Figure 1.4. A single-axis platform.

For navigation in three dimensions we must expand our platform to three axes by adding gimbals, Figure 1.5. Here we have shown gyros with two sensing axes each (*two degree of freedom* gyros) and also schematically show a group of three accelerometers. We have also provided a *pickoff* on each gimbal axis, a device that measures the angle between two gimbals; the torquers, slip rings, and three servo loops are there but are not shown. If the gimbals are aligned with the vehicle axes before setting out, then the gimbals, being gyro stabilized, define the attitude and heading of the vehicle during the journey.

The gyros alone will try to maintain the platform aligned in inertial space. If the platform is operating in local level coordinates, the navigation computer must keep the platform horizontal. It does this by sending command signals to the gyros which otherwise would fight the gimbal motion. The roll and pitch accelerometers can be used to level the platform if we know that it is not accelerating. The relationship between the sensors, the gimbals and the navigation computer is shown schematically in Figure 1.6.

Heading and Attitude Reference Systems

Rather than use an expensive inertial navigator, many aircraft use a simplified version that indicates the direction of flight (the heading) and the attitude in roll and pitch. This Heading and Attitude Reference System (HARS) is a platform with three stabilized axes and two horizontal accelerometers. The accelerometers

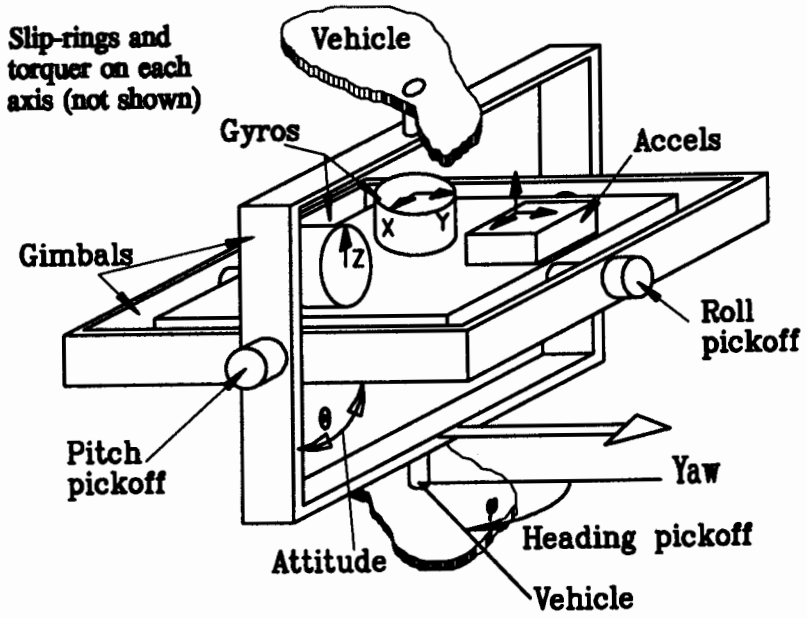

Figure 1.5. A three-axis platform.

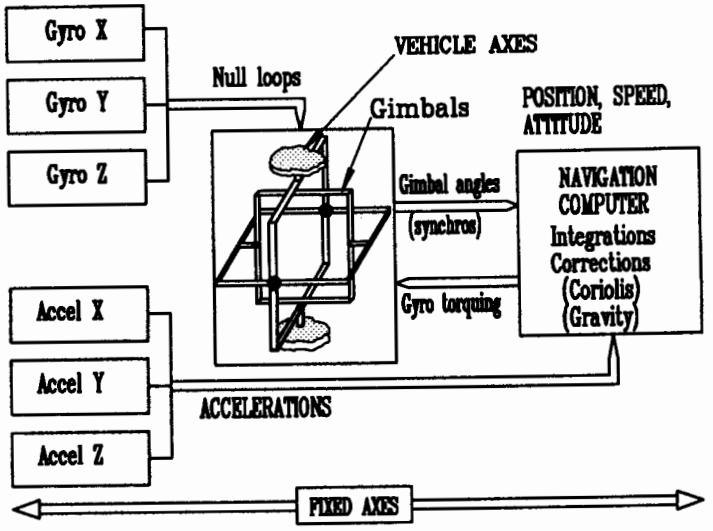

Figure 1.6. The basic platform system.

