CHAPTER 2

Large near-surface anomalies

In the analysis of a reflection seismogram recorded on land, two regions of the near-surface are of interest: the region beneath the source that generated the seismic signal, and the region beneath the receiver that recorded it. The estimation of near-surface anomalies thus depends directly on the geometry (physical layout) of a seismic survey.

In this chapter, I introduce the geophysical problems arising from large near-surface anomalies. Because analysis of the near-surface is interwoven with survey geometry, I begin this introduction by illustrating a typical reflection seismic experiment performed over an irregular near-surface. Residual statics estimation algorithms (including the one presented in this thesis) typically analyze seismic data in several domains. Using a field-data example with large statics, I illustrate three of the most important domains: field profiles, common midpoint gathers, and common midpoint stacks. Computational problems resulting from large near-surface variations are then described, followed by a brief introduction to the new technique presented in this thesis.

I define geophysical concepts and terms as they pertain to residual statics analysis, but this coverage is only brief. Further explanations of terminology can be found in Sheriff (1973). More details of the basic concepts in reflection seismology are in Claerbout (1985), Waters (1981), and Dobrin (1976). Taner et al. (1974) give an historical overview of the statics problem.

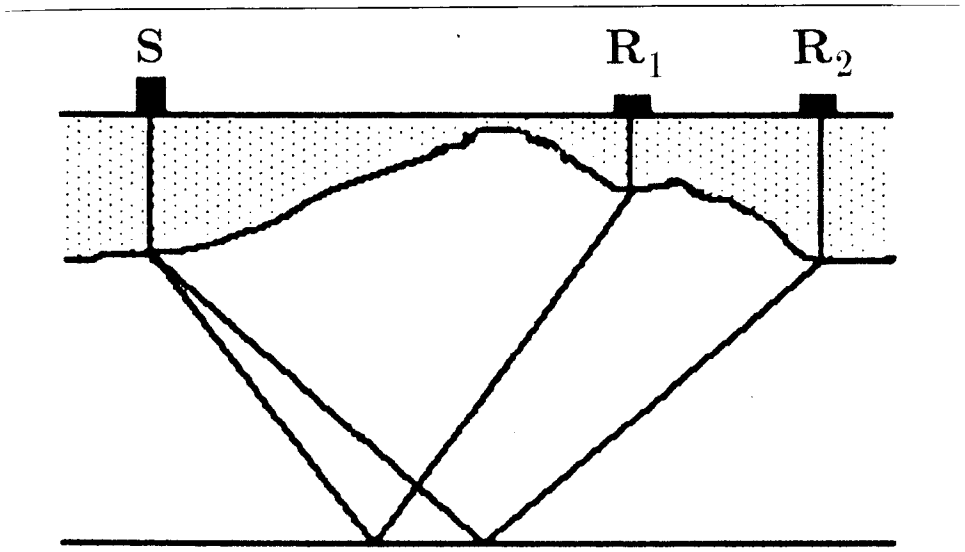

FIG. 2.1. A schematic diagram of a reflection seismic experiment. A shot, **S**, placed on (or near) the surface, excites seismic waves that propagate through the Earth. The waves are reflected upward at subsurface interfaces between different types of rock; the reflections are recorded by a line of receivers. Here, a single reflecting horizon causes reflections that are recorded at two receivers, $\mathbf{R}_1$ and $\mathbf{R}_2$. The near-surface, represented by the stippled region, is typically composed of inhomogeneous, unconsolidated soil and rock. Because the seismic velocity of the near-surface is usually low compared to that of underlying layers, raypaths in the near-surface are assumed to be vertical.

2.1 STATICS OBSERVED: FIELD PROFILE

Figure 2.1 shows a simplified view of a reflection seismic experiment. A source, or shot, excites seismic waves that propagate through the Earth. The waves are reflected upward at subsurface interfaces between different types of rock; the reflections are recorded by a line of receivers. In this figure, raypaths are drawn such that rays always travel vertically through the near-surface. This vertical-raypath assumption is commonly used in the analysis of near-surface anomalies, and has ample justification: the seismic velocity of the near-surface layer is usually lower than the velocity of underlying sediments; thus, because of refraction, raypaths are nearly vertical in the near-surface (Taner et al., 1974). The surface-consistent model for statics corrections, discussed in §2.3, is based on this assumption.

Throughout this thesis, two characteristics of the near-surface are of interest: lateral variations in thickness and lateral variations in seismic velocity. For each

