Electronics in the on-line control of railway movements: quantitative aspects
| AUTHOR | GELBSTEIN, E. E |
| COPY NO. | 020103/01 |
| VOL NO. | CLASS MARK |
| ARCHIVES | COPY |
| FOR REFERENCE ONLY |
ELECTRONICS IN THE ON LINE CONTROL OF RAILWAY MOVEMENTS
QUANTITATIVE ASPECTS

BY

EDUARDO ENRIQUE GELBSTEIN C.Eng., M.I.E.E.

A Doctoral Thesis

Submitted in partial fulfilment of the requirements for the award of

Doctor of Philosophy of the Loughborough University of Technology

November 1975

Supervisor: W. FORSYTHE
Department of Electrical and Electronic Engineering

© by Eduardo Enrique Gelbstein, 1975
ELECTRONICS IN THE ON-LINE CONTROL OF RAILWAY MOVEMENTS

by Eduardo Enrique Gelbstein

Doctoral Thesis

ABSTRACT

The present thesis is concerned with a quantitative examination of the on-line control of railway movements and develops a mathematical technique for the evaluation of safety based on the use of Markov processes, illustrated with examples.

In addition, the thesis presents a design methodology applicable to electronic safety systems. These systems are shown to be an essential element in the development of fully electronic railway signalling systems, as well as in the increased automation of railway movements.

An analysis of the limits of automation of railway movements is described and discussed together with a possible system configuration for the achievement of crewless train operation.

The research described herein has been carried out at the British Railways R & D division and the methods described have been successfully applied to real engineering problems.

The industrial R & D background of the present thesis is also reflected in the inclusion of a section on the socio-economic consequences of major innovation, particularly in the field of automation and in the consideration of costs and benefits.
Section 2 contains an approach evolved jointly with Mr. W.T. Parkman, also at the R & D Division of British Railways, and has been published as Reference 16.

Section 5 is a short description of the work carried out by the group under the direct responsibility of the author at the R & D Division of British Railways.

ACKNOWLEDGEMENT: The author wishes to thank Dr. K.H. Spring, Head of Research and the management of the British Railways R & D Division at Derby for their permission to prepare this thesis and their support.
LIST OF CONTENTS

1. SUMMARY 1

2. THE ON-LINE CONTROL OF RAILWAY MOVEMENTS 4
   2.1. Vehicle controls 4
   2.2. Safety controls 5
   2.3. Train regulation 7
   2.4. The planning loop 9

3. SAFETY SYSTEM PRINCIPLES 11
   3.1. Introduction 11
   3.2. Failure to safety 13
   3.3. Extension of the concept of failure to safety to include data 17
   3.4. Performance of safety systems 19

4. ANALYSIS OF SAFETY SYSTEMS 24
   4.1 Mathematical techniques 24
   4.2. Engineering considerations 28
   4.3. System structures for railway applications 30

5. DESIGN OF SAFE DIGITAL SYSTEMS 36
   5.1. General 36
   5.2. Data encoding 37
   5.3. Data transmission 39
   5.4. Data processing 41
   5.5. Brief review of relevant trends in electronic technology and conclusions 45
   5.6. Validation of safe systems 46
6. APPLICABILITY OF ELECTRONIC SAFE SYSTEMS
   PERCEIVED OPPORTUNITIES IN RAIL TRANSPORT
   6.1. Solid state techniques in railway signalling
   6.2. Automatic train operation

7. CREWLESS TRAIN OPERATION
   7.1. Formulation of the problem
   7.2. Crewless train operation tasks
   7.3. System structure

8. SOCIO-ECONOMIC IMPLICATIONS
   8.1. Automation and the user
   8.2. Automation and the operator
   8.3. Automation and the railway

9. CONCLUSIONS
APPENDIX I - MATHEMATICAL TECHNIQUES FOR THE ANALYSIS OF SAFETY SYSTEMS

Symbols

Definitions

I.1. General
I.2. Basic concepts
I.3. Steps in the evaluation of systems performance
I.4. Formulation of the probability of accident

APPENDIX II - CALCULATIONS OF AVAILABILITY, RELIABILITY AND SAFETY FOR SIMPLE SYSTEM STRUCTURES

II.1. Calculations for n/n
II.2. Calculations for 2/3 system with no notification of first fault
II.3. Calculations for 2/3 system with notification of first fault
II.4. The generality of the availability equation
II.5. Possible optimisation technique
II.6. Conclusions

APPENDIX III - EXAMPLE OF EVALUATION PROCEDURE OF A SAFETY SYSTEM (BRAWS)

III.1. Operation of BRAWS
III.2. Information in BRAWS
III.3. Performance of BRAWS
III.4. Analysis procedure
III.5. Sensitivity Analysis
III.6. Limitations of the study
APPENDIX IV - COST BENEFIT ANALYSIS FOR SOLID STATE TECHNIQUES IN SIGNALLING

APPENDIX V - RELATION OF THE PRESENT THESIS WITH OTHER RESEARCH ACTIVITIES AT LOUGHBOROUGH UNIVERSITY

REFERENCES

36 FIGURES.
1. SUMMARY

The present Thesis examines some of the problems and opportunities related to the use of electronics in railway control systems.

Although railways make considerable use of electronics, this is often in telecommunications and instrumentation, rather than in control.

Section 2 presents a short description of railway controls as feedback systems, and although simplified and inadequate for modelling purposes, these feedback systems are considered to be useful in identifying the fundamental functions, such as making the system safe by avoiding accidents.

Railway safety has been achieved over the years through both technology and disciplined use of Rules and Regulations — The concept of "failure-to-safety", discussed in Section 3, has been successfully applied to railway signalling and also vehicle braking systems.

When considering the use of electronics in systems with identifiable safety requirements, it becomes desirable to extend the concept of failure to safety to include data. This is done in Section 3 by introducing the concepts of "Failure" and "Error", as well as the hypothesis that the ability to detect the presence of faults in a system and the
ability to take action after detecting a fault are sufficient conditions for the implementation of a system exhibiting "Failure-to-Safety" characteristics.

In addition two new concepts are introduced: The first one is that Failure-to-Safety may not be the best way to describe the performance of a safety system. This is complemented by proposing the definition of "Probability of Accident", which takes into account other factors such as human error and the nature of standby systems where these are used.

The second concept, presented in Section 4, is that a mathematical description of the behaviour of a safety system can be achieved through the use of Markov processes and this is shown in greater detail in Appendix I.

Having established that a digital system can be used for safety and its behaviour quantified, Section 5 examines the problems of designing safe digital systems and the implications on encoding, data transmission and data processing. Appendix II analyses the behaviour of several processing configurations with regards to safety and reliability.

Section 6 presents a discussion on how the techniques described in preceding sections can be used in the design of a new generation of railway signalling equipment and in automatic train operation.
Section 7 presents a discussion on how electronic techniques can be used to achieve crewless train operation and the implications of doing so. This material is presented as an extension of the discussion in Section 2.

Section 8 contains a very short discussion on the socio-economic implications of innovation, particularly in an established transport technology. This discussion is not as extensive as the subject deserves, but it is introduced to indicate the author's awareness of the existence of the problem.

Appendices III, IV and V are included to describe practical aspects of the above work.

The work on Automatic Train Operation which is the background to Section 7, has been carried out by a team under the leadership of the author and the ideas on data processing presented in Section 4 have also been extensively discussed with members of the team at the Railway Technical Centre, British Railways Board, DERBY.

The remainder of the material presented herein is the work of the author.
2. THE ON-LINE CONTROL OF RAILWAY MOVEMENTS

This chapter presents a summary of the control system organisation of present on-line activities, mainly as a background to the original material presented in later chapters, although this form of description of a railway system has not — to the author's knowledge — been published elsewhere.

2.1. VEHICLE CONTROLS

The problem of having to control both the power and braking in a moving vehicle is undoubtedly familiar to everybody. Recent interest in new forms of public transport, such as the "personal rapid transport" (PRT), "minitrams" and "automated guideways", has resulted in a very large number of articles dealing with the dynamics of vehicles and the design parameters affecting the dynamics.

In the case of railways, the wide choice of motive power source, e.g. gas turbine diesel, DC electric, AC electric (and even steam), as well as the variety of possible braking systems, such as electronically controlled electropneumatic, vacuum, tread brakeblocks, disc brakes, etc., make this not only a very wide subject, but also a well researched one.

Because of this, this subject will not be examined in further detail in this Thesis.
2.2. SAFETY CONTROLS

In any form of transport, but particularly in the case of public transport, there is a need to avoid accidents, particularly those liable to produce injury. Railways have always been very conscious of this need, and are assisted in this function by an independent office, that of the Inspecting Officer of Railways, of the Department of the Environment.

It can be said that the signalling system carries out two fundamentally different functions:

- it selects the required route for each train movement.
  
  This involves the driving of a suitable device (the points) to the required position to allow a train to follow a particular path. It is essential to ensure that this device cannot move or change state while a train is traversing over it, as this situation could result in a derailment.

- it provides information to ensure that two trains travelling on the same line cannot collide. This spacing of trains is achieved by suitable logic and the information is transmitted to the train driver by means of lineside signals. In current practice, these display colour lights, and each indication contains information concerning the state of the track ahead. Figure 1 illustrates some basic relations as implemented by British Railways.
Both these functions depend on the knowledge of the position of each and all trains. This is achieved at present by dividing the track into finite length electric circuits, using the rails as conductors and relying on the low resistance of wheels and axles to short the circuit and interrupt the current flow through the relay coil, as shown in Figure 2. It is clear that this technique is only applicable to "steel on steel" technology, but it is a very simple - and therefore reliable - technique.

The track occupation, that is the position of trains, is one of the independent variables in a logic system - the other independent variable being the desired path or route - which after processing will drive and lock points and signals (a more detailed account of the functions performed by this logic and its organisation is given in Section 6). This logic is called INTERLOCKING, probably because in an earlier implementation the logic was performed by mechanical bars with slots, according to which position other bars could be moved and locked. Current technology uses electromechanical relays which have special design characteristics resulting in a very low probability of welded contacts - in the language of signalling technicians "it can't happen" - and this fact is used to design predictable failure response. This subject will be examined in further detail in Section 3.

Figure 3 shows a feedback loop in which the elements mentioned above are presented as inputs and outputs to the interlocking.
It will be noticed that the train driver, who is responsible for observing variable information (i.e. the signal aspect) and all fixed information, such as position and values of all speed restrictions associated with the track, the position and values of all track gradients (affecting braking performance), the timetable for his particular train, etc., is not a part of the feedback loop.

2.3. TRAIN REGULATION

Most known public transport systems other than taxis operate according to a schedule (rather than on demand), which is usually expressed in the form of a timetable. This timetable is usually designed as a compromise between resources, commercial viability, demand and social requirements, but the relationship between all of these exceed the scope of this Thesis.

Railways throughout the world have spent considerable effort and investment in the field of train regulation in order to smooth the effects of perturbations in the network, as there is abundant evidence to indicate that in many operating conditions the system can become "unstable" if perturbations exceed a given threshold. This "instability" becomes apparent by the propagation of delays and the inability to return to the original schedule without drastic action such as the cancellation of a number of services.
The use of the term "stability" may be inappropriate, as there is not sufficient understanding of the control system performance of a railway network to determine if there are "poles on the right half plane", but in the author's opinion it suitably describes the situation.

The equipment assuming this function is called a Train Describer and modern technology allows the use of minicomputers to perform the logic required to follow trains as they move along their route and display their identity on a suitable panel in the operating room of a signalbox (ref. 9).

A more detailed description of a railway signalling system is given in Section 6.1.

Figure 4 shows another feedback loop, called the "Train Regulation Loop", indicating in it the relations between the Safety Loop indicated in Figure 3, the basic role of the Train Describer, as well as some of the activities currently being researched within the Train Control Group, British Railways R & D Division, DERBY.

The top right hand corner of Figure 4 is a simplified version of the Safety Loop shown in Figure 3. (The box labelled I.D.F. represents an Internal Distribution Frame, equivalent to an electric junction box).

The full lines in Figure 4 indicate those elements and information paths already in use. Dotted lines are used to indicate developments now in progress in the Train Control Group
The following terms may not be self explanatory and are, therefore, defined:

Berth: Any position in which a train description may exist.
In current practice train descriptions are displayed by 25 mm x 50 mm cathode ray tubes on a signalling display giving the geographic layout of the area under control.

Fringe box: The most remote signalbox in an area where there is a transition in signalbox technologies, e.g. mechanical box to electromechanical box.

Train Reporting by Exception: The reporting of only those trains running other than according to the timetable.

Sequenced Route Commands: Automatic setting of routes according to timetabled moves (only possible when all trains are running as planned).

Target Speed/Distance Profile: Specification of an "optimum" speed/distance trajectory to minimise any specified cost function (delay, energy, etc.).

2.4. THE PLANNING LOOP

Figure 5 shows yet another feedback loop, this one concerned with the planning activities that take place in "almost real time", although the time constants may vary between minutes, as in the case of reporting empty or cripple wagons, to a week in the alteration of timetable information.
The details concerning the reporting and control of the vehicle fleet, exceed the scope of this Thesis and are not original, except for the case of "failure recovery" in crewless train operation, considered in Section 7.

This feedback diagram is included however, to show that the elements presented in Figures 3 and 4 are only a part of the overall control activities. It should be clear that data processing and simulation have a major role to play in the day-to-day operation of a railway network. A more detailed description of these loops has been published in Ref. 9.
3. SAFETY SYSTEM PRINCIPLES

3.1. INTRODUCTION

Electronic technologies have, over the last decade, undergone a major change, and offer the possibility of implementing systems that were impractical or uneconomic only a few years ago.

It has long been recognised that electronics could play an important part in railway signalling, as indicated by such non-safety applications as the computer based train describer and, perhaps even more important, safety systems such as the successful all-electronic interlocking at Henley-on-Thames built in the early 1960's.

It is felt, however, that with the advent of LSI (large scale integration), microprocessor subsystems and the increased knowledge of information and reliability theories, there is a need for a detailed review and also to establish a "Theory of Safety".

This Thesis sets out to examine and, where appropriate, extend the established concepts of safety and failure to safety. In addition the techniques available for the design of safe systems will be described in section 4.

Although these topics lend themselves to mathematical techniques it is considered that, wherever possible, mathematics and philosophy should be treated separately.
This section is divided in three subsections, each dealing with a separate aspect of the problem:

The first of these, 3.2, examines the concept of "failure to Safety" and analyses the necessary and sufficient conditions for a system to exhibit "failure to safety" properties.

The second subsection, 3.3, extends the concept of "failure to safety" to data, and in particular to data expressed in binary form, as it is considered that this extension will permit the development of a new generation of safety systems using digital integrated circuits.

The final subsection, 3.4, is both philosophical and mathematical, as, in order to quantify the "safety" of a system in a manner similar to that of the measurement of system reliability, a mathematical model of the Probability of Accident is developed and analysed by the use of Markov Processes.

The definitions of failure to safety, reliability, safety, etc. are presented in summary form in the text, where appropriate and are listed at the beginning of Appendix I.
3.2. FAILURE TO SAFETY

3.2.1. General.

The concept of failure to safety, although well established, is not used in a quantitative manner.

A system is said to be fail-safe if it reverts to a more restrictive state for any foreseeable fault condition or plausible combination of fault conditions.

In practice such behaviour is obtained by the combined use of suitable design principles and suitable components, as illustrated by the following two familiar examples:

Example 1: Failure-to-safety features in a D.C. track circuit.

Figure 6 shows a basic D.C. track circuit configuration. The system is so arranged that a current will flow through the relay coil at the receiver and only when the track is clear and there is no other probable fault present.

The relay is assumed to be fail-safe, in the sense that it is virtually impossible for its contacts to remain closed in the absence of current in the coil (see Example 2).

The track circuit in Figure 6 has all of the following properties:
- It will detect the presence of a train, as the axles will act as a shunt for the current normally flowing through the relay coil. The open contacts thus indicate the track is not free.
- 14 -

It will indicate the presence of faults such as: Broken rail, broken or disconnected feeds, open circuit relay coil, etc., by de-energising the relay, giving a "track occupied" indication, which is more restrictive than "track clear".

There are nevertheless situations which can lead to a dangerous situation, e.g. a train, the presence of which is undetected. Among others, a high resistance between wheelset and rail due to the presence of rust or insufficient contact pressure are likely examples.

Example 2: Special components.

An example of components with inherent failure to safety characteristics can be found in railway signalling relays. In these both the contact materials and the mechanical construction are chosen so as to make the undetected welding of front contacts virtually impossible.

In both examples it is the underlined "virtually" that is specially significant since it indicates that there is no absolute certainty that such a situation cannot occur, particularly when multiple fault situations are considered.

It is, therefore, necessary to accept that there is a probability, albeit small (with its associated confidence level), that "failure to danger" may occur.
3.2.2. Suitability of systems for safety applications - General philosophy

The flowchart in Figure 7 presents a means of classifying systems according to their characteristics.

Each stage in the flow chart represents a particular family of systems:

- **Infinite reliability systems**: Although no physical system can be said to have infinite reliability, in practice, there are systems of which the operational life is extremely short in relation to the time between failures. In virtually all the applications of this type of system, testing the system correctness is an essential part of the procedures until such a system becomes operational (good examples of this can be found in the space program).

This approach is obviously not applicable to railway signalling with typical operating lives ("mission time" in the reliability jargon) in excess of 25 years.

- **Ability to detect faults**: The lack of this ability implies that the operation of the system can be allowed to continue unchecked even when faults have occurred. While this may be of little importance in an audio amplifier, the obligations imposed on public transport operators with regard to safety make it essential to protect the system against faults, and the first logical requirement is to be able to detect that these have occurred.
- Ability to take action when faults have occurred: Knowledge of a failure condition is, per se, not sufficient to guarantee that the system is safe, and a suitable executive mechanism is required to take the system to a state defined as safe. Present signalling technology, i.e. relays, combines these last two properties in a single component. Thus, absence of current on a relay coil (detection) will open the contacts (action).

- Type of action taken: Assuming that the previous two conditions have been met, there is a choice of course of action:
  a. The system can be taken to a "safe" state where further action is inhibited until the system has been restored to normal operation, or
  b. The system reverts automatically to a standby system, known to be operational and ensuring full or partial operation.

It may be of interest to note that other systems assuming safety functions may require a different approach, due to fundamentally differing requirements.

For example, in automatic landing equipment for aircraft, the requirement is for correct performance over a very short period of time. This permits considerable testing of the equipment (and its standby units) prior to its use. This philosophy
allows the system to guarantee a system failure rate no worse than 1 in $10^7$ landings.

A somewhat similar situation exists in the protection of a nuclear reactor, where the requirements are that the system must operate correctly when needed, but where there is no a-priori knowledge of when it will be needed. Normally, the equipment is standing by.

In railway signalling, the requirement is for continuous operation in a large number of logical combinations, and each combination must be safe.

3.3 EXTENSION OF THE CONCEPT OF FAILURE-TO-SAFETY TO INCLUDE DATA

Figure 7 shows why the ability to detect the presence of faults is an essential component in fail-safe systems. These faults can be either the result of component failures (permanent) or interference (temporary).

In a data system, and regardless of the nature of the data (e.g. binary, decimal or plain language, etc.), the presence of faults in the system can lead to any of the following outcomes:

a. The data given by the system is the same as it would have been had the system not been affected by faults — in other words, the faults have no effect on the system for the particular data being considered.
b. The data given by the system can be identified as being incorrect by direct observation - in other words, the effects of the fault are obvious.

c. The data given by the system cannot be identified as being incorrect - in other words, it appears by observation to be plausible.

Conditions b) and c) can be labelled "Data Failure" and "Data Error" respectively, and are equivalent to "Failure to Safety" and "Wrong Side Failure".

