TECHNICAL CONDITION MONITORING METHODS TO MANAGE THE REDUNDANCY OF SYSTEMS. PART II: Classical Models

V. N. Bukov*, A. M. Bronnikov**, A. S. Popov***, and V. A. Shurman****

*Airborne Aeronautical Systems, Moscow, Russia
**Bauman Moscow State Technical University, Moscow, Russia
***Zhukovsky–Gagarin Air Force Academy, Voronezh, Russia
****Institute of Aircraft Equipment, Zhukovsky, Russia

*✉ v_bukov@mail.ru, **✉ bronnikov_a_m@mail.ru, ***✉ saga30@yandex.ru, ****✉ shurman@niiao.ru

Abstract. Redundancy management of a technical system involves a monitoring procedure (control of the current state of its components) to reconfigure the system as needed. Consisting of four parts, this survey presents modern and newly developed technical condition monitoring methods for redundancy management. Part I was devoted to a general description of built-in control, voting schemes, and fidelity rules; control codes and program execution control methods were briefly covered. Part II of the survey considers fault diagnosis methods based on the classical modeling of the system diagnosed in the discrete time and frequency domains. The Chow–Willsky fault detection scheme, as well as the definition of a residual and its generation procedures in the diagnosis problem, are presented. The main model-based diagnosis methods using equation errors, observers, parity equations, and redundant variables are described. In conclusion, the robustness problem of diagnosis methods of the corresponding type is briefly discussed.

Keywords: technical condition monitoring, redundancy management, diagnosis, residual, analytical models of systems, equation error method, observers, Beard filter, fault detection via parity equations, redundant variable method, robustness of methods.

INTRODUCTION

Mathematical model-based fault diagnosis methods[1] detect a mismatch between the real measurements of input and output signals of a system (on the one hand) and the corresponding signals of its mathematical model (on the other), expressed by the so-called *residual*. These methods started to be actively developed in the 1970s. In particular, static or parallel redundancy schemes [1, 2] were applied, which can be obtained directly from measurements (hardware redundancy) or analytical relations (analytical redundancy). Such methods were reviewed in [3–5].

In part II of the survey, we highlight the methods commonly used in fault diagnosis systems of various-purpose complex technical systems. Their primary indisputable advantage is engineering transparency, i.e., the intuitive clarity of the processes and causal relations for experts in the corresponding technical field. The effectiveness of the methods under consideration depends on particular conditions of application, and their choice is determined by the following factors: the functional purpose of the device, its structural organization and technological features of manufacturing and operation, and the required reliability and fidelity indicators.

1. GENERAL CONDITIONS OF MODEL-BASED METHODS

Figure 1 shows the Chow–Willsky scheme [4], which generalizes most fault detection methods for a system with actuators, an object, and sensors.

[1] In this context, we mean only analytical modeling in the “classical” setting: the central link is the reproduction (with certain level of detail) of the operation laws of a system diagnosed; approaches with other types of models will be presented in parts III and IV of the survey.

The real-time fault detection scheme includes two main blocks:

• Residual generation: a signal called the residual is formed by using the input and output signals of a monitored process. This residual should be independent of the input and output signals and should take a value equal to/close to 0 in the absence of faults.

• Residual estimation: the value of the residual is analyzed, and the presence or absence of a fault is decided accordingly.

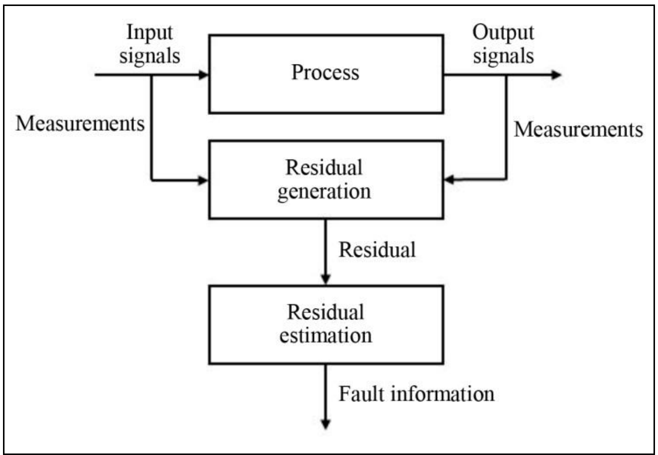

Fig. 1. The Chow–Willsky scheme.

Decision rules involve simple tolerance control or complex signal transformation methods, including statistical analysis.

In this case, the main attention of researchers is focused on residual generation since residual estimation procedures are largely universal.

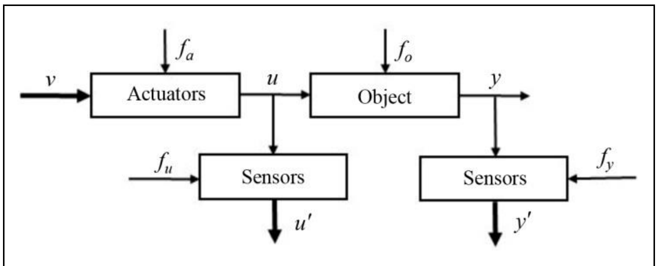

Fig. 2. The fault topology of a dynamic system.

Figure 2 explains the topology of faults affecting the system operation, with the following notation: v is an external input (disturbance); u is a control input applied to the object; y is the output signal of the object; u' and y' are measured signals (available for processing); f_a, f_o, f_u, and f_y are formalized representations of faults occurring in actuators, the object, and sensors of object's input and output signals, respectively. Only the signals v, u', and y' are known with some accuracy.

Despite the nonlinearity of most real systems monitored, modeling and identification methods for linear systems are often used to avoid the difficulties inherent in nonlinear models. In most cases, this is not a significant limitation since the dynamic system is monitored in the neighborhood of an expected operation mode, and deviations from the mode are described well by linear models.