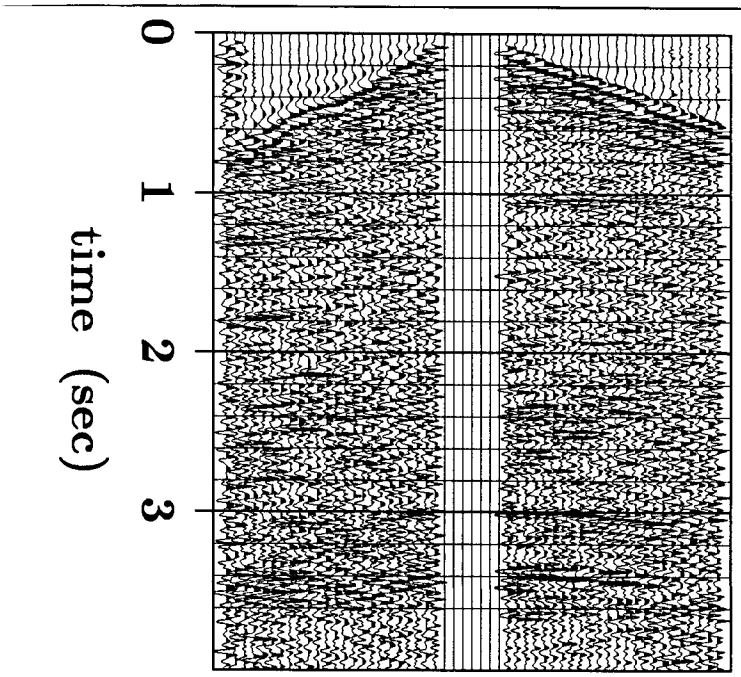

FIG. 2.2. Field profile 37 from a seismic survey performed in the Wyoming Overthrust belt. There were 24 receivers on each side of the source. The reflections are roughly hyperbolic on the right side, but the left side is distorted because of near-surface anomalies. These distortions can be characterized as simple time delays, or bulk shifts, of the individual seismograms.

recorded seismogram, the net effect of these two physical characteristics is a timing delay (either positive or negative) relative to the seismogram that would have been recorded if the near-surface were laterally homogeneous. The timing delay of each seismogram is due both to the source location at which the signal originated and the receiver location at which the seismogram was recorded.

A typical seismic survey is composed of hundreds of seismic experiments. Each experiment is performed with approximately 50-100 receivers spaced evenly along a line that is usually a few kilometers long. An example of a field profile, the result of a single seismic experiment, is shown in Figure 2.2. These data were acquired in the Wyoming Overthrust belt. In this experiment there were 24 receivers on each side of the source. On the right side of the profile there is little indication of near-surface irregularities: the first arrivals form a straight line and the reflections are roughly hyperbolic. On the left side, however, both the first arrivals and the reflections are distorted. The computational results in Chapter 4 show that these distortions are due to timing delays induced by a variable near-surface. Note now, however, that these distortions appear to be similar for each reflector, regardless of its depth. When distortions are constant

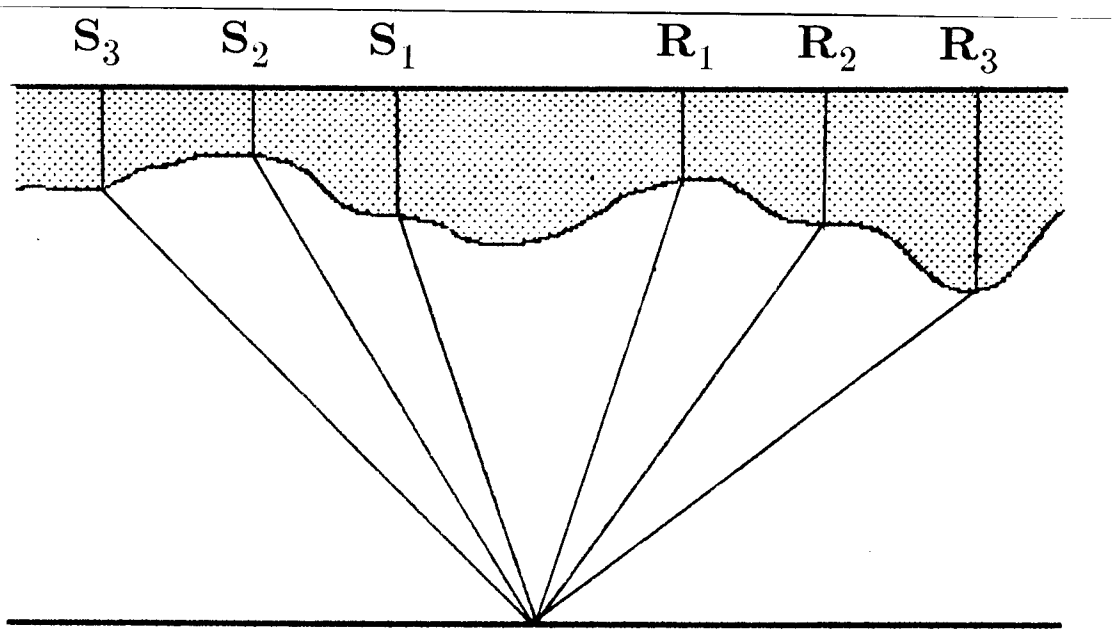

FIG. 2.3. Schematic diagram of a 3-trace common midpoint (CMP) gather. The three raypaths are taken from three different field profiles. Because reflection points are nearly the same for all traces in a CMP gather, the traces are later stacked to improve the signal-to-noise ratio. Stacking is effective, however, only if near-surface variations have been estimated and corrected.

with time, the near-surface is the likely culprit. The distortions are called *statics* because they are time-invariant.

2.2 STATICS OBSERVED: COMMON MIDPOINT GATHERS AND STACKS

Because many closely spaced profiles are collected in a typical survey, there is usually a high degree of redundancy in seismic data. Reflection seismologists exploit this redundancy by grouping all traces (seismograms) with the same midpoint between shots and receivers, and then averaging the traces in each group associated with a common midpoint (Waters, 1981). Figure 2.3 is a schematic diagram of a 3-trace common midpoint (CMP) gather. For a horizontally stratified Earth with laterally invariant velocities, subsurface reflection points are the same for all traces within a CMP gather; for gently dipping layers and mild lateral velocity variations this will also be approximately true.