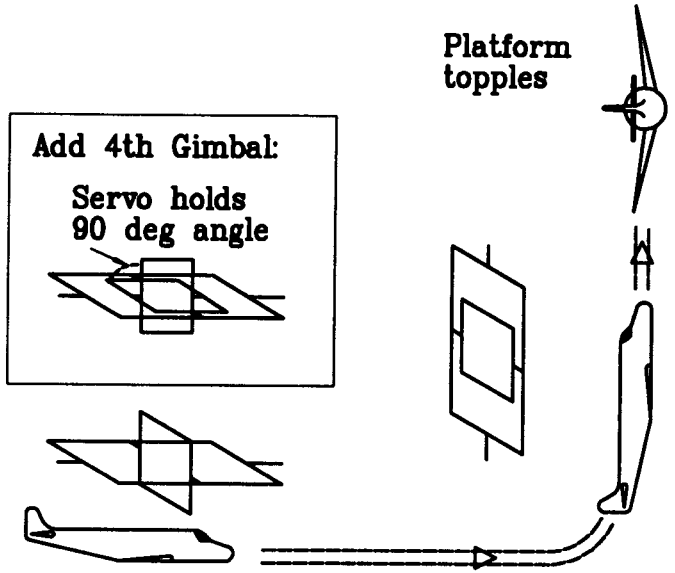

Figure 1.7. Gimbal lock.

provide a vertical reference to set the initial level before takeoff, whereas a flux valve (a kind of magnetic compass) can be used to set the magnetic heading. This heading can be corrected to True North using the local value of the magnetic variation. Alternatively, the heading can be set by gyrocompassing (page 15). Once under way, the HARS provides flight information under all visibility conditions.

Schuler Tuning

Imagine that we have a pendulum hanging in the vehicle that we are navigating, intended to provide a vertical reference. But as we accelerate, the pendulum tilts, giving a false vertical indication. Schuler showed that this would not occur with a pendulum of 84 minutes period, and we can make a compact compound pendulum with this period using an ordinary accelerometer and a servo loop with the correct characteristics. Correcting an inertial system so that it does not tilt when accelerated is known as *Schuler tuning* it.

Gimbal Lock

Three axis platforms have a limitation that can cause them to "lock up" during aerobatic maneuvers. This *gimbal lock* can be caused by a set of aircraft movements that cause two of the gimbal axes to become aligned, followed by a

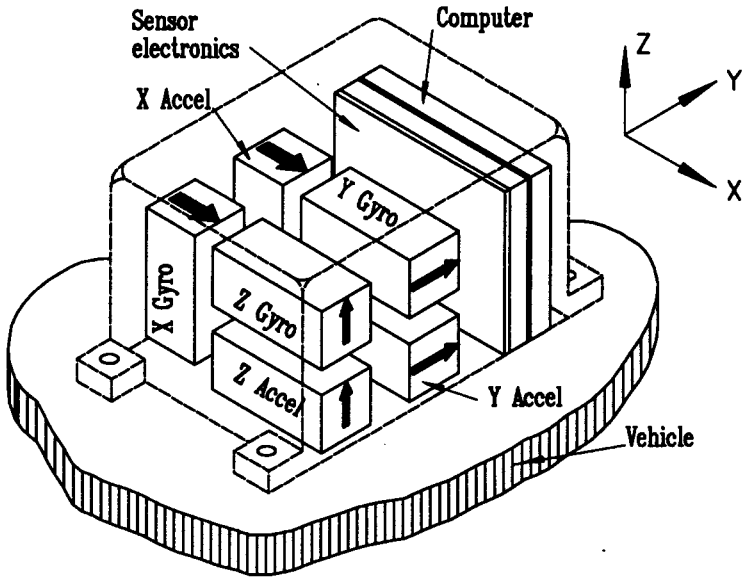

Figure 1.8. The strapdown system

rotation in the plane of the aligned axes. How this happens is much easier to see than to read, so Figure 1.7 shows a sequence that tumbles the platform. A plane flying straight and level climbs vertically. Now the roll gimbal is aligned with the pitch gimbal, and a turn over the wing, around the aircraft yaw axis, forces the pitch gimbal over and makes the system lose alignment. A fourth gimbal [10], a duplicate roll gimbal, can be added to overcome this limitation, although it adds considerably to the system's cost and size.