1.1. Description of Analytical Models

According to the linearity hypothesis of a dynamic process, its behavior in discrete time $t = 0, 1, 2, \ldots$ is described by the state-space model [6–10]

$$x_{t+1} = Ax_t + Bu_t, \quad y_t = Cx_t, \tag{1}$$

where $x_t \in \Re^n$ is the state vector of the system; $u_t \in \Re^r$ is the vector of true input signals; $y_t \in \Re^m$ is the vector of true output signals; A, B, and C are the numerical matrices of system coefficients.

The effect of faults on the object is described by the model

$$x_{t+1} = Ax_t + Bu_t + f_{ab.t}, \quad y_t = Cx_t + f_{c.t}, \tag{2}$$

where $f_{ab.t} = \Delta Ax_t + \Delta Bu_t$ and $f_{c.t} = \Delta Cx_t$; ΔA, ΔB, and ΔC are the variations of the coefficient matrices due to the occurred faults. (The fault vector f_o is represented by two vectors $f_{ab.t}$ and $f_{c.t}$, associated with time instant t). Sometimes faults are treated differently: $f_{ab.t} = A\Delta x_t + B\Delta u_t$ and $f_{c.t} = \Delta y_t$, where Δx_t is an external disturbance for internal variables; Δu_t is actuator errors (faults); Δy_t is sensor errors (faults).

State-space descriptions provide general and mathematically rigorous tools for modeling the system and robustly generating uncertainties in both the deterministic and stochastic cases (measurements without noise and noisy measurements, respectively). Therefore, the system matrices A, B, and C in canonical form for models (1) and (2) can be obtained using multivariate identification procedures.

The input and output measurements are affected by faults according to the formulas

$$u_t' = u_t + f_{u.t}, \; y_t' = y_t + f_{y.t},$$

where u_t' and y_t' are the measured values of the input and output signals.

Typically, step and ramp (gradient) signals are used to model sudden and evolving faults, representing a displacement (a fixed change during one time instant) and a drift (monotonic changes on a sequence of time instants), respectively.

Measurement noise affecting sensor output signals is often described by uncorrelated Gaussian processes.

The general model-based fault detection (FD) problem can be solved only with knowledge of the measured sequences u_t' and y_t'. In addition, a priori knowledge of the characteristics of the received signals u_t' and y_t' is widely used. Examples include the spectrum, dynamic range of the signal, and its variations. However, the need for a priori information about the signals monitored and the dependence of the signal characteristics on the unknown conditions of the system diagnosed are the main drawbacks of this class of methods.

Along with models (1) and (2), the operation of a system diagnosed can be described by frequency-domain models of the form

$$y_t = W_y^u(z)u_t + \Delta W_y^u(z)u_t, \tag{3}$$

where z is the forward shift operator (one time instant ahead)[2]; $W_y^u(z)$ is the transfer (rational polynomial) matrix relating the input u_t to the system output y_t; $\Delta W_y^u(z)$ is the variation of the transfer matrix $W_y^u(z)$ due to faults.

As a rule, model (3) is applied under known frequency characteristics of faults and disturbances: information in frequency spectra can be used as criteria for fault detection.

1.2. Residual Generation

The general structure of a residual generator [11, 12] is demonstrated in Fig. 3. Here, z_t is an auxiliary signal (as a rule, meaningful for the developer, e.g., the state vector of the system or its parts); r_t is the residual signal; $W_1(u_t, y_t)$ and $W_2(z_t, y_t)$ are the main blocks of the generator.

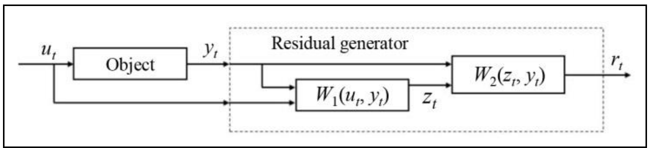

Fig. 3. The general structure of a residual generator.

Regardless of the method applied, residual generation is nothing but a linear mapping with the inputs consisting of the inputs and outputs of the process monitored. In the normal condition (no fault), the equality $r_t = 0$ holds.

The simplest residual generator is obtained when the transfer matrix W_1 coincides with model (3) of the object: $W_1(u_t, y_t) = W_y^u(z)$. In other words, this is a real process description yielded by a system identification procedure (e.g., an autoregressive exogenous model [13]).

The simplest, and most widespread, fault detection method is to compare the residual r_t or its some function $J(r_t)$ with a fixed threshold ε or a threshold function ε_t as follows:

$$J(r_t) = \begin{cases} \le \varepsilon_t & \text{for } f_t = 0 \\ > \varepsilon_t & \text{for } f_t \ne 0, \end{cases} \tag{4}$$

where f_t is the generalized fault vector. Thus, if the residual exceeds a fixed threshold, a fault is declared.

Such a test is particularly effective with fixed thresholds ε if the object operates predominantly in a steady state. It responds after either a sudden large fault or a persistent, gradually increasing fault.

In most practical cases, process parameters are completely unknown or known with insufficient accuracy. Then they can be determined using parameter estimation methods by measuring the input and output signals u_t and y_t if the main structure of the model is known.

2. THE EQUATION ERROR METHOD

For a *Single Input, Single Output* (SISO) object, the discrete-time model of order *n* is written in the vector form

$$y_t = \Psi^{\mathrm{T}}\Theta,$$

where $\Theta = \begin{bmatrix} a_1 & \cdots & a_n & b_1 & \cdots & b_n \end{bmatrix}^{\mathrm{T}}$ denotes the vector of object's parameters; $\Psi = \begin{bmatrix} y_{t-1} & \cdots & u_{t-n} \end{bmatrix}^{\mathrm{T}}$ is the vector of discrete-time sample data at n consecutive time instants (measurement points).