The traces within a CMP gather are generally assumed to be identical except for uncorrelated random noise and traveltime variations due to differences in raypaths. In practice, the latter variation is typically divided into two components: hyperbolically

increasing traveltime due to linearly increasing offset (the distance between a shot and a receiver), and time delays due to near-surface anomalies. In this thesis, it is assumed that traveltime differences due to changes in offset have been approximately removed by normal-moveout (NMO) correction (Waters, 1981), but that the effects of a variable near-surface, the residual statics, are unknown. In other words, it is assumed that the seismic velocity is approximately known everywhere except at the near-surface. This is also a typical assumption in the analysis of near-surface anomalies (Taner et al., 1974).

After the traces in CMP gathers have been NMO-corrected, the data are then stacked (summed) over offset to enhance the coherent signal relative to the noise. If near-surface effects have not been identified and corrected, however, stacking can actually degrade the signal, rather than enhance it.

Two NMO-corrected CMP gathers and a stacked section are shown in Figure 2.4. These data are from the same Wyoming survey from which the field profile in Figure 2.2 was taken. Each gather contains the seismograms from 24 different experiments; the stack represents the summed components of 171 different CMP gathers. Stacked sections should yield a readily interpretable view of the Earth. In this stack, a deep reflector appears at approximately 3.5 s. The reflection appears clearly on the far left side and at the center of the seismic section. Elsewhere there is a lack of spatial continuity. This discontinuous appearance might be indicative of the subsurface geology; evidence from the field profile in Figure 2.2, however, suggests that near-surface anomalies may be the cause of the reflector's distorted appearance. More compelling evidence of near-surface variations is seen in the CMP gathers. Both gathers were NMO-corrected with the same subsurface velocity function, but reflection events are horizontally aligned only in CMP gather 118. In contrast, the events in CMP 34 dip steeply upward with increasing offset. Because the traces in CMP 118 are horizontally aligned, trace 118 in the stack contains sharply defined reflection events; this result is expected from stacking. Trace 34, however, appears to contain only noise; this degradation of the signal is the effect of near-surface anomalies.

In Chapter 4, I present a detailed analysis of the influence of the near-surface on these data. Now, however, I will show only a result of that work. Figure 2.5 shows the data of Figure 2.4 after statics corrections have been estimated and applied. The reflections in the stack are now spatially continuous, and the reflections in both gathers are, as they should be, flat. Chapter 4 presents further evidence that the stack in Figure 2.5 is indeed more representative of the Earth than is the stack in Figure 2.4. Simply note now, however, that the stack after statics corrections presents a view of the subsurface that is clear and ordered relative to the disordered stack before statics

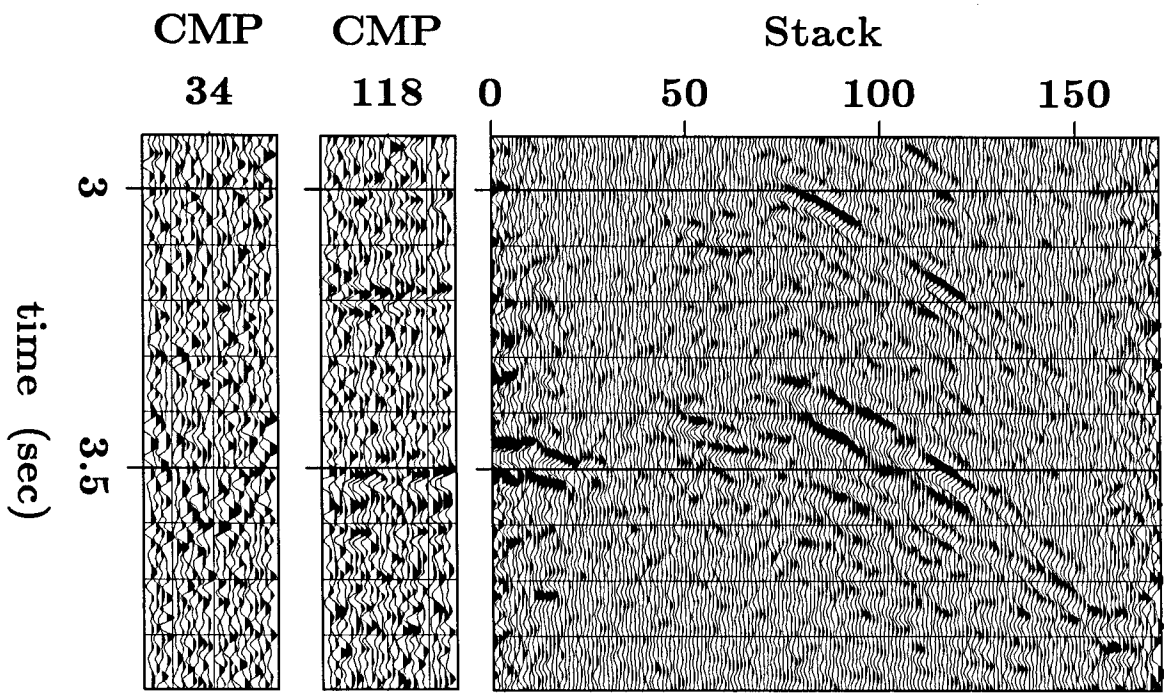

FIG. 2.4. NMO-corrected CMP gathers 34 and 118 and stacked section from the Wyoming Overthrust data. The data are displayed before statics corrections. In each gather, offset increases from the center outwards. The traces in CMP 118 are nearly horizontally aligned; thus trace 118 in the stack contains sharply defined reflections. Because of near-surface anomalies, however, the traces in CMP 34 are grossly misaligned; thus trace 34 in the stack appears to contain only noise.