The proposed definitions are general and are, in a way, more restrictive than those currently used in railway signalling, as illustrated by the example shown in Figure 8.

A lamp failure giving a single yellow instead of a double yellow indication is regarded as safe. Considering the data patterns however, such a failure would be classed as a "data error" and classified as a "wrong side failure", thus making the definition more restrictive, which is, in itself, safe.

The reason for this approach can be found in "Information Theory". In it, many techniques for protecting against the corruption of data have been developed and these rely on the knowledge of what random process the technique must protect against.

The success of the protection can be measured in terms of probabilities and the probability of detection of faults can be made arbitrarily high - at a cost. This will be discussed further in Section 4.
3.4. PERFORMANCE OF SAFETY SYSTEMS

3.4.1. General

While failure to safety, as defined in section 3.2., is a well established principle and indeed, a fundamental requirement in the design of railway signalling equipment, the performance of a safety system as a whole, depends on a number of other factors, to be described in detail in this chapter.

In addition, suitable methods for the numerical description of these factors - including safety - will be briefly described. These techniques are discussed in sufficient detail for their practical use.

3.4.2. Probability of accident

The real environment in which a safety system operates includes factors such as human intervention at various levels, events against which the system does not offer protection, etc. All these factors combine in practice to create different situations which may lead to an accident.

The most suitable tool for analysis is Probability Theory, since ultimately, accidents do not occur by design but rather by combinations of unexpected (but often predictable) events. The knowledge of an event's probability gives an indication of its average frequency (in space or time) and is useful to describe the overall performance of a complex system.
3.4.3. **Factors contributing to the "Probability of Accident"** $P_A$

The simplest - although practical - safety system is that in which only three conditions can exist, and only one at a time.

a. The system is working correctly.

b. The system has failed to safety.

c. The system has failed to danger.

Figure 9 shows a symbolic representation of these states and the possible transitions from one to another. In this simple system, it is assumed that it is not possible to pass from one failure state to another.

It is further assumed that, regardless of how much effort is placed into considering every conceivable source of failure and protection against all contingencies the probability of a failure to danger is not zero. After all, every designer must make some assumptions in his fail-safe philosophy.

The diagram in Figure 9 is not complete for the simple reason that it does not provide for a return path to state 'a' after failure has occurred, that is, reassume normal operation after a failure has been noticed, identified and repaired. Since signalling equipment is expected to provide continuous service, Figure 10 illustrates a somewhat more realistic situation in which the influence of both maintainability and the maintenance organisation is shown.
The need to provide continuous service just mentioned, introduces yet another factor, that is, the need to use some form of standby system able to offer at least a partial ability to operate the system while the repair takes place. This situation is illustrated in Figure 11, where a' is the temporary (while the repair lasts) state of operations. These concepts will be discussed in greater detail in Section 4.

In those cases where this standby system relies on human decisions, it can be expected that the human failure rate will be much greater than the corresponding equipment failure rate during normal conditions.

From these models, which can be developed to any degree of sophistication, it is possible to determine the possibility of being, at a given time, at states a, b or c. The techniques required for this purpose are given in Appendix I.

Having defined these three probabilities, namely $P_a$, $P_b$ and $P_c$, it is possible to list those situations that can lead to an accident. The following situations can be recognised as leading to accidents:

- The occurrence of an event for which the safety system gives no protection. In the case of railway signalling, an example of such an event could be the presence of an undetected obstacle on the track. This event can be associated with a probability $P_e$. 
The combination of events such that the correct operation of the safety system is insufficient to prevent an accident. An example of this could be found when a driver correctly applies the brakes to stop at a signal, but, due to local adhesion conditions, goes past the signal and derails. The conditional probability $P_{E1/1}$ will be used to represent these cases.

The combination of events such that the incorrect operation of the standby system (Figure 11) occurs while the system is in a safe failure state. An example of this can be found in authorising the wrong train to proceed past a signal at danger (Monmore Green, 1969). The conditional probability of such an event will be indicated by $P_{E2/2}$.

The combination of events such that an action which should be prevented occurs during the time when the system is in a dangerous failure state. The conditional probability of such a combination of events will be indicated by $P_{E3/3}$.

If $P_A$ denotes the "probability of accident" and it is assumed that the system can only be in one of its three states (normal, failed safe, failed dangerous) at any given time, then

$$P_A = P_E + P_1 \cdot P_{E1/1} + P_2 \cdot P_{E2/2} + P_3 \cdot P_{E3/3}$$

$$- P_E \cdot P_1 \cdot P_{E1/1} + P_2 \cdot P_{E2/2} + P_3 \cdot P_{E3/3}$$
To apply this approach in practice, all the values of these probabilities need to be identified. The three conditional probabilities can be determined by examining past accident records, analysing them by primary and secondary causes and expressing the results in the appropriate units. An example of a procedure that could be followed can be found in Appendix III.

$P_c$ will again be defined by historical records where these exist, or will need to be defined "a priori" if no other information is available and then revised at a later date.

The probabilities $P_1$, $P_2$ and $P_3$ depend on the actual design and implementation of the system and should, in principle, be calculable, as shown in Appendix I. These probabilities are identical to $P_a$, $P_b$ and $P_c$ mentioned above.

One possible procedure for carrying out these calculations is described in Section 4.

Before describing the calculation procedures it may be worth drawing attention to the fact that, as shown by the equation of Probability of Accidents, "Failure to Safety", in the ideal case $P_3 = 0$, does not necessarily make $P_A = 0$. 
4. ANALYSIS OF SAFETY SYSTEMS

4.1. MATHEMATICAL TECHNIQUES

In order to quantify the behaviour of a system in which its safety features are considered to be essential, it would be desirable to develop a mathematical model of their operation.

Section 3 has discussed the concept of "Failure to Safety" and also defined the terms "Data Failure" and "Data Error", associated with extending the concept of Failure to Safety to include data patterns.

The diagrams shown in Figures 9, 10 and 11 represent the simplest case, where a safety system (or subsystem) can exist in only one state at a time and in which the only possible states represent normal (correct) operation and fault states, one defined as being "safe" and the other as being "dangerous" - these definitions being a function of the total environment in which the safety system operates.

Transitions from one state to any other will be assumed to be caused by failures when moving from the normal state to a fault state and by repair procedures when returning to the normal states.

As it is common practice in reliability theory, the differential probability of transition from one state to another is labelled as a failure rate (or maintenance rate as appropriate). The
model studied in Appendix I does not depend on these rates being constant but assumes their mutual statistical independence.

These failure rates can be determined if the structure and constructional characteristics of the system are known by using established techniques, such as Failure Tree Analysis and/or Failure Mode and Effect Analysis (FMEA). A short description of the elements of such analysis is given in the first part of Appendix I, but it is not claimed to be original.

If all the transition rates can be defined, then Markov processes can be used to describe the (fault) behaviour of a system, and Appendix I examines several simple systems with two and three states to show how the whole of Reliability Theory can be shown to be a special case of Markov processes.

Although various reliability terms will be used in accordance to their "official" definitions in the British Standards and MIL-Specifications, some of the fundamental definitions are given here for reference purposes:

**Reliability:**

a) Reliability is the ability of the system/unit in question to perform a required function under stated conditions for a given period of time.
b) Reliability is the probability of survival, i.e. the probability that the system/unit in question of age \( x \) will not fail at time \( t \) or before, or, inversely, the probability that the product will, without failure, perform a required function under given conditions for a certain period of time.

**Availability:**

The probability of finding a system in its normal operating state at a given moment of time. Availability is a function of both failure rates and the repair time.

The calculation process requires the formulation and solution of a set of differential equations of the type

\[
\frac{dP_i(t)}{dt} = [C] P_i
\]

where \( P_i(t) \) is the probability function for each state in the system and \( C \) is a coefficient matrix for the system. As the examples in Appendix I illustrate, one attractive feature of this method is that it permits the direct evaluation of the probability of a Dangerous Failure state.

The next subsection, 4.2, will examine how the failure rates of a system can be modified by using engineering in its widest sense.
Safety (not defined in BS or MIL specifications).

A system is called "safe" if hazardous events on its operating environment, imposed by failures, are excluded with a sufficiently high probability, appropriate to its application.

Such a concept could be formulated as

\[ S = 1 - P_A \]

where

- \( S \) = System safety
- \( P_A \) = Probability of Accident (as defined in the preceding text).
4.2. ENGINEERING CONSIDERATIONS

4.2.1. Safety, Reliability, Availability - Trade-offs

While it is clear that the ideal system should never fail, the long service life expected from railway equipment makes the achievement of this ideal impossible with present technology.

As the previous section has shown, failure to safety is, by itself not sufficient to entirely avoid accidents and in fact, the lack of availability (due to failures to safety) not only affects the quality of service but also contributes to the probability of accident.

Failure to safety can be shown, from past accident records to make the signalling system so safe that less than 1% of all accidents can be directly attributed to signalling failures. Those that are, seldom occur in power box controlled areas and are due to mechanical component failures, such as failures of points locking bars, consisting of steel bars, nuts and bolts.

It can be further argued that when emergency procedures are used (pending restoration of normal operation) the severity of the accidents arising from human error is not as high as the likely consequences of a wrong side failure. However, this is not always the case (e.g. Monmore Green, 1969) and there is no doubt that the availability of the system should be as high as possible.
The need to incorporate Failure to Safety features in equipment can, in many cases lead to a loss of reliability in the overall system, as a certain degree of checking, and therefore, a form of additional equipment, is required. This must be accepted as unavoidable and in consequence there is a need for identifying the best compromise between all the following factors - (even though this compromise may well be different for each application).

Technology and standard of engineering: In addition to meeting the Failure to Safety requirements, equipment must be designed to such a standard (for any given technology) that the maximum possible reliability inherent in the particular technology is achieved (Reference 3).

This activity is mostly concerned with providing a suitable environment for the operation of the equipment.

System configuration: The way in which the equipment is organised, e.g. whether it is centralised or decentralised, whether it incorporates diagnostic aids, how accessible for repair it is, etc., has a fundamental influence on failure sensitivity. The consequences of a fault may not propagate beyond their local environment (e.g. failure of one filament in a two filament lamp) or they could, at the other extreme, be catastrophic (e.g. total failure of the power supply due to a lightning strike). System configuration is an important factor in defining its sensitivity to faults and, in turn, the standby requirements.
This will be examined further in 4.3.

Standby arrangements: Usually determined by the failure sensitivity, what is available in terms of equipment and manpower and what is economically viable.

All these factors and their interactions can be analysed by means of Operational Research techniques.

4.3. SYSTEM STRUCTURES FOR RAILWAY APPLICATIONS

Given a specification for a safety system, there will be, in addition to the basic operational specification, a number of constraints and operational conditions. These will have considerable influence on the choice of methods in which the system can be structured to meet the desired level of performance in terms of availability and safety. Those concerning failure performance have already been described and those which depend on functional requirements are presented in the flow chart in Figure 12.

It is assumed that consideration of new technologies is the main reason for this study, and this is indicated by the choice of starting point in the flow chart.
There is a very significant difference between those systems which are based on current signalling principles and practice and those which are not. In this context, the signalling principles in question are the use of fixed block and techniques such as track circuits or axle counters to detect the presence of vehicles.

It is implicitly assumed that failure to safety and the locking of points cannot be abandoned for railway applications. Whether or not lineside signals are retained depends on the proposed alternatives (e.g. the Tokaido Line (J.N.R.) does not have lineside signals, instead it uses continuous cab signalling).

If these principles are retained, it is also important to decide whether functional enhancements are provided or not. Examples of such functional enhancements are systems for track to train communications and the possibilities they offer.

When there are no enhancements, then a technological replacement is being considered (the Henley interlocking in Reference 1), and, as will be shown in subsequent paragraphs, this will require both open and closed information loops and the implementation can be either as a centralised or decentralised system.

When functional enhancements are proposed, these can be realised either with reporting (other than by a human operator) of the system (equipment performance) or without it. The former is an open loop system (within a closed loop – see 4.3.1.), and examples of this approach are found in the automatic operation of the Victoria Line, the BART syste, the Kiruna Mine railway.
Systems with reporting form closed loop systems, using two way transmission of information and centralised processing. An example of this approach can be found in the ORE A46 system for Train Control (Ref. 17).

Such a system approach is not only suitable but essential when the use of variable length or moving block signalling is desired, as shown in the flow chart.

4.3.1. Open and closed loop systems

Systems exhibiting failure to safety performance can be based on either a closed or an open loop approach - independently of whether they are centralised or not. Centralisation is required for data processing, rather than for failure to safety.

This section examines the fundamental features of open loop and closed loop systems with regards to failure to safety and develops examples of both types of loops as they are found in the existing BR Signalling system.

Examination of Figures 13 and 14 will reveal that in both cases an executive fail safe system, able to monitor the final system output is required (for example able to apply the train brakes if there is a fault in the system).

The main differences between the two approaches exist in the following two areas:
a. On the existence of a **DIRECT REMOTE CONTROL** of the validation unit associated with Logic II in the case of a closed loop.

b. On the existence of a validation unit at the Logic I level, this being dependent on the degree of centralisation and the amount of data processing in Logic I. In typical decentralised situations, such as the LT Victoria Line, there is no data processing as such at the Logic I level, but only the selection of pre-encoded and validated data. This selection must also be fail-safe (Ref. 18).

Sections 3.4.2. and 4 will make reference to other important features related to system structures.

Figure 15, intended to be self-explanatory, presents a complete view of the processes between signalman and driver as they are at present.

Under normal operation – that is, provided no signals are passed at danger and no emergency replacement of signals occurs – the signalman cannot tell if the driver has seen, and correctly acted upon, a signal.

Another well-established example of an open loop system is illustrated in figure 16 – BR automatic signalling.

4.3.2. **Discrete and Continuous Information**

Another factor playing an important role in the structure of safety
systems is whether or not it is essential for information to be continuous both in space and time.

It is also necessary to differentiate between information relevant to safety and information relevant to regulation. This study is primarily concerned with the former.

In the case of traditional signalling circuits, it has always been common practice to rely on the continuous presence of information to indicate the correct operation of equipment (based on the assumption that the undetected welding of signalling relay contacts constitutes an acceptable risk).

This concept has been extended to electronic circuits, such as those used in the Henley interlocking (Ref. 1), where an A.C. signal, namely square waves were used to carry information. Sampled data control systems and their theory can, therefore, be used to show that the *continuous* presence of information is not strictly required, as Figure 17a indicates that in a suitably chosen square wave the signal may be absent for half the period, and the failure to safety relies on average "information power", rather than on instantaneous values.

In data transmission systems, the same square wave can be used to illustrate a system where information is transmitted serially during $T_{on}$ and read and interpreted during $T_{off}$. It is not strictly necessary for a carrier signal to be present all the time, provided the system can recognise that the absence of signal has

* Except in signalling systems based on the use of axle counters.
exceeded a predetermined threshold value. This situation is typical in "Store and Forward" data communications systems.

It can be argued that exactly the same thinking can be applied to the spacial distribution of information: A train moving along a discrete channel, i.e. a communication channel with gaps in transmission, will see it, in time, as a "square wave", as shown in Figure 17b, with a non-zero information energy contents.

The tolerable gap for a given system can be determined from conventional control theory, and more specifically the theory of sampled data systems. (This theory is sufficiently well established not to have to be included in this study).

In the case of train control systems, the most important factor to take into account is the emergency replacement of a signal. (In the existing system, in the event of emergency replacement, if the line is occupied between the sighting point of the first warning (YY) and the replaced signal, the signalman cannot set a conflicting move until a time delay, intended to give the train sufficient time to stop or pass the signal).
5. DESIGN OF SAFE DIGITAL SYSTEMS

5.1. GENERAL

The preceding sections have shown how two major requirements lead to the design of safe systems:

- The ability to detect the presence of faults in the system, at various subsystem/component levels.
- The ability to take appropriate action following detection.

It has already been shown that fail-safe circuits based on established principles meet both these objectives.

This section is concerned with the techniques available to achieve sufficiently high probabilities of fault detection in systems with digital data transmission and processing.

The first relevant publications in this field were made in 1948 and 1952:

- 1948 - Wiener "Cybernetics" (Ref. 4)
- 1948 - Shannon - A mathematical theory of communication (Bell Sys. Tech. Jour.) (Ref. 5)
- 1952 - von Neumann - Reliable Automata constructed from less reliable modules (Ref. 6).

and these led to the development of Information Theory.

The field of information theory is primarily concerned with techniques for the handling of information in the presence of
random perturbations (i.e. failures and interference), and this section will concentrate only on those points which are strictly relevant to the problems under discussion.

There are three distinctly separate aspects in the field of protecting against random perturbations: Data encoding, Data transmission and finally, Data reduction (processing).

5.2. DATA ENCODING

In a quite general sense, the word encoding implies some transformation of input information prior to its entry to a communications channel and/or processing system.

Sometimes the encoding is absolutely essential, such as the encoding of plain language and numerical data when the information handling system is of a binary nature. Often, the purpose of the encoding is to improve the efficiency of the handling of information in a predetermined manner.

When "noise" (in the form of failures or interference) is present, the principal aim of encoding is to combat the effects of this noise, the idea being that of avoiding the incorrect interpretation of the information.

An extreme example of this can be found when, in order to transmit two binary messages $m_1$ and $m_2$ over a noisy channel, a long string
of zeros is attached to \( m_1 \) and a similarly long string of ones is attached to \( m_2 \). Thus, even if a high percentage of zeros (or ones) is incorrectly received as ones (zeros), there is still room for a good message detection. This is obviously not a very efficient way of organising a code, and coding theories, matching codes to types of interference have been evolved in order to provide adequate coding efficiency, necessary to make good use of the available channels.

The effectiveness of codes can be (at least partially) measured by their minimum "Hamming Distance" - this can be defined by considering the complete "dictionary" of code words and identifying the two words in it with the smallest number of different letters (bits) between them. This number is the minimum Hamming Distance (MHD).

The information redundancy inherent in the code can be used to implement a measure of error correction, but the theory shows that this is achieved at the cost of a reduced probability of successful fault detection*. It is, therefore, recommended that, for safety systems, a special analysis of the safety/availability ratio be undertaken before deciding in favour of error correction.

For train control systems the encoding process includes two stages:

- encoding numerical information in binary patterns
- providing adequate protection against random perturbations (failures and/or interference) by choosing a suitable code.

* N.B. This is because the check bits cannot be used for both purposes and an effective reduction in the Hamming distance results from their use for error correction.
5.3. DATA TRANSMISSION

Once data has been encoded as in 4.2, it is usually necessary to transmit this data from source to destination. The factors that need to be considered when selecting a technique for data transmission are the following:

- The nature of the transmission medium: This can be of two distinct types depending on whether conductive or radiating paths are provided (or a combination of both).

In conduction, wires, transmission lines, coaxial cables, etc. are used to carry the message from one point to another. In this case there is an attenuation per unit distance. They are also characterised by being directly connected from one specific point to another point so that privacy in the messages can be ensured more easily.

In radiation, the signals are fed to a transmitting antenna which converts them into propagating electromagnetic waves. A receiving antenna is used to intercept a portion of the radiated energy and two types of attenuation exist:

One proportional to an inverse power of the distance (often inverse square) and

one proportional to an exponential function of distance (negligible in many cases).

In the case of radio waves, it is more difficult to guarantee privacy due to the often limited spectrum available which can lead to interference between messages sent over the same frequency band.
The requirements of the application, such as whether it is necessary (or desirable) to transmit a multiplicity of messages over the same channel. In this case it is necessary to perform functions on the messages to combine in such a manner that their decomposition into the original messages is possible at the destination. Two important methods to achieve this are time-division and frequency-division multiplexing.

The nature of the information, e.g. whether it is analogue or digital, serial or parallel, etc.

