Chapter 2
The battle against noise in physics

In this chapter and in the next we discuss what we believe is the main difference between physics and the social sciences, namely the high level of noise which spoils and often altogether forbids the measurement of many social variables. This could seem a fairly unconventional view. Usually many other differences are underscored, for instance the role of individual freedom, the diversity and complexity of human behavior and so on. More specifically it is often stated that (i) social systems are inherently more complicated in the sense that societies comprise many different kinds of agents and interactions (ii) physical phenomena can be observed in the laboratory through experiments which can be repeated whereas social observations, it is argued, cannot be repeated. We do not mean that these differences do not exist but we believe that they are not as fundamental as one may have the impression at first sight.

The first point can be readily dealt with by noting that the level of complexity depends on how much detail one wants to take into account. Seen at molecular level, the dissolution of a piece of sugar in a cup of tea is a phenomenon of horrendous complexity in the sense that it involves many kinds of molecules and a great diversity of interactions. In contrast, for many social phenomena the precise nature of the agents is unimportant. For instance, the field of demography focuses on the numbers of birth and deaths, a perspective in which most other social features become irrelevant.

The second argument may at first seem to be rock solid but it must be considered in the light of the two following remarks.

- Strictly speaking, an experiment cannot be repeated even in physics. Consider for instance the swing of a pendulum which is one of the simplest experiments one can think of. If we measure its period of oscillation with a precision of 0.1 second, two successive experiments may indeed give identical results, say 1.2 second. However, if we are able to make high precision measurements with an accuracy of 10^{-6} second, successive measurements will lead to different results, say 1.200123, 1.200759 and 1.200023. There may be several reasons explaining these fluctuations. One can for instance invoke drafts, vibrations of the ceiling of the building to which the pendulum is attached. These vibrations are due to many factors: people walking at the same floor or at the floor above, vibrations due to trucks, trains or even planes in the vicinity of the building. These are high frequency shocks but there are also low frequency perturbations. Suppose for instance that the experiment is repeated a few hours later; due to the rotation of the Earth on its axis the positions of the Sun and Moon with respect to the pendulum will no longer be the same; this will modify the directions of their gravitational forces which will slightly change the period[1]. In short, the question of whether or not it is possible to reproduce an experiment really depends on the level of accuracy that one wishes to achieve.

- To complete our argument, it must now be shown that in the social sciences it is indeed possible to repeat quasi-experiments in a way which is similar to what is done in physics. For definiteness and in order to make the discussion more concrete we consider a specific phenomenon, the so-called Werther effect. According to this effect some people are induced to commit suicide when they learn of other people having committed suicide. This effect was first studied by David Phillips some thirty years ago in an influential paper (Phillips 1974)[2]. More specifically, Phillips analyzed the fluctuations in monthly numbers of suicides in the United States in the month following the announcement of a suicide on the first page of the *New York*

[1] In technical terms this is called a tidal effect.

[2] It should be noted that in spite of numerous (but non-converging and non-cumulative) studies, the reality of this effect has not yet been established with certainty. We will come back to this point in the next chapter. Here we use this effect as an example of a quasi-experiment and the reality of the effect is irrelevant for the present discussion.

Times. For the time being let us accept Phillips's result that on average there is a 2.5 percent increase in suicides following the publication of suicide stories as compared to months in which no suicide story was published. In a system theory perspective there is a close parallel between the Werther effect and the swing of a pendulum. This parallel is schematically illustrated in Fig. 2.1. Obviously, the Werther experiment can be repeated as often as one wishes. It can be repeated in any year between 1851 (when the *New York Times* began to be published) and now. It can also be repeated in any country in which there is a newspaper similar to the *New York Times*. The *Times* of London or many other major newspaper published in European countries can be used in the same way which means that the experiment can be repeated in several different countries.

There is however a major difference between the swing of a pendulum and the Werther effect. For the former the signal to noise ratio is much larger than one (ratios exceeding one hundred are not difficult to obtain), whereas for the Werther effect it is smaller than one.

Throughout the history of physics and chemistry, signal enhancing and noise reduction have been (and still are) permanent concerns. It is natural therefore to draw on this knowledge and this is why we first consider two examples of physical phenomena. The first one is the familiar pendulum that we selected for its simplicity. In spite of this simplicity, it is a source of useful principles and guidelines. The second example, the detection of gravitational waves, is of interest because it is still awaiting a satisfactory solution. For this reason it provides a useful parallel with social science problems.