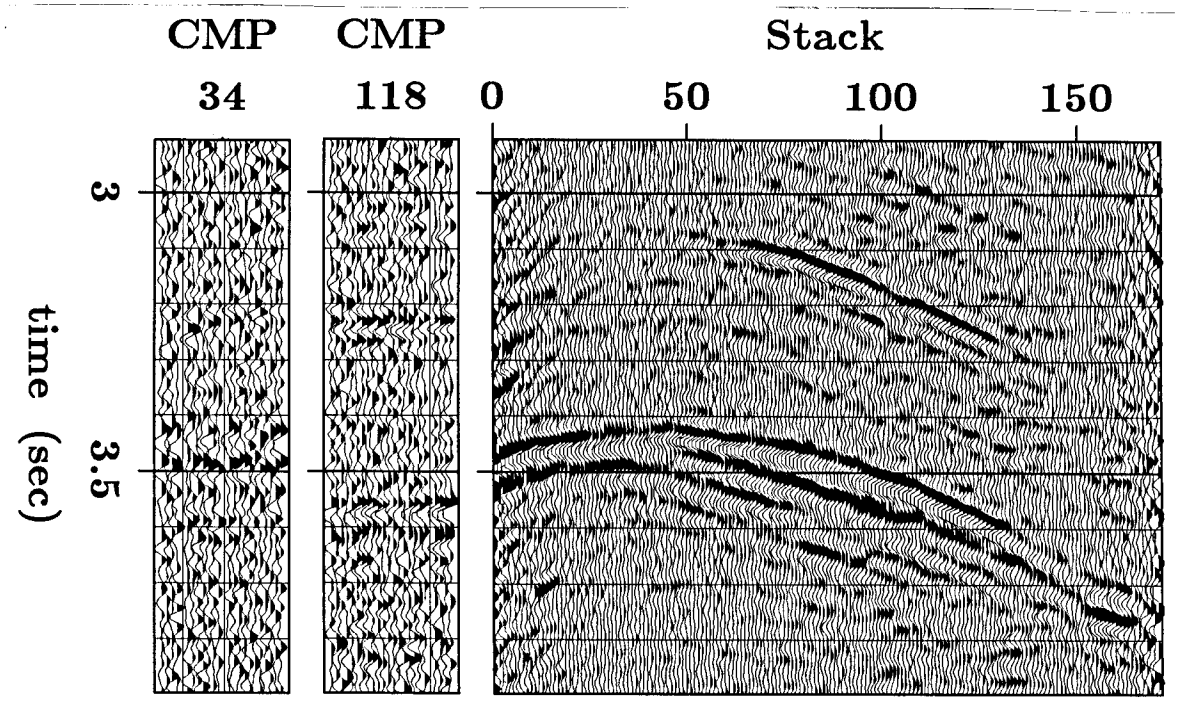

FIG. 2.5. NMO-corrected CMP gathers 34 and 118 and stacked section from the Wyoming Overthrust data. The data are displayed after statics corrections. The reflections in both gathers are now flat, and the reflections in the stack are sharply defined and spatially continuous.

corrections. This notion of disorder versus order arises later in analogy to crystal growth.

2.3 SURFACE-CONSISTENCY

Although the two pairs of CMP gathers in Figures 2.4 and 2.5 differ only by time shifts, the time delays are constrained to be *surface-consistent.* The concept of surface-consistency arises from the following model (Taner et al., 1974). As shown in Figures 2.1 and 2.3, all rays are assumed to travel vertically through the near-surface. Thus the effect of an anomalous region of the near-surface underlying a particular shot can be characterized as a simple time delay of the signal that emanated from that shot, regardless of where the signal was recorded. Likewise, the same characterization is true for all signals recorded at a given receiver, regardless of where the signals originated. Thus the static shift that is applied to any trace is regarded as the sum of a shot term, called a *shot static*, and a receiver term, called a *receiver static.* Other traces associated with the same shot or receiver are also shifted in accordance with the same shot or receiver static.

Surface-consistency is a powerful constraint: it insures that any statics solution conforms to a geophysically plausible model of the near-surface. Surface-consistent statics are essentially individual, trace-to-trace time delays that have been averaged. When a particular shot static is estimated, the averaging is performed over all receivers that recorded that shot. Likewise, the estimation of a receiver static implies averaging over all shots that initiated signals recorded at that receiver. In this thesis, all models of the near-surface are surface-consistent.

2.4 CYCLE-SKIPS AND THE ESTIMATION OF TIME DELAYS

Accurate estimation of shot and receiver statics depends on accurate estimation of trace-to-trace time delays. Although surface-consistent averaging reduces errors in individual estimates of time delays, this averaging is effective only when the errors are reasonably small.

Time delays between seismograms are usually estimated by crosscorrelation, a measure of similarity (Taner et al., 1974). The *normalized crosscorrelation* $R(\tau)$ between two discretely sampled traces $f(t)$ and $g(t)$ is defined by

$$R(\tau) = \frac{\sum_t f(t)\, g(t+\tau)}{\left(\sum_t f^2(t) \sum_t g^2(t) \right)^{1/2}} .$$

$R(\tau)$ varies between 1 and -1. The greater its value, the more similar $f(t)$ is to $g(t+\tau)$.

