UTOMATION AND
ONTROL IN TRANSPORT

REVISED EDITION

. BARWELL
NSEA, UK

pendent Opinions of the First Edition

ellent general reading for anyone interested in transportation ology.'
RNAL OF DYNAMIC SYSTEMS, MEASUREMENT & CONTROL

great merit of the book is the fact that it exists at all. An attempt mbine the practice of road and rail with the laws of dynamics he theory of control is a landmark in itself.'
TRANSPORTATION

author's concise and illuminating explanations..'
MODERN RAILWAYS

first comprehensive treatment in bookform of rail and road le and traffic control systems.'
METRON

NTENTS: Dynamic systems. Route capacity – laws for vehicle ving. Control of vehicle spacing – railway signalling. Problems ngestion – traffic regulation. Computer aids to operation – surveillance and control. Measurement of power – analogue uting. Vehicle detection. Vehicle identification. Commun- n of control signals to moving vehicles. Interlocking – ence control. Sorting and marshalling. Control of acceleration power. Control of braking. Steering – directional stability. matic railways. Some modern rapid transit lines. Possibilities for uture. Appendices: control theory – guide to literature. The SI Index.

BARWELL AUTOMATION AND CONTROL IN TRANSPORT 2nd REVISED EDITION Perga

F. T. BARWELL

AUTOMATION AND CONTROL IN TRANSPORT

2nd REVISED EDITION

Other Pergamon Titles on Interest

ANAND	Introduction to Control Systems
ELLOY & PIASCO	Classical and Modern Control with Worked Examples
HAIGHT	Transportation Research
MORRIS	Communication for Command and Control Systems
PATEL & MUNRO	Multivariable System Theory and Design
SINGH	Applied Industrial Control
STARKIE	Transportation Planning, Policy and Analysis
STRASZAK & TUCH	The Shinkansen: High-Speed Rail Network of Japan

Pergamon Related Journals
Free Specimen Copy Sent on Request

Automatica

Computers and Electrical Engineering

Computers and Industrial Engineering

Computers and Operations Research

Computers and Structures

International Journal of Engineering Science

Problems of Control and Information Theory

Solid State Electronics

Transportation Research

AUTOMATION AND CONTROL IN TRANSPORT

Second Revised Edition

By

F. T. BARWELL

Wh.Sch., B.Sc., Ph.D., F.I.Mech.E., F.I.E.E., F.I.C.E.
Swansea, U.K.

PERGAMON PRESS

OXFORD . NEW YORK . PARIS . FRANKFURT . TORONTO . SYDN

U.K.	Pergamon Press Ltd., Headington Hill Hall, Oxford OX3 0BW, England
U.S.A.	Pergamon Press Inc., Maxwell House, Fairview Park, Elmsford, New York 10523, U.S.A.
CANADA	Pergamon Press Canada Ltd., Suite 104, 150 Consumers Rd., Willowdale, Ontario M2J 1P9, Canada
AUSTRALIA	Pergamon Press (Aust.) Pty. Ltd., P.O. Box 544, Potts Point, N.S.W. 2011, Australia
FRANCE	Pergamon Press SARL, 24 rue des Ecoles, 75240 Paris, Cedex 05, France
FEDERAL REPUBLIC OF GERMANY	Pergamon Press GmbH, Hammerweg 6, D-6242 Kronberg-Taunus, Federal Republic of Germany

First Edition 1973
Second Revised Edition 1983

Library of Congress Cataloguing in Publication Data

Barwell, F. T. (Frederick Thomas), 1911—
Automation and control in transport.
Includes indexes.
1. Transportation—Automation. 2. Railroads—Automation. I. Title.
TA1230.B37 1983 629.04'9 82-18981

British Library Cataloguing in Publication Data

Barwell, F. T.
Automation and control in transport.—2nd ed.
1. Automatic control 2. Traffic engineering
I. Title
629.04 HE333

ISBN 0-08-026712-2

Printed and bound in Great Britain by

Foreword

THE march of science during this century has been no less daunting in the field of Tr
port than any other and those of us who grew up in the earlier days of rail transport,
instance, do not cease to be staggered by the changes in recent years that are now regar
as commonplace.