The error signal is given by

$$e_t = y_t - \Psi^{\mathrm{T}}\hat{\Theta},$$

where $\hat{\Theta}$ is the vector of parameter estimates. Based on the equations

$$J(\hat{\Theta}) = \sum_t e_t^2 = e^{\mathrm{T}}e, \quad \frac{dJ(\hat{\Theta})}{d\hat{\Theta}} = 0,$$

the least-squares estimates of the parameters are determined by the formula

$$\hat{\Theta}_t = (\Psi^{\mathrm{T}}\Psi)^{-1}\Psi^{\mathrm{T}}y_t. \tag{5}$$

[2] Such an interpretation ignores the initial conditions of the model. An alternative is to treat the operator as the *z*-transform, thereby considering the initial conditions.

As described in [14–16], the estimates (5) can also be computed using the recurrent[3] least-squares algorithm in real time $t = 0, 1, 2, \ldots$ with respect to the estimates at time instant t:

$$\hat{\Theta}_{t+1} = \hat{\Theta}_t + \gamma_t \left(y_{t+1} - \Psi^{\mathrm{T}}_{t+1} \hat{\Theta}_t \right),$$

where

$$\gamma_t = (\Psi^{\mathrm{T}}_{t+1} P_t \Psi_{t+1} + 1)^{-1} P_t \Psi_{t+1}, \; P_{t+1} = (I_{2n} - \gamma_t \Psi^{\mathrm{T}}_{t+1}) P_t,$$

and I_{2n} stands for an identity matrix of dimensions $2n \times 2n$. The initial conditions P_0 for the matrix P_t is either the identity matrix or a supposed covariance matrix of the initial errors of the estimates $\hat{\Theta}_0$.

It is possible to improve convergence using filtering methods. In particular, a Kalman filter can be applied for parameter estimation in the case of noisy measurements [17, 18].

Fault diagnosis is performed by the difference between the vector $\hat{\Theta}_t$ of parameter estimates and its "reference" values, corresponding either to the absence or presence of definite faults in the system. For this purpose, one may employ the approaches described earlier; see Section 5 in part I of the survey [19].

3. OBSERVER-BASED APPROACHES

The principal idea of observer/filter-based methods is to estimate the system outputs from measurements using Lüenberger observers (in the deterministic case) or Kalman filters (in the presence of random noises). The output estimation error (or its weighted value) is used as a residual.

Note that in the case of observer-based fault diagnosis, the outputs must be estimated and the estimation of the state vector is often not required [20]. In addition, the advantage of an observer is the flexible choice of its gains, which leads to a rich variety of fault diagnosis schemes [21–30].

In the linear dynamic model (1) with exactly known matrices A, B, and C, the variables are restored using the observer

$$\hat{x}_{t+1} = A\hat{x}_t + Bu_t + He_t, \; e_t = y_t - C\hat{x}_t, \qquad (6)$$

where $\hat{x}_t \in \Re^n$ denotes the state vector estimate; $e_t \in \Re^m$ is the output error. At the same time, the state error $e^x_t \in \Re^n$ satisfies the equations

$$e^x_t = x_t - \hat{x}_t, \qquad e^x_{t+1} = (A - HC)e^x_t,$$

asymptotically vanishing if the matrix $A - HC$ is stable.

In the presence of disturbances and faults, we have the equations

$$x_{t+1} = Ax_t + Bu_t + Qv_t + L_1 f_t, \; y_t = Cx_t + Rw_t + L_2 f_t,$$

where v_t is the unmeasured vector of input disturbances; w_t is the unmeasured vector of output disturbances; f_t is the reduced vector of faults additively affecting the object's operation; Q and R are disturbance effect matrices; L_1 and L_2 are fault effect matrices.

Under the conditions $v_t = 0$ and $w_t = 0$, the state estimate satisfies the equations

$$e^x_{t+1} = (A - HC)e^x_t + L_1 f_t - HL_2 f_t,$$
$$e_t = Ce^x_t + L_2 f_t,$$

which describe the effect of the generalized faults f_t on the errors e^x_t and e_t, taken as residuals. The matrix H influences the properties of the residuals. The requirements for choosing H are stability and sensitivity to the disturbances v_t and w_t. If the signals are subjected to noises, a Kalman filter [17] should be used instead of classical observations.

If faults manifest themselves in the form of parameter variations ΔA or ΔB, the process behavior is described by

$$x_{t+1} = (A + \Delta A)x_t + (B + \Delta B)u_t, \; y_t = Cx_t,$$

and the estimation errors are written as

$$e^x_{t+1} = (A - HC)e^x_t + \Delta A x_t + \Delta B u_t, \; e_t = Ce^x_t.$$

In this case, the variations ΔA and ΔB are multiplicative errors, and the changes in errors depend on both parameter variations and changes in the input and state variables. Therefore, the effect of parameter variations on errors is less unambiguous than in the case of additive faults f_t. Particular solutions can be found in [14, 25].

For *Multiple Input*, *Multiple Output* (MIMO) processes, the following decomposition is applied: either a single observer excited by a single output [26], or a bank of observers excited by all outputs with hypothesis testing [27], or a bank of observers excited by either separate outputs [26] or all outputs except one [28, 29]. This approach allows detecting faults of different combinations of sensors.

In the MIMO case, the gain matrix H of the state observer in equation (6) is chosen so that fault signals $L_1 f_t$ change in a definite direction and fault signals $L_2 f_t$ in a definite plane [29, 31].

[3] In programming, the equivalent term "recursive" is used.

In the presence of directional residual vectors, the fault isolation problem is reduced to determining the nearest fault signature direction for the residual vector. The original form of a “fault detection filter” was proposed by R. Beard [31] and H. Jones [32] for generating directional residual vectors.

Subsequently, many simpler methods were proposed, including a “robust fault detection filter” [33]. It represents a class of Lüenberger observers with a special feedback gain matrix. Another possibility is to use output observers or the so-called generalized observers, e.g., observers with an unknown input with output reconstruction, if the state estimate $\hat{x}_t$ is not of primary interest.

4. BEARD FILTER

We will illustrate the possibilities and limitations of approaches using fault signature directions with the Beard filter [31].