In conventional approaches to residual statics analysis (e.g. Wiggins et al., 1976; Taner et al., 1974), $g(t)$ is an NMO-corrected, unstacked trace and $f(t)$ is an artificially constructed "reference" trace. The reference trace is assumed to differ from the unstacked trace by only a time shift and uncorrelated noise. Thus the value of τ that yields the maximum $R(\tau)$ is considered to be the best estimate of the true time delay.

One of the easiest ways to construct a reference trace is to stack all traces in the CMP gather that includes $g(t)$, except for trace $g(t)$ itself. Figure 2.6 illustrates an example of such a construction. The trace labeled $g(t)$ is the twenty-fourth trace of CMP gather 118 in Figure 2.4. Trace $f(t)$ is the stack of the remaining 23 traces in that gather. The crosscorrelation of these two traces is shown in Figure 2.7. Because $f(t)$ and $g(t)$ are similar (except for a small time shift), one peak of the crosscorrelation function is considerably greater than the others. This peak occurs at –8 ms. Although the "true" time delay of $f(t)$ can still be debated, the dominance of a single peak suggests that a time delay of 8 ms is a reliable estimate.

Reliable estimates of time delays are difficult to obtain, however, when a CMP gather is severely affected by statics. This is illustrated in Figures 2.8 and 2.9. In Figure 2.8, the trace labeled $g(t)$ is the seventeenth trace of CMP gather 34 in Figure 2.4; $f(t)$ is the stack of the remaining 23 traces in that gather. Because CMP gather 34 is grossly distorted by statics, the stacked trace $f(t)$ is not similar to the unstacked trace $g(t)$; thus the crosscorrelation function in Figure 2.9 has several peaks of nearly equal height. Although the greatest peak occurs at –144 ms, this lag is not a reliable estimate of the true time delay. The peaks at –24 ms or 160 ms could be chosen with almost equal reliability; thus the risk of a gross error is considerable. Gross errors due to picking the wrong peak are called *cycle-skips*, because the magnitude of these errors typically occurs in multiples of the dominant period of the seismograms being crosscorrelated.

Figure 2.10 shows how cycle-skips distort seismic data. The data shown here resulted from the application of a statics solution that was obtained using an algorithm that iteratively chose the greatest peaks in surface-consistently averaged crosscorrelation functions. The input data were the Wyoming Overthrust data of Figure 2.4, the same data that led to the previous statics solution (using a different algorithm) in Figure 2.5. Because the reflection events in CMP 118 are nearly aligned before statics

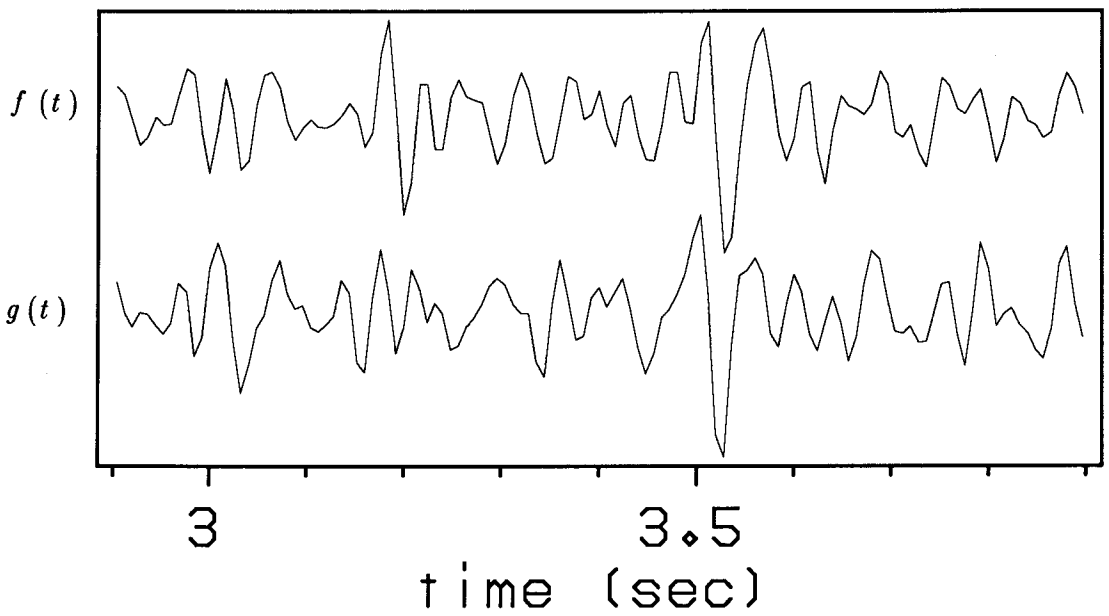

FIG. 2.6. Two seismograms. The trace labeled $g(t)$ is the twenty-fourth trace of CMP gather 118 in Figure 2.4; $f(t)$ is the stack of the remaining 23 traces in that gather.