In practice the maximum amount of information that can be transmitted over a noisy channel is defined by Shannon's formula

\[ C = B \log_2 \left(1 + \frac{S}{N}\right) \]

in which \( C \) is the information capacity of the channel, \( B \) is the channel bandwidth, \( S/N \) is the "signal to noise" ratio, expressing the ratio of signal power to noise power - assuming white noise.

which clearly demonstrates the interdependence between the variables.

Modulation is the technique for transforming information into a form suitable for transmission. A large number of techniques exists and these use various means to make the most effective use of the available signal power within the channel, thus maximising the S/N ratio. Some modulation techniques make it possible to increase
the S/N ratio in the demodulator (this is used in high quality FM radio transmissions).

In the transmission of binary data, the performance of the system is perhaps expressed better by the bit error rate than by signal to noise ratios, since the former also takes into account other sources of corruption such as nonlinearities, etc.

The combination of the techniques outlined in Section 5.2 and the suitable design of the data transmission link can be used to achieve not only a very low bit error rate (values of $10^{-7}$ being common) but also to reduce the probability of an undetectably incorrect message (a wrong side failure) to an arbitrarily small value.

5.4. DATA PROCESSING

To distinguish it from the processes described under 5.2 and 5.3 data processing will be associated with computation.

Computation is associated with the acts of selection and necessary destruction of information - Computation is always associated with data reduction.

Computation is, therefore, not only found in applications of "digital computers", but also in systems such as railway signalling where data reduction takes place, that is, where all input information cannot necessarily be reconstructed from the observed output.
An example of this is found when considering a railway signal showing a Red aspect. This can be due to any of the following situations: Track ahead occupied
Route not set
Route set but signal approach released
Track circuit failure in the section ahead
Lamp failure in the section ahead
Signalman requiring to contact driver, etc.

It is however due to the increasing use of digital computers and current technological developments that there is wide interest in fault-tolerant computing, that is, in achieving in computing results similar to those achievable in data encoding and transmission, as described in 5.2. and 5.3.

While this is a vast although relatively new subject, it has been recognised that the original ideas established by Shannon and von Neumann already mentioned (Ref. 5) and (Ref. 6) are valid for computing systems.

It is thus possible to define (Ref. 7) a "Computation Capacity" equivalent to the Channel capacity discussed in 5.3, in this case representing the maximum amount of information that can be computed by a noisy system, even though the noise may manifest itself as the effects of component failure.
It can be shown that codes can be found such that messages may be processed by a given system with arbitrarily small probabilities of error. Such coding may take forms as simple as replication, where \( n \) identical systems are used and checked.

This is mentioned because current technological developments make it attractive to use redundancy in the form of replication in order to achieve a sufficiently high probability of error detection in computing.

While journals such as the IEEE transactions on computers are beginning to publish articles on the use of arithmetic codes which can be used to achieve "low noise computation", from an expediency point of view, replication has a lot to recommend it and only one significant drawback: The possibility of common mode design faults, i.e. design faults which will produce the same (incorrect) result in all units.

A discussion on the validation of safe digital systems is given below, in Section 5.5.

The successful application of replication techniques to safe digital systems will depend, to a significant degree, on the methods employed to check the system outputs.

The concepts described in the preceding sections can be applied to the analysis of various replicated structures which could be put to practical use.
Three such possible configurations are studied in detail in Appendix II, and these are shown to be part of one general "family model". The cases analysed are those of a duplicated system and a triplicated system. In the latter case, a distinction is made between those in which the presence (or successful detection) of the first fault is announced, and those in which it is not.

It can be shown from the mathematical analysis carried out in Appendix II that a triplicated structure with announcement of a first fault, and which can continue to operate as a duplicated system until repair is completed, is a suitable configuration for a safe system.

The analysis in Appendix II assumes that the fault detection circuits are fault-free. This is justified in practice if the detection circuits are the same as those used in practice to carry out the logic functions of the systems.*

These analyses were originally intended for a study of the performance of microprocessor systems in which checking of their performance is achieved by crossconnecting inputs and outputs in such a way that, effectively, each microprocessor checks the other ones and, therefore, the conclusions drawn in Appendix II apply to this case.

* N.B. I.e. if the elements of the comparator external to the logic can be shown "infinite" reliability, as discussed in Appendix II.
5.5. **BRIEF REVIEW OF RELEVANT TRENDS IN ELECTRONIC TECHNOLOGY AND CONCLUSIONS**

With regards to the use of electronics for the design of safe and reliable systems, perhaps the microprocessor stands above all other current developments as it provides a perfect example of the marriage of solid state electronic circuits and computer technology in a way that substantiates their key role in the field of electronics.

Being eminently suitable for processing and digital communications systems they contribute to the trend towards distributed computing, where each computing activity is broken up into segments and dedicated, interconnected processors work on each segment.

The reasons for this trend can be understood from the design problems associated with large computer systems. The first of these is a **control** problem since the larger the system, the more time and resources are spent deciding what should be done.

A second problem is a **software** problem, since as a general rule, programs for large systems do not use hardware capabilities efficiently, specially so in multiprocessor systems.

A more serious problem in large systems is the almost unavoidable presence of "programming bugs" which significantly affect performance.
As microprocessors and their associated large scale integrated circuits can be identified as a bridge between engineering design and software programming, their use may accelerate the efforts to apply to programs the same kind of discipline as is applied in large engineering design efforts, such as the hierarchical organisation of design solutions and the precise specification of subsystems. This approach is known under the general name of "structured programming" and it aims to produce programs that are so well organised that they can be easily understood, analysed and modified - and naturally documented - in other words, it is just the use of sound engineering practices for software - even though for the time being ultrareliable software remains an R & D activity.

In conclusion, it can be stated that there are considerable technological opportunities arising from electronic data communications and processing.

5.6. VALIDATION OF SAFE SYSTEMS

The practical introduction of any system responsible for safety functions presents the designer and user with severe problems, namely:

- To ensure that the system is "initially correct", that is, that its design and implementation are free of design errors.

- To verify that the failure performance of the system is comparable with that expected of it.
The latter point is merely a part of reliability engineering and is more than adequately discussed in the standard works on this subject.

The first subject is more complex and is, strictly speaking, a management rather than an engineering problem. For illustration purposes only two schematics, Figures 18 and 19, are given to indicate the number and nature of steps involved in initial checking.

Figure 18 applies to a conventional railway signalling system. In it, the "Signalling Plan" represents the functional specification for the particular location and the "Control Tables", the way in which the Signalling Plan is expressed as a logic function.

Figure 19 applies to a new field of work in railways, related to the transmission of information between track and trains, and intended to provide Automatic Train Operation and, eventually, the option of Crewless Train Operation (see Sections 6 and 7).

Figure 19 applies specifically to the programming of Read Only Memories to be installed at specific locations on the track, and containing geographic track information required for processing by electronic equipment on the train.
6. APPLICABILITY OF ELECTRONIC SAFE SYSTEMS

PERCEIVED OPPORTUNITIES IN RAIL TRANSPORT

The techniques described in the preceding sections can be used in a number of areas. The present section will describe two further important areas of application:

- Railway signalling
- Automatic train operation,

both in the context of British Rail practice and constraints.

The information presented here does not constitute a completed R & D program, but is regarded instead as a scientific formulation of the problem with the proposal for a design philosophy.

6.1. SOLID STATE TECHNIQUES IN RAILWAY SIGNALLING

Railway signalling has existed and evolved for as long as railways have existed. Its main safety requirements can be summarised as:

- Preventing the movement of route selection mechanisms (points) while a train is passing over them. Failure to do this may result in a derailment.

The movement of points is necessary to allow the routing of trains to their intended destinations.
Avoiding collisions at railway intersections (junctions and crossings) as well as collisions between trains on the same line (it is appropriate to remember that the braking distance of a high speed train is in the order of 2 km).

A further requirement exists in that the above two conditions shall be met even when there are equipment failures in the signalling system.

The three main functional elements in the signalling system are:

a. The interface to signals (in modern systems, colour lights) and points (driving an electric machine) as well as the interface from various information sources such as track circuit receivers used for train detection, filament proving circuits, point locking proving, etc.

b. The communication of this information to the logic circuits which operate on it. These communications can take the following forms:
   - Direct cabling, usually used in all local circuits and automatic signals (those which do not require centralised logic).
   - Failsafe multiplexed links.
   - Non-failsafe multiplexed links.
The use of failsafe multiplexing currently relies on the use of mechanical oscillators of frequencies chosen so as to differ from supply harmonics under virtually all conditions and offer the possibilities of very high centralisation of the logic and significant reductions in cabling.

Two such signalboxes exist in British Railways, Trent and Saltley in the London Midland Region.

Non-failsafe multiplexing has been used successfully for many years and safety is assured through remote failsafe logic, "remote or local interlocking".

c. The interlocking is a logic function and can be regarded as a crude form of computer, performing combinatorial and sequential logic, according to the simplified flowchart in Figure 20. These logic functions are performed by specially designed relays (Reference 8) designed to exhibit a very high degree of failure to safety. The flowchart is considered to be self-explanatory.

The overall physical organisation of a railway signalling system is shown in Figure 21 and in it all the elements previously described are indicated (see Reference 19).

The current cost of railway signalling approaches £40,000 per mile if all elements, such as cable trenches and buildings,
are included. A cost breakdown, however, indicates that the major components are:

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labour</td>
<td>50%</td>
</tr>
<tr>
<td>Cabling</td>
<td>25%</td>
</tr>
<tr>
<td>Equipment</td>
<td>25%</td>
</tr>
</tbody>
</table>

The use of electronics in a way similar to that described in Section 5 can make a significant contribution to cost savings, not only because of present costs, but also because the future cost of special components, such as railway relays, will continue to rise, whereas the cost of solid state circuits will continue to decrease as the market size for them grows.

A further benefit will accrue from the inherent reliability of solid state techniques and from the possibility of using equipment redundancy to achieve a higher operational availability - which will then be limited by the performance of interface equipment, such as point machines.

A possible realisation of an electronic railway signalling system is shown in Figure 22.

As shown in Figure 21, railway signalling is a distributed logic system and logic functions at each location have an influence on other locations. This implies both data communications and data processing, as previously discussed.
The elements shown in Figure 22 are intended to show how digital integrated circuit technology can be used to advantage. The box labelled "Data Acquisition System" represents the processes of amplification and filtering, modulation and demodulation and finally coding and decoding with a check of the data validity as discussed in Section 5.

The box labelled "Processing System" actually performs the desired logic functions and, through suitable interfaces, drives the appropriate signalling equipment and reports on the state of elements such as points, signal lamp filament current, track circuit occupation, etc.

The availability of Read Only Memories offers considerable advantages as it allows for the pre-encoding of standard messages (in a manner similar to that shown in Figure 19) and it is, therefore, envisaged that pre-encoded messages could be used to reduce the probability of incorrect data (from processing) being given a valid transmission code and incorrectly interpreted.

6.2. AUTOMATIC TRAIN OPERATION (ATO)

Despite the fact that ATO has been in use for several years by several urban transport authorities, such as in London Transport's Victoria Line, the San Francisco BART system, etc., no such system is yet in operation in main line railways anywhere in the world (see list of references).
The main reasons for this can be found in the following:

- Railways operate a wide range of train types, e.g. locomotive hauled, multiple units, with varying lengths, different traction and braking systems and different braking distances. This effectively rules out all the "simple" ATO systems relying on programmed operation which only work well when the rolling stock has uniform characteristics.

- Main line railways operate a complex service with different station stopping strategies (all stations, semifast, express, etc.) which further complicates the operation of an ATO system.

- Main line railways in the busier areas of the network are very sensitive to perturbations and sophisticated techniques for train regulation strategies are a necessary element in an ATO package.

It is only very recently that both computer technology and the understanding of the problem have come sufficiently close together to allow for satisfactory solutions.

These reasons indicate that the simplest form of ATO system which would be applicable to main line railways will contain all of the following elements:
- A track to train communications link capable of transmitting sufficient information for safety, regulation and preferably with a speech option.

- The safety element should be able to determine the maximum permissible speed for given track and train characteristics and should never produce information which is not applicable to a particular railway movement.

- The regulatory element should provide for the selection of station stops information and to execute the stopping with a specified accuracy. In addition, regulatory information in the form of desired point-to-point timings or optimum speed profile should be processed and communicated to the appropriate train.

Figure 23 indicates how these objectives could be met taking the present day railway as the starting point and the elements in it relate to the feedback loops discussed in Section 2. The socio-economic arguments for ATO are presented in Section 8.
7. CREWLESS TRAIN OPERATION

The previous sections have been primarily concerned with the safety aspects of railway controls. Several railway administrations have shown that a certain degree of automation could be achieved using the existing infrastructure (Reference 17).

Looking further to the future it can be anticipated that public demand for a cheap, reliable, fast and frequent transport service may require a much higher degree of automation due to the difficulties of meeting all the above objectives using the present approach.

It is well-known that the use of computers and related technologies has revolutionised many process industries, and today there are fast, accurate sensors, computing facilities adequate for simulation and control to the point that "optimisation" of the process is a realistic possibility.

There are nevertheless many activities which have been uninfluenced by these developments because of labour relations implications, the high cost of advanced automation systems and/or some systems which are so complex that proper models for them have not yet been produced. Another limitation is the quality of transducers performing (human) sensing functions.

This does not preclude the possibility that a "machine" will be programmed to replace a human operator in all aspects and it is
suggested that the best way to combine the capabilities of the
human and electronic technology is through an evolutionary approach,
introducing automatic control with a definite and achievable
return on investment at each stage.

The four basic stages in this evolution are as follows:

- **Complete human control:** The human operator is the sole controller
  of the process, including measurement and/or estimation of
  parameters and the appropriate use of the controls. Electronic
  systems may be present but in a purely informative capacity.
  This situation is found in today's train driving.

- **Machine monitoring a human operator:** As a partial
  substitute of the human operator, a control system releases the
  operator from tedious or repetitive tasks.

  When the special abilities of the operator are needed, the system
  could be used "in parallel" with the operator, in an advisory
  capacity. The operator is still the decision maker and also
  activates the controls, but there is cooperation between man and
  machine.

  This stage permits the progressive development of models in order
  to mimic the human decisions in those situations not yet
  automated.

  Such a stage in the implementation of an automation system is
acceptable and economic, because the operator can work as usual and the system is working without interfering. Experience in fields other than transport has shown that when automation systems provide an acceptable level of control, operators are willing to delegate many functions to the automatic system.

The system described in 6.2. could be introduced as a "Speed Advisory System" in order to validate its operation prior to the next stage.

- **Operator supervising machine:** This is the next step in the evolution towards full automation, in which the operator acts as transducer, feeding his readings, specially those variables which are difficult to quantify or measure (qualitative or subjective).

- **Full automation:** At this level the automation system measures, decides and actuates the control. A human operator is now responsible for action in the case of system malfunction.

Such an evolutionary implementation has several advantages. From an economic and social point of view, the automation system which is introduced is not immediately aimed at removing the operator, but at helping him improve productivity by selecting a better operating mode. The problem of human acceptance of the automation system is much lighter than in the case of a drastic technological change.
No excessive investment is made in any step without having proven the success of the preceding step, thus justifying the cost, even if the ultimate objectives of full automation are never achieved.

The step by step nature of the approach simplifies the practical implementation. Instead of designing a new system to replace the old one, the target is to move from the old system to the new one by a series of stages. The role of the operator changes little by little and each step begins only when both the automation system and the operator are ready.

7.1. FORMULATION OF THE PROBLEM

For the successful design of an automation system it is desirable to classify the different classes of tasks which the system must cater for prior to allocating the different tasks to a particular structure within the system.

Functional tasks can be classified into NORMAL OPERATION tasks and FAILURE MODE tasks:

7.1.1. Normal operation tasks

These can be simple, complex and vigilance tasks:

Simple tasks: Simple tasks are those implying a sequence of well defined operations and involving a minimum of decision making.
Complex tasks: Complex tasks imply a sequence of moderately well-defined operations, but involving a significant degree of decision making. Complex tasks are characterised by greater variability in the sequences to be performed. Complex tasks account for a large proportion of the operator's job in most industrial processes.

Vigilance tasks: The process operator monitors the operation and integrity of the process and of the control system and is the main component in dealing with malfunctions. As the degree of automation increases the proportion of vigilance tasks concerned with monitoring tends to grow.

On most industrial processes, some automatic equipment exists which scans process variables and equipment status for "out of limit" conditions and gives alarm signals. Response to an ambiguous alarm signals is part of the process operator's job, but the overall vigilance tasks involve rather more, as there is a wide range of information to which he is required to respond, even though this information is complex, with ill-defined out of limit conditions, weak and noisy, infrequent and unexpected.

The vigilance tasks involve the recognition of the development of undesirable process conditions before actual alarms occur, the detection of malfunctions in equipment and instruments, even though there is no explicit alarm. This is clearly a problem in pattern recognition and decision making.
7.1.2. Failure mode tasks

These can be Control, Emergency, Isolated events and Total system collapse.

Control tasks: These usually follow from the above vigilance function and consist of determining whether a control intervention is required and/or possible, to compensate for malfunction.

Emergency tasks: The process system is sometimes required to deal with an emergency, and although shutdown systems (fail safe as in Section 3), which protect against the principal hazards are provided, emergency situations do occur and the system or its operator must deal with them.

The demands made on the operator by an emergency vary considerably. In some cases it will be sufficient for him only to execute a well formulated and practiced drill; in others he may have to develop a quite new strategy and there may be very many different situations.

Isolated tasks: The possibility exists that unusual acts may occur which affect the performance of the system – such as sabotage, or vandalism.

Total system collapse: This may occur due to a fortuitous combination of events and a form of standby system is required, which is totally independent of the main control system.
7.2. Tasks

Crewless train operation differs from Automatic Train Operation as described in Section 6.2. (ATO) in that the latter retains a "train operator" on board train, who also has the option - and is in some situations required - to change to a "manual" driving mode.

In recent years social and employment trends in large conurbations have indicated that there may well be a case for crewless train operation not only on economic grounds, but mainly based on the need to maintain a service of sufficient quality.

This section will examine the engineering problems and identify both the structure of a suitable control system as well as the technological areas of opportunity for electronic engineering.

Based on the analysis presented in Section 7.1 the following can be identified:

**Simple tasks:** All the following, concerned with the operation of vehicle controls and safe operation are considered to be simple tasks:

- Determination of maximum permissible speed $V_{\text{max}}$
- Determination of optimum desired speed $V_{\text{opt}}$
- Proving that $V_{\text{opt}} \leq V_{\text{max}}$ before execution
Acceleration and braking controls within permitted constraints.
Accurate stopping at stations
Door controls (open/close, obstructed, etc.)
Train starting.

**Complex tasks:**
- Permissive operation (more than one vehicle/train within one block section)
- Continuous evaluation of braking (adaptive adjustment)
- Obstacle detection
- Definition of train parameters (length, braking characteristics, etc.)

**Vigilance tasks:** These are concerned with the observation and reporting possibly be exception of vehicle status affecting, amongst others, the following:

**Braking system:**
- Status (adaptive or conventional)
- All failures and loss of performance
- Variations in nominal applications and release times
- Wheel slip and slide - wheel flats, etc.

**Traction system:**
- Fuel status (and/or electric traction voltage and current)
- Field Current (in electric traction)
- All failures in the traction controls or motors
- Cooling system performance, etc.
Auxiliary systems: Lighting, heating, ventilation, door controls, load sensing valves, emergency speech communication link, quality of ride, fire, smoke, etc.

Control tasks: In the case of a fully automated passenger vehicle, the following can be recognised as control tasks:

Braking: Remote switching of the adaptive braking Remote application of brakes.

Traction: Remote switching of emergency power source (e.g. batteries) in case of main supply failure or equipment failure.

Ancillary systems: Remote control to prevent/allow detraining of passengers at stops other than permitted (scheduled) stations; allowing manual operation takeover in certain circumstances Taking appropriate action if vehicle is overloaded (or empty).