Inevitably the speed of thought has outstripped the capacity for action. This
continues so that new methods have to be sought to maintain this momentum of deve
ment. No one today speaks in terms of individual forms of locomotion. It is all Trans
with a capital T, which must be considered in its broadest sense for the movemen
large numbers of passengers and/or quantities of goods from area to area by the
expedient methods.

The introduction of new techniques has called for a new terminology from whi
new language appears to be emerging. Such language is more difficult for the
generations to acquire, who have never heard of transport "modes", but no doubt it
be readily absorbed and used by the new computer-minded transport men of the pre
and future systems of mass movement as they come to grips with the new media.

An assembly of the appropriate developments in each of the respective field
transport is accordingly most timely and as a history of their evolution Professor Barv
following account makes fascinating reading.

Here one has to examine the qualifications of the author as to his right to record
correctly assess the available facts.

He came to me in the 1950s, from the Mechanical Engineering Research Laborato
the Department of Scientific and Industrial Research, when I was looking for assist
in future electrical research in the field of rail traction where many problems were
encountered on the newly proposed 25-kV, single-phase, 50-cycle electrificatio
British Railways.

He became responsible for producing the first comprehensive programme of elec
research in traction for British Railways which has now been taken over by the Rese
Centre at Derby. Therefore he can rightly be regarded as a proud father of projects
have resulted in further healthy offsprings.

In the process of this work he also acquired considerable kudos in the internat
railway field by means of his joint collaborations with them as will be evident fro
numerous references given, and of course he is a world authority on a number of
transport subjects.

Having posed the problems it becomes essential to express them mathematicall
herein this book fulfils its most important function, for the student must never accep
statement on its face value. The mathematical treatment given to all the presen
future forseeable modes of transport appears formidable but it is an essential require
for the researcher. Here formulae are given in abundance, for all the proven as well

"untried" systems of transport which must make this book a standard work on the subject. One need question Professor Barwell's qualifications no further.

I have long been an advocate of automation in any form and there are many other reasons which are perhaps outside the scope of this book that increase the justification for its application. The demand for it is bound to increase throughout the world where the resources can be found to pay for its installation. Electricity is the master agent and Professor Barwell's ability to produce the distilled essence of his knowledge into book form makes an admirable catalyst for the future expansion of automation in Transport.

In these days when the whole world can watch men being transported to the moon and back by automation, nothing is impossible. Such achievements have established automation and control in Transport as one of the most important present-day activities.

From what Professor Barwell writes it is evident that much has been done, but it is also crystal clear that much remains to be done if we are to cope with future world problems of heavily populated countries.

So long as we have men with the knowledge and foresight of Professors Barwell and Laithwaite, preferably located in the appropriate Chairs of Engineering in our universities, to provide the necessary stepping-stones to the future, we need have no fear of the contribution Engineering can make in the advancement of this particular science.

S. B. WARDER
F.I.E.E., F.I.Mech.E., Consultant,
Formerly Chief Electrical Engineer,
British Transport Commission

It was very sad that Mr. Warder died shortly after writing this Foreword. He had exerted a profound influence in the field of Railway Electrical Engineering and will be greatly missed.

Preface to the Second Edition

DURING the time which has elapsed since the publication of the first edition of this l much that was then embrionic has now become standard practice. Developmen transport technology have been paralleled by the introduction of new industrial prod notably the microprocessor, which have extended the range of possible further adv

As in the previous edition, rail technology has featured more prominently in the than road. This can partly be explained by the relative ease with which automatic de can be applied to guided or captive vehicles and partly because of the rapid grow urban transit rail systems in centres as far removed as Mexico City and Hong K There have also been major developments in inter-city transport such as the introdu of the 'High-Speed" and "Advanced Passenger" trains (HST and APT) in the U Kingdom, the Train à Grande Vitesse (TGV) in France and the magnetically-levi experimental vehicles in Japan.