Consider a linear dynamic system with discrete time $t = 1, 2, \ldots$, described by the equation

$$x_{t+1} = Ax_t + Bu_t + B\begin{bmatrix}0 & \cdots & f_i & \cdots & 0\end{bmatrix}^{\mathrm{T}}, \quad x_{t=0} = x_0, \tag{7}$$

where $x_t \in \Re^n$ is the state vector; $u_t \in \Re^l$ is the control vector; $A \in \Re^{n\times n}$ and $B \in \Re^{n\times l}$ are matrices of constant coefficients; the additional term $B\begin{bmatrix}0 & \cdots & f_i & \cdots & 0\end{bmatrix}^{\mathrm{T}}$ corresponds to an unknown disturbance (external signal) f_i at the ith input of the system, $i = \overline{1, l}$, $l \le n$. By assumption, all system faults under consideration are reduced to a set of unknown external signals.

The Beard filter allows detecting m faults of the specified type and finding the corresponding column b_i of the matrix B.

The filter involves an auxiliary dynamic system of the form

$$w_{t+1} = Aw_t + Bu_t + K(x_t - w_t), \; w_{t=0} = w_0, \tag{8}$$

which repeats the structure of the diagnosed system (7) with the additional term $K(x_t - w_t)$. The coefficient matrix K is assigned in a special way.

Subtracting equation (8) from equation (7) gives

$$x_{t+1} - w_{t+1} = A(x_t - w_t) - \\ -K(x_t - w_t) + B(u_t - u_t) + B\begin{bmatrix}0 \\ f_i \\ 0\end{bmatrix}. \tag{9}$$

By introducing the residual $r_t = x_t - w_t$, we write equation (9) as

$$r_{t+1} = (A - K)r_t + B\begin{bmatrix}0 \\ f_i \\ 0\end{bmatrix}. \tag{10}$$

Note that the residual is independent of the control vector u_t and state x_t of the system, being determined only by the presence of an external signal f_i and the properties of the matrix $(A - K)$.

Let the coefficient matrix K be chosen so that

$$A - K = \alpha I_n, \tag{11}$$

where I_n denotes an identity matrix of dimensions $n \times n$; α is a real number with norm $|\alpha| < 1$. Then the expression (10) takes the form

$$r_{t+1} = \alpha r_t + B\begin{bmatrix}0 \\ f_i \\ 0\end{bmatrix}, \quad r_{t=0} = r_0, \tag{12}$$

and the filter is stable. Due to this property and the diagonality of the matrix (11), the residual r_t gradually converges the vector

$$r_{\lim} = \frac{1}{1-\alpha}\begin{bmatrix}b_{1i} & \cdots & b_{ii} & \cdots & b_{ni}\end{bmatrix}^{\mathrm{T}} f_i.$$

The matrix B of equation (7) contains m columns, each defining a particular direction in the n-dimensional space.[4] Thus, the issue is to determine the columns of the matrix B on which the vector (12) is projected in the vector representation. All l external signals (faults) f_i can be monitored simultaneously.

The possibilities of the Beard filter are not exhausted by this. Similar calculations yield relations for the column-wise determination of variations in the elements of the matrix A in equation (7). In this case, it is necessary to consider the equation

$$x_{t+1} = Ax_t + Bu_t + A\begin{bmatrix}0 & \cdots & \mu_i & \cdots & 0\end{bmatrix}^{\mathrm{T}}. \tag{13}$$

Under the appropriate conditions, the residual converges to

$$r_{\lim} = \frac{1}{1-\alpha}\begin{bmatrix}a_{1i} & \cdots & a_{ii} & \cdots & a_{ni}\end{bmatrix}^{\mathrm{T}} \mu_i.$$

The remaining considerations are analogous to the above. All n external signals (faults) μ_i can be estimated simultaneously.

[4] The columns are supposed to be linearly independent.

The main properties of the Beard filter are as follows:

– Constant faults of the specified type are detected.

– Fault estimation requires time to complete the filter transients.

– When only part of the state vector is measured, $y_t = Cx_t$, where $y_t \in \Re^m$ and $C \in \Re^{m\times n}$, one can estimate only m faults described by m columns according to the schemes (6) and (12) from the set of all columns of the matrices A and B.

For further presentation, we recall some formulas from part I of the survey; see Section 5 in [19].

Let a syndrome be an m-dimensional vector $\Delta = \begin{bmatrix}\Delta_1 & \cdots & \Delta_m\end{bmatrix}^{\mathrm{T}}$, and let the possible faults of the object be formalized by an n-dimensional vector $F = \begin{bmatrix} f_1 & \cdots & f_n\end{bmatrix}^{\mathrm{T}}$. Then

$$\Delta = \Phi(Y, Z, F) = 0. \qquad (14)$$

For the syndrome Δ to respond to a single fault $f_i \neq 0$, it suffices to satisfy the nonzero sensitivity condition

$$\exists j : \frac{\partial \Phi_j}{\partial f_i} \neq 0. \qquad (15)$$

In the case of a multiple fault (when several faults f_i occur simultaneously), the additional condition

$$\begin{bmatrix} \partial\Phi_1 / \partial f_1 & \cdots & \partial\Phi_1 / \partial f_n \\ \vdots & \ddots & \vdots \\ \partial\Phi_m / \partial f_1 & \cdots & \partial\Phi_m / \partial f_n \end{bmatrix} \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix} \neq 0, \qquad (16)$$

(no mutual compensation (14) for the effects of these faults) must be valid.

The Beard filter is a special case of the approach with algebraic invariants, where redundancy is introduced artificially via the auxiliary system (8). The sensitivity condition (15) always holds for each time instant: the syndrome (14) is the residual r_{i+1} and , due to equations (12) and (13), $\partial r_{i+1} / \partial f_i = B$ and $\partial r_{i+1} / \partial \mu_i = A$. Condition (16) turns into the above property of a bounded number of detectable faults, as it prescribes the linear independence of the columns $\begin{bmatrix}\partial r_{1.i+1} / \partial f_k & \cdots & \partial r_{m.i+1} / \partial f_k\end{bmatrix}^{\mathrm{T}}$, making the corresponding matrix nonsingular and, as a necessary condition, the number of residuals $r_{j.i+1}$ equal to the number of detectable faults f_k.