Emergency tasks: The following are considered to be emergency tasks: Dealing with rolling stock failure, including determining whether failed vehicle can be assisted (i.e. has it derailed or are the brakes fully applied and cannot be released)
Dealing with fires, smoke, running into obstructions on the track, having passed the station without stopping, passenger operating emergency brake, etc.

Isolated events: System response and operation in the case of attempted suicide, vandalism, sabotage, accident (e.g. at a level crossing), etc.

7.3. SYSTEM STRUCTURE

A system structure similar to that described in Section 2 and shown in Figure 24, can be shown to be adequate for a crewless train operation system evolved from the automatic train operation system described in Section 6.2.

The system structure required for crewless train operation is shown in Figure 24, and is clearly based on the structure already shown in Figure 23. The most significant differences between ATO and CTO are the need for:

- Specially designed rolling stock to allow for the monitoring of variables and reporting of off-limit conditions in a fail-safe manner where appropriate.

The most important of these parameters are shown in Figure 25 and purely as an illustration of the magnitude of the problem, Figure 26 presents a simplified flow-chart of the requirements for the automatic door controls in a crewless rail vehicle.
A central organisation to which all malfunction or off-limits information is reported and which is also responsible for the overall management of the crewless system.

Among the most important activities which will have to be carried out with much tighter controls than at present, are the following:

- Track maintenance planning and updating of Permanent Speed Restriction values in the train control system (affecting the safety of operations)

- Track equipment test schedules and maintenance programs. In addition, also the real time location, status and fuel situation reporting.

- Coaching stock: Equipment test schedules and maintenance programs. Real time control of location, load status, auxilliary equipment, updating of train characteristics for train control purposes.

The above will supply the necessary information for statistical analysis on passenger flows, reliability, rolling stock utilisation, demand patterns, etc.

It is anticipated that such a system could be made operational before the end of this century.
8. **SOCIO-ECONOMIC IMPLICATIONS**

All the recent trends in housing, work and leisure patterns, manufacture and consumption, etc., contribute to highlighting the role of mobility in a modern industrial society.

In recent years, transportation has become increasingly controversial due to traffic congestion, pollution, energy considerations and last but not least, the economics of transport. The cost of providing new roads, bus services and rail services is continually increasing, reaching the point where the quality of service suffers. This inevitably reflects on both the quality of life and the national economy.

While the object of this study is not concerned with making a case for rail transportation, the systems discussed for possible future use do represent substantial investment and will, in the case of automation, introduce new factors. It is, therefore, important to assess the implications of these techniques in order to, at least, identify major parameters and to recognise, at an early stage, whether or not there are major obstacles that could inhibit these developments.

The three basic areas for investigation are: The user, the operator and the organisation and the country as a whole.

8.1. **AUTOMATION AND THE USER**

Automation and mechanisation are finding more and more applications
where the general public makes use of automatic equipment. Examples of this are found in coin-operated vending machines, washing machines, and more recently electronic cash dispensers and readouts in shops and libraries.

The main difficulties which have to be recognised by the designer are:

- That the equipment will be abused or misused by the user, in the shape of bent or foreign coins, disregard of instructions, use of force, etc.

- That the equipment will be subject to malicious damage if it is totally unattended.

In the case of railways both these situations exist already and may well be aggravated. On the other hand, the point is being reached where public demand for cheap, fast, reliable and frequent transportation cannot be satisfied without the use of automation.

In the case of crewless train operation (CTO) there will be a number of ergonomic engineering problems requiring attention, associated with the lack of railway personnel on the train to provide information and control both under normal circumstances and under failure conditions. This situation is not entirely unlike that found in automatic lifts.

Further automation and mechanisation aids can be expected in the
supply of information concerning train services as well as the
issue and verification of travel tickets.

In the case of freight movements, the ability of the railway to
offer genuine door-to-door services is limited to private sidings,
and thus only to customers placed relatively close to railway lines.

Provided the railway can reduce transit times between random pairs
of terminals, and this is very much subject to a change in operating
practice for wagonload and parcel freight services, there is
considerable scope for increasing the railways' share of the freight
market.

8.2. AUTOMATION AND THE OPERATOR

Is it really possible to design man completely out of an
automated system?

From the point of view of equipment and hardware, there are no
major obstacles, but the practicality of this does not appear
feasible in the very short term.

In general terms, a human operator is better able to anticipate
trends and interpret relationships which may not be easily
defined in a computer's program. He can also deal with small
adjustments for unprogrammed changes in operating conditions.
For as long as the human operator remains part of the system - during which the term "mechanisation" may be less misleading than "automation" - the man/machine interface should be designed taking into account man's limitations and strength - showing all that is relevant without overloading the operator.

The process of mechanisation must be viewed in conjunction with the problems of staff recruitment, availability and training, since these will greatly affect the degree of mechanisation which is compatible with job satisfaction and the operator's ability to deal with unexpected events.

It is in this area that there is potential for major social or industrial disruption, but this could be avoided by redesigning jobs that may not be eliminated, in order to make them less routine, less tedious and, where applicable, less dangerous.

Seen in the context of railway services, the problem of automation or mechanisation and the operator, leads to the following options:

a. To reduce the number of staff in the organisation without visible change in output
b. To redistribute staff the other jobs (e.g. there is a shortage of maintenance staff at the time of carrying out this study).

c. To offer a new form of business leading to increased revenue with lower unit costs.

Option c can be clearly seen to be the best choice, while elements of b could also be beneficial, subject to the constraints of implementing these decisions in a real world where both politics and organised labour exist.

To assist decision making, it is also desirable to examine the differences between Efficiency and Productivity, where the former depends on how resources are used and the latter largely on the tools available to carry out the job.

Examples of areas where efficiency can be improved in any large organisation are:

- Time required to reach a decision (delay = increased cost)
- Existence of long-term investment and policy plans.
- Avoidance of a "saving regardless of cost" attitude
- Overheads.
Increases in productivity should ideally be accompanied by an increase in the share of the market if there is a social need to avoid redundancies and consequent contributions to unemployment.

Many recent developments in British Railways, such as the commissioning of the TOPS (Total Operations Processing) System are productivity aids, in this particular case to improve wagon-load freight traffic.

8.3. AUTOMATION AND THE RAILWAYS

The railway as an organisation have, in addition to the day to day operation of the railway network, the responsibility for planning and marketing services and investment.

Since virtually all the investment capital is supplied by the Government (this situation being almost universal) it is necessary to investigate how this investment will affect both the railways and the nation as a whole, so that investment priorities can be allocated.

The effects on capital and manpower resources which are a consequence of innovation are many and complex, and Figure 27 represents a simplified analysis of cause and effect.

In order to briefly illustrate the nature and scope of such analysis, the following factors have been identified as being benefits arising from the implantation of automatic and crewless train operation:
8.3.1. Benefits of Automatic Train Operation

A study carried out by the author for the British Railways Board indicates that implementation of Automatic Train Operation in all the important suburban railways under British Railways operation would result in minimum savings, indicated below:

- **Reduced staff training:** By reducing route learning requirements by at least 50%, applicable to all drivers of suburban trains. At 1975 costs this represents £1.5 million per annum.

- **Improved staff utilisation:** By reducing stress, hence improving working conditions and permitting a longer shift. At 1975 costs this represents £2.5 million per annum.

- **Accident prevention:** By reducing the probability of human error. Considering equipment damage costs only, a nominal figure of £1 million per annum is assumed.

- **Energy conservation:** Studies carried out by other Administrations (References 20 and 21) as well as by British Railways revealed that energy savings of 15 to 30% could be achieved by automatic train operation by optimising the rate of acceleration and taking full advantages of the train's kinetic energy.
Since approximately £45 million were spent in 1975 on energy for suburban rail services, the possible savings represent between £7.5 and £15 million at current prices.

If the possibility of energy costs rising in real terms is also considered, it can be shown that ATD can be cost-effective on energy conservation grounds alone.

8.3.2. Benefits from Crewless Train Operation

In addition to the benefits listed under 1 above, the following apply:

- Increased revenue as a result of improved quality of service (i.e. increased frequency of service, more seats/hour, etc.).

  Considering that the 1974 season ticket revenue was £75 million, an increased revenue of £5 million is assumed.

- Reduced operating costs: As a result of a reduction in suburban driving staff, reduction in overtime payments, etc. This is conservatively estimated at £10 million per annum.

In addition other benefits can be listed, but these are not "economic" in the sense that they will not appear in the "Profit and Loss" accounts of the Railways Board. Examples of these are:
Social benefits from diverting traffic from road to rail.

- Reduced need to carry out suburban road improvement schemes

- Time savings to passengers, etc.

- Etc.

There are techniques and established procedures to quantify these benefits (Reference 22), and these are almost always used in the justification of road improvement schemes. This will not be attempted here as the resulting figures are greatly in excess of the economic benefits listed above and as such are likely to add to the source of long arguments.
9. **CONCLUSIONS**

Railways are already major users of electronic equipment (Reference 9) and this situation is clearly irreversible. The problems of electronic equipment design and reliability have been considered for several years and are reasonably well understood.

To date, however, electronic equipment has not been extensively used for applications where failure to safety is required.

This thesis contains the mathematical basis for the numerical evaluation of safety systems operating in the "real world", i.e. subject to both failures and maintenance. The mathematical techniques extend established reliability methods and have been applied in practice within British Railways.

The mathematical techniques extend to the well known concepts of "failure to safety" to include data, by introducing the definitions of "error" and "failure".

Two further new concepts are introduced: The first one is that Failure to Safety may not be the best way to describe the performance of a safety system, and a "Probability of Accident" is postulated, to take into account other factors such as human error and the nature of standby systems where these are used.
The second concept is that the mathematical description of a safety system can be achieved through the use of Markov processes.

This Thesis also contains guidelines for the application of these techniques and recommendations concerning the design of safe digital systems.

Two appendices included in this Thesis illustrate the more practical aspects of the techniques described as applied to new structures (Appendix II) and an existing system (Appendix III).

Apart from the electronic aspects, control is a fundamental part of railway operations and Section 2 shows the railway as a set of independent but hierarchically organised control systems.

It should be clear to the reader that this representation is an oversimplification and all attempts at developing a universal, comprehensive model of the railway network undertaken to date have failed. The purpose of presenting this view is, however, to indicate major interrelationships and their merging points.

This is illustrated in Section 7, where crewless train operation is discussed as an example. Section 6 formulates how the techniques described in this Thesis could be applied to the next generation of railway signalling equipment. Appendix IV gives a cost benefit analysis.
Section 8 presents a short discussion of the implications of innovation on this scale for the user, the operator and the organisation as a whole and is the result of experience gained over several years of industrial experience in the field of innovation.
APPENDIX I

MATHEMATICAL TECHNIQUES FOR THE ANALYSIS OF SAFETY SYSTEMS

List of Symbols:

∪ : Union
∩ : Intersection
f, γ, φ : Function of
ε : Belonging to

λ_{ij} : Failure rate related to the transition from State i to State j

μ_{ji} : Repair rate related to the transition from State j to State i.
APPENDIX I

MATHEMATICAL TECHNIQUES FOR THE ANALYSIS OF SAFETY SYSTEMS

DEFINITIONS

RELIABILITY

a. Reliability is the ability of the system/unit in question to perform a required function under stated conditions for a given period of time.

b. Reliability is the probability of survival, i.e. the probability that the system/unit in question of age x will not fail at time t or before, or, inversely, the probability that the product will, without failure, perform a required function under given conditions for a certain period of time.

AVAILABILITY

The probability of finding a system in its normal operating state at a given moment of time. Availability is a function of both failure rates and the repair time.

FAILURE TO SAFETY

A system is said to exhibit failure to safety characteristics, or simply to be "fail-safe", if it reverts to a more restrictive state for any foreseeable fault condition or plausible combination of fault conditions.

SAFETY

A system is called "safe" if hazardous events on its operating
environment, imposed by failures are excluded with a sufficiently high probability, appropriate to its application.

Such a concept could be formulated as

\[ S = 1 - P_A \]

where

- \( S \) = System Safety
- \( P_A \) = Probability of Accident.
APPENDIX I

MATHEMATICAL TECHNIQUES FOR THE ANALYSIS OF SAFETY SYSTEMS.

I.1. GENERAL

The problem of technological innovation always presents the problems of evaluating system reliability and safety, comparing specific characteristics of alternative designs objectively, evaluating and estimating the contributions to system unreliability and unsafety and identifying weak or critical elements in the design.

A number of techniques can be applied to obtain answers to all the above, and are based on the theory of Probabilistic Reliability. It will be assumed here that the reader is familiar with the fundamentals of these techniques. References 10 and 11 are recommended for more detailed information.

In any form of probability analysis, a systematic approach is essential. A distinction can be made between qualitative and quantitative analysis.

The qualitative analysis implies the decomposition of a system into subsystems and components, followed by the determination of possible systems operating and failed states. One possible analysis technique, well-known to signal engineers, is now called "Failure Mode and Effect Analysis" (FMEA). In it, the possible failure mode of each component is used to analyse the failure mechanism of the system. (Ref. 12). A short example is also given in Appendix III.
This appendix will examine how such a model can be enhanced by the use of Markov processes, to provide a quantitative analysis and how the latter, in turn provides information for estimating the "probability of accident".

Simple examples of application are given in Appendices II and III.

1.2. BASIC CONCEPTS

It will be assumed that the correct operation of the system depends on the correct operation of its component parts, and the operating behaviour of a component or system can be described by examining all possible operating and failed states and by considering the probability that a component or system will be in a particular state at a certain time.

The set of all observable system states $S_i$, assumed to be finite in number ($i = 1, \ldots, n$) is defined as $Z$ (see Figure 28). This set can be partitioned into two basic subsets: $Q$ (correct operational states) and $F$ (failure states). It has already been shown in Section 3 that subset $F$ can be further subdivided into safe and dangerous subsets. This study will return to this subdivision in the following pages.

In the following analysis, it is intended to relate the set of observable system states with the elementary states of each component, $E$.*

* Note. Each "component" may be a subsystem, and the subdivision can thus be repeated as many times as necessary.
For this purpose the following statements and stages need to be followed:

Subsets $Q$ and $F$ are related by the following equations:

$$Q \cup F = Z$$
$$Q \cap F = 0 \text{ (disjoint)}$$

For each $k$ (see Figure 35) ($i = 1, \ldots, n$), the set of all observable substate states $X_{ik}^k$ ($i_k = 1, \ldots, b_k$) is defined as $E^k$, and this can also be partitioned into two (or more) subsets:

- $E^k_Q$, subset of all operable component* states
- $E^k_F$, subset of all inoperable component* states (for all failure modes)

for which the relations

$$E^k_Q \cup E^k_F = E^k$$
$$E^k_Q \cap E^k_F = 0 \text{ (disjoint)}$$

Each possible state in a component can be described as an "elementary state" $\xi$, and there is a functional relationship between each possible state and each observable state:

$$x_{ik}^k = \Psi(\xi_{ik}^k) \quad k = 1, \ldots, m$$

$$x_{ik}^k \in \xi^k \quad i_k = 1, \ldots, b_k \quad r_k = 1, \ldots, g_k$$

* Could be a subsystem.
If \( g_k > b_k \), then clearly several elementary states \( \varepsilon_{rk} \) are assigned to one and the same observable subsystem state \( \chi_{1k} \) (this is analogous to the loss of information due to computation defined in Section 4.3).

Similarly, a functional relation exists between observable component states and observable system states, that is:

\[
S_j = \mathcal{P}(\chi_{i1}, \ldots, \chi_{im})
\]

\( S_j \in \mathbb{Z} \)

and a system function \( \mathcal{P} \) may be defined to relate the observable system states to the elementary states \( \varepsilon_{rk} \)

\[
S_j = \mathcal{P}(\psi(\varepsilon_{r1}), \ldots, \psi(\varepsilon_{rm})) = \mathcal{P}(\varepsilon_{r1}, \ldots, \varepsilon_{rm})
\]

(by extension of the previous argument, several elementary states \( \varepsilon_{rk} \) may be assigned to one and the same observable system state \( S_j \)).

As a system state is defined by \( m \) elementary states,

\[
(\varepsilon_{r1}, \ldots, \varepsilon_{rk}, \ldots, \varepsilon_{rm}), \quad r_k \in (1, \ldots, g_k)
\]

the transition from one system state to another, i.e. from operational to failed, is determined by elementary state transitions. An elementary state transition will be assumed to be the consequence
of a certain event \( \gamma \), occurring at discrete but random times, 
\( t_{\gamma} \), \( \gamma = 1, 2, \ldots \).

These events can be:
- component failure
- external (and statistically independent) interference
- maintenance time
- repair, etc.,

which will be characterised by their own probability distributions.

In order to carry out an "FMAE" analysis and the appropriate reliability calculations, (where "reliability" refers to the theory, and include numerical descriptions of time-to-safe failure and time-to-dangerous failure), the following information is necessary:

The number \( q_k \) of elementary states \( \mathcal{E} \) for each basic component
The number \( b_k \) of elementary states \( \mathcal{X} \) for each subsystem
The probability of occurrence of each \( x_{1k}^k \)
The system structure, in this case, defining all the system states \( S_j \) and allocating them as \( Q_j \) or \( F_j \).

A short example is given in the first part of Appendix III.
I.3. **Steps in the Evaluation of Systems Performance**

(reliability and safety)

The systems of interest to this study can be classified into three classes:

a. non-repairable systems (at least during their mission)

b. partially repairable systems (those which allow repair of certain subsystems/components without interrupting system operation by means of redundant parts)

c. repairable systems (as above but covering the whole system).

The evaluation steps are the following:

1. Analysis of the system structure:
   
   a. without considering repair - Comparison of various types of redundancy
   
   b. considering the repair of certain - specified - components
   
   c. all components repairable.

2. Sensitivity analysis leading to the identification of critical subsystems/components or parameters.

   (In systems such as in 1.c above, the analysis must include the maintenance organisation and procedures).

3. Determination of all the following:

   a. Basic component, event and maintenance data

   b. Choice of mathematical model (e.g. probabilistic or deterministic, etc.)

   c. Evaluation (numerical) of the model
From all the range of models available for analysis, those based on the use of Markov processes are considered to be the most suitable.

"Markov processes" became firmly established in statistical techniques in the mid-50's, and are, essentially a technique for dealing with the "equation of motion" in systems where a large number of states can exist and in which the transition from one state to another is random. This randomness is expressed by the requirement that the probability of transition from state \( i \) to state \( j \) depends only on \( i \) and \( j \).

The main reason why Markov processes (which can be easily shown to encompass the whole of modern probability reliability theory) have become so well established is the ease with which they can be used (by manual or computer methods.)