For further reading, books by Markey and Hovorka [11] and Savant [12] describe inertial systems in more detail.

Strapdown Systems

The platform system's complex mechanical construction goes against the trend in machine design. Since the 1960s, technical advances in electronics and optics have reduced their size and cost, while there has been little progress in reducing the cost of mechanisms, apart from robotic fabrication. Generally, lower costs and higher reliability have come from replacing mechanisms with electronics.

Navigation systems have benefitted from this technological culture; the strapdown system is the outcome. The strapdown system replaces gimbals with a computer that simulates their presence electronically. In the strapdown system the gyroscopes and accelerometers are rigidly mounted to the vehicle structure so that they move with the vehicle, as shown in Figure 1.8. Now, unlike the plat-

form gyros, the strapdown gyros must measure the angles turned, up to the maximum rotation rate expected. Airplanes can experience short-term rates up to 400 deg/s, so a gyro with 0.01 deg/h performance has a dynamic range of 10^8 - two orders above the platform gyro.

As the vehicle travels, the gyros measure the yaw, pitch, and roll angles turned in a short time (say, 0.01 s), and pass them to a computer that uses them to resolve the accelerometer outputs into the navigation axes set. Simplifying to two dimensions, we can easily express the coordinate transformation between the navigation axis set (x_{nav}, y_{nav}) and the body axis set (x_{body}, y_{body}), at a time when the z-axis gyro has measured that there is an angle θ between them. The accelerometers fixed in the body axes record accelerations a_x, a_y. Converting to navigation axes we get

$$(a_x)_{nav} = a_x \cos\theta - a_y \sin\theta \qquad (1.1)$$

$$(a_y)_{nav} = a_x \sin\theta + a_y \cos\theta$$

When adding the third (z) axis we must use more complicated transformations as the rotations are non-commutative. Typical systems use direction cosines, and the transformation between the inertial set [X_i, Y_i, Z_i] and the body axes set [X_b, Y_b, Z_b] is expressed as follows:

$$r^i = \begin{vmatrix} r_x \\ r_y \\ r_z \end{vmatrix} = C_b^i r^b$$

$$C_b^i = \begin{vmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{vmatrix}$$

where c_{ij} are the direction cosines between the j^{th} axis in the inertial frame and the k^{th} axis in the body frame; Britting [13] describes the necessary computations. The sensors and computer in a strapdown system are arranged as shown schematically in Figure 1.9.

System Alignment

The first step in inertial navigation is aligning the platform to the navigation axes or, equivalently, initializing the axes in the strapdown system's computer. We can level the platform or measure the attitude of the vehicle for a strapdown system, using accelerometers. We can find North using a magnetic compass (which is not very accurate), by astronomical sightings (which means we must be able to see stars), or by *gyrocompassing*. In some cases, such as that of missiles suspended under the wing of a fighter plane, we can align the missiles' strapdown systems

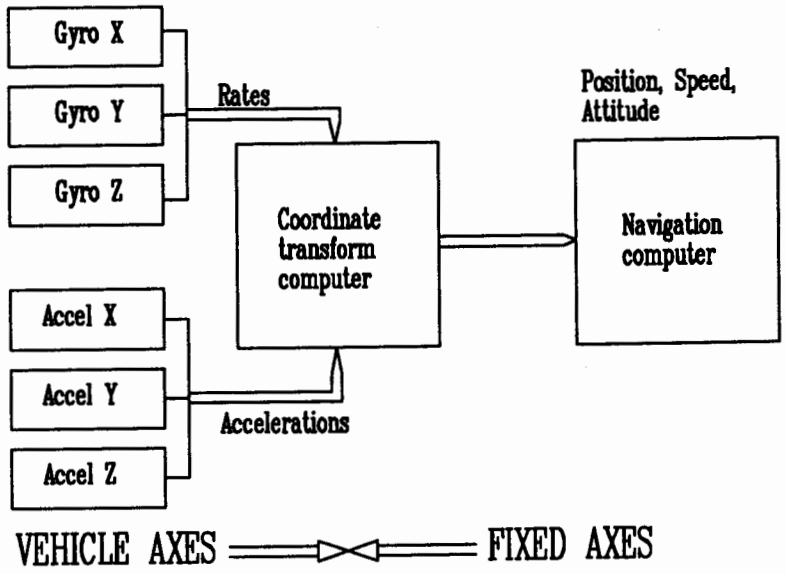

Figure 1.9. The basic strapdown system.

by referring their outputs to those of a master navigator in the fighter, called *transfer alignment*.

