CHAPTER 5

BIOMECHANICS OF HUMAN MOVEMENT

Kurt T. Manal and Thomas S. Buchanan
University of Delaware, Newark

5.1 WHY STUDY HUMAN MOVEMENT?

The biomechanics of human motion is a fascinating field. Who among us has never marveled at the graceful motions of a dancer or the rapid finger movements of a musician? From the time of Aristotle onward there have been countless books written on the topic of movement in animals and humans. Despite the great minds that have considered the topic, it is just recently that much advancement has been made experimentally. Historically, the study of human movement has been costly and very time consuming. This is because, in order to study them, movements are almost always discretized and then analyzed step by step, with the size of the steps determined by the speed of the movement (and the questions being asked). Whether it be frames of film from a movie camera or digitized samples from an electrogoniometer, most movements are recorded as series of static images that are then reassembled to provide kinematic and kinetic information.

There has been a tremendous growth in the study of human movement in the past 2 decades because of the low cost of digital data acquisition systems that make possible the storage and analysis of massive amounts of data that are required to accurately characterize complex motion. This growing interest in the study of human movement is coming from five predominate groups:

1. Basic scientists are interested in the control of human movement. How the nervous system controls the large number of degrees of freedom necessary to produce smooth, complex movements (or even simple ones!) is poorly understood. The study of the coordination of movement can be compared to the inverse problem faced by the roboticist. The roboticist develops computer programs to produce coordinated movements in a robot. On the other hand, the motor control researcher measures coordinated movements in order to understand what the "neural program" is.

2. Human movements are studied to understand and treat pathologies. For example, gait analysis is often used to help guide the physician contemplating surgery for children with cerebral palsy. The best choice for a tendon transfer or muscle lengthening surgery can be predicted by using combinations of movement analysis and biomechanical modeling (e.g., Delp et al., 1996). Gait analysis can also be used to monitor the progression of the disease and the efficacy of the treatment.
3. The study of human athletic performance has been revolutionized by motion analysis equipment and software that make it possible to readily analyze complex three-dimensional movements. From cricket bowling to figure skating to swimming to pole vaulting, the kinematics and kinetics have been examined with an aim to improve human performance.
4. There is substantial interest in human movement from those studying ergonomics and human factors related to military applications. Both the development of human-machine interfaces for high-tech weapons and the minimization of industrial injuries require knowledge of human kinematics and kinetics.
5. The kinematics of human movement has been studied by animators interested in making computer-generated characters move in realistic ways. By recording actors while they perform choreographed dances and movements, it is possible to get complex kinematic data into a computer, which can then be used to animate a computer-generated image.

5.2 FORWARD VERSUS INVERSE DYNAMICS

There are two fundamentally different approaches to studying the biomechanics of human movement: forward dynamics and inverse dynamics. Either can be used to determine joint kinetics (e.g., estimate joint moments during movements).

5.2.1 Forward Dynamics

In a forward dynamics approach to the study of human movement, the input to the system is the neural command (Fig. 5.1). This specifies the level of activation to the muscles. The neural command can be estimated by optimization models (Zajac, 1989; Pandy and Zajac, 1991) or from electromyograms (EMGs). The neural command is the sum of the neuronal signals from the a-motorneurons (that originate in the spinal cord) to the fibers of each muscle. This can be represented by a single value (at any given time) for each muscle, which we will call muscle activation a_i, and which we will represent mathematically as a value between zero and one. Hence, if it is to be estimated from EMGs, additional steps are needed to transform EMGs to muscle activation.

Musculotendon dynamics govern the transformation of muscle activation a_i, to muscle force F_i. Once the muscle begins to develop force, the tendon (in series with the muscle) begins to carry load as well. Depending upon the kinetics of the joint, the relative length changes in the tendon and the muscle may be very different. For example, this is certainly the case for a "static contraction." (This commonly used name is an oxymoron, as something cannot *contract*, i.e., shorten, and be *static* at the same time. Hence, the tendon must lengthen as the muscle shortens if the joint is not to move!).

The force in each musculotendonous unit contributes toward the total moment about the joint. The musculoskeletal geometry determines the moment arms of the muscles. (Since muscle force is dependent upon muscle length, i.e., the classic muscle "length-tension curve," there is feedback between joint angle and musculotendon dynamics.) It is important to note that the moment arms of muscles are not constant values, but change as a function of joint angles. Also, one needs to keep in mind the multiple degrees of freedom of each joint, as a muscle may have multiple actions at a joint, depending on its geometry. Finally, it is important to note that the joint moment T_i is determined from

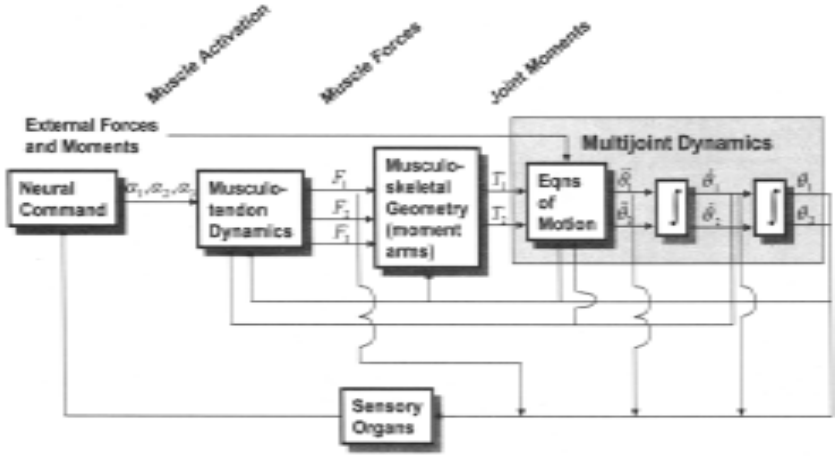

FIGURE 5.1 Forward dynamics approach to studying human movement. This simplified figure depicts the neural command and forces for three muscles and the moments and joint angles for a two-joint system. See text for details.

the sum of the contributions for each muscle. If not all muscles are included in the process, the joint moment will be underestimated. The output of this transformation is a moment for each joint (or, more precisely, each degree of freedom).

From the joint moments, multijoint dynamics can be used to compute the accelerations, velocities, and angles for each joint of interest. On the feedback side, the neural command is influenced by muscle length (via muscle spindles) and tendon force (via Golgi tendon organs). Many other sensory organs play a role in this as well, but these two are generally the most influential.

There are several limitations of the forward dynamics approach. First, it requires estimates of muscle activation. EMG methods have been used to this end, but the high variability in EMG signals has made this difficult, especially during dynamic conditions. Second, the transformation from muscle activation to muscle force is difficult, as it is not completely understood. Most models of this (e.g., Zajac, 1989) are based on phenomenological models derived from A. V. Hill's classic work (Hill, 1938) or the more complex biophysical model of Huxley's (Huxley, 1957; Huxley and Simmons, 1971), such as Zahalack's models (Zahalack, 1986; 2000). One way around the problem of determining force from EMGs is to employ optimization methods to predict muscle forces directly (bypassing these first two limitations). However, the choice of a proper cost function is a matter of great debate. Scientists doing research in human motor control find it surprising that biomechanical engineers replace their entire line of study (and indeed, the entire central nervous system) with a simple, unverified equation. Nevertheless, some cost functions provide reasonable fits of the data in addressing specific questions. Another limitation is that of determining musculoskeletal moment arms. These are difficult to measure in cadavers and even harder to determine with any accuracy in a living person. Finally, joint moments can easily be underestimated. Using forward dynamics, small errors in joint torques can lead to large errors in joint position.

5.2.2 Inverse Dynamics

Inverse dynamics approaches the problem from the opposite end. Here we begin by measuring position and the external forces acting on the body (Fig. 5.2). In gait analysis, for example, the position of tracking targets attached to the segments can be recorded by using a camera-based system and the external forces can be recorded by using a force platform.

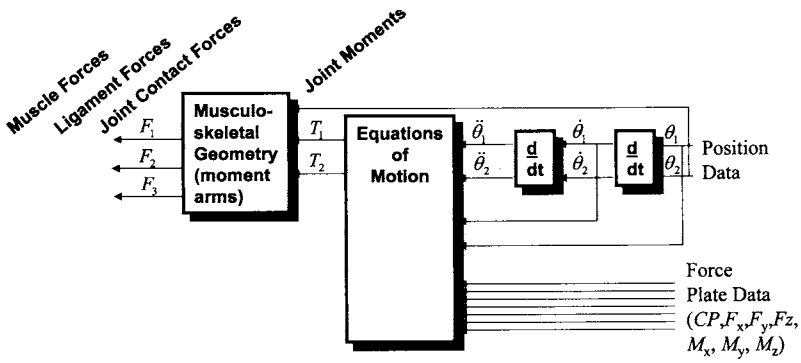

FIGURE 5.2 Inverse dynamics approach to studying human movement. This simplified figure depicts the angular position for two joints and the forces for three muscles. See text for details.

The relative position of tracking targets on adjacent segments is used to calculate joint angles. These data are differentiated to obtain velocities and accelerations.

The accelerations and the information about other forces exerted on the body (e.g., the recordings from a force plate) can be input to the equations of motion to compute the corresponding joint reaction forces and moments.

If the musculoskeletal geometry is included, muscle forces can then be estimated from the joint moments, and from these it may be possible to estimate ligament and joint compressive forces.

As with forward dynamics, inverse dynamics has important limitations. First, in order to estimate joint moments correctly, one must know the inertia of each body segment (this is embedded in the equations of motion). These parameters are difficult to measure and must be estimated. Typically, they are estimated by using established values from cadavers and scaled by simplistic scaling rules, the accuracies of which are rarely verified. Second, the resultant joint reaction forces and moments are net values. This is important to keep in mind if an inverse dynamics approach is used to predict muscle forces. For example, if a person activates his hamstrings generating a 30-N · m flexion moment and at the same time activates the quadriceps generating a 25-N · m extension moment, the inverse dynamics method (if it is perfectly accurate) will yield a net knee flexion moment of 5 N · m. Since the actual contribution of the knee flexor muscles was 6 times greater, this approach is grossly inaccurate and inappropriate for estimating the role of the knee flexors during this task. This is strongly stated because cocontraction of muscles is very common, yet this approach is widely used to estimate muscular contributions. Another limitation of the inverse dynamics approach occurs when one tries to estimate muscle forces. Since there are multiple muscles spanning each joint, the transformation from joint moment to muscle forces yields an infinite number of solutions. Choosing the proper solution requires some sort of optimization analysis, requiring the use of a cost function whose validity is sure to be challenged. Finally, if one wishes to examine muscle activations, there is no current model available that will do this inverse transformation. However, this is rarely the goal of an inverse dynamics analysis.