All that is required is the formulation and solution of a set of differential equations with constant coefficients in the form:

\[
\frac{dP_i(t)}{dt} = [C]P_i(t)
\]

where \( P_i \) represents the probability of the system being in state \( i \) at the time \( t \) and where \( [C] \) is a state-transition coefficient matrix. The following three examples will be used to illustrate the procedure:
Example 1 - Two state system (either it works or it doesn't) without repair - Figure 36a.

\[
\frac{dp_1(t)}{dt} = -\lambda_{12} p_1(t) - [C] = [-\lambda_{12}]
\]

\[ p_1 = e^{-\lambda_{12}t} \quad \text{provided} \quad p_1(t=0) = 1 \]

\[ \lambda_{12} = \text{constant} \]

the familiar exponential reliability formula, in this case

\[ p_1(t) = R(t). \]

Example 2 - Two state system with failures and repair - Figure 36b

In this case:

\[
\begin{bmatrix}
\frac{dp_1(t)}{dt} \\
\frac{dp_2(t)}{dt}
\end{bmatrix}
= \begin{bmatrix}
-\lambda_{12} & \mu_{21} \\
\lambda_{12} & -\mu_{21}
\end{bmatrix}
\begin{bmatrix}
p_1 \\
p_2
\end{bmatrix}
\]

and assuming \( \lambda_{12}; \mu_{21} = \text{constant} \)

\[ p_1(0) = 1 \quad \text{(the system is initially working)} \]

\[ p_2(0) = 0 \]

Using the Laplace transform to operate algebraically:

\[ s \psi_1(s) - 1 = -\lambda \psi_1(s) + \mu \psi_2(s) \]

\[ s \psi_2(s) = \lambda \psi_1(s) - \mu \psi_2(s) \]

\[ s \psi_1(s) + s \psi_2(s) = 1 \]
\[
\Psi_1(s) = \frac{\mu_{21} + s}{s(s + \lambda_{12} + \mu_{21})}
\]
\[
\Psi_2(s) = \frac{\lambda_{12}}{s(s + \lambda_{12} + \mu_{21})}
\]

and by transforming back

\[
P_1(t) = \frac{1}{\lambda_{12} + \mu_{21}} \left[ \mu_{21} + \lambda_{12}e^{-(\lambda_{12} + \mu_{21})t} \right]
\]
\[
P_2(t) = \frac{\lambda_{12}}{\lambda_{12} + \mu_{21}} \left[ 1 - e^{-(\lambda_{12} + \mu_{21})t} \right]
\]

As previously discussed in Section 3, \(P_1\) is the system's instantaneous availability. The steady state availability is found for \(t \rightarrow \infty\):

\[
P_1(\infty) = A_S = \frac{\mu_{21}}{\lambda_{12} + \mu_{21}} = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}}
\]

**Example 3:** Three state system with repair (Figure 36c).

This example is directly applicable to the discussion on Fail-Safe systems as discussed in Section 3.

In the interest of clarity, for this illustration it will be assumed that \(\mu_{31} = \mu_{21} = \mu\).

In this case:
\[
\begin{bmatrix}
\frac{dp_1(t)}{dt} \\
\frac{dp_2(t)}{dt} \\
\frac{dp_3(t)}{dt}
\end{bmatrix} =
\begin{bmatrix}
-(\lambda_{12} + \lambda_{13}) & \mu & \mu \\
\lambda_{12} & -\mu & 0 \\
\lambda_{13} & 0 & -\mu
\end{bmatrix}
\begin{bmatrix}
p_1 \\
p_2 \\
p_3
\end{bmatrix}
\]

\[p_1(t) + p_2(t) + p_3(t) = 1\]

which gives

\[p_1(t) = \frac{1}{\lambda_{12} + \lambda_{13} + \mu} \left[ \mu + (\lambda_{12} + \lambda_{13})e^{-(\lambda_{12} + \lambda_{13} + \mu)t} \right]\]

\[p_2(t) = \frac{\lambda_{12}}{\lambda_{12} + \lambda_{13} + \mu} \left[ 1 - e^{-(\lambda_{12} + \lambda_{13} + \mu)t} \right]\]

\[p_3(t) = \frac{\lambda_{13}}{\lambda_{12} + \lambda_{13} + \mu} \left[ 1 - e^{-(\lambda_{12} + \lambda_{13} + \mu)t} \right]\]

One attractive feature of this mathematical model is that it permits the direct evaluation of the probability of a Dangerous Failure State, which can in turn be used to calculate the MTDF (Mean Time to Dangerous Failure). The procedure in this case consists in removing the return path from state 3 back to state 1. The probability \(p_2(t)\) of being in this state at time \(t\) is therefore the probability of arriving there between 0 and \(t\).

It should be clear that this technique is applicable to systems with any degree of complexity, but beyond three states, it is desirable to resort to computer solutions.
I. 4. FORMULATION OF THE PROBABILITY OF ACCIDENT

As described in Section 3.4, a number of factors can lead to accident. The following formulation can be used if the probabilities $P_1$, $P_2$ and $P_3$ are calculated as above:

$$P_{\text{ACC}} = P_E + P_1 \cdot P_{E_{1/1}} + P_2 \cdot P_{E_{2/2}} + P_3 \cdot P_{E_{3/3}}$$

$$- P_E \left\{ P_1 P_{E_{1/1}} + P_2 P_{E_{2/2}} + P_3 P_{E_{3/3}} \right\}$$

where $P_E$ is the probability of an event against which protection is not provided.

$P_{E_{1/1}}$ is the conditional probability of an event leading to accident while the system operates correctly and against which some protection is provided.

$P_{E_{2/2}}$ is the conditional probability of an event leading to accident while the system has failed to safety and includes the performance of the standby system.

$P_{E_{3/3}}$ is the conditional probability of an event leading to accident while the system has failed to danger.

These conditional probabilities can, in principle be calculated, but in practice they can be determined from data from past accident records, analysis by primary and secondary causes, etc., using well-established statistical techniques.

Appendix III shows how $P_{E_{1/1}}$ has been calculated for the BR Automatic Warning System.
APPENDIX II

CALCULATIONS OF AVAILABILITY, RELIABILITY AND SAFETY FOR
SIMPLE SYSTEM STRUCTURES

The methods described in Section 3 and Appendix I are used here to derive expressions for the Availability, Reliability and Safety of the following system configurations.

- n/n replicated system with "unanimity" voting
- 2/3 majority voting redundancy without partial fault notification
- 2/3 majority voting redundancy with partial fault notification.

The resultant expressions will be shown to lead to a method for deciding which configuration will yield the best compromise between performance and cost.

DEFINITIONS OF TERMS USED

Each unit is assumed to have an output which can take two basic types of failure state (error and failure), which will be represented here as \( F_S \) (safe) and \( F_D \) (dangerous). A fault is defined as any event which causes the unit to assume a failed state permanently, regardless of input instructions.

"Degraded performance" is the state of the system in which not all the units are fault-free but a correct output from the system as a whole is assured. In a m/n majority voting system,
the maximum number of faults which will allow the state of
degraded performance is \((n-m)\).

"Partial fault notification" is received when the comparison
system is able to detect and announce a state of degraded
performance.

A "system error" occurs when sufficient units have developed
faults for prediction of the accuracy of the system output to
become impossible although the set of unit outputs would
be acceptable to the comparator system. A "system failure"
ocurs when an unacceptable set of unit outputs reaches the
comparator (e.g. different unit outputs in an \(n/n\) system).

Note: In these calculations it has been assumed that the
comparator system is infallible. If this cannot be assumed,
correction may be made to the probabilities of arrival in
each state by multiplying by the appropriate coefficients.
This assumption is permissible, because the approach that has
been taken at the British Railways R & D Division is such that
the comparator and the units are indistinguishable. This is
achieved by connecting all the logic units in a ring circuit,
so that each unit checks others and is, in turn, checked out
by other identical units. The only external components used
in these circuits are several orders of magnitude more reliable
than the logic units. This approach is now undergoing patent
protection procedures and cannot be disclosed further at this
stage.
PRELIMINARY CONSIDERATIONS

The choice of model for unit failure modes is affected by the proposed operational strategy. For these calculations it has been assumed that all units will be checked after time $T$ ($T<M.T.B.F.$), regardless of whether they have failed or not. (In this context "checked" means checked on site thoroughly enough for there to be no reason to doubt that the unit is in its original condition. It will then be indistinguishable from a replacement unit. This approach can only be used with units having no wear-out mechanism and it shows one of the associated benefits of solid-state circuit technology). It has also been assumed that whenever a fault occurs in the system, all units are checked when the faulty unit is replaced.

In the case of a 2/3 majority system a further question arises. "Is notification received if one of the three units becomes "faulty"?". Both answers to this question have been considered.

From the first paragraph, it may be seen that the cycle ends either after time $T$ or when a repair takes place.

When no comparison with a correct unit is possible, a fault in a single unit should be treated as an error unless there is information redundancy. Repair conditions are considered in terms of the status of the whole system rather than that of single units.
The basic unit model has a probability of developing a fault by time $T$ given by:

$$P(F',T) = R_u(T)$$

where $R_u(T)$, the unit reliability, is assumed to be:

$$R_u(T) = e^{-(\lambda_{1S} + \lambda_{1D})T}$$

- Constant failure rate for safe failure state.
- Constant failure rate for dangerous failure state.

The probabilities of safe and dangerous faults by time $T$ are given by:

$$P(S,T) = \frac{\lambda_{1S}}{\lambda_{1S} + \lambda_{1D}} (1 - R_u(T))$$

$$P(D,T) = \frac{\lambda_{1D}}{\lambda_{1S} + \lambda_{1D}} (1 - R_u(T))$$

This gives a basic statistical building block that will be used to analyse the following systems:

- $n/n$ replicated system with comparison of outputs
- 2/3 majority voting system without partial fault notification
- 2/3 majority voting system with partial fault notification

II.1. CALCULATIONS FOR N/N

The system model is shown in Figure 30.

The reliability of an $n/n$ system is the probability that no units break down in time $T$. 
For a system with a finite cycle duration Availability is defined:

\[ A_s(T) = 1 - \sum \frac{P(F_x, T)}{P_{x1}} \]  

(1)

where \( F_x \) is a particular type of fault and \( P_{x1} \) is the repair probability when that fault occurs, provided \( P(F_x, T) \) is small.

This definition is well established in reliability theory (see Reference 10).

For all fault conditions, except simultaneous similar faults in all units, the system will arrive in state 3 (system failure).

\[ \therefore \ P_s(3, T) = 1 - \left[ R_u(T) \right]^n - P_s(2, T) \]  

(2)

Where \( P_s(2, T) \) is given by:

\[
P_s(2, T) = \frac{\lambda_1S}{\lambda_1S + \lambda_1D} \left( 1 - \left[ R_u(T) \right]^n \right) \cdot \left[ \frac{\lambda_1S}{\lambda_1S + \lambda_1D} \left( 1 - R_u(T) \right) \right]^{n-1}
\]

\[ + \frac{\lambda_4S}{\lambda_1S + \lambda_1D} \left( 1 - \left[ R_u(T) \right]^n \right) \cdot \left[ \frac{\lambda_1S}{\lambda_1S + \lambda_1D} \left( 1 - R_u(T) \right) \right]^{n-1} \]

\[ \therefore \ P_s(2, T) = \frac{\lambda_1S^n + \lambda_1D^n}{(\lambda_1S + \lambda_1D)^n} \left( 1 - R_u(nT) \right) \left( 1 - R_u(T) \right)^{n-1} \]  

(3)
Where \( T \) is the time for the comparator system to operate. \( T \) may however be altered to show the effect of the relative frequency of simultaneous faults. For example, if \( T = 0 \), this implies that simultaneous faults are impossible and the last equation becomes: 
\[
P_s(2,T) = 0,
\]
as might be expected.

On the other hand, the situation where only simultaneous faults can occur is modelled by setting \( T = \infty \). This gives
\[
P_s(2,T) = \frac{\lambda_{1S}^n + \lambda_{1D}^n}{\lambda_{1S} + \lambda_{1D}} \cdot [1 - R_u(nT)] \tag{4}
\]
which is the correct expression for this situation. All intermediate probabilities of simultaneous faults can, therefore, be established by defining a positive value for \( T \).

The safety of the system may be defined as:
\[
S_s(T) = 1 - P_s(2,T)
\]
\[
\therefore S_s(T) = 1 - \frac{\lambda_{1S}^n + \lambda_{1D}^n}{(\lambda_{1S} + \lambda_{1D})^n} (1 - R_u(nT)) (1 - R_u(T))^{n-1} \tag{5}
\]

Rewriting (2) gives:
\[
P_s(3,T) = 1 - R_s(T) - (1 - S_s(T))
\]
\[
\therefore P_s(3,T) = S_s(T) - R_s(T)
\]
We can now write an expression for availability:

\[ A_s(T) = 1 - \frac{p_s(3,T)}{\mu_{31} T} - \frac{p_s(2,T)}{\mu_{21} T} \]

\[ \therefore A_s(T) = 1 - \left[ \frac{S_s(T) - R_s(T)}{\mu_{31} T} \right] - \left[ \frac{1 - S_s(T)}{\mu_{21} T} \right] \]

\[ \therefore A_s(T) = 1 - \left[ \frac{1 - R_s(T)}{\mu_{31} T} \right] - \left[ 1 - S_s(T) \right] \left( \frac{1}{\mu_{21} T} - \frac{1}{\mu_{31} T} \right) \]

where \( \mu_{21} \) and \( \mu_{31} \) are defined as in Appendix I.

If \( R_s(T) = 1, S_s(T) = 1 \), then

\[ A_s(T) = 1, \text{ which is obviously true.} \]

\[ R_s(T) = S_s(T) = 0 \text{ cannot be tried, as } \]

\( a) \ T < \text{MTBF} \)

\( b) \ P(F_x, T) \ll 1 \)

II.2. CALCULATIONS FOR 2/3 SYSTEM WITH NO NOTIFICATION OF FIRST FAULT

From the definitions of failure and error states, it may be seen that the first fault which occurs will put the system into a state of degraded performance. The second fault will cause a system error and the third fault will leave it in this state. The system model is shown in Figure 31.

\( \mu_{11} \) only exists if only one fault occurs by \( T \) and it is then infinite. Since it does not matter what type of fault occurs, the binomial expansion may be used.
\[
\left\{ R_u(T) + \left[ 1 - R_u(T) \right] \right\}^3 \\
= \left[ R_u(T) \right]^3 + 3 \left[ R_u(T) \right]^2 \left[ 1 - R_u(T) \right] + 3 R_u(T) \left[ 1 - R_u(T) \right]^2 + \\
\left[ 1 - R_u(T) \right]^3
\]
\[(7)\]

(The terms from left to right being associated with 0, 1, 2, 3 faults respectively).

In order to make this calculation compatible with that for the \( n/n \) system, the concept of \( T \) will be reintroduced, even though it will usually be possible to assume \( T = 0 \). This modification may be achieved by assuming that any fault but the first can occur in time \( (T + T) \). The first 3 terms may be modified easily (bearing in mind that if a fault occurs the reliability must then also be maintained for time \( (T + T) \)) and they become:

\[
\left[ R_u(T) \right]^3 + 3 \left[ R_u(T + T) \right]^2 \left[ 1 - R_u(T) \right] + 3 R_u(T + T) \left[ 1 - R_u(T) \right] \left[ 1 - R_u(T + T) \right]
\]
\[(8)\]

Since the full expansion must continue to add up to unity, the fourth term can be shown to become:

\[
1 - \left[ R_u(T) \right]^3 - 3 R_u(T + T) \left[ 1 - R_u(T) \right]
\]
\[(9)\]

Now \( P(2, T) \) is the sum of the third and fourth terms.

\[
\therefore \quad P(2, T) = 1 - 3 \left[ R_u(t + T) \right]^2 \left[ 1 - R_u(T) \right] - \left[ R_u(T) \right]^3
\]
\[(10)\]

The safety of the system \( S_g(T) = 1 - P(2, T) \)
Under these fault definitions there are two forms of reliability and these are defined here:

a) \( R_{s1}(T) = \) Probability of no fault occurring at all in \( T \)

b) \( R_{s1'}(T) = \) Probability of system still being operational at time \( T \).

In this system: \( R_{s1'} = S_s(T) \)

This does not mean that the definition is trivial.

In more complicated systems this direct relationship no longer holds.

The equation for \( R_{s1'}(T) \) is as follows:

\[
R_{s1}(T) = R_u(3T)
\]  

(12)

We may define Availability from above:

\[
A_s(T) = 1 - \frac{P(2,T)}{M_{21}^{21}}
\]

where \( P(2,T) \) may be obtained from the equation given above in (10).

\[
\therefore A_s(T) = 1 - \frac{1 - S_s(T)}{M_{21}^{21}}
\]  

(13)

The boundary conditions of \( T = 0; R_u(T) = 0 \), \( 1 \) may be seen to be correct. An interesting point is that when \( T = \infty \), \( P(1',T) = 0 \)
because the system passes straight from state 1 to state 2.

II.3. CALCULATIONS FOR 2/3 SYSTEM WITH NOTIFICATION OF FIRST FAULT

The first question is: "What action is taken when notification of the fault is received"?

Three of the possible answers are as follows:

(i) Take no action apart from replacing the faulty units as soon as possible.
(ii) Allow the system to reformulate the remaining two non-faulty units into a 2/2 system.
(iii) Adopt the second course and replace the faulty unit as soon as possible, resetting the system to 2/3 when this has been achieved.

Each of these answers is dealt with in a subsequent section.

II.3.1. Answer (i)

The model for this system is the same as the previous one shown in Figure 31.

To obtain state 1, one fault in time $T$ is required.

To obtain state 2 at least one fault in time $T$ and at least one fault in time before repair $\frac{1}{\mu_{111}}$ are required.

$\mu_{111}$ will be referred to in this type of system as $\mu_{31}$ since it is the probability of repair from a state in which a fault has been detected.

State 2 can be reached by at least one fault in time $T$ and at least one fault in time $\left(\frac{1}{\mu_{31}} + \tau\right)$. 
The assumption implicit in this statement is that the repair time is a constant \( \frac{1}{\mu_{31}} \) and not a probability function.

\[
\therefore R(2, T) = \left[ 1 - R_u(3T) \right] \times \left[ 1 - R_u \left\{ 2 \left( \frac{1}{\mu_{31}} + \tau \right) \right\} \right]
\]

(14)

\[
\therefore S_g(T) = 1 - \left[ 1 - R_u(3T) \right] \times \left[ 1 - R_u \left\{ 2 \left( \frac{1}{\mu_{31}} + \tau \right) \right\} \right]
\]

(15)

\[
R_{s1}(T) = 1 - R_u(3T)
\]

(16)

\[
R_{s1}'(T) = S_g(T)
\]

(17)

\[
A_g(T) = 1 - \frac{1 - S_g(T)}{\mu_{21} T}
\]

(18)

The last three equations follow directly from definitions and it will be noticed that they are of the same form as equations already derived. The difference lies in the value of \( S_g(T) \).

If \( \frac{1}{\mu_{31}} = T \) and \( T \ll \frac{1}{\lambda} \), the value for \( S_g(T) \) becomes the same for both cases. This may be seen to be correct.

(By definition \( \frac{1}{\mu_{31}} \ll T \).)

II.3.2. Answer (ii)

The model for this situation is shown in Figure 32.
In this case \( \mu_{1'}, \mu_{31} \neq \mu_{31} \) since no repair procedure is initiated by arrival in state 1'.

It is apparent that:

\[
R_{s1}(T) = R_u(nT)
\]  

(19)

The ways in which the system can arrive in states 2 and 3 will be considered next. Clearly state 2 can be achieved by three simultaneous similar faults. It can also be achieved by one fault followed by two simultaneous faults which are similar to each other. It is not quite so clear that state 2 can be achieved just by two simultaneous similar faults. This can be shown by the following argument:

In order to reformulate into a 2/2 configuration, the system must be able to decide which unit has a fault. If two similar simultaneous faults occur, it will decide wrongly and use the two faulty units for its 2/2 logic. This is clearly a system error.