Gyrocompassing

To gyrocompass with a platform, a gyro with its input axis horizontal is connected by a servo to a platform torquer that will rotate the gyro about the vertical (yaw) axis. The servo is designed to drive the sensed rotation to zero, so the gimbal will swing around until the gyro's axis points East. But if the gyro has a bias - an output signal when there is no rotation input - it will point a little off East, enough so that the earth's rotation rate component cancels the bias. Then the platform is slewed 180°, and again servoed to null. This time the bias has not changed sign but the earth's rotation component has, so, by combining the East and West readings of platform heading, both North and the gyro bias can be deduced, a big advantage for the platform system. (In passing, note that some systems rotate continuously, the "rotating azimuth" mechanization, so that instrument biases are modulated at the rotation rate and filtered out. This improves performance for given instrument errors and alignment accuracy.)

A strapdown system cannot rotate the gyros except by swinging the entire vehicle, usually impossible to do over the half-circle necessary for East-West gyrocompassing. All one can do is measure the three components of earth's rotation with the gyros and compute where North must be, assuming that the gyro bias has not changed since it was last measured. To find North to an accu-

directions (usually orthogonal) in the vehicle, they can be placed so that they best use the space available.
2. Ruggedness. The simpler structure better withstands shock and vibration, and, being lighter, is easier to shock mount than a platform.
3. Reliability. There are no gimbal magnetics, no slip rings, and no bearings. The electronics that replace them are inherently more reliable.

Disadvantages:

1. Alignment. Strapdown systems are difficult to align because they cannot be easily moved. Transfer alignment is suitable for tactical systems.
2. Sensor calibration. Again, the immobility means that the sensors cannot be calibrated in the system. Therefore, they must be stable, a burden on the sensor design. Strapdown systems rely on sensor models, using real-time compensation of inertial errors and thermal effects.
3. Motion induced errors. The body motions induce unique sensor errors (torquer errors, anisoinertia, output axis angular acceleration), which can be compensated to some degree.
4. Accelerometer errors. Bias errors accumulate, and strapdown accelerometers may be subjected to components of gravity as the vehicle rolls and pitches, reducing the accuracy of the vehicle acceleration measurement and exciting cross-axis errors.
5. The strapdown computer. Not needed in the platform system, the computer must be fast enough to do all the strapdown calculations in a few milliseconds. In a typical tactical system, a bandwidth of 100 Hz demands that sensor compensation and coordinate transformation must be done in less than 0.01 s. This requires well-crafted program code.

The particular benefits of strapdown systems are well illustrated in the following story. One of the first strapdown systems fielded was used in the Apollo Lunar Module (LM) moon lander, as a backup to the primary platform system. The Apollo 13 spaceship had just left the earth when a fuel cell in the main ship (the service module) exploded, drastically reducing the amount of electrical power available. The ship was committed to a trajectory to the moon; it could not change direction much once in space. Mission Control decided to let the ship pass around the moon and come back to earth without landing on the moon. Because of the shortage of electrical power, the astronauts switched off the command module and used the LM. Everything possible was switched off, so that the three astronauts could stay warm enough; even so, the cabin got unpleasantly cold, down to 4°C (40°F). The platform navigator was turned off, and the crew relied entirely on the lower-power strapdown system, operating below its design temperature, for navigation around the moon and back to earth. They then left the LM and used the command module's navigator through re-entry to a safe landing [14,15].