5. FAULT DETECTION VIA PARITY EQUATIONS

The main idea of *fault detection via parity equations* is to check the consistency of different (usually input and output) measurements obtained at an object diagnosed and a control scheme [34].

As for observers, the model parameters and the structure of the process observed must be known a priori.

This fault detection approach is as follows. For a SISO object with a known transfer function $W(z) = b(z)/a(z)$, one uses, in parallel, a model of the form $W_m(z) = b_m(z)/a_m(z)$, where $a(z)$, $b(z)$, $a_m(z)$, and $b_m(z)$ are known polynomials of the operator z. According to the developer's decision, any scheme can be implemented as shown in Fig. 4.

The scheme in Fig. 4a corresponds to the residual of outputs; the scheme in Fig. 4c, to the residual of inputs; the scheme in Fig. 4b, to the intermediate solution. In these schemes, when the dynamic properties (parameters of polynomials) of the object and its model coincide, the residual r_t is 0. The presence of faults causing variations in the parameters of the object's transfer function generates a non-zero residual.

The above schemes have different sensitivities to different faults. Moreover, they can be augmented with different filters to achieve acceptable sensitivity or coarseness [34, 35]. Here, by default, the measured signals are equal to their true values: $u'_t = u_t$ and $y'_t = y_t$.

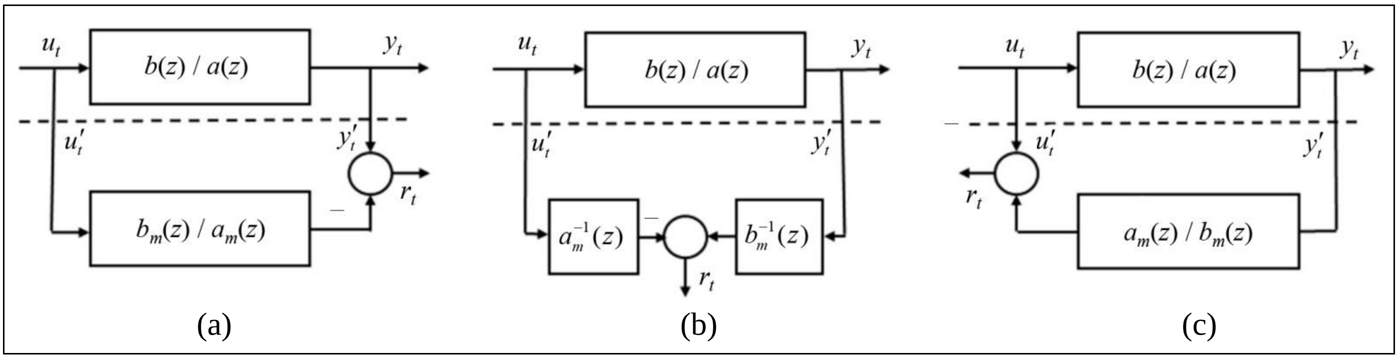

Fig. 4. Typical schemes of fault detection via parity equations: (a) by outputs, (b) by the intermediate signal, and (c) by inputs.

As shown in [36], fault detection via parity equations provides less design flexibility compared to observer-based methods without any constraints.

A disadvantage of this method is the generation of only one feature (residual) for an SISO object, which prevents from specifying the character and location of the fault: there is one equation, and a non-zero value of one residual indicates only the inconsistency of this equation. Nothing more can be extracted from this information. Various solutions were proposed to overcome this drawback [33, 37].

6. THE REDUNDANT VARIABLE METHOD

Pioneered by L.A. Mironovskii [38], this method is a most effective one. The method involves a dynamic system for variables $z \in \Re^k$, supplementing the model of an object diagnosed:

$$x_{t+1} = Ax_t + Bu_t, \tag{17}$$

where $x_t \in \Re^n$, $u_t \in \Re^l$, and A and B are constant matrices.

The redundant variables are introduced by satisfying two conditions

$$Mz_t = 0,\ Tz_t = x_t, \tag{18}$$

where $M \in \Re^{k\times(n+k)}$ and $T \in \Re^{n\times(n+k)}$ are known (given) constant matrices.[5] Substituting the expression (17) into conditions (18) and resolving for z_{t+1}, we obtain the following equation of order $(n+k)$:

$$z_{t+1} = \begin{bmatrix} T \\ M \end{bmatrix}^{-1} \begin{bmatrix} AT \\ 0 \end{bmatrix} Tz_t + \begin{bmatrix} T \\ M \end{bmatrix}^{-1} \begin{bmatrix} B \\ 0 \end{bmatrix} u_t .$$

It describes the behavior of the redundant system. This system, if stable, is equivalent to the original system (17) in terms of the output x_t and, at the same time, can be used to monitor its operation by the k-dimensional residual vector

$$r_t = Mz_t .$$

Various refinements of the redundant variable method are known [11, 39], both in terms of formulation and solution. They extend the original method to models with transfer functions and nonlinear models and minimize the redundancy of the system.

Like the Beard filter, the redundant variable method is a special case of the method with algebraic invariants. Therefore, the sensitivity conditions (15) and (16) apply to it as well.

[5] The rows of the matrices M and T are supposed to be linearly independent.