5.2.3 Comparing Forward and Inverse Dynamics Methods

Given the limitations of each method, which should be used: forward or inverse dynamics? That depends on the question being asked. If one's primary interest is in joint kinematics, it makes more sense to start with a measurement of position as in the inverse dynamics approach. If one is primarily interested in muscle forces, one could argue that forward dynamics has more advantages. For

estimating joint moments during movements, inverse dynamics is probably the best bet, depending upon the specific application.

For the remainder of this chapter, we will concentrate on the inverse dynamics approach for the study of human movement. Inverse dynamics is more commonly used than forward dynamics for studying human movement. A forward dynamics approach will be addressed in a subsequent chapter (Chap.6, "Biomechanics of the Musculoskeletal System," by Marcus G. Pandy and Ronald E. Barr).

5.3 TOOLS FOR MEASURING HUMAN MOVEMENT

In this section we will discuss three of the more common methods used to collect human movement data: electrogoniometers, electromagnetic tracking devices, and optoelectronic measuring systems. Of these distinctly different measuring tools, optoelectronic systems are the most common registration method, and therefore most of this section will focus on video-based motion analysis.

5.3.1 Electrogoniometers

Electrogoniometers are devices that convert joint angle to a voltage. The voltage can be sampled continuously, making electrogoniometers ideal for measuring dynamic movement. There are basically two designs, both of which fall under the category of resistive transducers. These devices, namely potentiometers and strain gauges, output a voltage related to the angular position of the joint. The voltage is converted to an angle by using a manufacturer supplied scale factor specific to each transducer. The joint angle can be displayed in real time and/or stored on a computer equipped with an analog-to-digital data acquisition card.

Potentiometers. A potentiometer is nothing more than a variable resistor that is sensitive to changes in angular position. Two arms, one fixed to the outer casing of the potentiometer and the other to the rotating shaft, can be used to mount the device to the segments on either side of a joint. The potentiometer is placed over the joint axis of rotation with the arms secured to the segments by medical tape or elasticized wraps. Changes in joint angle will cause the *wiper* (i.e., sliding contact) of the potentiometer to slide across the resistor, resulting in an output voltage linearly related to the joint angle. It is important that the potentiometer be placed over the axis of rotation; otherwise, movement of the joint will be restricted. The electrogoniometer is ideally positioned when the rotating shaft of the potentiometer and the joint axis of rotation are aligned. More elaborate mounting methods have been designed to house mutually perpendicular potentiometers in multiple degree of freedom exoskeletal linkages (e.g., Chao, 1980; Shiavi et al., 1987). These devices are no longer commonly used, but are mentioned for historical purposes since they have played an important role in many previous studies.

Strain Gauges. Strain gauges can also be used to detect changes in joint angular position. An example of a 1 degree of freedom electrogoniometer is illustrated in Fig. 5.3. Two and three degrees of freedom electrogoniometers of this type are also available. Strain-sensitive wires are mounted within a connecting element and electrically connected to form a Wheatstone bridge. Each strain-sensitive wire is, in effect, a resistor and is sensitive to strains in particular directions. Hence, when the electrogoniometer is forced to rotate about the axis of rotation as drawn in the figure, the bridge circuitry becomes unbalanced. This unbalance is noted as a change in the output voltage of the bridge and is proportional to the amount of rotation. The design is clever because pure rotation about axes perpendicular to the axis of rotation depicted in Fig. 5.3 will *not* unbalance the bridge. Another interesting and practical characteristic of this device is that it does not have to be positioned over the joint axis of rotation as is the case

for rotatory potentiometers. Note, however, that the base of the mounting blocks must lie in the plane of rotation without necessarily being centered over the axis of rotation.

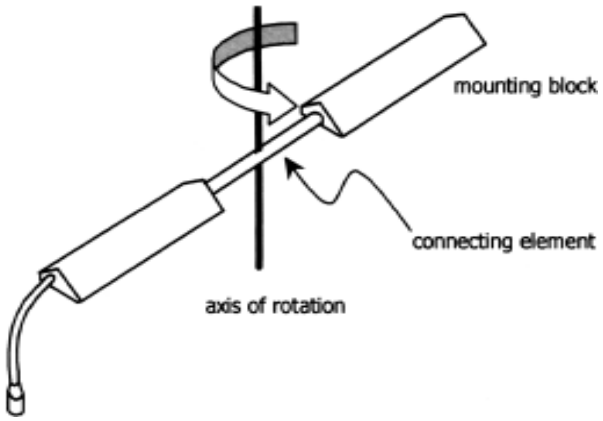

FIGURE 5.3 Single degree of freedom strain gauge for measuring joint angular position. Strain-sensitive wires are fixed to the connecting element between the mounting blocks. The mounting blocks are secured to both segments on either side of a joint. Changes in joint angle are output as a voltage proportional to the amount of rotation about the axis of rotation.

Electrogoniometers of this type can be configured to display joint angles in real time and/or interfaced with a computer for data storage. Additionally, data can be saved to a storage unit (i.e., data logger) strapped to the subject. The data logger is ideal for recording dynamic movements in the field. The stored data can be uploaded to a computer at a later time and converted to joint angles for analysis.

There are advantages and disadvantages associated with the use of electrogoniometers. In their favor are ease of use and cost. On the other hand, they are less accurate than other systems used to record movement. In addition, both designs (i.e., potentiometer and strain gauge) require placement over the joint, which may interfere with the natural kinematics because of cumbersome cabling and/or method of attachment. Another drawback of these devices is that while they provide a *relative* measure of joint angular position, the data do not lend themselves to an inverse dynamics analysis in which joint reaction forces and moments are of interest, the computation of which requires knowledge of the absolute positions of the body segments.

5.3.2 Electromagetic Tracking Systems

Electromagnetic tracking technology originated in the defense industry and has since become widely used in the entertainment industry (e.g., motion pictures, animation, and gaming). The use of electromagnetic tracking has become increasingly popular in the academic environment as evinced by the growing number of research publications using this technology.

Electromagnetic tracking is based on Faraday's law of magnetic induction. That is, electrons in a conductor experience a spontaneous magnetic force when moved through a magnetic field. The magnitude of the induced force (i.e., electromotive force or EMF) is proportional to the strength of the magnetic field through which the conductor is moved. The magnitude of the EMF (i.e., voltage) is also related to the speed the conductor is moving. If the conductor is in the shape of a loop, the same principles apply with an induced EMF proportional to the strength of the magnetic

field perpendicular to the cross-sectional area of the loop. The induced EMF is related to the magnetic flux Φ_B:

$$\text{EMF} = -\frac{d\Phi_B}{dt} \tag{5.1}$$

Conceptually, the strength and direction of a magnetic field can be thought of as the density and direction of magnetic field lines. The magnetic flux will vary as the conductor moves closer to/farther from the source of the magnetic field and also as it rotates relative to the magnetic field lines. The more field lines passing through the loop of the conductor, the greater the induced EMF. This principle forms the basis for electromagnetic tracking. That is, the general idea is to move a conducting sensor through a magnetic field and record the induced voltage.

The basic components of an electromagnetic tracking system are an active transmitter and passive sensors. The transmitter is stationary and contains three orthogonal coils (i.e., antennas) that are activated in sequence, with only one antenna generating a magnetic field at a time. Interestingly, if the subject (and therefore the sensors attached to the subject) stops moving within the magnetic field, we might think the induced voltage in each sensor would remain constant. However, revisiting Eq. (5.1) shows that this is not the case, because the magnetic flux must change with respect to time or the EMF goes to zero. There are two things that can be controlled to ensure a changing flux: (1) make sure the subject never stops moving or (2) change the strength and direction of the magnetic field. Electromagnetic tracking systems use the latter strategy to ensure the magnetic flux changes. The transmitter not only emits a magnetic field, but also serves as a fixed reference about which position and orientation of each sensor is reported.

Each receiving sensor contains three orthogonal coils used to detect the magnetic field emitted by the transmitter according to the principles of magnetic induction. The receiving coils are contained within a 1-cubic-inch plastic housing that protects and provides a convenient method for attaching the sensor to the subject. The sensors are generally secured by double-sided tape and wrapped with an elasticized band. Proprietary signal processing takes place in real time, compensating for the strength of the earth's magnetic field. Individual coil signals can be used to determine the orientation of the sensor relative to the antenna generating the magnetic field. Each coil within the sensor detects three signals from the transmitter (i.e., one for each antenna of the transmitter) for a total of nine signals. These nine signals suffice to locate the position and orientation of the sensor relative to the transmitter. For example, the receiving coil most parallel to the currently active transmitting antenna will experience the largest EMF, while the more orthogonal the coil, the smaller the induced voltage. Because each coil within a sensor is the same distance from the transmitter, it is possible to determine the distance and orientation of the sensor relative to the currently active antenna by comparing the strength of the induced EMF in each coil to the strength of the emitted magnetic field.

There are two types of electromagnetic tracking systems that are used for the study of human movement. The biggest difference between these systems is that one (e.g., from Polhemus Incorporated) uses an ac magnetic field, while the other type (e.g., from Ascension Technology Corporation) uses a pulsed dc magnetic field. The precision and accuracy of electromagnetic tracking systems is affected by metallic objects, low-frequency electronic noise, and the distance of the sensor from the transmitting antennas. The radius within which precise and accurate data are sampled depends on the particular system and strength of the transmitter. However, when used in an ideal environment, the precision and accuracy of both systems is more than adequate for studying human movement. With proper setup, static accuracy of less than 2 mm rms and 0.5° rms for both systems is possible. Real-time data can be sampled at a rate of up to 144 Hz, depending on the system, the number of sensors tracked, and the type of data communication interface with the computer. These systems do not suffer from line-of-sight problems typical of optoelectronic systems and are therefore ideal for capturing complex movements.