1 Improving the signal to noise ratio in the pendulum experiment

The signal to noise ratio Y/N can be raised either by increasing the signal Y or by reducing the noise N[3]. This could seem a fairly trivial remark but it has far reaching consequences especially for social phenomena. In the case of the pendulum the step of increasing the signal is done almost without thinking about it. Nobody would give the pendulum an initial deviation of only two or three degrees. It is clear that with such a small amplitude, the movement would be much influenced by friction at the articulation point or by air turbulence. On the contrary, if the pendulum oscillates with an amplitude greater than, say 40 degrees, its energy is large with respect to the energy of perturbing factors, its movement will be more regular and the measurement of its period will be more accurate. This method of raising the signal to noise ratio plays an important role in different chapters of this book; for the sake of brevity we call it the *extreme value technique*. Naturally, it has inherent limitations. In the case of the pendulum the initial deviation cannot be made larger than 180 degrees[4], or if one wishes to study the variations of the period as a function of the initial angle, small oscillations cannot be avoided. This raises the question of whether there are other noise reduction techniques. From our previous discussion of the pendulum experiment we know that there are high-frequency as well as low-frequency sources of noise. How is it possible to reduce their impact on the movements of the pendulum? In the following lines we examine several kinds of means which can serve this purpose.

- If the experiment is performed in a vacuum container the mass of the pendulum will no longer be subject to high frequency fluctuations due to the air turbulence.

- If in addition the experiment is performed overnight when the building and its

[3] Strictly speaking, the signal to noise ratio is rather (in the notations of Fig. 2.1) $H(X)/\sigma(N)$. However, when the noise N is of the same magnitude as the signal $H(X)$ it is difficult to know if there is really a signal; this is why we often replace $H(X)/\sigma(N)$ by $Y/\sigma(N)$. Such a substitution is particularly welcome for the Werther effect or for the detection of gravitational waves. For a system subject to high levels of noise $H(X)/\sigma(N) \ll Y/\sigma(N) \sim 1$; for a system subject to a low level of noise, $Y/\sigma(N) > H(X)/\sigma(N) \gg 1$.

[4] Initial deviations of more than 90 degrees require of course a pendulum with a rigid arm.

surroundings are much quieter than during day time, the fluctuations due to exogenous vibrations will be greatly reduced.

- The exogenous influence of the Sun and Moon cannot be eliminated; however, if the experiment is performed at times when the Sun and Moon [5] are in the same configuration, this effect can be reduced. Alternatively, as we know the strength of the gravitational forces and the way in which they influence the movement of the pendulum it is possible to compute appropriate corrections. Naturally, these corrections can only be made because we know the laws which govern this system. In the time of Galileo it would have been impossible to perform such calculations.

All these improvements will raise the signal to noise ratio and reduce the error bars[6]. Naturally, such high accuracy experiments have little in common with standard class room experiments. What lessons can be drawn from this simple example? First, it must be stressed that the noise has been cut off at its source. This is an important point. In economics and in the social sciences the standard methodology is to select the data without giving much attention to the issue of noise reduction, and then to subject these noisy data to statistical analysis in the hope that a signal will emerge[7]. Unfortunately, no statistical treatment can substantially improve the signal to noise ratio; it is only by enlarging the data set and by making use of additional information pertaining to the system under consideration that the signal to noise ratio can be raised markedly. The pattern matching technique that we examine below is an example of this kind. More generally, the pendulum example makes us realize that

[5] The planets, especially Jupiter, may also have an influence but it is much smaller.

[6] In the previous enumeration we mentioned only hardware improvements in the experiment itself. Naturally, for any experimental device the precision can be increased by making a large number of measurements and taking the average; the averaging technique will be considered in more detail later on.

[7] This assertion can be illustrated by the example of Phillips's paper about the Werther effect. Among the 33 suicide stories that he selected are cases as different as American actresses (Carole Landis and Marilyn Monroe) and various political figures including a Soviet defector (Victor Kravchenko), an Egyptian Field Marshall (Abdel-Hakim Amer) and a Chinese Army leader (Lo Jui Ching). All these events are considered indiscriminately which results in a high level of noise. Naturally, it is only through a better understanding of the phenomenon that a selection can be made which would raise the signal to noise ratio. The fact that an improved selection is possible is suggested by the observation that only a few cases contribute significantly to the signal.

the battle against noise can only be won if we have a good knowledge about the system under investigation and the sources of noise. Because they were unaware of the turbulence produced in the wake of the pendulum by the drag due to air resistance, physicists of the seventeenth century would have been unable to identify this source of noise. Similarly, building a device that can shield the pendulum from the vibrations of the building requires a good knowledge of the physics of vibrations. Incidentally, this example illustrates the fact that all subfields of physics are connected: to make an accurate measurement of a phenomenon in the field of mechanics one needs a good understanding of fluid mechanics, astronomy or electroacoustics. The fact that in the social sciences many subfields are patently under-developed helps to explain why it is so difficult to perform accurate measurements.

