Introduction to Fault Diagnosis

This document is an abstract of the book by author Rolf Isermann [1].

The term **supervision** of technical process describes a set of actions, whose aim is to take appropriate actions to forgo damage or accidents. To obtain information about the state of the watched system the states of the system have to be closely monitored. Deviation from normal process behaviour are caused by *errors* and *faults* which with time may cause malfunction or failures. Faults in the system originate from following two sources:

- **external:** mainly environmental influences, like humidity, dust, chemicals, electromagnetic radiation
- **internal:** higher friction or wear (as a result of missing lubrication), overheating, leaks, shortcuts

These faults first influence the internal process parameters Θ (e.g. resistance, capacitance, stiffness) and/or internal state variables (e.g. mass flows, current, temperatures) and get propagated through the dynamic process transfer behaviour to the system's output Y as measurable change ΔY.

Faults and errors manifest themselves differently based on the type of the system:

- **Open-loop:** fault detection is simpler as faults directly affects internal variables, states and outputs of the system (as they cause permanent offsets) as no regulator is present which would try to compensate for the changes
- **Closed-loop:** fault detection is more complex as regulator compensates for the faults and only if the fault grows in size and causes the manipulated variable to reach saturation or permanent deviation can be noticed, therefore not just system outputs but inputs have to be also monitored.

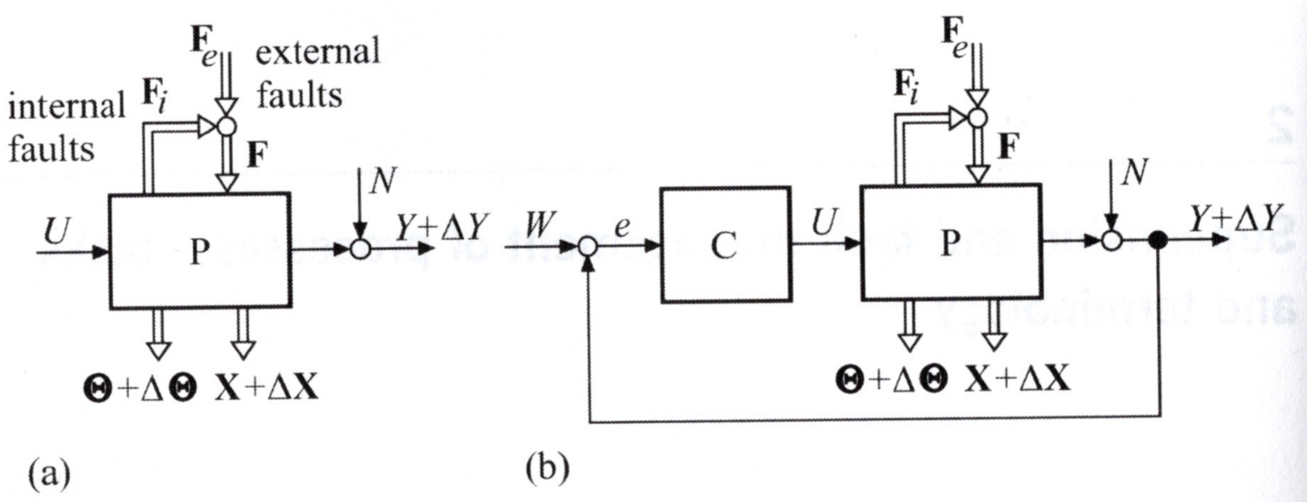

Figure 1. Schematic representation of the influence of external and internal faults on the internal process parameters, states of (a) open-loop, (b) closed-loop system

For monitoring and automatic prediction the following two approaches are used:

- **Classical methods**: This group of methods are based on monitoring the output variables and comparing it to a set of thresholds. If the output variables exceeds the given thresholds the system is deemed to be faulty.
- *Pros*:
  - simple
- *Cons*:
  - only static changes can be monitored
  - some fault states are not measurable or the boundaries between faulty and functional states are not easily distinguishable
  - in-depth faults diagnosis is not possible
- **Advanced methods**: They utilise the available measurements and relate them together in the form of mathematical process models, in this way maximising the amount of useful information on the system. These methods have to meet the following requirements:
  - early and fast detection of even small faults
  - diagnoses of process parts and their manipulating devices (actuators) and measurement equipment (sensors)
  - detection of faults
  - supervision of the process in transient states

Supervision with fault detection consists of the following steps:

1. *Feature generation*: - e.g. special signal processing, state estimation, identification and parameter estimation
2. *Fault detection and generation of symptoms*
3. *Fault diagnosis:* identifying faults from extracted features
4. *Fault evaluation*: - classification of faults into different hazard classes
5. *Decision on actions*: - automatic/manual intervention

In fault management loop feature and symptom generation work continuously, while fault diagnosis, decision making and repair act as discrete event.