5.3.3 Optical Methods: Camera-Based Systems

The most common method of recording human movement involves "filming" the motion of interest. Historically, images were stored on conventional media such as 16-mm film or on videotape. Today's

standard is based on digital technology, bypassing physical media per se, sending the data directly to the computer. Data are commonly sampled at a rate between 50 and 240 frames per second, depending on the movement of interest. For example, natural cadence walking is often sampled at a rate of 60 Hz, while running and arm movements tend to be sampled at 100 Hz or faster. There are two types of high-speed video-based systems that are used for studying human movement. The fundamental difference between designs is related to their use of active or passive tracking of targets.

Active-tracking target systems use infrared light-emitting diodes (LEDs) to indicate the position of the target in space. The diodes are pulsed so that only one target is illuminated (i.e., active) at a time. Thus, if a target is not detected by a camera and then suddenly reappears in the camera's field of view, it will automatically be identified by its order in the pulse sequence. Active target systems are subject to line-of-sight problems common to all optical-based tracking systems. That is, the target must be "seen" by a camera to be detected. Active targets emit a restricted angle of light that may not be detected by the camera if the target rotates relative to the segment to which it is attached.

A limitation of active target systems is that on-board electronics must be strapped to the subject with leads to each of the diodes. These wires, combined with the subject being tethered to the cameras, can interfere with certain movements. Tetherless systems (i.e., telemetered systems) are available; however, wires to each diode are still necessary.

In contrast, passive tracking targets merely reflect projected light and do not actively communicate their position in space. It is therefore important that the tracking targets reflect more light than surrounding objects. To promote this, tracking targets are covered with a highly reflective material, most often in the form of retroreflective tape; however, reflective ink or paint can also be used. In addition, a ring of stroboscopic LEDs mounted around the camera lens housing is used to illuminate the tracking targets (see right panel of Fig. 5.4).

Passive tracking targets typically range between 10 and 40 mm in diameter, with the size of the target usually related to the field of view in which the movement takes place and the accuracy of the experimental setup. There is a trade-off in target size, since overly large targets may obscure the detection of other targets, while too small a target may not reflect sufficient light to be detected by the cameras. A reasonable rule of thumb is that the diameter of the tracking targets should be between 1 and 2 percent of the largest dimension of the calibrated workspace. The workspace may be thought of as the region in which the movement will take place, and is generally defined during a system calibration process (discussed in "Camera Calibration," below).

The Role of the Video Camera. The determination of the three-dimensional coordinates of tracking targets from multiple two-dimensional camera views is often taken for granted or treated as a *black box*. In the sections that follow, we will discuss the basic principles of reconstructing three-dimensional target coordinates from multiple camera images. It is advantageous to describe how the process works for a single tracking target prior to discussing how three-dimensional kinematics of segmental motion are calculated.

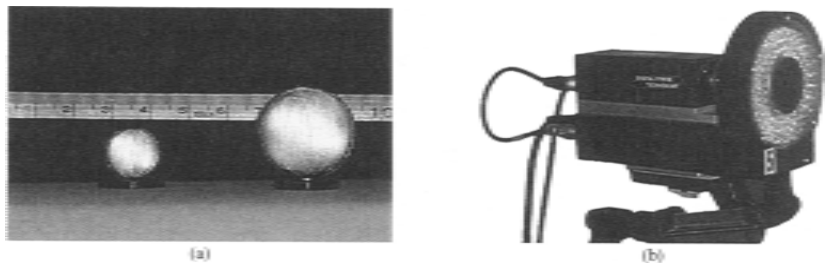

FIGURE 5.4 (*a*) 10- and 25-mm retroreflective-tape-covered tracking targets. Note how the targets are mounted on lightweight plastic pedestals. The pedestals make it easier to attach the targets to the segment. (*b*) High-speed digital video camera used to "film" the position of the tracking targets. Note the stroboscopic ring of LEDs around the lens of the camera.

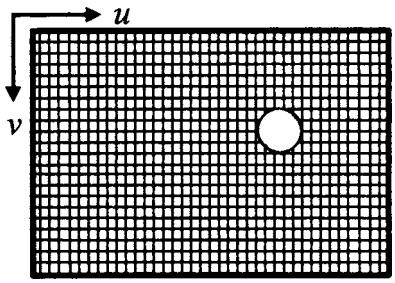

FIGURE 5.5 Schematic representation of the light-sensitive elements of the imaging sensor. The circle is the projection of a target onto the image plane. Note the direction of the *u*, *v* reference axes.

For the purposes of this discussion, we assume the image plane of our high-speed video camera is a charge-coupled device (CCD) sensor. The image plane may be thought of as the *exposure medium* onto which real-world *object-space* is projected. The term *object-space* will be used to describe the *X*, *Y*, *Z* inertial reference system in which the tracking targets move. Individual elements of the sensor are arranged in a series of rows and columns, with each element responsible for converting the intensity of light to a voltage such that the greater the intensity of light striking the element, the greater the voltage. This is particularly relevant because the tracking targets should reflect more light than all other objects detected by the camera. The matrix arrangement of light-sensitive elements is illustrated schematically in Fig. 5.5. Note the internal *u*, *v* coordinate system of the image plane.

Consider the case where a single tracking target is detected by a camera and no other light-reflecting objects are visible. The circle in Fig. 5.5 is used to depict the projection of the target onto the imaging sensor. Clearly, other elements of the sensor would be excited to varying degrees depending on the ambient lighting, but have been turned off for the sake of this example. Figure 5.5 is also helpful in depicting the effect of using binary thresholding to suppress background light sources or reflective objects other than the tracking targets. The idea is to suppress all voltages below a user-specified threshold, which has the effect of moving the light-reflective tracking targets to the foreground. This is ideal since we are ultimately concerned with locating the center of a prospective target in *u*, *v* coordinates and do not want other sources of light affecting the location of the computed center. One method of determining the center of a target is to scan the matrix of light-sensitive elements for transitions in voltage (i.e., edge detection) and fit a circle of best fit to the resulting "edges." While this approach is conceptually straight forward, binary thresholding as described here would eliminate useful information that could otherwise be used to determine the center of the projected target at greater, subpixel accuracy. For example, rather than simply treating each element of the sensor as being on or off, we could use a weighted average of sensor element voltages and a geometric constraint that the active elements form a circle. The center of the target in the image plane is assumed to lie at the center the circle.

If one were to draw a line from the center of the target in the image plane to the *X*, *Y*, *Z* coordinates of the target in object-space, it would be clear that the mapping between these spaces is not unique, since all targets lying on this line would map to the same *u*, *v* coordinates. From this, it is evident that the location of a target in object-space cannot be determined with only one camera. This raises a subtle but important distinction regarding the role of the camera in video-based motion analysis. The camera does not actually record the location of a target in object-space, but rather the role of the camera is to define a ray in the direction of the target. When multiple cameras view the same target, the location of the target in object-space is assumed to lie at the intersection of the directed rays from each camera. The cameras must first be calibrated before the intersection of these rays can be calculated.

Camera Calibration. Each camera must be calibrated before it can contribute to locating a target in object-space. Camera calibration defines a *mapping* from three-dimensional object-space into the two-dimensional *u*, *v* coordinates of the camera. This mapping is expressed in Eq. (5.2) in terms of homogeneous coordinates:

$$\begin{pmatrix} \lambda u \\ \lambda v \\ \lambda \end{pmatrix} = \mathbf{A} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{5.2}$$

where λ = scale factor relating the spaces
u, v = image plane coordinates of a target
$\mathbf{A}$ = 3 × 4 transformation matrix
X, Y, Z = coordinates of a target in object-space

Expanding the right-hand side of equation 2 results in the following set of equations.

$$\lambda u = \alpha_{11}X + \alpha_{12}Y + \alpha_{13}Z + \alpha_{14} \tag{5.3}$$

$$\lambda v = \alpha_{21}X + \alpha_{22}Y + \alpha_{23}Z + \alpha_{24} \tag{5.4}$$

$$\lambda = \alpha_{31}X + \alpha_{32}Y + \alpha_{33}Z + \alpha_{34} \tag{5.5}$$

Substituting λ into Eq. (5.3) and (5.4), and introducing the following equation:

$$\beta_{ij} = \alpha_{ij}/\alpha_{34} \tag{5.6}$$

leads to two convenient expressions relating the coordinates of the center of the target in the image plane and the location of the target in the object-space.

$$u = \beta_{11}X + \beta_{12}Y + \beta_{13}Z + \beta_{14} - u\beta_{31}X - u\beta_{32}Y - u\beta_{33}Z \tag{5.7}$$

$$v = \beta_{21}X + \beta_{22}Y + \beta_{23}Z + \beta_{24} - v\beta_{31}X - v\beta_{32}Y - v\beta_{33}Z \tag{5.8}$$

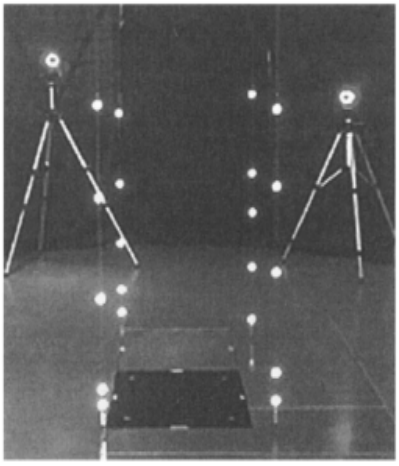

FIGURE 5.6 The X, Y, Z coordinates of the control points (i.e., reflective targets) are known relative to the origin of the object-space. Once the cameras have been calibrated, the hanging strings with the control points are removed from the field of view of the cameras. The black rectangle flush with the floor is a force platform (see Sec. 5.4.5).