Let the possible faults of an object diagnosed be formalized by the equation

$$x_{t+1} = Ax_t + Bu_t + F ,$$

where $F = \begin{bmatrix} f_1 & \cdots & f_n \end{bmatrix}^{\mathrm{T}}$ is the vector of unmeasured signals representing changes (faults) in the object. According to conditions (18), at the first time instant, for each formalized fault f_i we obtain

$$\frac{\partial r_{t+1}}{\partial f_i} = \begin{bmatrix} \partial r_{1.t+1} / \partial f_i \\ \vdots \\ \partial r_{k.t+1} / \partial f_i \end{bmatrix} = M \begin{bmatrix} T \\ M \end{bmatrix}^{-1} \begin{bmatrix} \left.\begin{matrix} 0 \\ 1_i \\ 0 \end{matrix}\right\} (n \times 1) \\ 0_{k\times 1} \end{bmatrix}_{(n+k)\times 1}, \tag{19}$$

where 1_i is the unit standing in the ith row of the last cofactor column. At each subsequent time instant $t+q,\ q = \overline{2, N}$, in view of formula (17), we have

$$\frac{\partial r_{t+q}}{\partial f_i} = \begin{bmatrix} \partial r_{1.t+q} / \partial f_i \\ \vdots \\ \partial r_{k.t+q} / \partial f_i \end{bmatrix} = \sum_{s=0}^{N-1} A^s M \begin{bmatrix} T \\ M \end{bmatrix}^{-1} \begin{bmatrix} \left.\begin{matrix} 0 \\ 1_i \\ 0 \end{matrix}\right\} (n \times 1) \\ 0_{k\times 1} \end{bmatrix}_{(n+k)\times 1}, \tag{20}$$

where A^s is the sth degree of the matrix A from formula (17).

In the presence of faults, non-zero values of the components (19) and (20) are not obvious and depend on the choice of the matrices T and M, which requires developer's art. For example, as the reader can easily verify, in the simplest case of a diagonal or codiagonal matrix $\begin{bmatrix} T^{\mathrm{T}} & M^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$, the sensitivity of all components of the residual r_{t+1} to any fault f_i has zero value.

7. ROBUSTNESS OF MODEL-BASED METHODS

The robustness of fault diagnosis algorithms to various uncertainties, including modeling errors and the active noise of meters, is a problem deserving special attention.

As a rule, in theoretical studies and practical applications, all uncertainties are generalized into additive disturbances affecting a system (an object, sensors, and actuators). Although the disturbance vector is unknown, by assumption, it can be estimated through identification procedures.

The objective of disturbance elimination approaches in a fault diagnosis system is to completely eliminate the effect of disturbances on the residual, which may be generally impossible. There is a trade-off between sensitivity to errors and robustness to modeling uncertainty, so robust residual generation can be viewed as a multicriteria optimization problem [4, 40]. It consists of maximizing the effect of faults and minimizing the effect of uncertainty.

One active way to achieve robust fault detection [1] is based on an approximate representation of the modeling errors in model (1):

$$\Delta W_y^u(z)u_t \approx W_y^d(z)d_t,$$

where d_t is an unknown vector and $W_y^d(z)$ is the estimated transfer function. Robust fault detection algorithms can be obtained by using this approximate structure in the design of disturbance-eliminating residual generators.

Another approach, called passive robustness, involves a residual generator with an adaptive threshold. A simple example is a robust fault detection method with a threshold adaptor or selector [41]. This method requires no effort to develop a robust residual generator.

In this case, the residual generation uncertainty is represented as

$$r_t = H(z)\Delta W_y^u(z)u_t.$$

Under the assumption of small modeling errors,

$$\left\|\Delta W_y^u(z)\right\| \le \delta,$$

the adaptive threshold can be produced by a linear system of the form

$$\varepsilon_t = \delta H(z)u_t.$$

In this case, the threshold ε_t is no longer fixed but depends on the input signal u_t, thus being adaptive to the operating mode of the system. A fault is detected if

$$\|r_t\| > \|\varepsilon_t\|.$$

A more detailed approach to robust residual generation proceeds from a discrete object description in terms of transfer matrices:

$$y_t = \left(W_y^u(z) + \Delta W_y^u(z)\right)u_t + W_y^d(z)d_t + W_y^f(z)f_t, \quad (21)$$

where u_t and y_t are the input and output of the object; f_t is the vector of faults to be detected; d_t is the vector of disturbances; $\Delta W_y^u(z)$ is the representation error of the transfer matrix $W_y^u(z)$; $W_y^d(z)$ is the effect of the modeling disturbance; the matrices $\Delta W_y^u(z)$ and $W_y^d(z)$ together describe the modeling uncertainties.

The disturbance generator r_t in Fig. 4a is described by the equation

$$r_t = H_y(z)\left(y_t - y_t^m\right), \quad (22)$$

where y_t^m denotes the output of model (1) without faults, errors, and disturbances; $H_y(z)$ is the transfer matrix of residual processing. Substituting the expression (21) into equation (22) yields

$$\begin{aligned} r_t &= H_y(z)W_y^f(z)f_t \\ &+ H_y(z)W_y^d(z)d_t + H_y(z)\Delta W_y^u(z)u_t. \end{aligned} \quad (23)$$

Extracting the first term in formula (23) against the background of the other two terms is a very difficult task. Therefore, robust residual generation is commonly reduced to satisfying the condition

$$H_y(z)W_y^d(z) = 0. \quad (24)$$

Here, various approaches are applied (e.g., the ones with observers, optimization, given structures, identification schemes, etc.). Often, such a problem cannot be completely solved, as sensitivity to faults is lost. However, compromise solutions are known.

Decoupling from disturbances can also be achieved using design methods in the frequency domain. An example is the robust detection of faults using standard H_∞ filtering.

For instance, when condition (24) fails, an approximate estimate can be obtained, e.g., in the form of the efficiency index [42]:

$$J_d = \frac{\left\|H_y(z)W_y^d(z)\right\|}{\left\|H_y(z)W_y^f(z)\right\|} \text{ or } J = \left\|\frac{\partial r}{\partial \varepsilon}\right\| \Big/ \left\|\frac{\partial r}{\partial f}\right\|.$$

More elegant and advanced H_∞ optimization methods are based on the algebraic Riccati equation [43]. Often, a slightly modified H_∞ filter is used to form the residual, i.e., the objective of design is to minimize the effect of disturbances and modeling errors on the estimation error and, consequently, on the residual. However, robust residual generation differs from robust estimation as not only disturbance attenuation is required: the residual must remain sensitive to faults while minimizing the effect of disturbances.