Star Trackers

In applications where a missile may not have time for an accurate ground alignment before firing, or in cases where flight time is very long, the guidance system may be provided with a star tracker to update its alignment. This is a type of telescope with an optical detector at its focus, which can be precisely pointed at a spot in the heavens where a star is known to be. Once flying high enough that the atmospheric turbulence, smoke, and haze do not obscure the view, the tracker locks onto the pre-selected star and uses the star's known position to correct for misalignment in the inertial navigator.

As star trackers are expensive and require a window in the vehicle, they tend to be used only in long range missiles, reconnaissance planes, and long range bombers. They also need a clear view of the sky. Radio systems are all-weather navigation aids. Omega gives 2-4-nautical mile position accuracy but no velocity information, whereas satellite systems provide 15-m position accuracy and velocity data.

Satellite systems have the advantage over ground-based transmitters that they are radiating almost vertically down, so that satellite signals are much less affected by hills and the curvature of the earth. The newest United States satellite system is the Navstar *Global Positioning System* (GPS) and is accompanied by the similar Russian GLONASS system.

The Global Positioning System

The Navstar Global Positioning System (GPS) is a space-based position and navigation system that can provide three-dimensional position to an accuracy a few meters, anywhere on or near the earth, to those with the proper receivers. Although it is primarily a military system and provides the highest accuracy only to those who know its secret codes, it provides a lower accuracy (15-30 m) signal to all users. The precise encrypted "P" code and the coarser clear/acquisition "C/A" signal are transmitted on separate frequencies.

When it is complete, GPS will consist of 18 satellites, 3 each in 6 orbital planes. They will orbit at a height of 20,200 km (10,900 nautical miles) with a 12-h period; the orbits are inclined at 55°, as illustrated in Figure 1.11. The satellites already in orbit are demonstrating that the system will easily meet its design objectives.

In operation, each satellite broadcasts both the P and C/A signals at precisely known times, timed by each satellite's on-board atomic clock. Each receiver has a clock, too, so it can measure the time at which it receives the signal from the satellite. As the speed at which the radio wave travels is known, the distance between the satellite and the receiver is found by multiplying that speed by the time difference between satellite transmission and user reception. Each satellite's

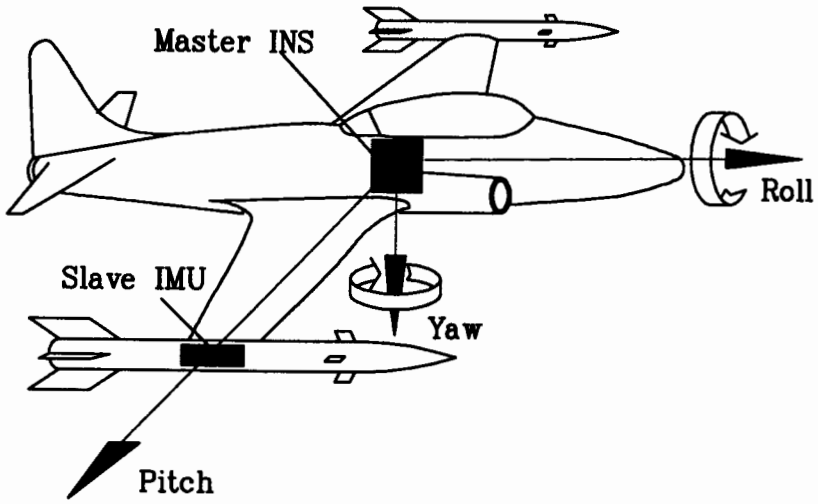

Figure 1.10. Transfer alignment.

racy of 1 mr, one must measure one thousandth of the horizontal earth's rate. At Boston, Massachusetts, for example, HER = 10 deg/h, so for 1 mr alignment accuracy we must know the gyro bias to 0.01 deg/h, a demanding requirement met only in expensive gyros.

Transfer Alignment

Another way to align a strapdown system is to transfer alignment from a master system. A fighter aircraft can initialize wing-carried missiles, for example, by carrying out maneuvers designed to allow the fighter's master navigator to send alignment information to the missile, as shown in Figure 1.10. If the fighter flies straight and level, the master and slave velocity outputs can be matched by adjusting the accelerometer bias, and controlled rolls or turns can calibrate the slave system's gyros.