Note that the u, v coordinates of the target are known. Therefore, if the X, Y, Z coordinates of the target also are known, we are left with 11 unknowns (i.e., transformation parameters) in two equations. The unknown betas (i.e., β_{ij}) can be determined if the X, Y, Z coordinates of at least six *control points* are detected by the camera. That is, each control point provides two equations that can be used to solve for the 11 unknown betas. The term *control point* is used to make clear that the X, Y, Z coordinates for these targets are known, having been accurately measured relative to the origin of the object-space. The control points are used solely for the purpose of calibrating the cameras and are removed from the field of view once the cameras have been calibrated. The distribution of the $n \geqslant 6$ control points must not be colinear, and the control points should encompass the volume within which the movement will take place. This volume is often referred to as the *workspace*. One method of defining the workspace is to hang four strings with a number of control points attached to each string, as shown in Fig. 5.6.

The direct linear transformation (DLT) proposed by Abdel-Aziz and Karara (1971) is perhaps the most well known method of calibrating the cameras among those conducting video-based motion analysis. The unknown betas for each camera are related to internal and external camera parameters. Examples of internal parameters include the principal distance from the center of the camera lens to the image plane and the u, v coordinates of the principal point. (The principal point lies at the intersection of the principal axis and the image plane.) Although the number of internal parameters can vary depending on the accuracy of the geometric representation of the camera, the number of external parameters remains fixed at six. The six external parameters (i.e., three position and three orientation) describe the relationship between the internal camera coordinate system and the object-space.

Prior to development of the DLT method, the six external parameters were measured manually. This was a painstaking process and subject to errors. The simple act of bumping a camera or repositioning the cameras for a new experimental setup involved remeasuring the external parameters. The DLT greatly facilitated video-based motion analysis, providing a convenient method of solving for the external camera parameters and determining the mapping from object-space to the *u, v* coordinates of the image plane.

This discussion on camera calibration is not meant to be comprehensive. However, it does provide the basic background for understanding how and why cameras are calibrated. Additional terms can be added to the basic 11-parameter DLT model to correct for symmetrical and asymmetrical lens distortions. These errors can be treated, in part, during camera calibration, and may also be accounted for by using lens correction maps provided by the manufacturer.

In recent years, the so-called wand or dynamic calibration method has become widely used in place of hanging strings with control points. In a dynamic calibration, two retroreflective targets attached to a wand are moved throughout the entirety of the volume in which the movement will take place. The targets on the wand are not control points *per se* because their locations in object-space are not known *a priori*. However, since the coordinates of the two wand targets can be measured with respect to the *u*, *v* coordinates of each camera and because the distance between targets should remain constant, the cameras can be calibrated in an iterative manner until the length of the wand as detected by the cameras matches the true length of the wand (i.e., distance between targets). Although the length of the wand can be reconstructed very accurately by this method, the direction of the object-space reference axes does not have to be known for determining the length. A static frame with a predefined origin and control points arranged to define the object-space reference axes is placed in the field of view of the "calibrated" cameras to establish the direction of the *X*, *Y*, *Z* object-space axes.

Calculating Object-Space Coordinates. Once the cameras have been calibrated and a set of betas for each camera is known, the opposite approach can be used to locate the position of a target in object-space. The term *reconstruction* is often used to describe the process of calculating three-dimensional coordinates from multiple ($n \geq 2$) camera views. Consider the example illustrated in Fig. 5.7, where two cameras have a unique perspective of the same tracking target. The *u*, *v* coordinates of the target in each camera view are known, as are the betas for both cameras as a result of the calibration. The unknowns in this case are the *X*, *Y*, *Z* coordinates of the target in the object-space. Rearranging Eq. (5.7) and (5.8) and adding two more equations for the second camera leads to the following:

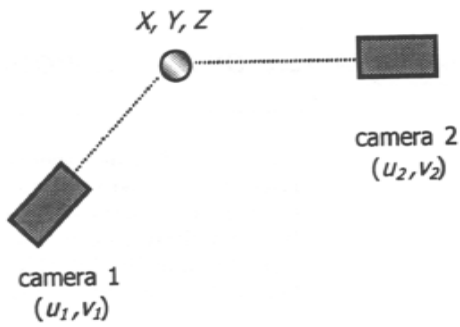

FIGURE 5.7 Cameras 1 and 2 each have a unique perspective of the tracking target in object-space (i.e., silver circle). The *X*, *Y*, *Z* coordinates of the target can be calculated by using the *u*, *v* coordinates and the betas determined during calibration.

$$u_1 = (\beta'_{11} - u_1\beta'_{31})X + (\beta'_{12} - u_1\beta'_{32})Y + (\beta'_{13} - u_1\beta'_{33})Z + \beta'_{14} \tag{5.9}$$

$$v_1 = (\beta'_{21} - v_1\beta'_{31})X + (\beta'_{22} - v_1\beta'_{32})Y + (\beta'_{23} - v_1\beta'_{33})Z + \beta'_{24} \tag{5.10}$$

$$u_2 = (\beta''_{11} - u_2\beta''_{31})X + (\beta''_{12} - u_2\beta''_{32})Y + (\beta''_{13} - u_2\beta_{33}'')Z + \beta''_{14} \tag{5.11}$$

$$v_2 = (\beta''_{21} - v_2\beta''_{31})X + (\beta''_{22} - v_2\beta''_{32})Y + (\beta''_{23} - v_2\beta''_{33})Z + \beta''_{24} \tag{5.12}$$

where the subscript on *u* and *v* indicates camera 1 or 2, with and used to identify the betas for cameras 1 and 2, respectively. We can express Eq. (5.9) to (5.12) compactly if we let C_{ij} be the terms in parentheses, where *i* indicates row and *j* column [see Eq. (5.13) for an example], and let L_i be the

combination of the left-hand side and the lone beta on the right-hand side of Eqs. (5.9) to (5.12) [see Eq. (5.14) for an example]. Equation (5.15) reflects this compact notation.

$$C_{11} = (\beta'_{11} - u_1\beta'_{31}) \tag{5.13}$$

$$L_3 = (u_2 - \beta''_{14}) \tag{5.14}$$

$$\mathbf{L} = \mathbf{C}\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{5.15}$$

Since **C** is not a square matrix (it is 4 × 3), the unknown *X*, *Y*, *Z* coordinates can be solved by using the Moore-Penrose generalized inverse, as follows:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = (\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{L} \tag{5.16}$$

which, in essence, yields a least-squares solution for the *X, Y, Z* coordinates of the tracking target. The solution is easily expanded to account for $n>2$ cameras.

While only two cameras are necessary to reconstruct the three-dimensional coordinates of a tracking target in object-space, more than two cameras are recommended to help ensure that a minimum of two cameras see the target every point in time. The cameras should be positioned so that each has a unique perspective of the workspace. Ideally the cameras should be placed at an angle of 90° with respect to one another. In practice this may not be possible, and every effort should be taken to maintain a minimum separation angle of at least 60°.

5.4 ANALYSIS OF HUMAN MOTION: AN INVERSE DYNAMICS APPROACH

The inverse dynamics approach is the most commonly used method to solve for unknown joint reaction forces and moments. The analysis begins with the most distal segment, moving upward through the kinematic chain, requiring that all external forces acting on the system are known. A free-body diagram appropriate for a two-dimensional inverse dynamics analysis of the foot and shank is illustrated in Fig. 5.8. This can be expressed mathematically in a generalized form suitable for a two- or three-dimensional analysis of *n* segments.

$$\Sigma\mathbf{M}_i = I_i\, d\boldsymbol{\omega}_i/dt \tag{5.17}$$

$$\Sigma\mathbf{F}_i = m_i\, d\mathbf{v}_i/dt \tag{5.18}$$

where $\Sigma\mathbf{M}_i$ = sum of the moments acting on segment *i*
I_i = inertia tensor for segment *i* about its center of mass (COM)
ω_i = angular velocity of the segment

Forces acting on segment *i*, mass, and linear velocity of the segment correspond to $\mathbf{F}_i$, m_i and $\mathbf{v}_i$, respectively.

The three types of measurement *tools* described in Sec. 5.3 all provide kinematic data of some form. For example, goniometers provide an estimate of joint angular position, while electromagnetic tracking systems output the relative position and orientation of the sensors attached to the segments. In the case of video-based motion analysis, output data are in the form of target coordinates expressed in the object-space. It is important to note that output from all of these devices contains

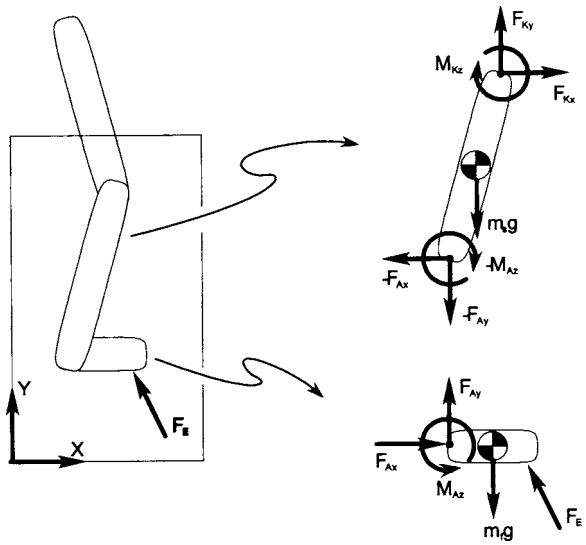

FIGURE 5.8 Two-dimensional FED of the foot and shank segments. Note how the forces and moment acting at the distal end of the shank are equal in magnitude but opposite in direction to the forces and moment acting at the ankle. F and M are used to indicate force and moment respectively, with A and K used to distinguish between the ankle and knee joints. The foot and shank segments are represented by f and s, while F_E is used to denote the external force acting on the foot.

some degree of error. That is, the sampled signal is actually a combination of "true" signal and "noise." This is an important consideration because the equations of motion contain linear and angular acceleration terms, values that are obtained by numerically differentiating the position data. Differentiating the raw data will have the undesirable effect of magnifying the noise, which can severely compromise the integrity of the results. An excellent discussion of this topic can be found in Winter (1990). The point we wish to make is that the raw data should be treated to reduce or eliminate the amount of contaminating noise before the data are used in subsequent calculations. This process of treating the raw data is commonly referred to as *data smoothing*.