There may be some change in the readership since the publication of the first edi At that time a number of the author's colleagues were engineers who had received training on steam or direct-current electric railways and who were then engaged o introduction of high-voltage a.c. electrification associated with modern signalling. It thought that "control theory" would provide a unifying discipline which would these men to co-operate on their new tasks. Thus, Appendix One was provided, not treatment in itself, but as a source book from which guidance could be obtained re ing the selection of books which would be of greatest value to any particular reade

At the present time a number of well-qualified graduates are joining the industry w engineering training has been general and academic and who may be unfamiliar the problems and indeed the terminology of transport in general or of railwa particular. I hope that this book will provide them with a useful introduction t industry.

As systems grow the opportunities for interaction between the different mode increased. Thus, in a major city, real time optimisation of traffic flow by computer trol of traffic-signals may be coupled to bus and train operation so that the com transport system becomes demand responsive.

The writer has travelled on all the novel transport systems mentioned and, i majority of cases, has discussed the control and maintenance aspects with the respon engineers. He would like to record his thanks to these persons for their invariable wi ness to devote their time to such discussion. He would also like to thank the Leverh

Trust for the award of an Emeritus Fellowship and his colleagues at the University College of Swansea, notably Messrs. M. J. Clarke and D. J. Leech and Dr. D. J. Oborne, for their continued co-operation.

1982 F. T. BARWELL

Preface to the First Edition

THERE can be few people today who, in some form or another, do not experienc growing problems of congestion in transport situations, problems for which no solution is in sight. Experience, particularly in the U.S.A., has shown that the d approach, extensive construction of new roads, is rapidly overtaken by increased dem This presents engineers with the challenge of examining all technical means at pr available for easing the flow of passengers and goods from place to place as we innovating in areas where existing technology cannot provide adequate solutions.

The problem is so immense that reliance can no longer be placed on the developme any single technique such as road or rail, but all methods must be studied and integ into a complete transport system. Individual methods are referred to by transport m "modes" and the critical requisite for success is the ease with which transfer can place between one mode and another.

Given adequate solutions to problems of inter-modal transfer, each mode ca studied with the object of evaluating and optimising its contribution to the o problem. The principal factor limiting the capacity of any transport mode is the degr control which can be exercised over the individual vehicles and the extent to which control can be made automatic, and the purpose of this text is to examine, fron standpoint of elementary control theory, the requirements of ground transport sys and the related technical devices so far developed. The study of transport as a contr system can provide a common theme in the education of men preparing to ente various disciplines required in the design, construction and operation of tran systems.

Some explanation may be needed for the prominence afforded to railway techn in this text. Justification for this is twofold; firstly, the continuing importance o contribution made by this mode to the conveyance of high traffic flows and secondly to its long history, control principles have been developed and applied to a much gr extent than is the case with modes of more recent origin. Such principles are, how fundamental so that the study of their application to railways in the past may he point the way to their application in an updated technological guise to the soluti new problems.

The mathematical demands placed on the reader have been restricted to those c sponding to the "A" level examination, but some results of more advanced control tl have been given in an appendix together with references to more specialised texts.

The S.I. system of units (Système International d'Unités) has been used throughout because of the simplification of calculation which arises from the use of a consistent system of units. Conversion tables are included for the use of those who prefer to work in the traditional British system.

The author wishes to acknowledge permission, given by the following bodies, to reproduce material from earlier contributions—The Institutions of Mechanical, Electrical and Railway Signal Engineers, The International Railway Congress Association, the *Journal of Science and Technology* and the Organisation for Economic Co-operation and Development.

Finally he wishes to thank his erstwhile colleagues on the Staff of the former British Transport Commission for much helpful discussion.