The main point is the fact that in physics, whatever their cost, the required improvements *can* be made. This leads us to the crucial question: how can one increase the accuracy of observations in the social sciences? Before coming to this point, we wish to discuss an example of physical measurement in which the problem of noise reduction is more serious than in the case of the swing of a pendulum.

2 Noise reduction in the detection of gravitational waves

Gravitational waves are ripples in the fabric of space-time. When they pass through a detector they will decrease the distance between two test masses. This effect is very small however: 10^{-18} meter over a distance of 4,000 meters between the test-masses. In order to detect the effect of a gravitational wave the distance between the test-masses must be measured with a precision of $2.5\ 10^{-20}$ percent. In spite of being shielded in a concrete cover, the vacuum pipe which connects the test-masses is subject to many sources of vibrations: trucks on nearby roads, micro-earthquakes, sound waves due to supersonic planes, thermal variations and so forth.

In order to achieve this daunting objective two methods are used which it is worthwhile to detail because they constitute two major techniques of signal detection. The first technique relies on pattern identification. It is a common experience that even in a very noisy place such as a pub we are able to identify a familiar melody however softly played. We do this by pattern matching. Our ears and brain are wonderful pattern matching organs. Catching gravitational wave signals is not unlike what our ears do routinely. Of course, the pattern of the signal depends upon the source. A supernova and a spinning neutron star have very different signal patterns. Moreover, for a spinning star the period is specific to each pair of stars. As there are many other mechanisms which can trigger gravitational waves, it is easy to realize that the number of patterns to search can become very large making pattern matching a formidable computational task. The second technique relies on using several detectors. As we already noted micro-earthquakes or fluctuations in the measurement devices can cause disturbances that simulate gravitational events. Such factors are site-dependent and are unlikely to happen at two widely separated sites. This is why there are 5 different sites: one in Germany, one in Italy, one in Japan and two in the United States.

Several lessons can be drawn from this example.

- It emphasizes the importance of having *different realizations* of the same phenomenon. If the sources of noise in these realizations are not positively correlated they will to some extent cancel one another, thus improving signal identification. An even more favorable case is when the sources of noise are negatively correlated for in this case noise reduction can be much greater. This case which unfortunately is not very common is discussed in the next chapter.

- It shows how crucial it is to have some information about the expected shape of the signal. In the case of gravitational waves we have seen that there are many possible candidates which means that a broad range of pattern matching tests must be performed. In the social sciences the situation is even more difficult in the sense that usually one knows very little about the shape of the signal that is to be expected.

For instance, in the case of the Werther effect we do not know the timing of possible excess-suicides. Actually, even if somehow we were able to know that the time scale of the phenomenon is a matter of days it would be difficult to take advantage of this knowledge because in most countries only monthly suicide data are available.

- For the detection of gravitational waves it is of critical importance that the signal is as big as possible. Only phenomena of cataclysmic proportions have a chance of being detected. We are in this case in the same situation as in the social sciences in the sense that (i) the signal cannot be changed at will (ii) big events are rarer than small events. Thus, the detection of the phenomenon is conditioned by favorable circumstances which are out of the control of the researcher.

3 Pattern matching: a simulation

How can pattern matching be used in the social sciences? Fig. 2.2 shows how the response of a system to an impulse looks like when the level of noise is progressively increased. This simulation was performed by using the following linear stochastic equation:

$$Y_t = aY_{t-1} + N_t + \delta_{t,t_1} \quad a < 1 \tag{2.1}$$

where N_t is a Gaussian random variable of mean zero and standard deviation σ; the Kronecker symbol δ_{t,t_1} which is non-zero only for $t = t_1$ describes the input shock. The first graph in the panel of Fig. 2.2 corresponds to $\sigma = 0$ and shows the purely deterministic response, a steep increase followed by an exponential fall. When σ is equal to 8 we see that the exponential fall is very distorted and no longer recognizable. However it can be noted that the steep increase is still visible. Let us now forget for a moment that the graphs in Fig. 2.2 were generated by a simulation and assume instead that they correspond to a social phenomenon that one wishes to identify. The identification is highly conditioned by what we know about the shock. Parameters which are of particular importance are the time when the shock occurred and the sign of the shock. If these two parameters are known the shock can still roughly be identified on the three graphs in the second line of Fig. 2.2. This example clearly shows how important it is to have as much information as possible about the shock.