5.4.1 Data Smoothing

As previously stated, all motion capture data contains some degree of noise. For example, consider the plot in Fig. 5.9. The dotted line is the trajectory of a tracking target attached to the shank of a subject walking at a natural cadence. The data are limited to the stance phase of gait and have not been treated (i.e., these data are "raw"). The thick line passing through the raw data is a smoothed form of the original signal. The purpose of this section is to introduce several methods that are commonly used in biomechanical studies to smooth motion capture data.

For the purpose of this discussion, consider a single tracking target whose position has been determined by video-based motion analysis. We begin by assuming the data were collected at an adequate sampling rate to prevent aliasing of the signal. The sampling theorem states that the data should be sampled at a rate of at least 2 times greater than the highest-frequency component in the signal being sampled. This minimum rate is commonly described as the *Nyquist limit*. It is not unreasonable to sample the data at 5 times the Nyquist limit to ensure the integrity of the data in both the frequency and time domains. Historically, sampling at such a high rate was prohibitive because of

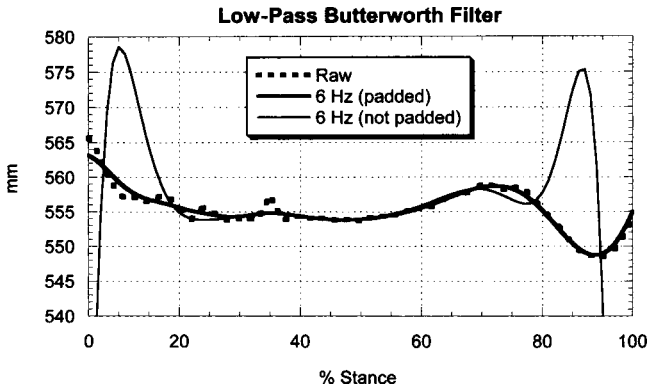

FIGURE 5.9 The bold black squares are raw data for the *X* coordinate of a tracking target attached to the shank of a subject during the stance phase of natural cadence walking, where *Y* is the forward direction and *Z* is the vertical direction. The thick solid line is the result of filtering the raw data in the forward and reverse direction using a fourth-order Butterworth low-pass digital filter set at a cutoff frequency of 6 Hz. The front and back ends of the raw data were padded prior to filtering. Note how the thick line fits through the original raw data. The thin line is the result of filtering the raw data using exactly the same Butterworth filter, with the only difference being that the raw data were not padded prior to filtering. Clearly, the smoothed data represented by the thin line are not suitable for analysis.

constraints on disk space. Storage space, however, is no longer an issue and sampling data well above the Nyquist limit is recommended.

Two general approaches of smoothing data include curve fitting and digital filtering. Both approaches can yield similar results; however, the underlying theory behind each approach is different.

Curve Fitting. Curve fitting, as the name implies, involves fitting a function or a series of functions through the raw data with a goodness of fit generally based on a least-squares difference. For example, polynomial regression and piecewise polynomial approximation are methods of curve fitting, the latter of which is more commonly used in studying human movement. Cubic and quintic splines are the most common of the piecewise approximation methods. These splines require that a smoothing parameter be specified to determine how closely the smoothed data fit through the original data points. The goal is to select a smoothing parameter that does not over- or undersmooth the raw data. In practice, it may be difficult to determine an ideal smoothing parameter. Algorithms have been created in which an ideal smoothing parameter can be determined by a statistical procedure known as *generalized cross validation* (GCV). The GCVSPL package (Woltring, 1986) is one such program that uses GCV to identify an ideal smoothing parameter for the spline. A description of the GCVSPL package and source code is available for download from the International Society of Biomechanics (1999).

Digital Filtering. Digital filtering is another method that is used to smooth biomechanical data. The concept is based on the fact that any signal, if sampled at an adequate rate, can be recreated from a series of sine and cosine waveforms of varying frequency. This principle can be used to reduce the amount of noise, if the frequency content of the noise is known. For the sake of this example we assume the acquired data are contaminated with high-frequency noise. Although a variety of digital filters exist, we focus our attention on the Butterworth filter because it is perhaps the most widely used filter in biomechanics research.

A low-pass Butterworth filter is designed to attenuate frequencies above a specified cutoff frequency, while allowing frequencies below the cutoff to pass through the filter unattenuated. Butterworth filters are not infinitely sharp. The *order* of the filter characterizes the sharpness, or how much the signal is attenuated in the vicinity of the cutoff frequency. The higher the order, the sharper the filter response. Computer implementation of a Butterworth filter is straightforward, which may be in part why it is so widely used. Although we described the role of the Butterworth filter as extracting particular frequencies, the raw signal is actually filtered in the time domain as seen below.

$$Y(t) = a_0X(t) + a_1X(t - 1) + a_2X(t - 2) + b_1Y(t - 1) + b_2Y(t - 2) \tag{5.19}$$

where $Y(t)$ and $X(t)$ are the filtered and raw data at time (t). The (t - 1) and (t - 2) notation is used to indicate data at 1 and 2 samples prior to the current time. Equation (5.19) is for a second-order recursive filter; higher-order filters require additional recursive terms. The a and b coefficients for a Butterworth low-pass filter can be determined from the following equations:

$$\omega_c = \tan\left(\frac{\pi f_c}{f_s}\right) \tag{5.20}$$

$$K_1 = \sqrt{2}\omega_c \tag{5.21}$$

$$K_2 = \omega_c^2 \tag{5.22}$$

$$K_3 = \frac{2a_0}{K_2} \tag{5.23}$$

$$a_0 = a_2 = \frac{K_2}{1 + K_1 + K_2} \tag{5.24}$$

$$a_1 = 2a_0 \tag{5.25}$$

$$b_1 = -2a_0 + K_3 \tag{5.26}$$

$$b_2 = 1 - 2a_0 - K_3 \tag{5.27}$$

where f_c and f_s are the cutoff frequency and the sampling frequency expressed in hertz, respectively.

Several practical considerations should be noted when using a Butterworth digital filter. First, we see from Eq. (5.19) that the filtered data at time t are related in a recursive manner to raw and filtered data at times t - 1 and t - 2. This can cause problems at the beginning of the data set unless the front end of the data is padded with extra data points. The dotted line in Fig. 5.9 illustrates the consequence of not padding the data set prior to filtering. Clearly the smoothed data at the beginning (and also at the end!) of the data set are erroneous. Two methods of padding the front end of the data involve reflecting the first n data points (10 or more generally work well for a second-order filter) about data point 1, or simply by collecting more data than is actually needed. It should be noted that this type of digital filter introduces a phase lag in the smoothed signal. The easiest method of correcting for this phase lag is to refilter the already filtered data in the reverse direction. This will shift the data in an equal and opposite direction, realigning the raw and filtered data temporally. Note that filtering the already-filtered data in the reverse direction will increase the sharpness of the filter response. If the data are filtered in the reverse direction, it is advisable to pad the back end of the data set for the reasons cited earlier.

The thick line in Fig. 5.9 is the result of filtering the raw data in the forward and reverse directions using a fourth-order, low-pass Butterworth filter set at a cutoff frequency of 6 Hz (note that the front and back ends of the raw data were padded). This raises an interesting question, that is, how do we identify an appropriate cutoff frequency for the filter? There are a number of methods that can be used to help select an appropriate cutoff frequency. A fast Fourier transform (FFT) can be used to

examine the content of the signal in the frequency domain, or one of several residual analysis methods can be used (Jackson, 1979; Winter, 1990).

5.4.2 Tracking Motion of the Segment and Underlying Bone

We continue with our example of how motion data collected with a video-based tracking system is used in an inverse dynamics analysis. Calculating joint kinetics from the observed kinematics and the external forces acting on the body requires knowledge of how the bones are moving. In this section we describe how tracking targets attached to the segments can be used to track motion of the underlying bones. We assume the target coordinates have been smoothed by an appropriate method.

The first step in calculating joint and segmental kinematics is to define orthogonal anatomical coordinate systems (ACSs) embedded in each segment. Because it is the kinematics of the underlying bones that is most often of interest, we must define a set of reference axes that are anatomically meaningful for the purposes of describing the motion. An anatomical coordinate system (i.e., ACS) is constructed for each segment in the kinematic chain. Retroreflective targets positioned over anatomical sites (hence the term *anatomical targets*) are used to define the ACS for each segment. Consider the case in Fig. 5.10*a* where anatomical targets are positioned over the malleoli and femoral condyles. These targets are used to define an ACS for the shank (ACS_{shank}). The frontal plane of the shank is defined by fitting a plane through the four anatomical targets (Fig. 5.10*b*). The next step is to define the ankle and knee joint centers, which are assumed to lie midway between the malleoli and femoral condyle targets, respectively. The longitudinal axis of the shank lies in the previously defined plane, originating at the distal joint center (i.e., the ankle joint) and pointing in the direction of the knee joint center. The origin of the ACS_{shank} is set at the COM of the segment. The COM lies along the longitudinal axis, at a location generally determined by using anthropometric lookup tables (see Sec.

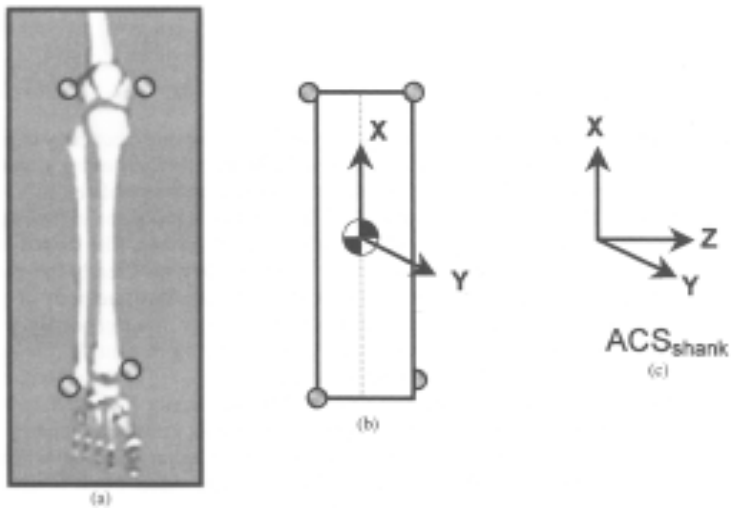

FIGURE 5.10 Retroreflective targets are placed over the medial and lateral malleoli and femoral condyles (*a*). These anatomical targets are used to define the frontal plane of the shank (*b*). The *X* axis projects from the ankle joint center toward the knee joint center. The *Y* axis lies perpendicular to the frontal plane, with the *Z* axis given by the cross product of *X* and *Y*. The orthogonal axes in the right panel represent the ACS_{shank}, which is located at the COM of the segment.