Advantages and Disadvantages of Platform Systems

Advantages:

1. Simpler gyros. Because the sensor platform rotates only at the small rates needed to keep it level, the gyros need only a small dynamic range. A maximum rate of 3 deg/s would suffice for a gyro of 0.01 deg/h performance (e.g., for an aircraft navigator), a range of 10^6. Further, gyro torquer errors do not lead to attitude error. The lack of gyro rotations means that there are no anisoinertia and output axis angular acceleration errors to minimize in the design.
2. Higher accuracy. Because the accelerometer axes are always well defined, the platform navigator can be very accurate; the North and East accelerometers see no component of gravity and measure only the vehicle accelerations. The vertical accelerometer, though, measures the vehicle's vertical motion in the presence of $1g$, therefore less accurately. In an aircraft this causes altitude errors, which can be compensated with a barometer signal.
3. Self-alignment by gyrocompassing.
4. Sensor calibration by platform rotations. The other sensor biases are obtained by orienting the platform with each major axis vertical in turn, provided there is enough time.

Disadvantages:

1. Complexity and cost. The gimbal structure and its bearings must be stiff so that the accelerometer axes remain defined even under vehicle vibrations, but the bearings and slip rings must have as little friction as possible. As a result, the gimbal structure is an elaborate, precisely made mechanism which cannot avoid being expensive.
2. Gimbal magnetics. Each gimbal must have a pickoff and torquer. The pickoffs (synchros) measure the intergimbal angles with arc-sec resolution over a full revolution, a range of 10^6. When the system is first switched on, the torquers need to provide enough torque to accelerate the platform inertia so that it can gyrocompass and align itself to Level and North in a reasonable time. The torquers also must not leak magnetic flux, for that could upset the sensors as the gimbals move around them.
3. Reliability. The bearings and slip rings tend to wear, degrading alignment and performance.

Advantages and Disadvantages of Strapdown Systems

Advantages:

1. Simple structure, low cost. Strapdown systems are lighter, simpler, cheaper, and easily configured for odd-shaped spaces. As it is only necessary to mount the sensors so that their sensing axes point in known

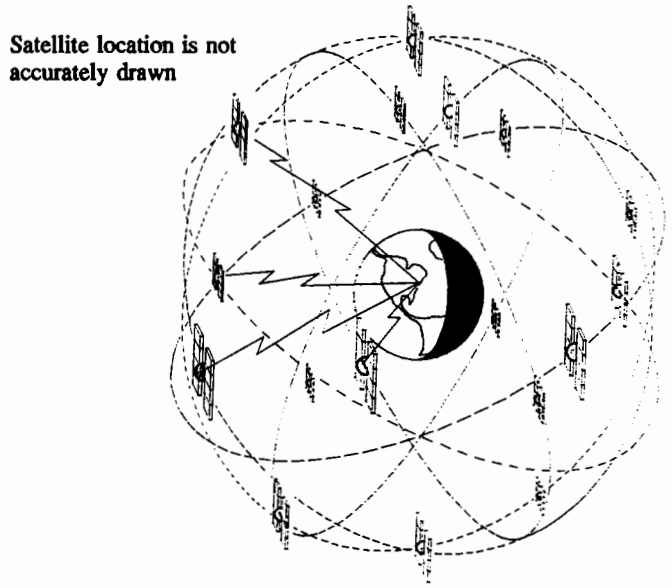

Figure 1.11. The Global Positioning System satellites.

position is known, so the user now knows position along the line of sight to the satellite. Users take data from three satellites at different angles in the sky to fix their position in three dimensions.

Now the user's clock might not be very accurate, and it is to everyone's advantage to allow it to be inexpensive, so each satellite also broadcasts the precise time in its message. As we now need a separate piece of data to decode time, we read data from a fourth satellite and compute position and accurate time. The system has been designed so that four satellites will always be visible from anywhere on the earth.