How does the simulation of Fig. 2.2 compare with the situation that one faces in the case of the Werther effect? Fig. 2.3 shows the monthly fluctuations of suicide rates in the United States. The short spiky fluctuations are seasonal variations; it is well known that in the northern hemisphere suicide rates go through a maximum in May-June and through a minimum in December[8]. Because of these huge seasonal fluctuations the identification of the Werther effect is very problematic even if one knows the date of the suicide story published in the *New York Times*.

The techniques which have emerged from our discussion of physical experiments, namely the extreme value technique, the multi-observation approach and the various forms of pattern matching, will be examined in more detail in the next chapter in a social science context. They will be illustrated by several examples through which we will get a better feeling of their potential and limitations.

[8]The seasonal pattern of suicides has been studied by Morselli (1879, pp. 136-142) and by Durkheim (1897, chapter 3, section 4).

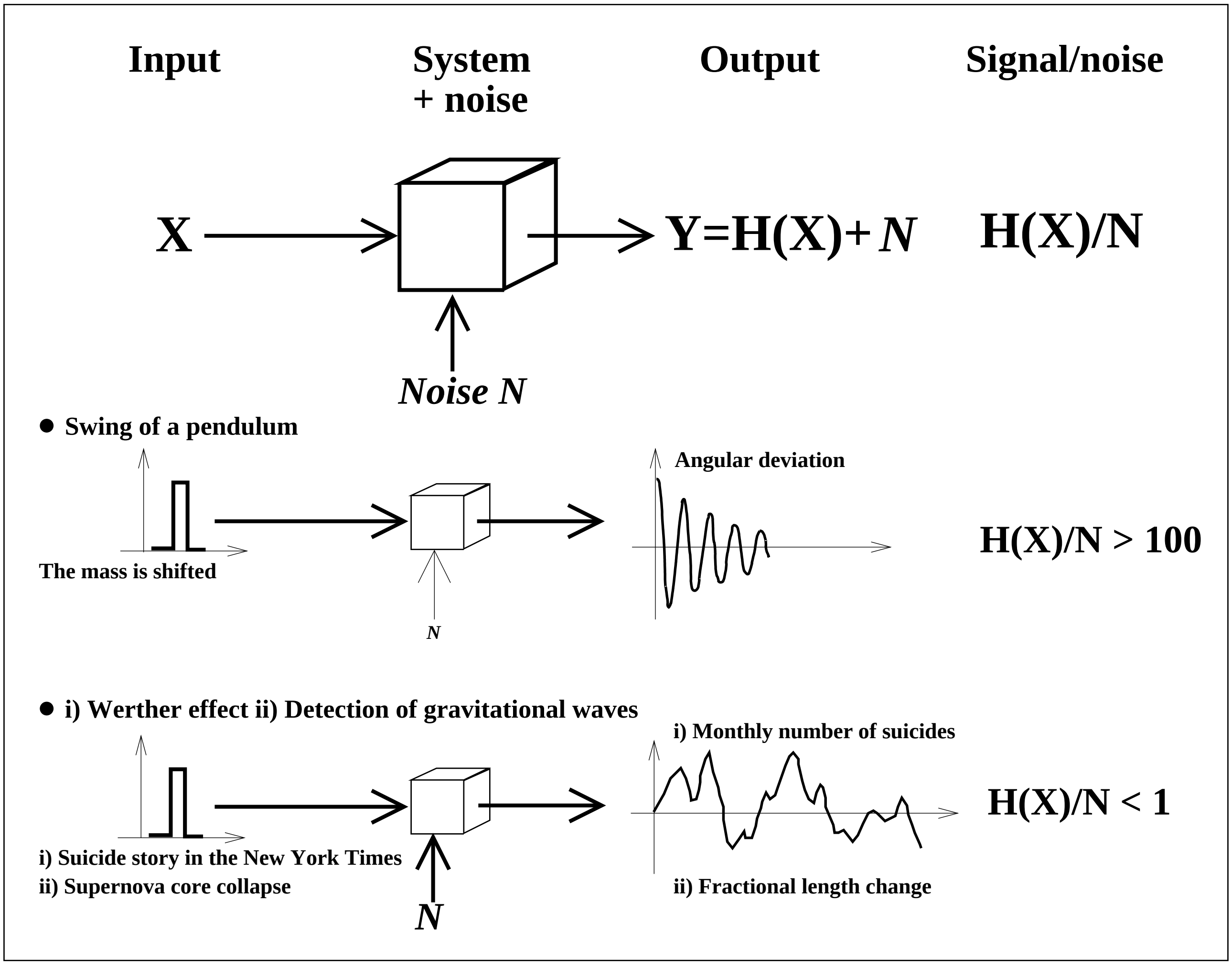