5.4.4). The unit vector X_s, originating at the COM, will be used to define the direction of the longitudinal axis of the ACS_{shank}. An anteroposterior (AP) axis lies perpendicular to the frontal plane of the shank, with a mediolateral (ML) axis formed by the cross product of the longitudinal and AP axes. Unit vectors Y_s and Z_s originating at the COM of the shank are used to define the direction of the AP and ML axes, respectively. The X_s, Y_s and Z_s unit vectors are orthonormal by way of construction and form the basis of the ACS_{shank} (Fig. 5.10*c*).

A similar approach can be used to construct ACSs for other segments in the kinematic chain, given suitable placement of anatomical targets (see Cappozzo et al., 1995). Figure 5.11 illustrates ACSs for the shank and thigh without specifying the exact details of how the ACS_{thigh} was constructed.

Although anatomical targets are used to construct the ACSs, and it is motion of the ACSs that is of interest, it is not practical to track motion of the anatomical targets because they are prone to being knocked off, and in many cases pose line-of-sight problems. The medial malleolus and femoral condyle targets are especially prone to these problems. From a data collection perspective, it is easier to track targets attached to a segment that have been positioned for optimal viewing by the cameras than it is to track targets over anatomical sites. If we define a relationship between the tracking targets and the ACS, we can estimate how the bones are moving by tracking motion of targets on the segment. The easiest way to do this is to construct a set of orthogonal axes using three tracking targets and a series of vector cross products [Eqs. (5.28) to (5.30)]. The resulting orthogonal axes are illustrated in Fig. 5.12. These axes will be referred to as a *local coordinate system*.

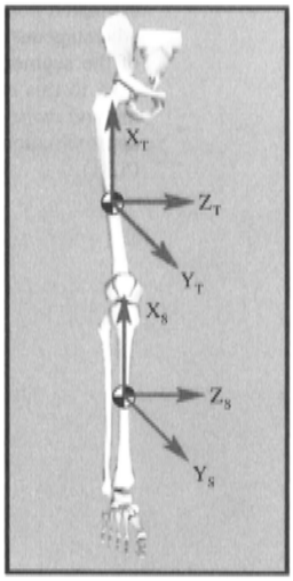

FIGURE 5.11 ACSs for the shank (subscript *S*) and thigh (subscript *T*) segments. The ACSs originate at the COM of each segment. Note that changes in the knee angle will cause the relative orientation between the ACS_{shank} and ACS_{thigh} to change.

$$\mathbf{i} = A - C \tag{5.28}$$

$$\mathbf{j} = (B - C) \times (A - C) \tag{5.29}$$

$$\mathbf{k} = \mathbf{i} \times \mathbf{j} \tag{5.30}$$

where the *X*, *Y*, *Z* coordinates of tracking targets *A*, *B*, and *C* are known in the object-space.

The relative position and orientation of the local coordinate system and the ACS can be represented as a translation vector relating their origins and a rotation matrix of direction cosines. Moreover, the relative position and orientation between the local coordinate system and the ACS should not change if we assume the segment is rigid. This relationship can be used to estimate the position and orientation of the ACS at any point in time by tracking motion of targets attached to the segment. This idea is easily expanded to multiple segments and forms the basis for comparing relative motion between adjacent bones (i.e., ACSs).

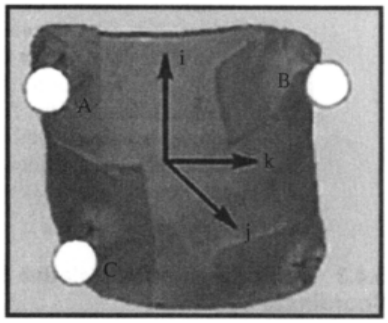

FIGURE 5.12 Three tracking targets *A*, *B*, and *C* are fastened to a contour molded shell. The shell is attached to the segment by elasticized wrap or some other convenient method. Targets *A*, *B*, and *C* are used to construct an orthogonal local coordinate system as in Eqs. (5.28) to (5.30). The origin of the local coordinate system has been drawn in the middle of the figure for convenience.

Constructing a local coordinate system for the purposes of estimating motion of the ACS is a straightforward and convenient method. However, the position and orientation of the local coordinate system is generally sensitive to errors in the coordinates of the tracking targets, and therefore, the

estimated position and orientation of the ACS will also be affected. For this reason, it is generally advantageous to use more then three targets per segment and a least-squares method to track motion of the segment and underlying bone. The singular value decomposition (SVD) method has been used to this end with good success (Soderkvist and Wedin, 1993; Cheze et al., 1995). The SVD method *maps* all of the tracking targets ($n \geqslant 3$) from position a to position b using a least-squares approximation. This is illustrated schematically in Fig. 5.13 and represented algebraically in Eq. (5.31).

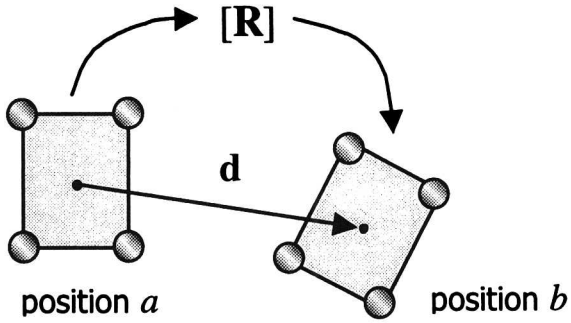

FIGURE 5.13 Least-squares mapping of the tracking targets from position a to position b. **R** is a 3 × 3 rotation matrix and **d** is a displacement vector.

$$\min \sum_{i=1}^{n} \| Ra_i + d - b_i \|^2 \tag{5.31}$$

where n represents the number of targets attached to the segment, with a_i and b_i used to indicate the object-space coordinates of the individual tracking targets. R is a 3 × 3 rotation matrix, while d is a displacement vector that, when combined with R, maps all targets in a least-squares sense from their position in a to their position in b. Because the coordinates of the tracking targets are also known relative to the ACS, the same least-squares approach can be used to determine how the ACS moved between position a and position b. Note that although this example maps the targets on the same segment, this idea can also be used to determine relative kinematics between adjacent segments (cf. Soderkvist and Wedin, 1993).

5.4.3 Joint Kinematics: Relative Motion between Adjacent Anatomical Coordinate Systems

It is clear from Fig. 5.11 that changing the knee angle will affect the relative orientation between the ACS_{shank} and the ACS_{thigh}. The orientation at any point in time can be represented by a 3 × 3 matrix of direction cosines. The nine elements of the direction cosine matrix are related to an ordered sequence of rotations about a particular set of axes. This can be visualized by starting out with the ACS_{shank} and ACS_{thigh} initially aligned, moving the ACS_{shank} into its final orientation relative to the ACS_{thigh} by rotating about the Z, Y', X'' axes of a moving reference frame. The ACS_{shank} is the moving reference in our example. The prime superscripts indicate that the orientation of the primed axes is related to a previous rotation. The first rotation in the Z, Y', X'' sequence takes place about the ML axis of the ACS_{thigh} (or equivalently about the Z axis of the ACS_{shank} because both ACSs are aligned at the onset!). The Y' axis about which the second rotation occurs is perpendicular to both the ML axis of the thigh and the longitudinal axis of the shank. This mutually perpendicular axis is often called the *line of*

nodes (or *floating axis* in joint coordinate system terminology). The line of nodes is formed by the vector cross product of the ML axis of the thigh and the longitudinal axis of the shank. The final rotation takes place about the longitudinal axis of the shank (i.e., X''). Note the double superscript indicating the orientation of the longitudinal axis has been influenced by two previous rotations about the Z and Y' axes. These ordered rotations are known as Euler Z, Y', X'' angles. The Euler angles described here are commonly reported in biomechanical studies (e.g., Grood and Suntay, 1983) because the rotations take place about clinically meaningful axes corresponding to joint flexion-extension (Z), abduction-adduction (Y'), and internal-external rotation (X?). The Z, Y', X'' sequence of rotations is expressed in matrix notation in Eq. (5.32), with the elements of the individual matrices shown in Eq. (5.33).

$$\mathbf{R} = [R_z][R_{y'}][R_{x''}] \tag{5.32}$$

$$R_{x''} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & -\sin\psi \\ 0 & \sin\psi & \cos\psi \end{bmatrix} \quad R_{y'} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$$

$$R_z = \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.33}$$

where ϕ, θ, and φ are the Euler angles about the Z, Y', X'' axes, respectively. Expanding Eq. (5.32) by using the matrices from Eq. (5.33) leads to the rotation matrix $\mathbf{R}$ in Eq. (5.34).

$$\mathbf{R} = \begin{bmatrix} \cos(\phi)\cos(\theta) & \cos(\phi)\sin(\theta)\sin(\psi) - \sin(\phi)\cos(\psi) & \cos(\phi)\sin(\theta)\cos(\psi) + \sin(\phi)\sin(\psi) \\ \sin(\phi)\cos(\theta) & \sin(\phi)\sin(\theta)\sin(\psi) + \cos(\phi)\cos(\psi) & \sin(\phi)\sin(\theta)\cos(\psi) - \cos(\phi)\sin(\psi) \\ -\sin(\theta) & \cos(\theta)\sin(\psi) & \cos(\theta)\cos(\psi) \end{bmatrix} \tag{5.34}$$

It is easy to show that a different sequence of rotations can be used to move the ACS_{shank} from its initially aligned position to its final orientation relative to the ACS_{thigh}. Because matrix multiplication is not commutative in general, the terms of $\mathbf{R}$ in Eq. (5.34) will differ depending on the sequence of rotations selected. Equations (5.35) to (5.37) can be used to determine the Euler angles for this particular sequence of rotations:

$$\phi = \arctan\left(\frac{r_{21}}{r_{11}}\right) \tag{5.35}$$

$$\theta = \arctan\left(\frac{-r_{31}}{\sqrt{r_{11}^2 + r_{21}^2}}\right) \tag{5.36}$$

$$\psi = \arctan\left(\frac{r_{32}}{r_{33}}\right) \tag{5.37}$$

where r_{ij} is the element in the ith row and jth column of matrix $\mathbf{R}$.