It can be concluded that any two or more simultaneous similar faults will cause a system error.

\[
P(2,T) = P(\text{two s.s. faults, } T) + P(\text{three s.s. faults, } T)
\]

where s.s. = simultaneous similar.
Now:
\[
P(3 \text{ s.s. faults, } T) = \left[ 1 - R_u(3T) \right] \left[ 1 - R_u(T) \right] \frac{\lambda_{1S}^3 + \lambda_{1D}^3}{(\lambda_{1S} + \lambda_{1D})^3}
\] (20)

which is, of course, \(P(2, T)\) for a 3/3 system.

and \(P(2 \text{ s.s. faults, } T) = 3 \left[ 1 - R_u(3T) \right] \left[ 1 - R_u(T) \right] R_u(T) \frac{\lambda_{1S}^2 + \lambda_{1D}^2}{(\lambda_{1S} + \lambda_{1D})^2}\) (21)

If \(T\) is small, then:

\[
P(3 \text{ s.s. faults, } T) \ll P(2 \text{ s.s. faults, } T)
\]

and \(R_u(T) = 1\)

\[
\therefore \quad P(2, T) = 3 \left[ 1 - R_u(3T) \right] \left[ 1 - R_u(T) \right] \frac{\lambda_{1S}^2 + \lambda_{1D}^2}{(\lambda_{1S} + \lambda_{1D})^2}
\] (22)

\[
\therefore \quad S_s(T) = 1 - 3 \left[ 1 - R_u(3T) \right] \left[ 1 - R_u(T) \right] \frac{\lambda_{1S}^2 + \lambda_{1D}^2}{(\lambda_{1S} + \lambda_{1D})^2}
\] (23)

Turning to the Modified Binomial shown in Section 2.2., it may be seen that:

\[
R_{s11}(T) = \left[ \text{First and Second Terms of M.B.} \right]
\]

and \(P(3, T) = \left[ \text{Third and Fourth Terms of M.B.} \right] - P(2, T)\)

\[
\therefore \quad R_{s11}(T) = \left[R_u(T)\right]^3 + 3 \left[R_u(T + \tau)\right]^2 \left[1 - R_u(T)\right]
\] (24)
And thus:

\[ P(3,T) = 1 - R_{s11}(T) - \left[ 1 - S_s(T) \right] \]

\[ \therefore P(3,T) = S_s(T) - R_{s11}(T) \] (25)

The expression for Availability is:

\[ A_s(T) = 1 - \frac{P(3,T)}{\mu_{31}T} - \frac{P(2,T)}{\mu_{21}T} \]

\[ = 1 - \frac{S_s(T) - R_{s11}(T)}{\mu_{31}T} - \frac{1 - S_s(T)}{\mu_{21}T} \]

\[ \therefore A_s(T) = 1 - \left[ 1 - \frac{R_{s11}(T)}{\mu_{31}T} \right] - \left[ 1 - S_s(T) \right] \left( \frac{1}{\mu_{21}T} - \frac{1}{\mu_{31}T} \right) \] (26)

Now it can be seen that all equations for Availability look very similar. Remarks on this fact will be found in Section 3.

II.3.3. Answer (iii)

The model for this system is the same as that shown in Figure 32.

For this system:

\[ R_{s1}(T) = R_u(3T) \] (27)

State 3 may be reached by:

a) two single faults, one in Time T, the second in \( \frac{1}{\mu_{31}} \)
b) two simultaneous dissimilar faults.

\[ p(3, T) = \left[ 1 - R_u(3T) \right] \left[ 1 - R_u(\tau) \right]^2 \cdot \left[ 1 - R_u\left( \frac{3}{\mu_{31}} \right) \right] R_u(\tau) + \]

\[ \begin{align*}
&+ \frac{3 \times 2 \lambda_{1S} \lambda_{1D}}{(\lambda_{1S} + \lambda_{1D})^2} \left[ 1 - R_u(3T) \right] \left[ 1 - R_u(\tau) \right] \\
&+ \frac{6 \lambda_{1S} \lambda_{1D}}{(\lambda_{1S} + \lambda_{1D})^2} \left( 1 - R_u(\tau) \right)
\end{align*} \]

(28)

State 2 may be reached as in the previous system except that if a single fault occurs, the simultaneous similar faults have to occur within time \( \frac{1}{\mu_{31}} \) (i.e. "before the system is repaired", instead of "before the end of period \( T \)", which was the case in the previous system) Referring again to the graph, the assumption is made that a single fault will occur, on average, after time \( T/2 \).

Therefore, simultaneous similar faults occurring in time \( \left( \frac{T}{2} - \frac{1}{\mu_{31}} \right) \) when a single fault has already occurred do not cause a system error.

From the equation giving \( p(2, T) \):

\[ p(2, T) = 3 \left[ 1 - R_u(3T) \right] \left[ 1 - R_u(\tau) \right] \frac{\lambda_{1S}^2 + \lambda_{1D}^2}{(\lambda_{1S} + \lambda_{1D})^2} - \]

\[ - \left[ 1 - R_u(3T) \right] \cdot 3 \left[ 1 - R_u\left( \frac{3T}{2} - \frac{3}{\mu_{31}} \right) \right] \left[ 1 - R_u(\tau) \right] \frac{\lambda_{1S}^2 + \lambda_{1D}^2}{(\lambda_{1S} + \lambda_{1D})^2} \]

(29)
\[ P(2, T) = \frac{3(\lambda_1 + \lambda_2)}{(\lambda_1 + \lambda_2)^2} \left[ 1 - R_u(3T) \right] \left[ 1 - R_u(2T) \right] \]

\[ A_s(T) = 1 - \frac{1 - R_{s11}(T)}{\mu_{31}T} - \left[ 1 - S_s(T) \right] \left( \frac{1}{\mu_{21}T} - \frac{1}{\mu_{31}T} \right) \]

\[ R_{s11}(T) = 1 - P(2, T) - P(3, T) \]

II.4. THE GENERALITY OF THE AVAILABILITY EQUATION

The fullest version of the Availability Equation used here is (32)

\[ A_s(T) = 1 - \frac{1 - R_{s11}(T)}{\mu_{31}T} - \left[ 1 - S_s(T) \right] \left( \frac{1}{\mu_{21}T} - \frac{1}{\mu_{31}T} \right) \]
All other Availability Equations used are simplified versions of this

<table>
<thead>
<tr>
<th>Equation</th>
<th>Simplification</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6)</td>
<td>$R_s(T) = r_{s1}(T)$</td>
<td>No state 1'</td>
</tr>
<tr>
<td>(13)</td>
<td>$r_{s1}(T) = s_s(T)$</td>
<td>No state 3</td>
</tr>
<tr>
<td>(18)</td>
<td>&quot;</td>
<td>&quot;</td>
</tr>
<tr>
<td>(26)</td>
<td>No simplification</td>
<td>possible</td>
</tr>
<tr>
<td>(32)</td>
<td>&quot;</td>
<td>&quot;</td>
</tr>
</tbody>
</table>

Note that the term $(\frac{1}{\mu_{21}} - \frac{1}{\mu_{31}})$ is unlikely to vanish because the time $\frac{1}{\mu_{21}}$ includes time for detection of an error. An error is an "undetectable" fault and so it is almost certain that:

$$\frac{1}{\mu_{21}} > \frac{1}{\mu_{31}}$$

Therefore the term cannot be assumed to vanish.

II.5. POSSIBLE OPTIMIZATION TECHNIQUE

All the equations given above are in terms of six variables:

$$\lambda_{12}, \lambda_{13}, \tau, \mu_{21}, \mu_{31}, T.$$  

These can be divided into two sets of three.
The first three, \( \lambda_{12}, \lambda_{13}, \) and \( \zeta \), may be called the "design variables" and the second three, \( \mu_{21}, \mu_{31}, T \), the "strategic variables".

Supposing the basic problem is as follows:
"Given 1 or more circuits, to find the best system configuration and strategy in use for each and thus determine which is the best".

Now if the strategic variables are set up on mutually orthogonal axes, each point in this space will have associated with it:

a) a cost for the strategy

b) for each circuit design and system configuration values of \( A_s(T), R_s(T) \) and \( S_s(T) \) which will have associated costs.

In addition the cost of the design will be represented as a constant potential in this space.

Therefore, for each circuit design and system configuration, there is a cost potential in the strategic space. It is possible therefore, to pick the minimum cost point which has satisfactory values of \( A_s(T), R_s(T) \) and \( S_s(T) \) for each design and configuration and thus pick the best.

N.B. If the cost association figures for \( A_s(T), R_s(T) \) and \( S_s(T) \) are realistic, then the values will automatically be satisfactory.
Simplifications will probably be available. For example, it may be felt that $\mu_{21}$ and $\mu_{31}$ are virtually fixed in which case the graphs merely become cost vs. time between replacements.

Since values of $\mu_{21}$ and $\mu_{31}$ will not be constant throughout the country, it may prove that one design or configuration is better in one place than another. Such facts must be borne in mind when using this process.

II.6. CONCLUSIONS

It is not possible to make any sweeping conclusions from these equations without putting numbers in. Some limited conclusions may be drawn, however.

(i) It is apparent that the Availability of the system is considerably enhanced by the existence of state $1''$, the state of degraded performance.

(ii) Availability, Reliability and Safety are increased if some useful action is initiated by the arrival in state $1''$.

(iii) Safety is considerably increased if this action is to reformulate the remaining units into a 2/2 system, thus opening the path to state 3 (inaccessible to a 2/3 system). This arrangement could be referred to as a 2/2 system with an on-line spare which is used first.

(iv) When the above arrangement is adopted, the advantage obtained by initiating a repair procedure in addition to reformulating the system is strictly second order. Whether the advantage repays the extra stress on maintenance could
be deduced from the optimisation technique in Section 4 for particular designs.

(v) For systems in which simultaneous similar faults cause system errors, the probability of a system error will be minimized for constant \((\lambda_{1S} + \lambda_{1D})\) when \(\lambda_{1S} = \lambda_{1D}\)

This means that "preferred" fault-types should be avoided if both high and low faults are legitimate outputs, i.e. electronic circuits are preferrable to, say, "fail-safe" relays.
APPENDIX III

EXAMPLE OF EVALUATION PROCEDURE OF A SAFETY SYSTEM

PERFORMANCE OF THE BRITISH RAILWAYS AUTOMATIC WARNING SYSTEM

The partial study presented in this Appendix is part of a wider study currently in progress at the R & D division at Derby, related to costs and benefits of Train Control Systems.

The information used in the material presented here is based on published data available to the general public in the form of Annual and Accident reports.

While it is appreciated that the information available is neither complete, nor sufficient to achieve statistical rigour, it is the author's opinion that the resulting margin of error is acceptable, and that it suitably illustrates the procedure.

The study is concerned with determining the accident rate attributable to, or otherwise not prevented by, the British Railways Automatic Warning System (BRAWS), but only a part of the study is given here as an example.

III.1. OPERATION OF BRAWS

System description

Trackside equipment: Permanent magnet

Electromagnet (energised only when the signal to which the equipment is associated is showing clear).
The permanent magnet is the first one to be found in the normal direction of travel and is placed, on average, 200 m before the signal.

Trainborne equipment: Magnetic receiver

Logic to determine whether signal is showing clear.

If the signal is showing clear the trainborne equipment is automatically reset and a bell rings in the cab.

If the signal is showing other than clear, a horn is sounded and an acknowledgement by the driver is necessary to cancel an otherwise automatic brake application. A special (sunflower) display will be activated after the acknowledgement to be reset by the next permanent magnet.

BRAWS came into being during the modernisation of motive power and was originally designed to operate on steam locomotives and semaphore signals (and AWS indications were given at all distant signals) and its installation only began in earnest following the accident at Harrow on 8th October 1952 (High speed double collision).

Since then, there has been little change in BRAWS design to make it compatible with multi aspect colour light signalling or to update its technology (it is in fact an overlay system designed for semaphore signalling).
III.2. INFORMATION IN BRAWS

It needs to be recognised that (see British Railways General Appendix, Rules and Regulations) there is a distinction between what a driver thinks things mean and what he is told they mean. Thus, if the driver is to get useful information (as against correct but useless information), additional information is required, e.g. is he in BRAWS territory, what happened when he opened up the cab, etc.

Strictly speaking BRAWS can give the following information:

a. At the instant of passing over the magnet assembly, the signal is showing clear (except in situations such as outlined below - blank display and bell).

b. If the display is blank but there is no bell, the display indicates: That the previous signal was showing clear (section between magnets), or that the train is running over non-equipped territory (see also the second example below).

c. If the display is showing a sunflower: An acknowledgement has been made after passing a magnet. The display represents one condition out of the following:

At the instant of passing over the last magnet:
The signal was showing double yellow, or
" " " " yellow, or
" " " " red, or
" " " " red and subsidiary, or
the electromagnet circuit had failed, or
the magnet is associated with a Speed Restriction Warning (Morpeth magnet),
or train is running over non-equipped territory.

EXAMPLES OF MISLEADING INDICATIONS

a. Running through a non-equipped zone

\[ \begin{align*}
\text{AWS} & \quad G \\
\Delta & \quad t^\circ \\
& \quad t^\circ \\
& \quad t^\circ \\
& \quad Y \\
& \quad YY
\end{align*} \]

DERBY STATION

train display will be blank ("clear")
despite having passed a YY aspect.
b. Premature cancellation of Speed Restriction Warning magnet reminder.

<table>
<thead>
<tr>
<th>BURTON</th>
<th>PSR</th>
<th>Aws</th>
<th>DERBY</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td></td>
<td>60</td>
<td>10</td>
</tr>
</tbody>
</table>

250 yds in "special cases"

200 yds

50 yds

worst case

200 yds

worst case 250 yds

approx. 6 seconds at 90 mph

BRAWS warning for PSR (Morpeth magnet)

3 - 4 seconds Acknowledgement and reminder

Display cleared by electromagnet

no reminder of PSR warning

---

but neither of these cases is a fault state.

III.3. PERFORMANCE OF BRAWS

Despite considerable technical innovation in the field of railway signalling (colour light signals, BR. 930 relays, geographic circuits, remote control, etc.) BRAWS has remained largely unchanged.
In re-examining the concepts and objectives of track to train communication systems, it became essential to determine the performance of BRAWS particularly with:
- Four aspect colour light signals
- Complex operating practice (speed restriction warnings, 2-way working, complex track layouts, coupling and splitting of trains at stations, etc.)
- High traffic densities (fast dynamics in changing signal aspects while the train is still in section).

III. 4. ANALYSIS PROCEDEURE

Appendix I presents the mathematical basis necessary for the analysis of safety systems. The techniques in this Thesis fall into two categories:
- Reliability analysis techniques, designed for the purpose of identifying the failure rates applicable to different failure modes. It should be clear that these techniques are considerably more detailed, and therefore more involved than the simple calculation of MTBF, based on adding the failure rates of all components.
- Calculation of the conditional probability functions which define the probability of accident $P_A$, which determine how well the system under consideration interfaces with its operating environment.

The main part of the study carried out for British Railways and presented in this Appendix, dealt with the evaluation of the
conditional probability of an event leading to an accident even when the system is working correctly. $P_{1/1}$ and more specifically the study concerned itself with the comparison of the value of this probability under "with" and "without" BRAWS conditions.

This example indicates that this evaluation is best carried out by the examination of records and, in the study of new systems, where considerable departure from established practice in the interface between man and machine, considerable caution in the choice of coefficients will need to be exercised.

In the interest of clarity, however, a brief example of the reliability analysis techniques, as applied to BRAWS will be given, although not in detailed numerical form.

Figure 28, of Appendix I, has been modified in Figure 34 to illustrate how this method of analysis can be applied in practice and it shows that, depending on the nature of the system, the subdivision into subsystem, component, etc., may need different numbers of steps.

The division shown takes as its obvious starting point the full system and its interaction, and defines two $(m = 2)$ subsystems, trackside and trainborne. These are in turn subdivided into components $\xi$, namely $\xi_1$ and $\xi_5$ for the trackside subsystem and $\xi_6$ to $\xi_{12}$ for the trainborne subsystem and some of these components can in turn be subdivided again, should it be necessary.
Thus, considering the system level, the n observable system states are defined as follows:

**Q<sub>j</sub>** : Correct operating states:

**Q<sub>1</sub>** : Signal aspect: Green
   Cab indication: 1 sec. Bell, at a point between 200 yds in front of the signal and the signal itself.

**Q<sub>2</sub>** : Signal aspect: Double yellow, Single yellow, or Red.
   Cab indication: Horn followed by a brake application if not cancelled within 4 seconds at a point between 200 yds in front of the signal and the signal itself.
   Following cancellation the indicator will show Yellow and Black in a sunflower pattern.

**F<sub>j</sub>** : Failure states:

**F<sub>1</sub>** : Signal aspect: Green
   Cab indication: Horn and Bell simultaneously.

**F<sub>2</sub>** : Signal aspect: Green
   Cab indication: Horn instead of Bell

**F<sub>3</sub>** : Signal aspect: Green
   Cab indication: None.

**F<sub>4</sub>** : Signal aspect: Double Yellow, Single Yellow or Red
   Cab indication: Bell and Horn simultaneously.

**F<sub>5</sub>** : Signal aspect: Double Yellow, Single Yellow or Red
   Cab indication: Bell instead of Horn.

**F<sub>6</sub>** : Signal aspect: Double Yellow, Single Yellow or Red
   Cab indication: Brake applied without Horn warning.

**F<sub>7</sub>** : Signal aspect: Double Yellow, Single Yellow or Red
   Cab indication: None.
\[ F_8 \] : Not at signal
  Cab indication: Horn.

\[ F_9 \] : Not at signal
  Cab indication: Bell.

\[ F_{10} \] : Horn and brake not cancellable

\[ F_{11} \] : Indicator "stuck" in warning state.

All these observable Failure states can be classified into Safe and Dangerous, the context being that of an Automatic Warning System:

Safe Failure States: \( F_1, F_2, F_3, F_4, F_6, F_8, F_9, F_{10}, F_{11} \).

Dangerous Failure States: \( F_5, F_7 \).

The next stage in the analysis could be the study of the two major subsystems, namely, the trainborne and the trackside equipments.

The detailed description of BRAWS, including circuit diagrams, etc., can be found in Booklet No. 24, issued by the Institution of Railway Signal Engineers, and need not be repeated here (Ref. 19). It is sufficient, therefore, to complement the brief description given earlier in this Appendix by postulating that each of the subsystems can be further split into a number of components (which could, if desired, be further subdivided into their elements, etc.).

Concerning the two subsystems, their observable states can be described as follows:
Trackside equipment observable states \( x_{11} \)

- Permanent magnet present and of adequate strength \( x_{11} \)
- Permanent magnet absent or of inadequate strength \( x_{21} \)
- Electromagnet correctly energised (signal aspect Green) \( x_{31} \)
- Electromagnet permanently energised \( x_{41} \)
- Electromagnet permanently de-energised \( x_{51} \)

These five states can again be identified with correct safe and dangerous system observable states:

\[
\begin{align*}
E_{Q}^1: & \quad x_{11}, x_{31} \\
E_{F \text{ safe}}^1: & \quad x_{51} \\
E_{F \text{ dangerous}}^1: & \quad x_{21}, x_{41}
\end{align*}
\]

Trainborne equipment observable states \( x_{12} \)

- 1 second Bell at a point between 200 yds in front of the signal and the signal itself \( x_{12} \)
- Horn followed by Brake application if not cancelled within 4 seconds, at a point between 200 yds in front of the signal and the signal itself \( x_{22} \)
- Horn and Bell simultaneously \( x_{32} \)
- Horn instead of Bell \( x_{42} \)
- Bell instead of Horn \( x_{52} \)
- No indication \( x_{62} \)
- Brake applied without Horn warning \( x_{72} \)
- Horn not at signal \( x_{82} \)
Bell not at signal $x_{92}$
Horn and Brake not cancellable $x_{10,2}$
Indicator permanently on "warning" $x_{11,2}$

and, as before, these 11 states can be identified with correct, safe and dangerous system observable states:

$E_2^Q: x_{12}, x_{22}$
$E_2^F_{safe}: x_{32}, x_{42}, x_{72}, x_{92}, x_{10,2}, x_{11,2}$
$E_2^F_{dangerous}: x_{52}, x_{62}$

At this stage it becomes possible to identify the particular role of each subsystem and component. For example, it can now be shown that all the fault states $F_1, F_4, F_6, F_8, F_9, F_{10}$ and $F_{11}$ can only be attributed to failures in the trainborne equipment.

The next stage in the analysis concerns itself with the analysis of the individual components. It can be noted from Figure 34 that several of these components are of a sufficiently simple nature to permit the analysis of their contribution to system performance without further subdivision.