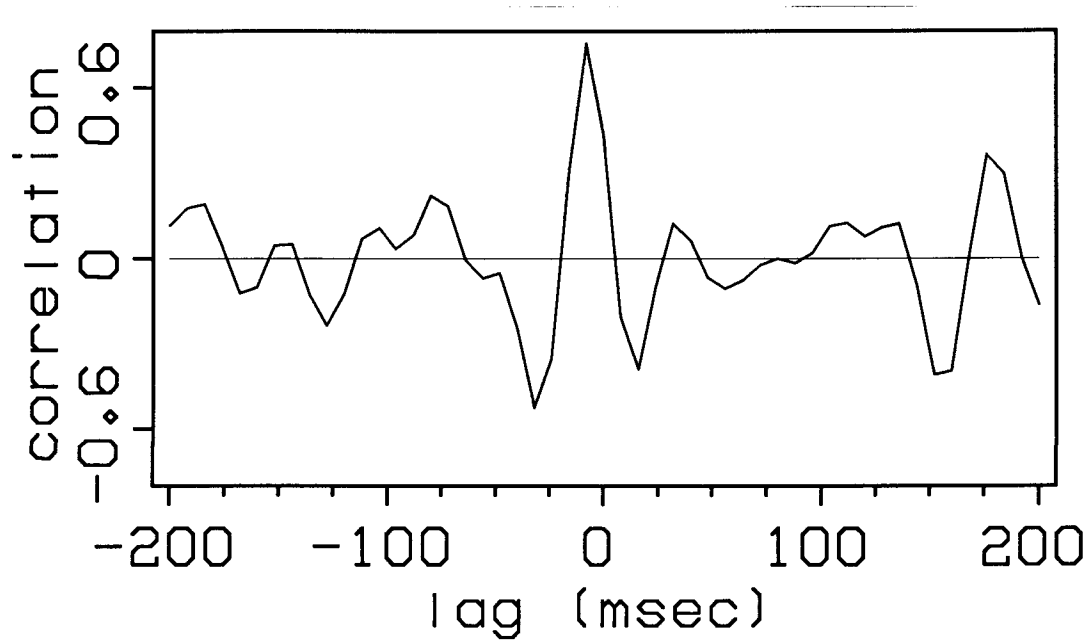

FIG. 2.7. Normalized crosscorrelation of the two traces in Figure 2.6. The lag interval is 8 ms. Because the two traces are similar (except for a small time shift), one peak in the crosscorrelation function is considerably greater than the others. The greatest peak occurs at the first negative lag, –8 ms. Because this peak strongly dominates the crosscorrelation function, 8 ms is a reliable estimate of the time delay.

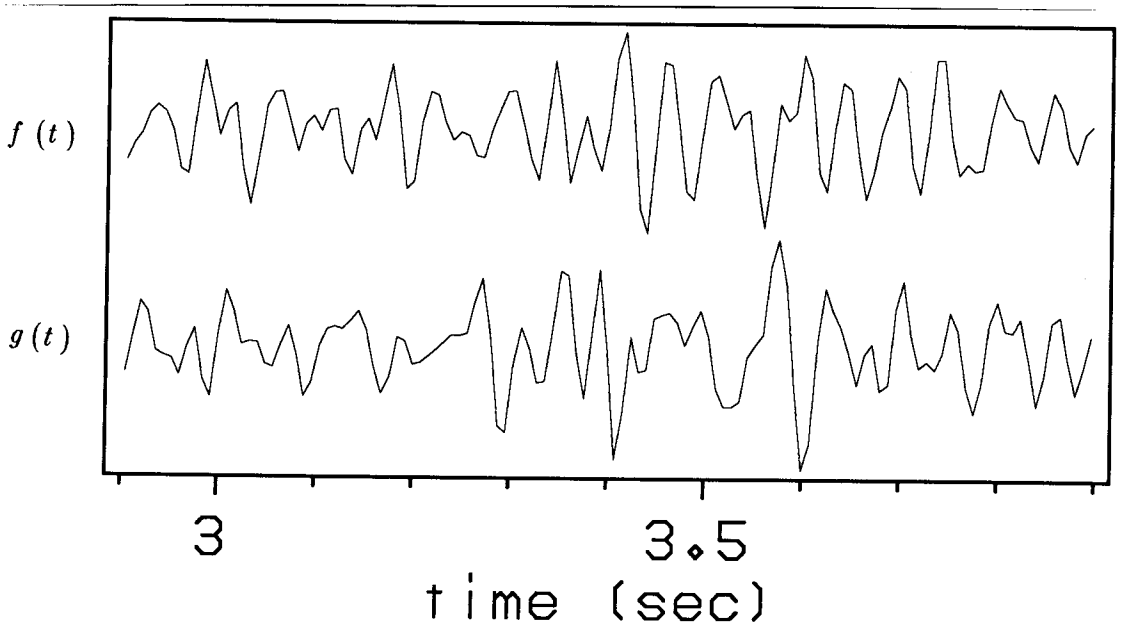

FIG. 2.8. Two seismograms. The trace labeled $g(t)$ is the seventeenth trace of CMP gather 34 in Figure 2.4; $f(t)$ is the stack of the remaining 23 traces in that gather. The two traces are not similar because the statics in CMP gather 34 caused the stacked trace to be distorted.