Supervision actions and fault management depending on the hazard classes of the diagnosed faults.

- *safe operation:* - imminent shut down in case of danger
- *reliable operation*: - hindering further fault through changes in operation state
- *reconfiguration*: - using redundant parts to maintain system operation
- *inspection*: - detailed diagnosis by additional measures
- *maintenance*: - exchange of worn parts, it can take place instantaneously or at next possibility
- *repair*: - fault removal

Useful terms:

- **Fault:**
  - unpermitted deviation of at least one property (feature) from acceptable condition

- it may result in loss or reduction of capability to perform required function
- they are independent wether a system is in operation or not
- frequently, they are hard to detect, especially if they are small or hidden
- fault may initiate failure or malfunction

- **Failure**:
  - permanent interruption of system's ability to perform required function (it is an event)
  - results from one or more faults
  - following categorisation can be applied to failures:
    - number:
      - single
      - multiple
    - predictability:
      - random
      - deterministic (predictable under certain condition)
      - systematic (or causal) failure (depends on known conditions)
- **Malfunction**:
  - intermittent irregularity in the fulfilment of system's desired function (temporary abruption)
  - results from one or more faults
- **Reliability:**
  - ability of a system to perform required function under stated conditions, during given scope of time
  - can be affected by malfunction and failures
  - its measure is called *mean-time-to failure* (or MTTF) , which is defined as:
    $$MTTF = \frac{1}{\lambda},$$
    where λ is the *rate of failures per time unit*.
  - the improvement of reliability are oriented towards reducing faults, failures and malfunctions
  - the improvement of reliability leads to improvement of safety, however the opposite is not always true
- **Safety:**
  - ability of a system not to cause any danger to person or equipment or the environment
  - it is concerned with dangerous effects of faults, failures, malfunctions
  - a system or equipment is considered safe when the risk of its operation is not larger than a specified limit
- **Availability:**
  - the probability that the system or equipment will operate satisfactory and effectively at any period of time
  - the *availability measure* is the ratio between the amount of time the system is operational and time it is under repair, it is given by:
    $$A = \frac{MTTF}{MTTF + MTTR},$$
    where $MTTR$ is the abbreviation for mean-time-to repair
  - availability can be improved by:
    a. *large operation time*
       - *perfection of parts*: high reliability of components
       - *tolerance*: tolerable faults through reliable structure
    b. *small repair time*

    - fast and reliable fault diagnosis
    - fast and reliable fault removal
- **Dependability**
  - form of availability that has the property of always being available when required (at any randomly chosen time)
- **Integrity**
  - is the ability to detect faults in its own operation and to inform human operator
- **Fault tolerance**
  - by design not all faults can be totally avoided, therefore the system has to be designed to be fault tolerant, i.e. faults are compensated in such way that they do not lead to failures
  - it can be improved by improving redundancy by addition of redundant function modules (either software or hardware) either identical or diverse

Knowledge based fault detection

Fault detection in general is based on:
- measured variables
- observed variables and states by human operator

The automatic processing of measured variables requires:
- analytical process knowledge
- evaluation of observed variables requires human expertise, also known as *heuristic knowledge*

Symptom generation

The analytical knowledge about the process is used to produce quantifiable, analytical information. To do this data processing based on measured process variables is performed to generate the characteristic values by:
- limit value checking
- signal analysis using signal models (correlation function, frequency spectra , ARMA)
- process analysis by using mathematical process models (which include parameter estimation, state estimation, parity equations)

In later stages of processing special features can be extracted from the characteristic values (e.g. process coefficients, filtered or transformed residuals), which are then used to identify faults. For this, methods of change detection and classification are applied.

Analytical symptoms appear as a result of changes (discrepancies) in the measured signals, signal models, process models. In addition to the analytical symptoms, *heuristic symptoms* can be generated by using qualitative information from human operators. For this reason the following sources of information can be used:
- *process history*: information about maintenance, repair and former faults
- *statistical data*: obtained from same or similar processes

Fault diagnosis

The task of fault diagnosis consists of determining the type, size, location of most possible faults, as well as its time of appearance. Fault detection methods use analytic and heuristic symptoms, which should be presented in a unified way (confidencenumbers, membership functions, fuzzy sets etc. …) Then classification or methods or clusters can be applied to identify the type and source of the faults.

Reference:

[1] ISERMANN, Rolf, 2006. Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance. 1. Springer-Verlag Berlin Heidelberg. ISBN 978-3-540-24112-6.