F. T. BARWELL

Contents

NOTATION

1. **Dynamic Systems**

1.1 System concepts
1.2 Open-loop systems
1.3 "Feedback"—closed-loop systems
1.4 The phase-plane diagram

2. **Route Capacity—Laws for Vehicle Following**

2.1 Lane capacity
2.2 Car-following theory
2.3 Automatic vehicle systems

3. **Control of Vehicle Spacing—Railway Signalling**

3.1 Necessity for signalling on railways
3.2 The "train-order" system
3.3 The "block" system
3.4 Axle-counting—radio link
3.5 Single-line working
3.6 Lock and block
3.7 Multiple-aspect signalling

4. **Problems of Congestion—Traffic Regulation**

4.1 Random events
4.2 Queues—Poisson arrivals—constant service times
4.3 Exponential service times
4.4 Effect of delays on headway of signalled systems
4.5 Statistical aspects of car-following behaviour
4.6 Traffic waves

5. **Computer Aids to Operation—Traffic Surveillance and Control**

5.1 Application of digital computers
5.2 Continuous progress control (C.P.C.), Dynamic programming
5.3 Total operations processing system (TOPS)
5.4 Optimum train sequence
5.5 Junction optimisation technique (JOT)
5.6 Traffic surveillance and control
5.7 Split cycle and offset optimisation technique (SCOOT)
5.8 Control of cascaded vehicles

5.9 System flow charts 69
5.10 Train describers 69
5.11 Microprocessors 69
5.12 Simulation of train-following behaviour 71

6. **Measurement of Power—Analogue Computing** 75

6.1 Mechanical manipulation of data 75
6.2 Equivalence of mechanical and electrical quantities 78
6.3 Potential and flow 78
6.4 Operational amplifiers 78
6.5 Vehicle suspension analogy 80
6.6 Application to strings of vehicles 83
6.7 Hybrid computers 85
6.8 Simulation of service environment 86

7. **Vehicle Detection** 88

7.1 Presence detectors 88
7.2 Track circuits 88
7.3 Jointless track circuits 90
7.4 Guided radar 91
7.5 Obstacle detection by radar 95
7.6 Ultra-sonic train detection 96
7.7 Two-way communications at ratio frequency 96

8. **Vehicle Identification** 98

8.1 The "Identra" system 98
8.2 Bus electronic scanning indicator 98
8.3 Automatic wagon-recording system (A.W.R.S.) 101

9. **Communication of Control Signals to Moving Vehicles** 104

9.1 Automatic warning systems—cab signalling 104
9.2 The train stop 107
9.3 The "Indusi" system 107
9.4 The "Signum" system 108
9.5 Beacon devices 109
9.6 Coded track circuit 110
9.7 Professor Poupé's system of coded track circuits 113
9.8 Use of continuous conductors in the track 114
9.9 Combination of magnetic and inductive loop systems 117

10. **Interlocking—Sequence Control** 120

10.1 Mechanical interlocking 120
10.2 Boolean algebra 122
10.3 Control of intersections—automatic half barriers 125

11. **Sorting and Marshalling** 128

11.1 The hump yard 128
11.2 Automatic retarders 129

11.3 Hydraulic retarders
11.4 Control of wagon movement using rope-hauled mules
11.5 Linear motor propulsion along classification tracks
11.6 Simulation of throughput of an automated flat yard

12. **Control of Acceleration and Power**

12.1 Limitations
12.2 Equations of motion
12.3 Values of resistance coefficient
12.4 Estimation of distance–time relationships
12.5 Coasting
12.6 Control of engine speed
12.7 Automatic transmissions
12.8 Control of electric motive power
12.9 Use of transducers in power control
12.10 Thyristor control of a.c. locomotives
12.11 Chopper control
12.12 Application of induction motors
12.13 Wheel-slip control—torque reaction
12.14 Control of diesel power
12.15 Adaptive control
12.16 Shock factors in acceleration