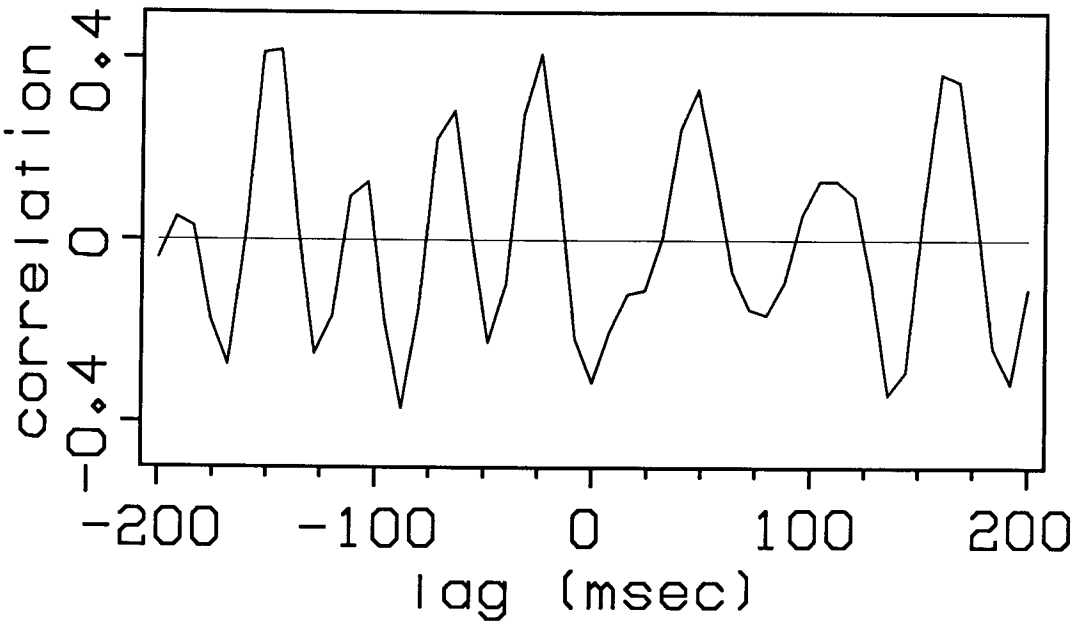

FIG. 2.9. Normalized crosscorrelation of the two traces in Figure 2.8. Because the two traces are not similar, this crosscorrelation function contains several peaks of nearly equal height. Although the greatest peak occurs at -144 msec, this lag is not a reliable estimate of the true time delay. The peaks at -24 ms and 160 ms can be chosen with almost equal reliability; thus the risk of a gross error is considerable. Gross errors due to picking the wrong peak are called *cycle-skips*.

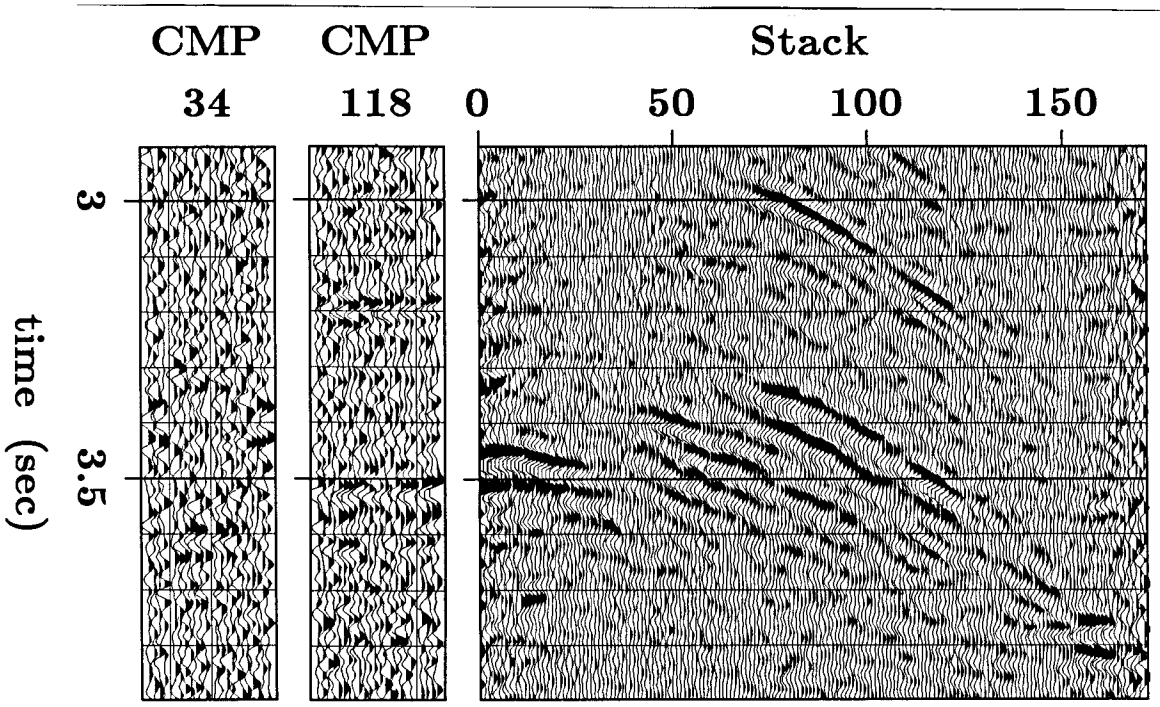

FIG. 2.10. CMP gather 34 and stacked section from the Wyoming Overthrust data. The data are displayed after statics corrections; the statics were estimated using an algorithm that iteratively chose the greatest peaks in surface-consistently averaged crosscorrelation functions. Because the near-surface anomalies are large, many of the crosscorrelation functions were similar to the one in Figure 2.9; thus many gross errors, or cycle-skips, occurred. The cycle-skips are evident in the gather, where a reflection that should be at 3.5 s has been split: the reflection in the outer traces is at approximately 3.4 s, and the reflection in the inner traces is at approximately 3.6 s. Because this behavior is repeated in many gathers, the resulting stack is a confused combination of many discontinuous reflections.