CONCLUSIONS

According to part II of the survey, fault diagnosis methods based on the classical modeling of an object monitored still attract many researchers. Fault diagnosis methods based on the identification approach, observers, and parity equations are actively applied.

A widely known solution is the Beard filter, representing a special case of the approach with algebraic invariants, where redundancy is artificially introduced via an auxiliary system. The redundant variable method is an effective approach to fault diagnosis. The robustness of fault diagnosis algorithms to various uncertainties, including modeling errors and the active noise of meters, deserves special attention.

In part III of the survey, we will analyze diagnosis methods based on neural networks, fuzzy and structural models, and models in the form of sets. Finally, part IV will be devoted to new approaches to fault diagnosis and combinations of different models and methods.

REFERENCES

1. Potter, I.E. and Suman, M.C., Thresholdless Redundancy Management with Array of Skewed Instruments: Technical Report, *AGARDOGRAPH. Integrity in Electronic Flight Control Systems*, 1977, no. 224, pp. 1511–1525.
2. Ivanov, N.V., Krasnov, I.A., Mozgalevskii, A.V., and Yudin, Yu.V., USSR Inventor's Certificate no. SU 475604 A1, *Byull. Izobret.*, 1975, no. 24.
3. Ray, A. and Luck, R., An Introduction to Sensor Signal Validation in Redundant Measurement Systems, *IEEE Contr. Syst. Mag.*, 1991, vol. 11, no. 2, pp. 44–49.
4. Chow, E.Y. and Willsky, A.S., Issues in the Development of a General Algorithm for Reliable Fault Detection, *Proceedings of the 19th Conf. on Decision & Control*, Albuquerque, NM, 1980.
5. *Osnovy tekhnicheskoi diagnostiki. T. 1: Model ob"ektov, metody i algoritmy diagnoza* (Fundamentals of Technical Diagnosis. Vol. 1: Models of Objects, Diagnosis Methods and Algorithms), Parkhomenko, P.P., Ed., Moscow: Energiya, 1976. (In Russian.)
6. Mozgalevskii, A.V., Technological Diagnostics. Continuous Objects (Survey), *Automation and Remote Control*, 1978, vol. 39, no. 1, pp. 118–136. (In Russian.)
7. Shumsky, A.E. and Zhirabok, A.N., *Metody i algoritmy diagnostirovaniya i otkazoustoichivogo upravleniya dinamicheskimi sistemami* (Methods and Algorithms for Diagnosis and Fault-Tolerant Control of Dynamic Systems), Vladivostok: Far-Eastern State Technical University, 2009. (In Russian.)
8. Kulida, E.L. and Lebedev, V.G., Modern Approaches to Prognostics and Health Management of an Aircraft Electromechanical Actuator, *Control Sciences*, 2024, no. 3, pp. 2–15.
9. *Diagnosis and Fault-Tolerant Control 1: Data-Driven and Model-Based Fault Diagnosis Techniques*, Puig, V. and Simani, S., Eds., Hoboken: John Wiley & Sons, 2021.
10. Mironovsky, L.A., Solov'eva, T.N., and Shintyakov, D.V., Fault Detection Optimization for Controllable Dynamic Systems, *Information and Control Systems*, 2019, no. 6, pp. 12–21. (In Russian.)
11. Mironovsky, L.A., Functional Diagnosis of Dynamic Systems, *Avtomatika i Telemekhanika*, 1980, no. 8, pp. 96–121. (In Russian.)
12. Basseville, M., Detecting Changes in Signals and Systems: A Survey, *Automatica*, 1988, vol. 24, no. 3, pp. 309–326.
13. Billings, S.A., *Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*, Hoboken: Wiley, 2013.
14. *Issues of Fault Diagnosis for Dynamic Systems*, Patton, R.J., Frank, P.M., and Clark, R.N., Eds., London: Springer-Verlag, 2000.
15. *Fault Detection, Supervision and Safety for Technical Processes. Safeprocess'91*, Isermann, R. and Freyermuth, B., Eds., Oxford: Pergamon Press, 1992.
16. Chernodarov, A.V., *Kontrol', diagnostika i identifikatsiya aviatsionnykh priborov i izmeritel'nykh vychislitel'nykh kompleksov* (Control, Diagnosis, and Identification of Aircraft Instruments and Measuring Computing Systems), Moscow: Nauchtekhlitizdat, 2017. (In Russian.)
17. Jazwinski, A.H., *Stochastic Processes and Filtering Theory*, New York: Academic Press, 1970.
18. Totiev, D.A. and Vinichenko, S.N., Kalman Filter Application for Signal Processing from Measuring Devices, in *Sbornik nauchnykh trudov kafedry avtomatiki i promyshlennoi elektroniki* (Collection of Scientific Papers of the Department of Automation and Industrial Electronics), Moscow: The Kosygin State University of Russia, 2024, pp. 165–168. (In Russian.)
19. Bukov, V.N., Bronnikov, A.M., Popov, A.S., and Shurman, V.A., Technical Condition Monitoring Methods to Manage the Redundancy of Systems. Part I: Built-in Control and Partition into Classes, *Control Sciences*, 2025, no. 2, pp. 2–10.
20. Chen, J. and Patton, R.J., *Robust Model-Based Fault Diagnosis for Dynamic Systems*, Norwell: Kluwer Academic, 1999.
21. Frank, P.M., Enhancement of Robustness on Observer-Based Fault Detection, *Int. J. Control*, 1994, vol. 59, no. 4, pp. 955–983.
22. Frank, P.M. and Ding, X., Survey of Robust Residual Generation and Evaluation Methods in Observer-Based Fault Detection System, *J. Process Control*, 1997, vol. 7, no. 6, pp. 403–424.
23. Liu, G.P. and Patton, R.J., *Eigenstructure Assignment for Control System Design*, Chichester: John Wiley & Sons, 1998.
24. Zhirabok, A.N., Shumsky, A.Y., and Suvorov, A.Y., Observer-Based Robust Fault Diagnosis: Logic-Dynamic Approach, Proceedings of the 10th International Conference on Informatics in Control, Automation and Robotics (ICINCO), Reykjavik, Iceland, 2013, pp. 239–244.
25. Isermann, R., Estimation of Physical Parameters for Dynamic Processes with Application to an Industrial Robot, *Int. J. Control*, 1992, vol. 55, pp. 1287–1298.
26. Clark, R.N., Instrument Fault Detection, *IEEE Trans. Aero. & Electronic Systems*, 1978, vol. 14, no. 3, pp. 455–465.
27. Willsky, A.S., A Survey of Design Methods for Fault Detection in Dynamic Systems, *Automatica*, 1976, vol. 12, no. 6, pp. 601–611.
28. Wünnenberg, J. and Frank, P.M., Sensor Fault Detection Via Robust Observer, in *System Fault Diagnosis, Reliability, and Related Knowledge-Based Approaches*, Tzafestas, S., Singh, M., and Schmidt, G., Cham, Switzerland: Springer, 1987.
29. Frank, P.M., Advances in Observer-Based Fault Diagnosis, *Proceedings of TOOLDIAG'93 Conference*, Toulouse, France, 1993, pp. 817–836.
30. Zhirabok, A.N. and Zuev, A.V. Interval Observers for Estimating Unknown Inputs in Discrete Time-Invariant Dynamic Systems, *Automation and Remote Control*, 2024, vol. 85, no. 8, pp. 747–760.