Taking, for example, the components of the trackside subsystem:

$E_1$, Permanent magnet:

$E_{11}$: Permanent magnet of adequate strength
$E_{12}$: Permanent magnet of inadequate strength
$E_{13}$: Permanent magnet missing or damaged.
$E_{11}$ is the correct state, and an analysis of performance (or a guess when the system is new) will give a probability of this being the case. In BRAWS this probability is estimated as $1 - 5 \cdot 10^{-4}$ (from failure reports dated 1970 - 73).

$E_{12}$ has now been virtually eliminated by design, as High Strength Ferrites with a Square Histeresis Loop are used.

The probability of a damaged magnet, despite its protective steel ramp and case is thus in the order of $5 \cdot 10^{-4}$.

$E_{2}$, Electromagnet:

- $E_{21}$: Electromagnet energised to adequate strength
- $E_{22}$: Electromagnet energised to insufficient strength
- $E_{23}$: Electromagnet permanently energised
- $E_{24}$: Electromagnet not energised.

$E_{21}$ is the correct operating state.

$E_{22}$ and $E_{24}$ are safe failure states, but $E_{22}$ constitutes a failure in the power supply circuit (see $E_{5}$) and $E_{24}$ can occur due to either of four conditions:

a. Open circuit electromagnet coil.

b. Cable failure between the signal and the electromagnet (see $E_{3}$).

c. Circuit not completed by the signal relay (see $E_{4}$).

d. No power supply (see $E_{5}$).
Thus, \( \mathcal{E}_{22} \) is quantified under \( \mathcal{E}_{5} \), and \( \mathcal{E}_{24} \) in this list will be associated only with the failure rate of the electromagnet coil.

\( \mathcal{E}_{23} \) is the dangerous failure state and can occur only as a result of a failure in the signal relay (see \( \mathcal{E}_{4} \)).

\( \mathcal{E}_{3} \), Cable to Electromagnet.

- \( \mathcal{E}_{31} \): Cable correct
- \( \mathcal{E}_{32} \): Cable open circuited
- \( \mathcal{E}_{33} \): Cable missing (vandalism, etc.).

These two latter conditions represent the situation described under \( \mathcal{E}_{24-b} \), and are safe failure states.

\( \mathcal{E}_{4} \), Signalling relay.

- \( \mathcal{E}_{41} \): All contacts and coil operate normally
- \( \mathcal{E}_{42} \): Open circuit coil
- \( \mathcal{E}_{43} \): Short circuit coil
- \( \mathcal{E}_{44} \): Welded contacts.

\( \mathcal{E}_{42} \) prevents the relay from being energised and thus leads to the condition described as \( \mathcal{E}_{24-c} \).

\( \mathcal{E}_{43} \) is virtually eliminated by design and, in the worst case, could result in some turns being shortcircuited and thus reduced coil inductance and longer operating time.
\( \varepsilon_{44} \) is also virtually eliminated by the design of carbon-silver contacts. This is potentially a dangerous failure, but long years of experience with this relay design indicate that the probability of such a failure is within acceptable values (no such incident has been recorded to date).

**5, Power Supply.**

- \( \varepsilon_{51} \): Normal output
- \( \varepsilon_{52} \): Reduced output
- \( \varepsilon_{53} \): No output,

where \( \varepsilon_{52} \) and \( \varepsilon_{53} \) are safe failure conditions associated with \( \varepsilon_{22} \) and \( \varepsilon_{24} \).

The study of the power supply could be taken further by examining the effect of each basic component, e.g. battery, trickle charger, fuse, etc.

Examining now the trackside equipment observable states, these can be associated with their component states as follows:

- Permanent magnet present and of adequate strength \( x_{11} \leftarrow \varepsilon_{11} \)
- Permanent magnet of inadequate strength (*) \( x_{21} \leftarrow \varepsilon_{12} \)
- Permanent magnet absent (*) \( x_{21} \leftarrow \varepsilon_{13} \)
- Electromagnet correctly energised \( x_{21} \leftarrow \varepsilon_{21}, \varepsilon_{31}, \varepsilon_{41}, \varepsilon_{51} \)

(* dangerous states)
Electromagnet permanently energised (*)

Electromagnet permanently de-energised

\[
\begin{align*}
\xi_{41} & \rightarrow \xi_{44} \\
\xi_{51} & \rightarrow \xi_{22}, \xi_{24}^a, \\
& \quad \xi_{32}, \xi_{33}, \xi_{42}^b, \\
& \quad \xi_{52}, \xi_{53}
\end{align*}
\]

thus illustrating the mapping of component observable states into observable subsystem states.

The analysis of the trainborne equipment can be carried out exactly along the same lines. Its somewhat more complex nature makes it necessary to subdivide it into a number of component parts.

The diagram in Figure 35 shows, with the exception of the power supply, the circuit diagram of the trainborne BRAWS equipment, as used on BR locomotives and multiple units.

For the purpose of the analysis, the following will be considered as components:

- The receiver
- The reset coil
- The EP valve
- The indicator
- The reset button
- The reminder of the logic unit
- The power supply.
$\xi_6$, Magnetic receiver.

Three conditions can be identified with the magnetic receiver:

$\xi_{61}$: Correct operation, the receiver can respond normally and will provide electric contacts at both N and S positions.

$\xi_{62}$: Receiver will not commutate from its rest position ("stuck at N").

$\xi_{63}$: Receiver cannot be returned to position N ("stuck at S").

$\xi_{62}$ can be shown to be a dangerous condition as it effectively prevents the detection of a permanent magnet.

$\xi_{63}$ is a safe state in which the brakes cannot be released (as far as this analysis is concerned, this fault is local to the receiver – logic faults are discussed in $\xi_{11}$).

$\xi_7$, Reset coil.

The reset coil is placed near the receiver and is driven by the logic in such a way that the automatic brake application initiated by the equipment may be cancelled by the driver in charge of the train.

$\xi_{71}$ is the normal operating state of this coil.

$\xi_{72}$ represents an open circuit reset coil. It will be accepted that due to its constructural features, a short circuited coil cannot occur under any forseeable operating conditions.
The inability to reset the receiver would also result in a safe state in which the brakes cannot be released.

\( \mathcal{E}_8 \), Electropneumatic valve (EP valve)

The electropneumatic valve is an integral part of the braking system, and this study will only be concerned with electrical failures.

\( \mathcal{E}_{81} \) represents the correct operation of the EP valve
\( \mathcal{E}_{82} \) represents an open circuited EP valve coil
\( \mathcal{E}_{83} \) represents a shortcircuited (welded) EP valve contact

\( \mathcal{E}_{82} \) is a safe state in which the brakes cannot be released.
\( \mathcal{E}_{83} \) is a dangerous state in which the brakes cannot be applied (The probability of this event is reduced by both valve design and the use of redundancy (twin and triple valves)).

\( \mathcal{E}_9 \), The indicator.

In addition to its function as a visual reminder, the indicator performs a logical function in the reset coil circuit (contacts 4 and 5 in the bottom part of Figure 35), and the reset coil will only be energised if the indicator has moved to the "sunflower" position and closed the contact.

\( \mathcal{E}_{91} \) contact correct
\( \mathcal{E}_{92} \) infinite resistance contact
\( \mathcal{E}_{93} \) welded contact.
\( \xi_{g2} \) results in a safe state in which the brakes cannot be released.  
\( \xi_{g3} \) does not by itself lead to a fault condition (other than a "stuck at sunflower" indicator.

\( \xi_{10} \), The reset button.

The possible conditions are:

\[ \begin{align*}
\xi_{10,1} & : \text{Correct operation} \\
\xi_{10,2} & : \text{Infinite resistance contacts} \\
\xi_{10,3} & : \text{Welded contacts} \\
\xi_{10,3a} & : \text{in position 2} \\
\xi_{10,3b} & : \text{in position 3}.
\end{align*} \]

\( \xi_{10,2} \) and \( \xi_{10,3a/b} \) will both result in the inability to energise the relay coil, and hence in a situation where the brakes cannot be released.

\( \xi_{11} \), The logic unit.

The logic functions of BRAWS are carried out by a simple circuit of four relays, labelled in the circuit diagram in Figure 35 as BR, SR, EPR and NCR.

The proper operation of the (sequential) circuit also relies on certain time delays, being achieved by the use of capacitors.

The basic operation of the logic can be summarised as follows:

When the receiver passes over a permanent magnet, the S contact is made. This energises the SR relay, which latches (contact \( SR_3 \)).
The indicator's "8" coil is energised, resetting it to the blank position.

The circuit is arranged in such a way that the EPR relay will remain energised for 1 second after opening the contact $SR_1$.

Two basic situations are now possible:

a. Energised electromagnet present.

In this case the EPR relay does not drop out and the returning of the receiver to the N contact causes the BR relay to be energised (sounding the Bell for 1.5 sec.). The SR relay drops and the system returns to its original state when the BR relay is de-energised (through contact $SR_1$).

b. No electromagnet.

After a 1 second delay, the EPR relay becomes de-energised (thus releasing the EPR$_1$ latch contact). EPR$_2$ opens the EP valve and EPR$_3$ closes to permit the completion of the NCR circuit.

If and when the reset button is operated, the NCR relay becomes energised and the receiver's reset coil returns the receiver to the N contact.

Contact NCR$_1$ provides an alternative path for energising the EPR relay. NCR$_2$ drops the SR relay and this, together with contact NCR$_3$ energises the "y" coil of the indicator. Contact NCR$_4$ latches the NCR relay, and the trainborne unit will remain in this state until a new permanent magnet is encountered, in which case the cycle begins again.
Failure analysis

The logic unit represents a good example of a subsystem for which further subdivision into simpler elements, e.g. the relays, does not provide the desired answers, due to the interdependence between these elements.

In situations such as these, systematic analysis, point by point, appears to be the best answer. Logic circuits lend themselves to computer simulation and, in fact these techniques have been used in 1968/69 by the British Railways R & D Division, working in conjunction with the National Computing Centre (NCC) of Manchester.

In the case of the logic unit, this analysis would be carried out by considering each relay in turn and considering the cases of open circuit coil and welded contacts. The type of relay used in the logic unit does not exclude the possibility of a welded contact while the other contacts (excluding the back contact) are operating correctly.

Because such an analysis would be neither original nor particularly useful in illustrating the procedure, it will not be included in this analysis.

12 Power supply.

Although the power supply can also be subdivided into various elementary parts, for the purpose of this simple analysis, it will be treated as a single unit.
The circuit has been arranged in such a manner that

12,1 Normal power supply: System operates correctly

12,2 Power supply failure (zero output voltage)

results in the operation of the EP valve, and thus, a brake application.

As already shown in the case of the trackside equipment, all these failure states could now be mapped onto the subsystem failure states, and these in turn with the system observable states.

This will lead to failure equations listing all the events which can cause a particular failure state, and if the nature of the statistical distribution of these failures is known, a numerical description of the probability of transition to any particular state can be obtained.

These probabilities can then be used in the Markov process models of the system described in Appendix I, in order to calculate the probabilities $P_1$, $P_2$ and $P_3$.

The second part of this analysis will be concerned with the determination of one of the conditional probabilities $(P_{1/1})$ in the equation leading to the probability of accident $P_A$. 
For the second part of the analysis it was considered that the most suitable method of evaluation would be to correlate accident records, since BRAWS is intended as an aid to safety with train miles under different sets of conditions. The analysis procedure involved the following two areas:

**Accident Identification**

For the purpose of this study the following steps were followed:

a. Listing of all accidents resulting from Signals Passed at Danger (SPD) which were reported between 1.1.1969 and 31.7.1974 (a total of 120).

These dates were chosen on the following grounds:

- The period under examination was long enough to even out statistical fluctuations.
- Signalling and traffic conditions have been sufficiently stable to permit comparison.

b. From the above listing, the following were deleted:

- Those where the primary cause was other than human error.
- Those that occurred in non-colour light signal territory.
c. The remainder were classified into:

- BRAWS fitted territory and fitted trains.
- BRAWS fitted trains over unfitted portions within BRAWS fitted territory (e.g. complex parts at Derby, Birmingham, Glasgow Central, etc.).
- Territory not fitted with BRAWS (Mainly the Southern Region).

These accidents are listed in Tables 1, 2 and 3 respectively.

**TABLE 1.**

**MAJOR ACCIDENTS IN COLOUR LIGHT MULTIPLE ASPECT SIGNALS AND BRAWS TERRITORY 1969 - 1974. PRIMARY CAUSE: HUMAN ERROR LEADING TO SPD**

<table>
<thead>
<tr>
<th>Date</th>
<th>Location</th>
<th>Line</th>
<th>Type</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.4.69</td>
<td>Monmore Green</td>
<td>L.M.</td>
<td>Collision</td>
<td></td>
</tr>
<tr>
<td>27.5.70</td>
<td>Olbury</td>
<td>L.M.</td>
<td>Collision</td>
<td></td>
</tr>
<tr>
<td>16.12.71</td>
<td>Lenton Junction</td>
<td>L.M.</td>
<td>Collision</td>
<td></td>
</tr>
<tr>
<td>27.4.73</td>
<td>Kidsgrove</td>
<td>L.M.</td>
<td>Collision - AWS isolated by driver</td>
<td></td>
</tr>
<tr>
<td>29.6.73</td>
<td>Euston</td>
<td>L.M.</td>
<td>Signalman able to avoid accident by diverting to empty platform</td>
<td></td>
</tr>
<tr>
<td>18.5.73</td>
<td>Reading Station</td>
<td>W.R.</td>
<td>Collision (W.R. Ramp)</td>
<td></td>
</tr>
<tr>
<td>21.10.73</td>
<td>Bethnal Green</td>
<td>E.R.</td>
<td>Collision and Derailment</td>
<td></td>
</tr>
</tbody>
</table>
**TABLE 1.** (continued)

<table>
<thead>
<tr>
<th>Date</th>
<th>Location</th>
<th>Nature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>19.10.70</td>
<td>Glasgow Central Sc.R.</td>
<td>Collision after passing</td>
<td>two signals at danger</td>
</tr>
<tr>
<td>30.8.73</td>
<td>Shields Junction SC.R.</td>
<td>Collision into train</td>
<td>stopped to extinguish fire</td>
</tr>
<tr>
<td>11.6.74</td>
<td>Muirhouse Junction Sc.R.</td>
<td>Collision between 2 EMUs</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 2.**

**MAJOR ACCIDENTS IN COLOUR LIGHT MULTIPLE ASPECT SIGNALS. UNFITTED AREAS IN BRAWS TERRITORY 1969 - 74. PRIMARY CAUSE: HUMAN ERROR LEADING TO SPD.**

<table>
<thead>
<tr>
<th>Date</th>
<th>Location</th>
<th>Nature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>21.5.69</td>
<td>Manchester Piccadilly</td>
<td>Collision</td>
<td></td>
</tr>
<tr>
<td>25.12.69</td>
<td>Manchester Piccadilly</td>
<td>Run through points</td>
<td></td>
</tr>
<tr>
<td>17.7.70</td>
<td>Crewe South Junction</td>
<td>Run through points</td>
<td></td>
</tr>
<tr>
<td>17.7.70</td>
<td>Birmingham New Street</td>
<td>Run through points</td>
<td></td>
</tr>
<tr>
<td>5.10.70</td>
<td>Manchester Piccadilly</td>
<td>Collision</td>
<td></td>
</tr>
<tr>
<td>19.10.70</td>
<td>Glasgow Central</td>
<td>Collision</td>
<td></td>
</tr>
<tr>
<td>9.7.71</td>
<td>Birmingham New Street</td>
<td>Derailment</td>
<td></td>
</tr>
</tbody>
</table>

Study not completed

Glasgow area reports 12 SPD in 1969) in unfitted zone* 16 SPD in 1970)

<table>
<thead>
<tr>
<th>Date</th>
<th>Location</th>
<th>Nature</th>
</tr>
</thead>
<tbody>
<tr>
<td>29.3.71</td>
<td>Glasgow Central</td>
<td>Collision</td>
</tr>
</tbody>
</table>

* Further information on these incidents is not available and these will, therefore, not be introduced into the calculations.
TABLE 3.

MAJOR ACCIDENTS ON SOUTHERN REGION COLOUR LIGHT MULTIPLE ASPECT SIGNAL TERRITORY 1969 - 70.

PRIMARY CAUSE: HUMAN ERROR LEADING TO SPD.

<table>
<thead>
<tr>
<th>Date</th>
<th>Location</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1.69</td>
<td>Paddock Wood</td>
<td>Collision</td>
</tr>
<tr>
<td>18.11.69</td>
<td>Portsmouth</td>
<td>Derailment</td>
</tr>
<tr>
<td>12.11.70</td>
<td>Bexley</td>
<td>Collision - Driver claimed unable to brake.</td>
</tr>
<tr>
<td>27.11.71</td>
<td>Portsmouth</td>
<td>Collision</td>
</tr>
<tr>
<td>7.9.72</td>
<td>Clapham Junction</td>
<td>Derailment</td>
</tr>
<tr>
<td>12.10.72</td>
<td>Wimbledon</td>
<td>Freight Collision</td>
</tr>
<tr>
<td>20.12.72</td>
<td>Copyhold Junction</td>
<td>Collision</td>
</tr>
<tr>
<td>25.6.73</td>
<td>Cannon Street</td>
<td>Derailment</td>
</tr>
<tr>
<td>25.10.73</td>
<td>Waterloo Station</td>
<td>Derailment</td>
</tr>
</tbody>
</table>

Although the Waterloo-Bournemouth line of the Southern Region is fitted with BRAWS, none of the incidents listed in Table 3 took place over BRAWS lines.

Table 4, below, summarises the results of this part of the analysis.

TABLE 4.

<table>
<thead>
<tr>
<th>BR total</th>
<th>COLOUR LIGHT SIGNALS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BRAWS fitted territory</td>
</tr>
<tr>
<td>NUMBER OF ACCIDENTS</td>
<td>120</td>
</tr>
</tbody>
</table>
Train-miles evaluation

All the published data necessary for this part of the analysis has been listed in Tables 5, 6 and 7. This information is, however, not sufficient to give a fine enough breakdown as required.

To overcome this problem, assumptions based on extrapolations and information given verbally have been made. To estimate the possible error, a sensitivity analysis of the results will follow.

Assumptions:

a) It will be assumed that 70% of all train miles occur in colour light signal territory.

Data: Colour light signals are fitted on \( \frac{9,273}{22,561} \approx 41\% \) of all running track miles.

All the busiest lines have colour light signals.

b) It will be assumed that there is no BRAWS on the Southern Region (excluding Waterloo-Bournemouth) and furthermore that the SR carries:

- 60% of all EMU traffic
- 10% of all BR freight

These are taken as conservative estimates. SR achieves 48,500,000 train miles p.a. This amounts to approximately 20% of all BR train miles, although it is felt that this figure is too conservative and that 30% would be more appropriate. However, to avoid biasing the study, the figure of 20% will be used.
c) It will be assumed that the portion of BRAWS territory not fitted with track equipment, notably in complex layouts, accounts for 5% of all train miles.

c) There is no published data on track miles fitted with BRAWS. The route-mile figure of 3,759 miles (1973) can be converted by multiplying by a factor of 2.4 (BR average ration of 

\[
\frac{T. \text{ Miles}}{R. \text{ Miles}} : 2 \quad \text{BRAWS track miles approx. 7,500 m).}
\]

\[
\text{TABLE 5.}
\]

\[
\text{TRACK AND SIGNALLING (IN 1973)}
\]

\[
\text{SOURCE: BR ANNUAL REPORT AND ACCOUNTS}
\]

<table>
<thead>
<tr>
<th>ROUTE MILES</th>
<th>Passenger Traffic only</th>
<th>525</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freight and Passengers</td>
<td>8,407</td>
</tr>
<tr>
<td></td>
<td>Freight only</td>
<td>2,394</td>
</tr>
<tr>
<td>of which:</td>
<td>Fitted with BRAWS</td>
<td>3,759</td>
</tr>
<tr>
<td></td>
<td>Electrified AC overhead</td>
<td>979</td>
</tr>
<tr>
<td></td>
<td>DC overhead</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>DC 3rd rail</td>
<td>1,100</td>
</tr>
</tbody>
</table>

| TRACK MILES | Running lines | 22,561 |
|            | Sidings       | 6,826 |
| of which:  | Equipped with colour light signals | 9,273 | *1 |
|            | Electrified AC overhead | 2,542 |
|            | DC overhead      | 193  |
|            | DC 3rd rail      | 2,561 |
|            | Sidings all systems | 465  |
It is not known whether this figure includes mixed semaphore and colour lights (e.g. East Coast M.L.).