The methods outlined above can also be used to calculate segmental kinematics. For example, instead of calculating the relative orientation between the shank and thigh at time 1, we can use Euler angles to determine the relative orientation between the ACS_{shank} at times 1 and 2.

5.4.4 Body Segment Parameters

Re-examining Eqs. (5.17) and (5.18), we see that estimates for mass (m) and the inertia tensor (I) for each segment are required to determine the right-hand side of the equations. Several other terms,

including the location of the center of mass and the distances from the distal and proximal joint centers to the COM, are embedded in the left-hand side of Eq. (5.18). The term *body segment parameters* (BSPs) is used to describe this collection of anthropometric information.

There are essentially two approaches that are used for estimating BSP values. The more exact approach is to measure the BSP values experimentally. In practice, this is rarely done because the process is tedious, subject to error, and certainly not practical to perform for every subject. Because the BSP values are difficult to measure accurately, they are generally estimated by using anthropometric lookup tables and/or regression equations (e.g., Dempster, 1957; Winter, 1990; Zatsiorsky and Seluyanov, 1985).

For the case of a two-dimensional analysis, when motion is assumed planar, the moment of inertia in Eq. (5.18) takes on a single value. In the case of a three-dimensional analysis, I becomes a 3×3 inertia tensor. The main diagonal of the inertia tensor is constant and the off-diagonal elements vanish when the principle axis of inertia is aligned with the axes of the ACS. The diagonal matrix in Eq. (5.38) reflects this alignment and is the form used in Eq. (5.18) for a three-dimensional analysis in which the moments are expressed in the ACS of the segment.

$$\mathbf{I} = \begin{bmatrix} I_{xx} & 0 & 0 \\ 0 & I_{yy} & 0 \\ 0 & 0 & I_{zz} \end{bmatrix} \tag{5.38}$$

If we assume that each segment is a homogeneous solid of known geometry, we can use standard formulas for calculating mass moment of inertia about the X, Y, and Z axes.

5.4.5 Force Transducers

The role of a force transducer is to record external forces acting on the body. Force plates used in gait and postural studies to measure ground reaction forces are perhaps the most familiar type of force transducer used in biomechanics. A force platform is sensitive to the load a subject applies to the plate, with the plate exerting an equal and opposite load on the subject (hence the term *ground reaction forces*). Although we will limit our discussion to force plates, we wish to point out that other types of force transducers are used in the study of human movement. For example, multiaxial load cells are used to investigate the motor control of arm movements (e.g., Buchanan et al., 1993; 1998).

Commercially available force platforms use one of two different measuring principles to determine the applied load. The first type of force plate uses strain gauge technology to indirectly measure the force applied to the plate (e.g., AMTI and Bertec), while the second type uses piezoelectric quartz (e.g., Kistler). Piezoelectric materials produce an electrical charge directly proportional to the magnitude of the applied load. In this section we focus on how the output of a force platform is used in an inverse dynamics analysis, without considering how the forces and moments detected by the plate are calculated.

Force platforms are used to resolve the load a subject applies to the ground. These forces and moments are measured about X, Y, and Z axes specific to the force platform. In general, the orientation of the force platform axes will differ from the orientation of the reference axes of the object-space. This is illustrated schematically in Fig. 5.14. Thus, it is necessary that the ground reaction forces be transformed into the appropriate reference system before they are used in subsequent calculations. For example, the ground reaction forces acting on the foot should be transformed into the foot coordinate system, if ankle joint forces and moments are expressed in an anatomically meaningful reference system (i.e., about axes of the ACS_{foot}).

Another variable that must be considered is the location of the external force acting on the system. For the case of a subject stepping on a force platform, the location of the applied load is assumed to act at the center of pressure (COP). The term is aptly named, since the subject really applies a

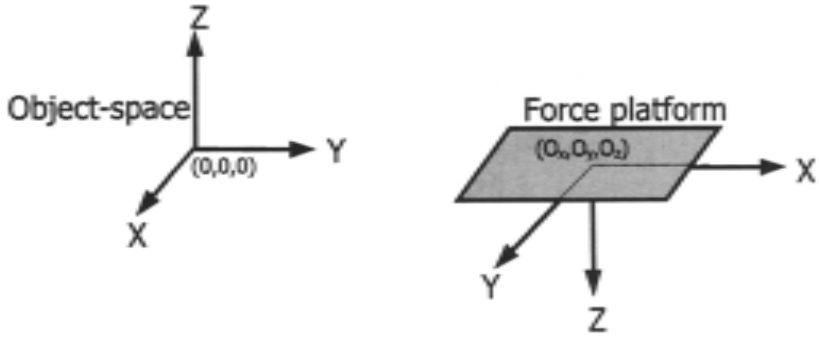

FIGURE 5.14 Relationship between the reference axes of the force platform and the reference axes of the object-space. In general, the reference axes will not be aligned. The force platform is located at a distance O_x, O_y, O_z from the origin of the object-space.

distributed pressure to the top surface of the force plate that is treated as an equivalent point force. As with the forces and moments, the location of the COP in the force platform system should be transformed into the appropriate reference system. Other devices such as pressure insoles and mats can measure pressure distributions, but are not suitable for three-dimensional motion analysis because they do not provide a complete 6 degree of freedom history of the applied load. If the data from a force platform are used in an inverse dynamics analysis, the data must be synchronized with the kinematic observations. Generally, analog data sampled from a force platform are collected at an integer multiple of the video collection rate.

5.4.6 Example: Results from an Inverse Dynamics Analysis

We close this section by presenting several examples of joint kinematic and kinetic data calculated using the methods outlined above. The ground reaction forces for a subject walking at a natural cadence are illustrated in Fig. 5.15. The vertical component of the ground reaction force (GRF) is by far the largest, with the peak AP component of the GRF next largest in magnitude. Notice how the AP force component has both a negative phase and a positive phase corresponding to *braking* and *propulsive* phases during stance. The first hump of the vertical component of the GRF corresponds to a deceleration of the whole body COM during weight acceptance (note how this corresponds with the AP braking force). The second hump in the vertical component of the GRF and the positive phase of the AP force component accelerate the body COM upward and forward as the subject prepares for push-off at the end of stance.

The curve in Fig. 5.16 is the sagittal plane ankle moment that was calculated using the data from the force platform and Eqs. (5.17) and (5.18). At the ankle, we see a small dorsiflexion moment shortly after contact. This moment prevents the foot from slapping down during initial contact with the ground (i.e., the dorsiflexion moment "pulls" the toes toward the shank). As the subject moves into the latter half of stance, a sizable plantarflexion moment is generated as a main contributor to the body's forward progression. This increase in plantarflexion moment is due to the gastrocnemius and soleus muscles contracting, essentially "pushing" the foot into the ground.

Also toward the end of the stance phase, the knee joint flexion angle increases in preparation for push-off. (Think what would happen if the knee did not flex as the leg begins to swing.) The weight acceptance period shortly after initial contact is mediated, in part, by the knee joint, which undergoes a brief period of flexion (Fig. 5.17). During this initial period of stance, the knee acts as a spring, resisting the force of impact. Hence, in Fig. 5.18 we see that a substantial knee extension moment is generated by the quadriceps muscle group to control knee flexion during this time.

Reporting the results of an inverse dynamics analysis in graphical form as we have done here demonstrates the interdependence of the kinematic and kinetic variables. The figures are helpful

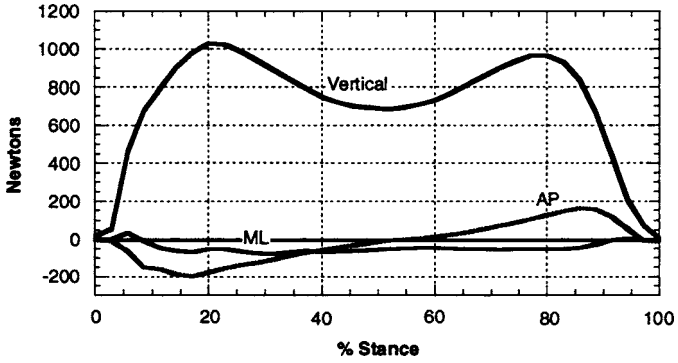

FIGURE 5.15 Ground reaction forces during the stance phase of natural cadence walking. The stance phase begins at foot strike and ends when the foot leaves the ground. ML = mediolateral, AP = anteroposterior.

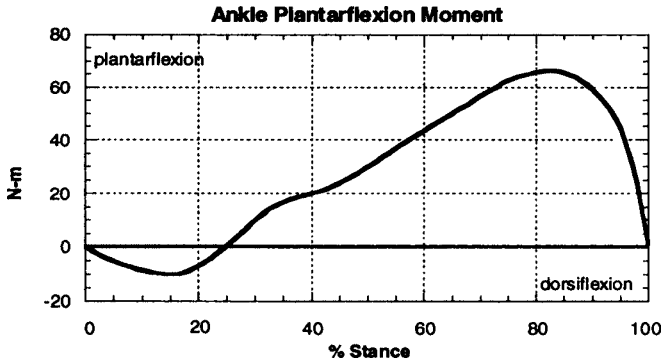

FIGURE 5.16 Sagittal plane ankle moment during the stance phase of natural cadence walking. Notice the small dorsiflexion moment during the first 20 percent of stance. This prevents the foot from slapping the ground shortly after contact. The large plantarflexion moment during the latter half of stance helps propel the body upward and forward.

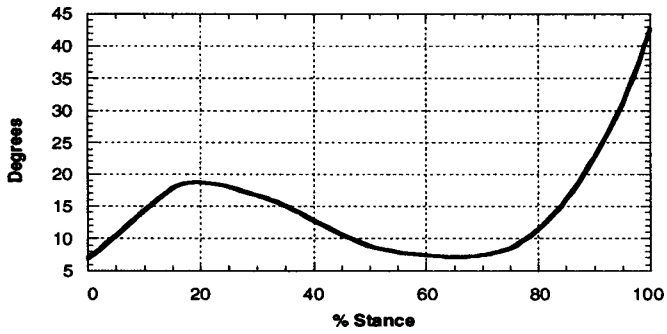

FIGURE 5.17 Knee flexion angle during the stance phase of natural cadence walking. The initial period of flexion (0 to 20 percent stance) helps absorb the shock of impact when the foot hits the ground. A second period of flexion begins at approximately 70 percent of stance, increasing rapidly in preparation for the swing phase.