Fig. 2.1 Signal to noise ratio in three experiments. The figure illustrates the response of a system to an input X in the presence of background noise N. In the first experiment (second line) the mass of a pendulum is displaced and one observes the angular deviation. In this experiment it is easy to get a signal to noise ratio $H(X)/N$ which is higher than one hundred. The third line illustrates two situations in which the signal to noise ratio is smaller than one: the detection of a possible Werther effect (explained in the text) and the detection of possible gravitational waves produced (for instance) by the collapse of the core of a supernova. At the time of writing both effects are still hypothetical. This situation will last until the signal to noise ratio can be substantially increased. In the ratio $H(X)/N$ the denominator must be replaced by a suitable estimate of the noise; when the sources of noise are clearly identified the amplitude of the noise can be estimated directly; when the sources of noise are not known, which is usually the case in the social sciences, a possible estimator of N is the standard deviation $\sigma(Y)$ of Y.

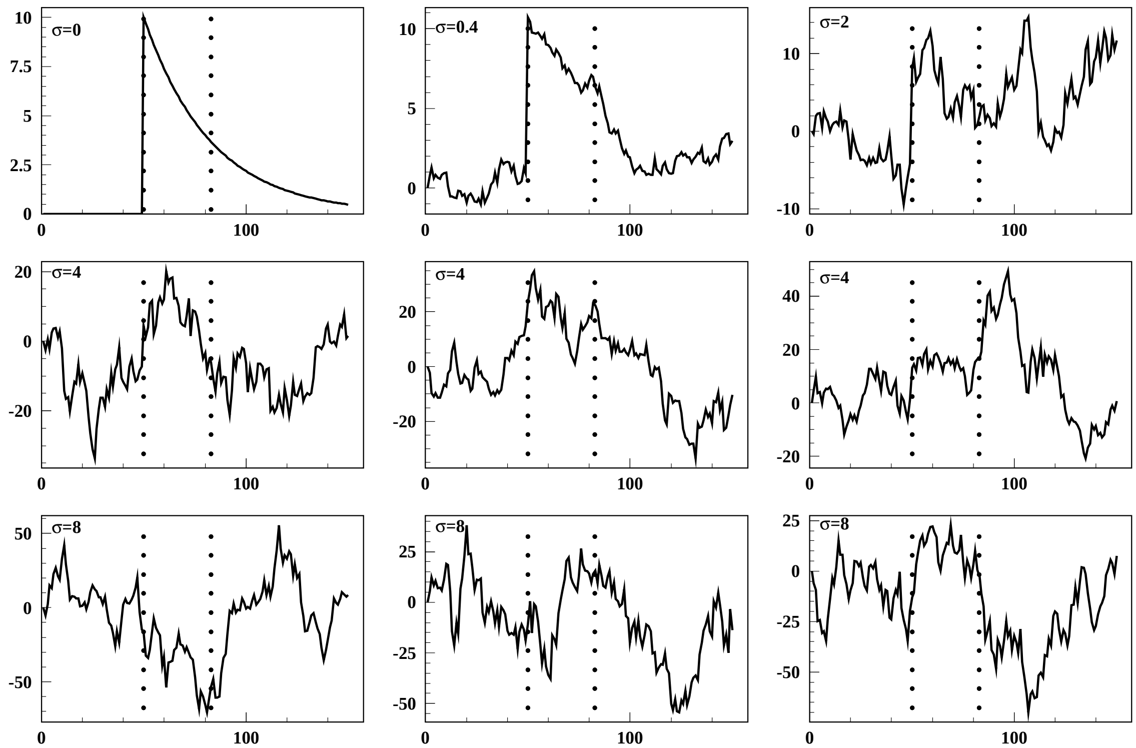

Fig. 2.2 Response of a first order system to a shock in the presence of increasing levels of noise. The figure represents the output Y_t of a system defined by the following stochastic recurrence equation:

$$Y_t = aY_{t-1} + \sigma\epsilon_t + \delta_{t,t_1} \quad a = 0.97$$

where ϵ_t is Gaussian noise of mean zero and standard deviation 1 and δ_{t,t_1} represents a shock occurring at time $t_1 = 30$; σ takes the values: $0,\ 0.4,\ 2,\ ,4,\ ,8$. As σ is increased it becomes more and more difficult to identify the signal. In fact signal identification crucially depends on how much information we have about the signal. If one knows that a sharp vertical discontinuity is to be expected it can be recognized even in the graphs with $\sigma = 8$. In contrast, the exponential relaxation process is much more difficult to identify in the last graphs. The vertical dotted lines indicate the relaxation time τ defined by $Y_t = e^{-t/\tau}$. Note that this relaxation time can also be observed very roughly in the overall shape of the broad peaks; it can be seen that their duration is approximately of the order of τ. This is of course not surprising because the response of the system to the shocks generated by the noise is of the same nature as its response to the deterministic shock δ_{t,t_1}.

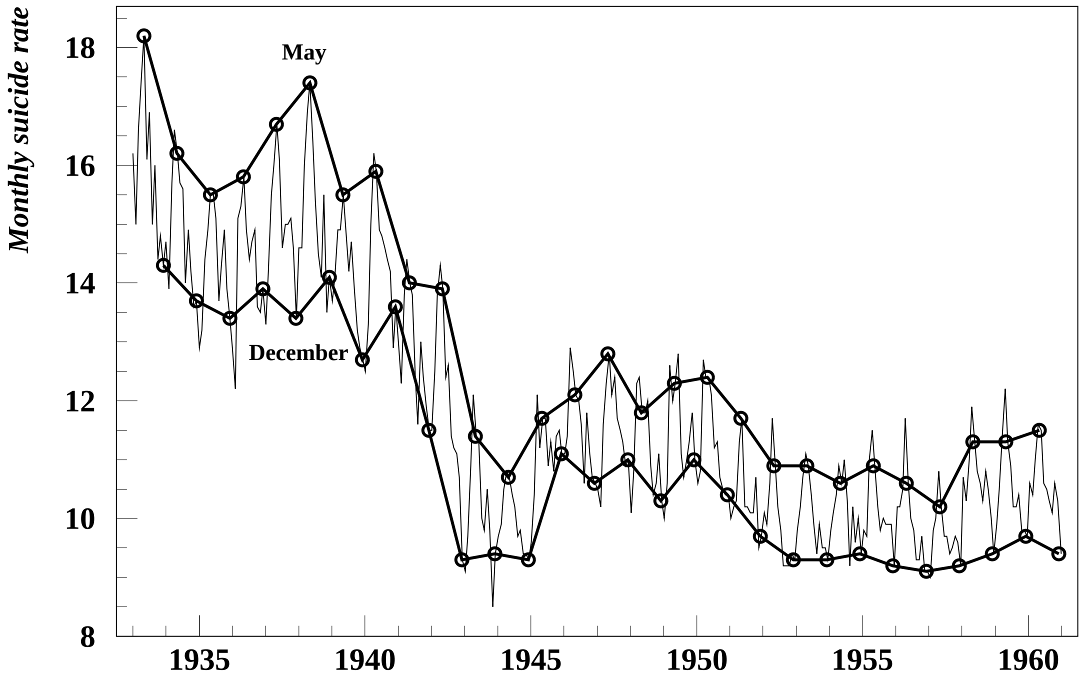

Fig. 2.3 Monthly variations of the suicide rate in the United States. The vertical scale refers to annualized monthly rates per 100,000 population. Monthly fluctuations (thin line) are seen to exhibit a seasonal pattern with a maximum usually occurring in May (represented by the upper curve) and a minimum in December (lower curve). The difference between seasonal maxima and minima was 27% in the 1930s but fell to 9% in the 1950s. Surimposed on this seasonal changes are substantial annual changes both in the seasonal pattern and in the average annual suicide rate. These changing features make the identification of any signal very difficult. *Source: Linder and Grove (1943), Grove and Hetzel (1968).*