corrections, the statics in this region are small, and the crosscorrelations of these traces yielded reliable estimates of time delays. Thus the traces in CMP 118 are well aligned after statics corrections. In CMP 34, however, the traces are poorly aligned before statics corrections. Because there are large statics, crosscorrelations of these traces were ambiguous; thus reliable estimates of time delays could not be obtained. The results of these ambiguities are cycle-skips. In CMP 34, cycle-skips have caused the deep reflection to lie on two different planes: the near-offset traces align at 3.6 s, while the far-offset traces align at 3.4 s. The stack in the region near trace 34 also displays confusingly discontinuous events.

From both geologic and geophysical points of view, the stack in Figure 2.10 is inferior to the stack in Figure 2.5. In the next chapter, I show how the stack-power criterion of Ronen and Claerbout (1985) can be used to quantitatively distinguish

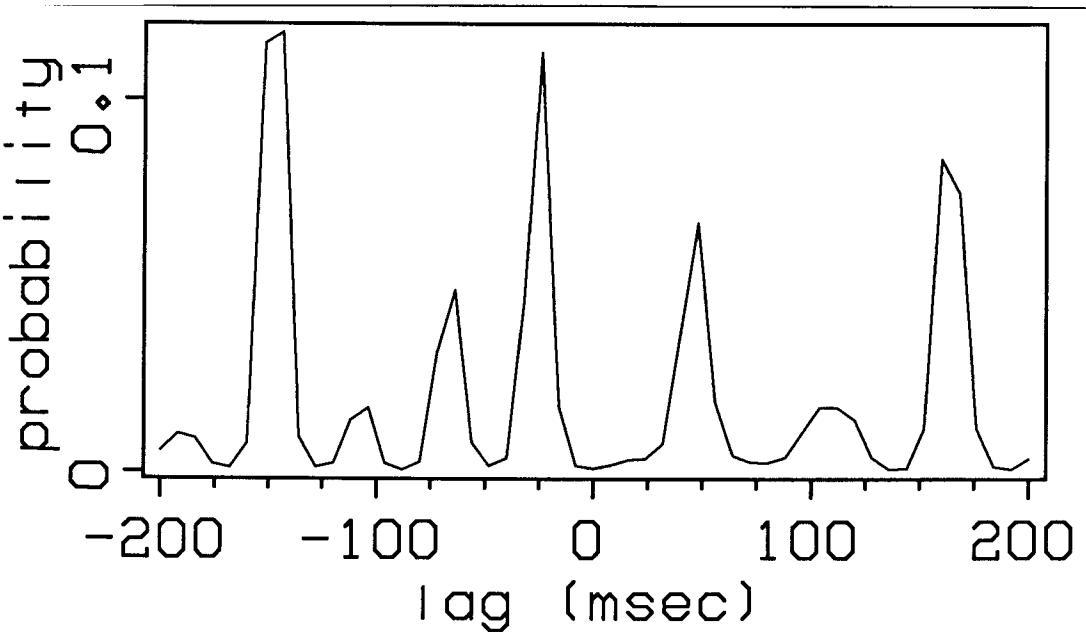

FIG. 2.11. A probability distribution, obtained by a transformation of the crosscorrelation function in Figure 2.9. The modes (maxima) of this distribution correspond to the peaks of the crosscorrelation function.

between these two stacks. Note now, however, the qualitative differences among the stacks in Figures 2.4, 2.5, and 2.10. The input stack in Figure 2.4 presents the most confused or most disordered view of the Earth, while the stack in Figure 2.5 presents the simplest, most ordered view. Figure 2.10 falls somewhere in the middle of this continuum from disorder to order.

2.5 A PROBABILISTIC PERSPECTIVE

Figures 2.9 and 2.10 raise a key question addressed by this thesis: If confronted with an ambiguous crosscorrelation function, which peak should a statics algorithm pick? I argue from a probabilistic point of view. Because an ambiguous crosscorrelation function yields little definitive information, I use it only as an indicator of the probability of the time delay, rather than letting the greatest peak wholly determine that time delay.

Figure 2.11 shows a probability distribution obtained by a transformation of the crosscorrelation function in Figure 2.9. The mathematical transformation is described in detail in the next chapter; note now, however, the similarity of the peaks and troughs in both figures. Random numbers chosen from this probability distribution are most likely to be the lags that yield the peaks in the crosscorrelation function; other

lags are assigned lower probabilities.

The most efficient form of the statics estimation algorithm I describe in subsequent chapters transforms surface-consistently averaged crosscorrelation functions to probability distributions. Estimates of time delays are randomly drawn from these distributions, and reference traces are constructed using the latest estimates of the shot and receiver statics. New crosscorrelation functions and probability distributions are then computed, and the process is repeated iteratively. As the reference traces improve, the probability distributions assign greater probabilities to the true time delays. Convergence is achieved when a stable solution is obtained. The result in Figure 2.5 was obtained using this method.