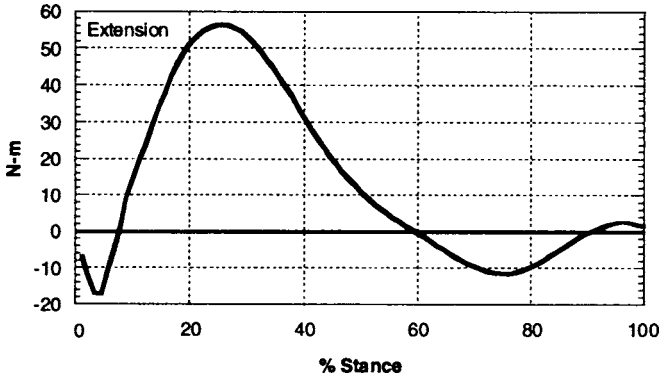

FIGURE 5.18 Knee extension moment during the stance phase of natural cadence walking. The net extension moment reveals that the quadriceps muscles are the dominant group during stance. The initial small flexion moment is caused by the vertical ground reaction force when the foot hits the ground.

when communicating with an athlete or trainer in breaking down a movement pattern to determine how performance might be improved. Inverse dynamics is also a valuable tool that is used to plan surgical treatment and assess the outcome. For example, consider the case in Fig. 5.19. The right panel depicts a varus aligned knee (i.e., *genu varum*). Genu varum is more commonly known as bow-leggedness. Because the net moment is the result of all muscles acting about the joint, it is a reasonable assumption that the medial compartment forces will be greater for the genu varum knee than forces for the normally aligned knee. These increased forces coupled with reduced joint contact

FIGURE 5.19 Examples of a normally aligned (left) and a genu varum (right) knee. A larger abduction moment and reduced joint contact area for the genu varum knee will lead to higher than normal stresses during stance.

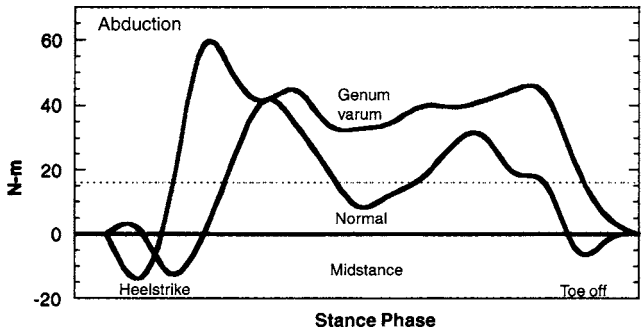

FIGURE 5.20 Stance phase knee abduction moment for a normally aligned and a genu varum knee. Note the difference during midstance. The dashed line is the postsurgical mean data for patients undergoing high tibial osteotomy to correct the genu varum. The mean postsurgical midstance abduction moment (mean = 16 N · m) has returned to a near normal level. Data taken from Weidenhielm et al. (1995).

area (the area being substantially less, as it is mostly over the medial side) will lead to greater stresses in the joint, which may predispose an individual to knee osteoarthritis. The abduction moments for an unimpaired and a genu varum knee are shown in Fig. 5.20. Note how the midstance abduction moment is significantly greater for the presurgical genu varum knee. The dashed line is the average postsurgical midstance moment for patients who underwent high tibial osteotomy surgery, which is a surgical procedure in which a wedge of bone is removed to better align the joint. Note that the midstance knee abduction moment for these subjects has returned to near normal following surgery.

These examples demonstrate how, by using human movement analysis, we can calculate the kinematics and kinetics during complex motions and how these, in turn, can be used to provide information about clinical efficacy and athletic performance.

5.5 CONCLUDING REMARKS

In this chapter we have examined forward and inverse dynamics approaches to the study of human motion. We have outlined the steps involved in using the inverse approach to studying movement with a particular focus on human gait. This is perhaps the most commonly used method for examining joint kinetics. The forward or direct dynamics approach requires that one start with knowledge of the neural command signal, the muscle forces, or, perhaps, the joint torques. These are then used to compute kinematics.

Before concluding, a brief word might be said for hybrid approaches that combine both forward and inverse dynamics approaches to meet in the middle. These methods record both the neural command signal (i.e., the EMG) and the joint position information using standard motion analysis methods as described in Sec. 5.4. The EMG is processed so as to determine muscle forces, which are then summed together to yield joint moments. These same joint moments can also be computed from the inverse dynamics. This provides a means by which to calibrate the EMG to muscle force relationships. This method has been shown to work well with gait studies (Bessier, 2000) and has great potential for studying altered muscle function associated with pathological gait, which cannot be readily examined by optimization techniques.

The biomechanics of human movement is a growing field, spanning many disciplines. As new techniques are developed and shared across these disciplines, the field will continue to grow, allowing us to peer deeper into the mechanics of movement.

REFERENCES

Abdel-Aziz, Y. I., and Karara, H. M. (1971). "Direct linear transformation from comparator coordinates into object-space coordinates," *Close-range photogrammetry*. American Society of Photogrammetry, Falls Church, Va.

Bessier, T. F. (2000). "Examination of neuromuscular and biomechanical mechanisms of non-contact knee ligament injuries," doctoral dissertation, University of Western Australia.

Buchanan, T. S., Moniz, M. J., Dewald, J. P., and Zev Rymer, W. (1993). "Estimation of muscle forces about the wrist joint during isometric tasks using an EMG coefficient method," *Journal of Biomechanics*, **26**:547–560.

Buchanan, T. S., Delp, S. L., and Solbeck, J. A. (1998) Muscular resistance to varus and valgus loads at the elbow. *Journal of Biomechanical Engineering*, **120**(5), 634–639.

Cappozzo, A., Catani, F., Della Croce, U., and Leardini, A. (1995). "Position and orientation in space of bones during movement: anatomical frame definition and determination." *Clinical Biomechanics*, **10**(4), 171–178.

Chao, E. Y. S. (1980). "Justification of triaxial goniometer for the measurement of joint rotation," *Journal of Biomechanics*, **13**:989–1006.

Cheze, L., Fregly, B. J., and Dimnet, J. (1995). "A solidification procedure to facilitate kinematic analyses based on video system data," *Journal of Biomechanics*, **28**(7), 879–884.

Delp, S. L., Arnold, A. S., Speers, R. A., Moore, C. A. (1996). "Hamstrings and psoas lengths during normal and crouch gait: implications for muscle-tendon surgery," *Journal of Orthopaedic Research*, **14**:144–151.

Dempster, W. T. (1955). "Space Requirements of the Seated Operator: Geometrical, Kinematic, and Mechanical Aspects of the Body with Special Reference to the Limbs," Technical Report (55–159) (AD 87892), Wright Air Development Center, Air Research and Development Command, Wright-Patterson Air Force Base, Ohio.

Grood, E. S., and Suntay, W. J. (1983). "A joint coordinate system for the clinical description of three-dimensional motions: application to the knee," *Journal of Biomechanical Engineering*, **105**:136–144.

Hill, A. V. (1938). "The heat of shortening and the dynamic constants of muscle," *Proceedings of the Royal Society of London Series B*, **126**:136–195.

Huxley, A. F. (1957). "Muscle structure and theories of contraction," *Progress in Biophysical Chemistry*, **7**:255–318.

Huxley, A. F., and Simmons, R. M. (1971). "Proposed mechanism of force generation in striated muscle," *Nature*, **233**:533–538.

International Society of Biomechanics (1999). *ISB software sources*, http://isb.ri.ccf.org/software.

Jackson, K. M. (1979). "Fitting of mathematical functions to biomechanical data," *IEEE Transactions on Biomedical Engineering*, BME, **26**(2):122–124.

Pandy, M. G., and Zajac, F. E. (1991). "Optimal muscular coordination strategies for jumping," *J. Biomech.*, **24**:1–10.

Shiavi, R., Limbird, T., Frazer, M., Stivers, K., Strauss, A., and Abramovitz J. (1987). "Helical motion analysis of the knee—I. Methodology for studying kinematics during locomotion," *Journal of Biomechanics*, **20**(5): 459–469.

Soderkvist, I., and Wedin, P. (1993). "Determining the movements of the skeleton using well-configured markers," *Journal of Biomechanics*, **26**(12), 1473—1477.

Weidenhielm, L., Svensson, O. K., and Brostrom, L-A. (1995). "Change of Adduction Moment About the Hip, Knee, and Ankle Joints After High Tibial Osteotomy in Osteoarthrosis of the Knee," *Clinical Biomechanics*, **7**:177–180.

Winter, D. A. (1990), *Biomechanics and Motor Control of Human Movement*, 2d ed., John Wiley & Sons, New York.

Woltring, H.J. (1986). "A FORTRAN package for generalized, cross-validatory spline smoothing and differentiation," *Advances in Engineering Software*, **8**(2):104–113.

Zahalak, G. I. (1986). "A comparison of the mechanical behavior of the cat soleus muscle with a distribution-moment model," *Journal of Biomechanical Engineering*, **108**:131–140.

Zahalak, G. I. (2000). "The two-state cross-bridge model of muscle is an asymptotic limit of multi-state models," *Journal of Theoretical Biology*, **204**:67–82.

Zajac, F. E. (1989). "Muscle and tendon: properties, models, scaling and application to the biomechanics of motor control," in Bourne, J. R. (ed.), *Critical Reviews in Biomedical Engineering*, **17**, CRC Press, pp. 359–411.

Zatsiorsky, V., and Seluyanov, V. (1985). "Estimation of the mass and inertia characteristics of the human body by means of the best predictive regression equation," in Winter, D., Norman, R., Wells, R., Hayes, K., and Patla, A. (eds.), *Biomechanics IX-B* (pp. 233–239), Human Kinetics, Champaign, Ill.