13. **Control of Braking**

13.1 Forms of braking
13.2 Magnetic track and electrodynamic braking
13.3 Physiological aspects
13.4 Control of slip (slow up)
13.5 "On-tread" braking
13.6 "Off-tread" braking
13.7 Servo actuation
13.8 The compressed-air brake
13.9 The electropneumatic (E.P.) brake
13.10 The vacuum brake

14. **Steering—Directional Stability**

14.1 Steering
14.2 Directional stability
14.3 Hertzian contact
14.4 Running of coned wheels
14.5 Inscription within sharp curves—steering by flanges
14.6 Vehicle ride quality
14.7 Motion on curves at speed
14.8 Radial-type bogies (trucks)
14.9 Effect of oscillation on passengers

15. **Automatic Railways**

15.1 General principles and early history
15.2 Analysis of human contribution under the present system
15.3 Possible systems for intensively used passenger lines
15.4 Existing installations
15.5 Systems for high-speed working

15.6 Moving block systems 250
15.7 The relative cost of different systems of signalling and control 251

16. Some Modern Rapid Transit Lines 253

16.1 Need for guided transport 253
16.2 Recently constructed underground railways 253
16.3 Intermediate systems 256
16.4 Light electric vehicles—Tyne and Wear system 256
16.5 Pneumatic tyred systems 257
16.6 Innovative—fully automated systems 264

17. Possibilities for the Future 284

17.1 The need for development of new transport modes 284
17.2 Possible improvements in control on the highway 286
17.3 High capacity systems—effect of station stops 290
17.4 Small vehicles under computer control (CVS) 293
17.5 Inter-city transport—high-speed rail 298
17.6 Studies of new forms of urban and inter-city transport 309
17.7 Air-cushion support 314
17.8 Linear induction motors 315
17.9 Linear synchronous motors 336
17.10 Electro-magnetic support and guidance (attraction mode) 337
17.11 Magnetic support and guidance—the repulsion (electrodynamic) system 346
17.12 Unexplored possibilities 348

Appendix I 350
Control Theory—Guide to Literature 350

1. Equivalence of dynamic and active systems 350
2. Literature available 355
3. The Laplace transform–transfer function 356
4. Effect of time delay 360
5. Criteria for stability 360
6. Root locus plots 361
7. Nyquist and Bode representation 364
8. Common forms of non-linearity 368
9. The describing function 370
10. Inverse Nyquist or Whitely diagram 373
11. Liapunov's second method 373
12. Random inputs 374
13. On-line control—sampled data 375

Appendix II 377
The S.I. Units (Système International d'Unités) 377

NAME INDEX 381
SUBJECT INDEX 385

Notation

A	frontal area of vehicle.
a	acceleration or deceleration, length of semi-axis of ellipse of contact, constant as defined in text.
B	magnetic flux.
b	gauge/2 defined as half distance between contact ellipses, constant as defined in text, length of semi-axis of ellipse of contact.
C	capacitance (electrical), concentration of vehicles, capacity of route.
C_D	drag coefficient.
c	distance of c.g. from leading axle.
d	distance of trailing axle from c.g., density of traffic.
D	differential operator, dimension.
E	modulus of elasticity.
F	force.
f	frequency, acceleration.
$\not{f}$	velocity distribution in a string of vehicles, function defined in text.
G	shear modulus (of rigidity), transfer function.
g	gravitational acceleration, effective air gap.
H	magnetic field strength, power, transfer function.
h	height, headway in seconds.
I	polar moment of inertia.
$\mathcal{J}$	current (steady value).
i	current (instantaneous value).
K	constant as defined in text.
k	stiffness, constant.
L	inductance.
l	length.
M	mass.
N	$\dfrac{\text{spacing}}{\text{jam spacing}}$.
$\mathcal{N}$	ratio defined in text, control or "forward path" transfer operator.
n	coefficient of viscous damping. any number.
o	length of overlap.