**TABLE 6.**
ROLLING STOCK *(1973 BR ANNUAL REPORT AND ACCOUNTS)*

<table>
<thead>
<tr>
<th>Locomotives:</th>
<th>22 shunting</th>
<th>333 electric</th>
<th>2,517 diesel</th>
</tr>
</thead>
<tbody>
<tr>
<td>of which:</td>
<td></td>
<td>2,742 fitted with BRAWS</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>FITTED WITH BRAWS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diesel Multiple Units:</td>
<td>2,044 + 1,424</td>
<td>1,786</td>
</tr>
<tr>
<td></td>
<td>power cars</td>
<td>coaches</td>
</tr>
<tr>
<td>Electric Multiple Units:</td>
<td>2,665 + 4,508</td>
<td>1,459</td>
</tr>
<tr>
<td></td>
<td>power cars</td>
<td>coaches</td>
</tr>
</tbody>
</table>

**TABLE 7.**
TRAIN MILEAGES *(LOADED TRAIN MILES)* *(1973 BR ANNUAL REPORT AND ACCOUNTS)*

<table>
<thead>
<tr>
<th>Coaching stock:</th>
<th>Diesel</th>
<th>55,853,000</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Electric</td>
<td>12,987,000</td>
</tr>
<tr>
<td></td>
<td>DMU</td>
<td>55,451,000</td>
</tr>
<tr>
<td></td>
<td>EMU</td>
<td>70,239,000</td>
</tr>
<tr>
<td></td>
<td>Freight Diesel</td>
<td>48,933,000</td>
</tr>
<tr>
<td></td>
<td>Electric</td>
<td>4,957,000</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>244,420,000</td>
</tr>
</tbody>
</table>
BRAWS is also fitted in semaphore territory and it will be assumed that all colour light signal territory, with the exception of the Southern Region, is fitted.

**EVALUATION:**

<table>
<thead>
<tr>
<th></th>
<th>Non-colour light with or without BRAWS</th>
<th>Colour light with BRAWS</th>
<th>Colour light and BRAWS un-fitted areas</th>
<th>Colour light not fitted with BRAWS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train miles %</td>
<td>30</td>
<td>45</td>
<td>5</td>
<td>20</td>
</tr>
</tbody>
</table>

**Accidents per % of Train Miles.**

**TABLE 8.**

**SUMMARY**

<table>
<thead>
<tr>
<th>Non Colour Light</th>
<th>Colour light signals</th>
<th>BRAWS TERRITORY</th>
<th>Non BRAWS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Fitted areas</td>
<td>Non-fitted areas</td>
</tr>
<tr>
<td>Number of accidents</td>
<td>93</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>% of Train Miles</td>
<td>30</td>
<td>45</td>
<td>5</td>
</tr>
<tr>
<td>Accident Rate (% TM)</td>
<td>3.1</td>
<td>.22</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.36</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.39</td>
</tr>
</tbody>
</table>
III.5. SENSITIVITY ANALYSIS

**TABLE 9.**
NON BRAWS TERRITORY CARRIES 25% OF TRAIN MILES

<table>
<thead>
<tr>
<th>Light Signals</th>
<th>Non Colour Light</th>
<th>Colour light signals</th>
<th>Non Colour Light</th>
<th>Colour light signals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BRAWS Territory Fitted areas</td>
<td>Non-fitted areas</td>
<td>Territory</td>
<td>Non Colour Light</td>
</tr>
<tr>
<td>Number of accidents</td>
<td>93</td>
<td>10</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>% of Train Miles</td>
<td>30</td>
<td>40</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>Accident Rate (%TM)</td>
<td>3.1</td>
<td>.25</td>
<td>1.6</td>
<td>.36</td>
</tr>
</tbody>
</table>

**TABLE 10.**
BRAWS ACCOUNTS FOR 50% OF TRAIN MILES.

<table>
<thead>
<tr>
<th>Light Signals</th>
<th>Non Colour Light</th>
<th>Colour light signals</th>
<th>Non Colour Light</th>
<th>Colour light signals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BRAWS TERRITORY Fitted areas</td>
<td>Non-fitted areas</td>
<td>Territory</td>
<td>Non Colour Light</td>
</tr>
<tr>
<td>Number of Accidents</td>
<td>93</td>
<td>10</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>% of Train Miles</td>
<td>30</td>
<td>50</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>Accident Rate (%TM)</td>
<td>3.1</td>
<td>.2</td>
<td>1.6</td>
<td>.45</td>
</tr>
</tbody>
</table>
Thus indicating that colour light signals have made a much greater contribution to safety than BRAWS.

III.6. LIMITATIONS OF THE STUDY

The following factors are recognised as having an effect on the validity of this study:

- The effects of using permanent magnets to indicate the braking point to severe (and also to temporary) speed restrictions (these are the so called "Morpeth" magnets).

- Empty train miles have not been considered due to the lack of reliable data (although it is believed they may account for as much as 15% of Train Miles).

- Single Manning practice in the Southern Region (non BRAWS territory with the exception of the Waterloo-Bournemouth line), which make it difficult to separate the effects of single manning and A.W.S.

- The use of train miles instead of signal passed, as in busy areas there are more signals per mile than on plain track in low density lines.
APPENDIX IV (Supplement to Section 6.1)

COST BENEFIT ANALYSIS

Solid State Techniques in Signalling

Objectives:
To investigate the use of microprocessors and allied electronic technologies in fail-safe railway signalling systems. The high and ever increasing costs of labour, cabling and special components such as signalling relays, could be significantly avoided by the use of the new technologies.

The study will involve the design of fail-safe telemetry systems and the design of solid state interlocking techniques – both of them applicable to other systems, such as chemical processes and nuclear reactor protection, having therefore spinoff potential.

The study will consist of the following phases:

<table>
<thead>
<tr>
<th></th>
<th>man-</th>
<th>mater-</th>
<th>comput-</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>power</td>
<td>ials</td>
<td>ing</td>
</tr>
<tr>
<td>I:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study phase, duration one year, problem formulation and possible structures for the solution (1 year)</td>
<td>4</td>
<td>£2.5k</td>
<td>£2k</td>
</tr>
<tr>
<td>II:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hardware development, validation of principles and laboratory testing (2 years)</td>
<td>4</td>
<td>£5k</td>
<td>£3k</td>
</tr>
</tbody>
</table>
Present estimated cost of signalling per track mile £40,000,
of which: Labour £20,000
           Cabling £10,000
           Remaining Equipment £10,000.

These figures apply to an average cost for a major resignalling
scheme and include the cost of buildings, train describer, signalling
panel in the signal box, etc.

The cabling costs given include both signalling and telecommunications
cables and the actual interlocking amounts to 10% of the cost of a
scheme.

It is estimated that in the worst case the following savings can be
realised:

    Labour: Nil

* Signalling Cabling: 15%
** Interlocking : 30%

*Note: This saving is achievable through the use of fail-safe
digital telemetry (a likely saving would be 20%).

**Note:** The installed cost for a safety relay (type 930) is in the order of £60 and this provides 8 front and 8 back contacts. The logic capability of such a device is, therefore, limited and large numbers of them are required.

In terms of the total cost of a signalling scheme, the following worst case savings result:

<table>
<thead>
<tr>
<th>Labour:</th>
<th>Nil</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cabling:</td>
<td>5% (only signalling cables)</td>
</tr>
<tr>
<td>Equipment</td>
<td>3% (only interlocking)</td>
</tr>
</tbody>
</table>

Thus, the worst case average track mile saving is:

<table>
<thead>
<tr>
<th>Labour:</th>
<th>£ —</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cabling:</td>
<td>£ 500</td>
</tr>
<tr>
<td>Interlocking:</td>
<td>£ 300</td>
</tr>
</tbody>
</table>

£ 800 net saving per track mile.

In practice, it is expected that costs of both cabling and signalling relays will rise faster than the cost of electronic equipment.

These figures cannot be substantiated at the present time and are based on engineering judgement. The study reveals, however, that even marginal savings result in a favourable benefit-cost ratio and would appear to justify the study.
An additional incentive may be an increase in reliability, inherent in the microprocessor technology, which may help to reduce the stock of spares (not quantified in the study).

There are 9,400 track miles yet to be implemented in the Signalling Plan. The current rate of implementation is 550 track miles per annum, with a target fitment figure of 800 track miles per annum. In addition, assuming a 40 year life for signalboxes, commencing in 1990, a renewal plan will commence; i.e. the London - Crewe line, giving the following planned investment:

Outstanding track miles in signalling plan
(based on 1981 total BR track miles: 20,500
and 1,500 track miles not qualifying for colour light signalling)
: 9,400 track miles

Case I: Current fitment rate (worst case analysis)
: 500 t.m.p.a.

Outstanding track miles in 1985 (new technology available)
: 3,900

Completion of signalling plan and beginning of replacement of existing installations: 1992

Saving in installation costs: (constant)
: £800 per track mile
: £440,000 p.a.
Case II: Desirable fitment rate

\[ (= \text{target rate}) \quad : \quad 800 \text{ t.m.p.a.} \]

Outstanding track miles in 1985 (new technology available)

\[ : \quad 1,400 \]

Completion of signalling plan and beginning of replacement of existing installations: 1897
COST BENEFIT ANALYSIS (WORST CASE)

Fitment rate = 500 track miles p.a.

DISCOUNTED BENEFIT = (1985 - 2000) (10% rate)

(Discounted (cost of new technology - cost of

present technology) ........ £ 1,619,200

DISCOUNTED RESEARCH AND DEVELOPMENT COSTS (1976 - 1984) £ 201,220

Assumption: Probability of technical success: 0.60
Probability of investment capital
being available (signalling
being vital to railway operation
and equipment being life-expired):

This work is a direct spinoff from the Train Control developments
under trial at Wilmslow.

WORST CASE BENEFIT TO COST RATIO FOR THE R & D PROGRAM:

\[ 0.6 \left( \frac{\text{£1,619,200}}{\text{£201,220}} \right) = 4.8 : 1 \]

In view of the favourable ratio obtained for the worst case
analysis and in addition, the fact that the modernisation of
signalling equipment remains an important priority as far as BR
is concerned, it is considered that neither an "optimistic" nor
a "likely" study will contribute any useful information.
APPENDIX V.

RELATION OF THE PRESENT THESIS WITH OTHER RESEARCH ACTIVITIES AT LOUGHBOROUGH UNIVERSITY

The Department of Electrical and Electronic Engineering at Loughborough University has collaborated with the British Railways R & D Division since 1969.

A major effort in the study of moving block signalling systems was carried out by L. Pearson, Ph. D.

Present research by the author, Mr. P.D. Thomas and Mr. K. Gavin is concerned with analysing on-line control strategies with the objective of deriving quantitative techniques to measure performance, and the areas covered in the respective studies are safety and full automation, the problems of operating single vehicles and traffic regulation.
REFERENCES


GENERAL REFERENCES.

13. Gallagher - Information Theory


15. IEEE Transactions on Computers (special issues on fault tolerant computation).


Signals 1 and 2 cannot show "Line clear" simultaneously.

---

**Basic Railway Signalling**

*Figure 1.*
Simple Track Circuit

Fig. 2
SafetY Loop

Figure 3
Train Regulation Loop - Extended Use of Train Describers

Figure 4.
Railway Information, Planning and Control (Cybernetics)
Simple Track Circuit

Figure 6
Failure Performance of Systems

Is the reliability of the system infinite?

NO

Can the presence of faults be detected?

YES

Can the system take any action?

YES

Continue full or partial operation through automatic connection of standby

NO

Inhibit operations by reverting to safe state

YES

Safe system designed for high availability

NO

Ideal system or system where operational life is << MTBF

Wrong side failure or system not suitable for safety application

Figure 7
B.R. Signal Aspects

Figure 8
\( \lambda_{ij} \): failure rates

\( \lambda_{ij}(t) \): probability of transition from state \( i \) to \( j \) during interval at \( t \)

\( i, j = a, b, c \); \( i \neq j \)
\[ \lambda_{ij} = \text{failure rates} \]
\[ \mu_{ji} = \text{repair rates} \]
\[ \lambda_{ij} \Delta t = \text{probability of transition from } i \text{ to } j \text{ during interval } \Delta t \]
\[ \mu_{ji} \Delta t = \text{probability of transition from } j \text{ to } i \text{ during interval } \Delta t \]
\[ i,j = a,b,c ; i \neq j \]
\[ \lambda_{ij} = \text{failure rates} \]
\[ \mu_{ij} = \text{repair rates} \]
\[ \lambda_{ij} \, \Delta t = \text{as in figure 10} \]
\[ \mu_{ij} \, \Delta t = \text{as in figure 10} \]
\[ i,j = a, a', b, c ; i \neq j \]
\[ \text{except } a a' \]

Figure 11
Opportunities arising from new technologies

Current principles?  

Approach unsuitable

Reporting of system performance?

Functional Enhancements?

Closed Loop System

Replacement of Technology

Open Loop System

* Define standby system etc. (see text)

* Closed and open loops (see text)

Centralised  
either Centralised or decentralised  
Centralised

Figure 12

System Structures
Logic I may be either centralised or decentralised.

Validation unit required if real-time data processing takes place in Logic I (see text).

Validation unit must be fail safe (and executive by disabling the output).

**Safety Features in an Open Loop System**

Figure 13
Logic I may be either centralised or decentralised.

Validation unit must be fail-safe (and hence by disabling the system).

Input 1.
Safe Information

Enable

Input 2

Logic I

Information either safe or to be validated

Enable

Data processing

(Data transmission channels (protected against interference)

Direct control of remote validation unit

Depending on system design

Output
Safe Information

SafetY FEATURES IN A CLOSED LOOP SYSTEM

Figure 14
Open and Closed Loops in Signalling Systems

Figure 15
Example of a Safe Open Loop System: Automatic Signals
Continuous and Discrete Information Processes

Figure 17
(Simplified) Requirements for Validation of a New Signalling Scheme

Figure 18
Safety and Validation of Lineside Speed Supervision Information Figure 19
Simplified Flow Chart of Interlocking

Note: For simplicity, interlocking between points has been ignored.
ARRANGEMENT OF B.R. SIGNALLING EQUIPMENT

Figure 21
Two wire system in simple cases

Solid State Techniques in Signalling,
Possible Configuration at a Signal Location Figure 22
Functions: - to determine optimum regulation strategy if different from timetable
- to recommend (and enforce) an optimum speed-distance curve
- to recommend (and enforce) a minimum energy trajectory for a specified speed-distance curve or point-to-point timing
- to control stopping at stations

Functions: - to route the route information from the signal box to the area to which it applies (and back if necessary)

Function: To ensure that under all conditions the maximum permissible speed is not exceeded

Elements in an ATO system for general use

Figure 23
ON LINE CONTROL

Human Intervention for management of malfunction, emergency etc.

- Emergency commands
- "off-limits" reports and emergency information
- Direct to trains (relaying cables)

COMMUNICATIONS
Speech & Data

COMMUNICATIONS
(DATA CONTROL)

STANDBY & BATCH PROCESSOR

<table>
<thead>
<tr>
<th>files</th>
<th>history &amp; statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ON LINE

- Coaches
- Trains
- Network parameters (location)
- Traction
- Schedule
- Personal
- Spares & maintenance
- etc.

ALL THE ELEMENTS IN
Figure 30

- Vehicle instrumentation
- "Within limits" validation
- Fail-safe status report
- Adaptive Brakes
- Door Controls
- Fail-safe Tacho
- & Zero speed detector

traction and coaches in use
train characteristics
(length, braking, etc)

updated schedule
updated speed restrictions

CREWLESS TRAIN OPERATION:
SYSTEM CONFIGURATION

Figure 24
Vehicle Controls

- Zero speed condition at authorised station
- Zero speed not at station interlock
- Remote command to open * for emergency debraking
- Door jammed or locked *

- Lighting
- Battery emergency supply
- Heating and/or air conditioning
- Passenger communication link
- [charged/changing, etc]

- Auxiliary Equipment
- Brakes
- Train Control
- Doors
- Track
- Track to train communications
- \( x_{max}; x_{opt}; \) station stops, speed distance trajectory
- \( \text{from safety loop and Train Regulation Loop} \)
- \( \text{(electric) power supply} \)
- \( V; I; \frac{dI}{dt} \)
- \( \text{(or fuel tank where applicable)} \)
- \( Q; \frac{dQ}{dt} \)
- \( \text{Absolute position, speed, acceleration and jerk,} \)
- Traction: Power Controls *
- Wheel slip and wheel slide
- Regenerative mode *
- Automatic application
- Passenger initiated application
- Adaptive mode
- Normal or seized

* to System Management Loop

Figure 25
Simplified Door Control Requirements

Figure 26
INNOVATION

Technical or otherwise with or without development potential.

Will it give a better Quality of Service?

Will it attract a greater share of the market?

% No Yes

Will it provide enough return on investment?

Return Yes No

Will it increase productivity?

% No Yes

Will it overcome staff shortages?

Staff No Yes

Is there surplus staff?

Surplus Yes No

Any better alternatives?

Yes (start again) No

to stay in business

HOPE CAPITAL WILL BE AVAILABLE

INDUSTRIAL RELATIONS

FORGET IT

MANPOWER RESOURCES AND INNOVATION Figure 27
The figure illustrates a system observable states, partitioned into subsystems.

- **System observable states**: $Z$
  - Subsystems: $S_j$  
  - Components: $Q_j$, $F_j$

- **Subsystem $S_j$**: $S_j = E(X_{ik})$
  - Components: $k$
  - Observables: $E_k$

- **Component ($k$)**
  - Observables: $X_{ik}$, $b_k$

- **Component ($\epsilon_k$)**
  - Observables: $X_{uk}$ = $\Psi(\epsilon_{rk})$

- **Component ($\epsilon_{rk}$)**
  - Observables: $1_k$, $T_k$, $g_k$

*depending on the size of the system*

**Figure 28 (Appendix I)**
Figure 29

Diagram a: Transition from system working to system failed.

Diagram b: Failure and repair paths between states 1 and 2.

Diagram c: Normal operating state transitioning to dangerous failure state via states 2 and 3.
Figure 30
Graph 1. Mean Time To Fail In T vs λT
$k_1$ Trackside equipment

$k_2$ Trainborne equipment

$Q_1$ and $Q_2$ correct states

$F_1$, $F_2$, $F_3$, $F_4$, safe failure states

$F_5$, $F_6$, $F_7$, $F_8$, $F_9$, $F_{10}$, states

$F_{11}$

$F_{12}$, $F_{13}$ dangerous failure states

$E_1$ Permanent magnet

$E_2$ Electromagnet

$E_3$ Cable

$E_4$ Signalling relay

$E_5$ Power supply

$E_6$ Magnetic receiver

$E_7$ Reset coil

$E_8$ EP valve

$E_9$ Indicator

$E_{10}$ Reset button

$E_{11}$ Logic Unit

$E_{12}$ Power supply

Relays (see §35)

Wiring

Connector

Figure 34
B. R. A. W. S. Logic

Figure 35