The satellites' positions can drift slightly in orbit, from irregularities in gravity, from solar pressure and so on, so fixed ground stations observe them and send orbit data to a master control center in Colorado Springs. From there a message is sent to each satellite, telling it how much it is deviating from its predicted orbit, and providing small time corrections to its clock. The satellite incorporates these correction messages (*ephemeris* data) into its transmissions. All of this is naturally done in a computer in the user's receiver, and some companies make receivers small enough to carry in the hand [16,17]. It is possible to make a receiver for both GPS and GLONASS [18,19]. The C/A code is used as a coarse acquisition signal for those privileged with access to the P code, but it is the only signal available to commercial users. However, the C/A code gives 30 m accuracy, so it is still a valuable aid.

GPS will allow more accurate navigation, but will existing charts and maps be accurate enough to realize the benefit? Roeber [20] thinks not. However,

GPS can itself be used for map making and geodetic surveys. It will be used on test ranges to check out inertial navigation systems and will facilitate air traffic control and search and rescue missions. It will change the inertial navigation field for all except strategic weapon guidance, for such a device is exactly what an inertial navigation system needs for a partner.

The INS gives very accurate attitude and distance measurements for a short time, but gyro and accelerometer errors accumulate after a while and navigation errors grow. The GPS system has exactly the opposite characteristic, for it can give accurate position fixes periodically to correct the INS, but cannot tell you quickly if you're turning or rolling; the two together complement one another [21]. Should the vehicle lose the satellite signals for a while, perhaps because the vehicle has banked and a wing hides the antenna from the satellite or because a ground vehicle has entered a tunnel, the INS can carry on until the receiver reacquires the satellite signal. For military, use the INS can keep a missile on target even if the GPS signal is obliterated by powerful jammers.

Applications of Inertial Navigation

The most accurate systems, with heading error less than 0.001 deg/h (one revolution in 40 years!) are found in submarines and ICBMs, which need navigation with errors measured in tens of feet over long mission times. We might call these "1 mile a day" systems. Their accelerometers resolve gravity to a few μg while measuring vehicle accelerations of the order of $10g$, a dynamic range of 10^7.

Next, aircraft systems, with errors of less than 1 mile an hour, are found in all civil aircraft that cross the oceans, and in long range military aircraft. While all the 1-mile-a-day systems are platform systems using mechanical sensors, laser gyro strapdown systems are dominating the 1-mile-an-hour market now. Less expensive heading and attitude reference systems (HARS) are used in short range military aircraft, and, frequently, in long range planes, to back up the INS. HARS often use mechanical gyros, but laser gyro versions are appearing as laser gyros get cheaper.

Tactical weapons such as short range missiles, Harpoon, Tomahawk, and Phoenix being U.S. examples, exclusively use strapdown systems. They presently use only mechanical sensors, but that will change as the new optical (laser) gyros and micromachined accelerometers complete development. These are described in later chapters.

Conclusions

For a deeper study of inertial systems, the books by Broxmeyer [22], Draper et al. [23], Leondes [24], McClure [25], O'Donnell [26], and Parvin [27] could be

consulted. The guidance of spacecraft uses special techniques, described by Hynoff [28].

GPS will have a tremendous impact on the navigation field, for GPS and strapdown systems go together well. The strapdown processor can support the GPS, and the INS can tell the GPS receiver where to look for satellites and maintain navigation while satellites are obscured.

There is one very important exception, though, and that is the guidance of strategic weapons (including aircraft and submarines). In a war serious enough for strategic weapons to be used, one could not rely on the constellation of GPS satellites to survive. They would either be destroyed or their signals would be jammed. So strategic systems must remain self-contained (perhaps aided with a star-tracker), and the accuracy demanded of these systems is difficult to meet with strapdown gyros.

In the 1960s, when inertial navigation systems began to be used, their instruments were electro-mechanical and very sophisticated; they were expensive, large, and fairly fragile [29]. In the 1980s the gyros were more likely to be optical, using lasers; they were less expensive, smaller, and more rugged. In the 1990s gyros and accelerometers will continue to get cheaper, smaller, and more robust, based on technologies described in the later chapters of this book.