31. Beard, R.V., Fault Accommodation in Linear Systems Through Self-Reorganisation, *Technical Report MVT-71-1*, Cambridge: Massachusetts Institute of Technology, Man Vehicle Lab., 1971.
32. Jones, H.L., Fault Detection in Linear Systems, *PhD Thesis*, Cambridge: Massachusetts Institute of Technology, Department of Aeronautics, 1973.
33. Chen, J., Patton, R.J., and Zhang, H.Y., Design of Unknown Input Observer and Robust Fault Detection Filters, *Int. J. Control*, 1996, vol. 63, no. 1, pp. 85–105.
34. Gertler, J., Generating Directional Residuals with Dynamic Parity Equations, *Proceedings of the IFAC/IMACS Symp. SAFEPROCESS'91*, Baden-Baden, 1991.
35. Patton, R.J. and Chen, J., A Review of Parity Space Approaches to Fault Diagnosis for Aerospace Systems, *AIAA J. of Guidance, Contr. & Dynamics*, 1994, vol. 17, no. 2, pp. 278–285.
36. *Fault Diagnosis in Dynamic Systems, Theory and Application*, Patton, R.J., Frank, P.M., and Clark, R.N., Eds., London: Prentice Hall, 1989.
37. Chow, E.Y. and Willsky, A.S., Analytical Redundancy and the Design of Robust Detection Systems, *IEEE Trans. Automatic Control*, 1984, vol. 29, no. 7, pp. 603–614.
38. Britov, G.S. and Mironovskii, L.A., The Redundant Variables Method to Check Links with Fractional-Rational Transfer Functions, in *Informatsionnye materialy NS po probleme "Kibernetika" AN SSSR* (Information Materials of the Scientific Council on the Cybernetics problem of the USSR Academy of Sciences), 1971, no. 3. (In Russian.)
39. Ignatyev, M.B. and Katermina, T.S., The Redundant Variables Method for Checking and Correction of Computing Processes in Real Time, *SPIIRAS Proceedings*, 2013, vol. 3, no. 26, pp. 234–252. (In Russian.)
40. Zhirabok, A.N. and Kucher, D.N., A Robustness Improvement Method in Fault Diagnosis, *Vologdinskie Chteniya*, 2008, no. 69, pp. 6–9. (In Russian.)
41. Emami-Naeini, A., Akhter, M., and Rock, M., Effect of Model Uncertainty on Fault Detection: The Threshold Selector, *IEEE Trans. on Automatic Control*, 1988, vol. 33, no. 2, pp. 1106–1115.
42. Frank, P.M., Ding, S.X., and Köpper-Seliger, B., Current Developments in the Theory of FDI, *SAFEPROCESS'00: Preprints of the IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes*, vol. 1, London: Elsevier, 2000.
43. Zhou, K., Doyle, J.C., and Glover, K., *Robust and Optimal Control*, Upper Saddle River: Prentice Hall, 1996.

This paper was recommended for publication by V.G. Lebedev, a member of the Editorial Board.

Received November 10, 2024,
and revised April 23, 2025.
Accepted June 16, 2025.

Author information

Bukov, Valentin Nikolaevich. Dr. Sci. (Eng.), JSC Airborne Aeronautical Systems, Moscow, Russia
✉ v_bukov@mail.ru
ORCID iD: https://orcid.org/0000-0002-5194-8251

Bronnikov, Andrei Mikhailovich. Dr. Sci. (Eng.), Bauman Moscow State Technical University, Moscow, Russia
✉ bronnikov_a_m@mail.ru
ORCID iD: https://orcid.org/0009-0009-1216-3521

Popov, Aleksandr Sergeevich. Cand. Sci. (Eng.), Zhukovsky–Gagarin Air Force Academy, Voronezh, Russia
✉ saga30@yandex.ru

Shurman, Vladimir Aleksandrovich. JSC Research Institute of Aviation Equipment, Zhukovsky, Russia
✉ shurman@niiao.ru

Cite this paper
Bukov, V.N., Bronnikov, A.M., Popov, A.S., and Shurman, V.A., Technical Condition Monitoring Methods to Manage the Redundancy of Systems. Part II: Classical Models. *Control Sciences* **3,** 2–11 (2025).

Translated into English by *Alexander Yu. Mazurov*,
Cand. Sci. (Phys.–Math.),
Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ alexander.mazurov08@gmail.com