The development of electronic devices for recording auscultative data along with the new techniques for data processing supplemented by computer analysis provides an opportunity to considerably increase the efficiency of the diagnostics of respiratory disorders. This is connected not only with high sensitivity of the devices but also with the possibility of a simultaneous processing of acoustic data received through several channels and its recording in the form of power spectra, coherence functions, respiratory sonograms, etc. Today, the broad opportunities offered by the computer processing of breath sounds allow one to consider digital auscultation as an approved clinical technique for investigating the respiratory system [1–3].

However the computer analysis of auscultative data brings about several problems that are nontrivial from the point of view of the data processing. One of these problems is the separation of lung and cardiac sounds in the total signal recorded on the thorax surface. The importance of this problem is connected with the fact that the level of the sound signals generated by one’s heart is very high (as compared to breath sounds). This leads to the masking of important auscultative diagnostic signs by the cardiac sounds, which in this case play the role of interference. This phenomenon is aggravated by the fact that, in performing the frequency analysis, one deals with averaged spectra. As a result, the respiratory diagnostic signs appearing once or twice during the total realization under analysis can become undetectable.

The problem of separating the breath and cardiac sounds is also of importance for cardiologists when the detection of a weak diastolic and systolic murmur is necessary (in this case, the respiratory sounds play the role of interference).

There are several approaches that allow one to solve this problem to a certain degree. First, one can use the signals obtained when the patient holds his breath. Taking into account the absence of correlation between breath and cardiac sounds and performing the procedure of subtraction of uncorrelated interference, it is possible to obtain the desired statistical characteristics of the signal. Unfortunately, this technique is often inadmissible, because, even for healthy people, holding one’s breath leads to a change in the character of the cardiac activity and the statistical characteristics of cardiac sounds can be changed. Moreover, in the case of children and some categories of patients, this method cannot be realized.

In connection with the aforesaid, the techniques most popular in practice are those in which the effect of cardiac sounds and sounds connected with muscle vibration are suppressed in the frequency range below 100 Hz by the filters eliminating (or reducing) the low-frequency part of the analyzed sound signal [4]. However, such a direct solution of the problem can lead to a considerable loss of diagnostically meaningful information [5, 6], because it is difficult to apply a standard approach to a human body that is characterized by a large variability in both normal and pathological states. If it is necessary to retain the low-frequency part of the spectrum, other more sophisticated methods are used: recording of breath sounds within the interval between cardiac tones [7] or utilization of adaptive filtering techniques in the digital processing [8, 9]. The advantages and shortcomings of the indicated algorithms are analyzed in more detail in the paper by Iyer, Ramamoorthy, and Ploysongsang [10], who used a modified Kalman filter to solve the problem of separating the breath and cardiac sounds.

Below, we suggest a technique for solving this problem on the basis of the two-channel processing of sig-
nals. The technique allows one not only to solve the main problem (the separation of additive components from the total signal), but also to calculate additional diagnostic parameters.

It is evident that the problem can be solved most effectively when the corresponding algorithm contains the maximal amount of \textit{a priori} information on the nature of the signals to be separated. Therefore, the development of adequate models describing the formation of the signals recorded at the thorax surface is of primary importance.

Now, there is no consensus on the mechanisms that govern the formation of the breath sounds. The model suggested in [11–13] seems to be the most adequate one. In this model, it is assumed that the breath sounds detected at the thorax have a twofold nature. First, they are the sound generated in the trachea due to the turbulent airflow passing through it and then transformed in the course of its propagation through the lung and thorax tissues. The second source of breath sounds is the minor air-carrying channels with diameters of 2–3 mm. The sounds arising there have a weaker intensity. However, they are formed in the immediate vicinity of the thorax surface where the signal is recorded.

Taking the model suggested in [13] as the basis, we assume that the sound signal received by a sensor positioned at the thorax is the sum of three uncorrelated components: the cardiac sounds, the breath sounds, and the background noise. This model is schematically represented in Fig. 1.

Here, \( s_1 \) represents the lung sounds, which, according to the majority of researchers, are generated in the trachea and bronchi [11, 12]; \( s_2 \) represents the cardiac sounds; \( H_{mp} \) is the frequency characteristic depending on the properties of the lung tissue \((p = 1, 2)\); \( u_{m1} \) and \( u_{m2} \) are the lung and cardiac sounds transformed in the course of their propagation through the lung and thorax tissues; \( n_f \) is the background noise; and \( y_m \) is the signal received by the \( m \)th sensor.

The frequency characteristic \( H_{mp} \) reflects the transformation of the sound signals generated in \( p \)th source in the course of their propagation to the \( m \)th sensor.

The problem is to estimate the signal \( u_{m1}(t) \) from the available realization of the additive sum \( y_m(t) = u_{m1}(t) + u_{m2}(t) + n_f(t) \) of the signal \( u_{m1}(t) \) and the interference \( u_{m2}(t) + n_f(t) \). It is clear that the choice of what to take as a useful signal and what to consider as interference depends on whether a cardiologist or a lung doctor performs the examination of the patient.

In solving the problem, we proceed from the fact that it can be solved as soon as it will be possible to determine the spectra of the additive components of the recorded signal. This will provide an opportunity to apply the theory of optimal linear filtering and, in so doing, to solve the problem of interest.

It should be noted that each of the additive components of the signal \( y_m(t) \) has a complex structure. Both breath and cardiac sounds are nonstationary random processes, which exhibit not only periodic intensity variations, but also a frequency modulation.

The point is that, even in the case of normal breath sounds, the spectrum of inhalation slightly differs from the spectrum of exhalation [14], not to mention the breath sounds in the case of lung pathology.

As for cardiac sounds, they are divided into tones and noise. Tones I and II always accompany the normal operation of the heart. Tones III and IV appear less often. The noise (systolic and diastolic murmurs) occur mainly in the case of heart defects. Each of these components has its own characteristic frequency band (30–120 Hz for tone I, 70–150 Hz for tone II, 50–600 Hz for systolic murmur, and 120–800 Hz for diastolic murmur).

In the described conditions, the process at the output of the \( m \)th sensor is described by the equation

\[
y_m(t) = \sum_{p} h_{mp}^{(f)} s_p(t - \tau) d\tau + n_f(t), \quad p = 1, 2; \quad m = 1, N,
\]

where \( N \) is the number of sensors, \( p \) corresponds to the source number, and \( h_{mp}^{(f)} \) is a pulsed transition function satisfying the equation

\[
H_{mp}^{(f)} = \int h_{mp}^{(f)} \exp(2\pi f\tau) d\tau.
\]

We solve this problem using the signals recorded by two sensors. We denote the hypothetic transition functions between the sensors for each of the sources by \( \alpha(f) \) and \( \beta(f) \), i.e., \( \alpha = |H_{12}|H_{11}^{*} \) and \( \beta = |H_{12}|H_{22}^{*} \). We denote the phase shift between the signals at the sensors as \( \theta_1(f) \) for the first source and \( \theta_2(f) \) for the second one. We note that the functions \( \alpha(f) \), \( \beta(f) \), \( \theta_1(f) \), and \( \theta_2(f) \) are time independent; i.e., it is assumed that the properties of the lung tissue through which the sound signals propagate remain constant within different phases of breath and heart cycles. We also assume that the background noise recorded by each sensor is stationary and mutually uncorrelated. Taking into account the mutual uncorrelation of the cardiac and breath sounds, which is determined by the difference in their origin, we obtain