References

1. Duncombe, R.L., R.F. Haupt, "Time and navigation", Navigation, J. Inst. Nav., 17, 4, pp. 381-386, Winter 1970-71.
2. Dunlap, G.D, H.H. Shufelt, *Dutton's Navigation and Piloting*, United States Naval Institute, Annapolis, MD, 1970.
3. Savet, P.H. (Ed.), *Gyroscopes: Theory and Design*, McGraw-Hill, New York, 1961.
4. Allington, P.J.S., "The Sperry Mk 19 gyro-compass", Symposium on Gyros, Proc. Inst. Mech. Eng. (London), 1964-65, Vol. 179, 3E.
5. Draper, C.S., "Origins of Inertial Navigation", AIAA J. Guidance and Control, 4, 5, pp. 449-463, Sept.-Oct. 1981.
6. Draper, C.S., "Guidance is forever", Navigation, J. Inst. Nav., 18, 1, pp. 26-50, Spring 1971.
7. Wrigley, W., "History of inertial navigation", Navigation, J. Inst. Nav., 24, 1, pp. 1-6, Spring 1977 (37 references).
8. Hildebrant, R.R., K.R. Britting, S.J. Madden, "The effects of gravitational uncertainties on the errors of inertial navigation systems", Navigation, J. Inst. Nav, 21, 4, pp. 357-363, Winter 1974-75.
9. Paik, H.J., J-S. Leung, S.H. Morgan, J. Parker, "Global gravity survey by an orbiting gravity gradiometer", Eos, 69, 48, pp. 1601, 1610-1611, 1988.
10. Fernandez, M., G.R. Macomber, *Inertial Guidance Engineering*, Prentice Hall International, London, 1962.

11. Markey, W.R., J. Hovorka, *The Mechanics of Inertial Position and Heading Indication*, John Wiley and Sons, New York, 1961.
12. Savant, C.R. Jr., R.C. Howard, C.B. Solloway, C.A. Savant, *Principles of Inertial Navigation*, McGraw-Hill, New York, 1961.
13. Britting, K.R., *Inertial Navigation Systems Analysis*, Wiley-Interscience, New York, 1971.
14. Lewis, R.S., *The Voyages of Apollo*, Quadrangle, New York, 1974. Pages 166-168 relate to Apollo 13's return.
15. Kayton, M., "Avionics for manned spacecraft", IEEE Trans. on Aerospace and Electronic Systems, 25, 6, p. 802, Nov. 1989.
16. "Collins demonstrates first hand-held Global Positioning System receiver", *Aviation Week and Space Technology*, 19 June 1989, p. 153.
17. Magellan Hand-held GPS Receiver, *Defense News*, 4 Sept. 1989, p. 21; also *Aviation Week and Space Technology*, 30 Oct. 1989, p. 51.
18. Eastwood, R.E., "An integrated GPS/GLONASS receiver", Navigation, J. Inst. Nav., 37, 2, pp. 141-151, Summer 1990.
19. Klass, P.J., "GPS, Glonass and Glasnost", *Aviation Week and Space Technology*, 5 Oct. 1987, p. 11.
20. Roeber, J.F, "Where in the world are we?", Navigation, J. Inst. Nav., 33, 4, Winter 1986-87.
21. Buechler, D., M. Foss, "Integration of GPS and strapdown inertial subsystems into a single unit", Navigation, J. Inst. Nav., 34, 2, pp. 140-159, Summer 1987.
22. Broxmeyer, C., *Inertial Navigation Systems*, McGraw-Hill, New York, 1964. Good for IN systems mathematical background.
23. Draper, C.S., W. Wrigley, J. Hovorka, *Inertial Guidance*, Pergamon Press, New York, 1960.
24. Leondes, C.T. (Ed.), *Guidance and Control of Aerospace Vehicles*, McGraw-Hill, New York, 1963.
25. McClure, C.L., *Theory of Inertial Guidance*, Prentice Hall, Englewood Cliffs, N.J., 1960. Mathematical treatment, covers Euler's equations, etc.
26. O'Donnell, C.F. (Ed.) *Inertial Navigation Analysis and Design*, McGraw-Hill, New York, 1964.
27. Parvin, R.H., *Inertial Navigation*, Van Nostrand, New York, 1962. Good, readable account of the basics.
28. Hynoff, E., *Guidance and Control of Spacecraft*, Holt, Reinhart and Winston, New York, 1966.
29. Slater, J.M., *Inertial Guidance Sensors*, Reinhold, New York, 1964.