P normal load per unit breadth, pull, probability, pressure.
p pressure, length of periphery of cross-section.
q spring constant,
pneumatic trail.
R resistance (electrical), also tractive resistance, constant defined in text,
radius of curve (of track).
r radius (e.g. of wheel), constant defined in text.
S shape factor, transfer operator.
s distance,
complex variable (in transfer function).
T torque, time lag.
t time (elapsed).
V potential, electromotive force (steady value), velocity.
v voltage (instantaneous value).
W applied load, work done.
X axis as defined in text.
x variable—usually linear displacement,
input to servo system,
most probable time.
Y axis as defined in text.
y dependent variable,
output from servo system,
most optimistic time.
Z axis as defined in text, ratio.
Z Z transform.
z most pessimistic estimate,
lateral displacement.
α $\sin^{-1} W/F$, i.e. gradient expressed as angle subtended with horizontal in radians.
β coefficient defined in text.
γ angle of conicity of tyres.
ϵ base of Napierian logarithms,
error.
η coefficient—usually velocity dependent, viscosity.
θ slip angle of tyre,
deviation.
λ rate of arrivals expected or average numbers of occurrences,
wave length, slope of trajectory.
μ coefficient of friction,
$1/\mu$ = time occupied by service,
μ_0 = permeability of free space.
ν Poisson's ratio.
ξ creep ratio.
ρ momentum, density.
ρ_0 surface resistivity.
τ time constant.
ϕ angle of superelevation.
ψ a function.
ω angular velocity, pulsatance.

CHAPTER 1

Dynamic Systems

1.1. System concepts

Being concerned with the movement of masses, either animate or inanimate, tran is subject to the Newtonian laws governing the action of masses in motion. The t tories assumed by these masses are required to have a defined starting and finishing and collisions between them must be avoided. This necessitates the communicati information in some form or another to the moving mass so that the rate and dire of its progression can be controlled at all times. The technology of transport, there embraces the provision of some form of vehicle or container to support the pay- means for accelerating, decelerating and overcoming any resistance to motion i desired direction, means for determining the rate of progress and direction in relati the desired transit and capability for avoiding undesirable interactions with other tra

When a number of elements, each having prescribed properties, is combined int aggregation intended to fulfil a certain purpose, this is known as a *system*. The elen of any system selected for study will often contain sub-elements which can themselv regarded as systems. A system can be isolated and treated as an element of a l system provided that its behaviour can be described adequately in terms of the beha observed externally to the system in response to stimuli arising from outside that sys Thus, what to one engineer is a system, to another is an element. For example, an a mobile engineer may be concerned with an automatic system which adjusts the mecl cal advantage between engine and driving wheels in accordance with the variatio speed and tractive effort demanded. The traffic engineer, on the other hand, will re the automobile as a whole as an element. He is not particularly concerned with the m anical means whereby the driver's intentions are translated into tractive effort or bra force but only with the overall consequences of the driver's action.

In the study of transport, the different professional men will be concerned different levels of system. Each will be concerned, however, with at least two levels, i. those systems which can be regarded as the elements making up his system and the br er system within which his own efforts are intended to contribute. It is fortunate, how that the general types of behaviour, i.e. response to external stimuli of most system ments, can be described by relatively few laws and that the interaction between elen to determine system behaviour is similarly amenable to study from the genera viewpoint.[1, 2]

The dynamic behaviour of a system, i.e. how it responds to changing situations, is considered from the point of view of control engineering. It is convenient to descr system by means of block diagrams wherein each block, as shown in Fig. 1.1(a), represents an element the external behaviour of which is our concern whereas the technical means whereby this is achieved within the element, i.e. whether it is a true element or a sub-system, are matters of indifference. The system as presented by a succession of block diagrams may well constitute an element, a single block in a larger system to be considered by another engineer.

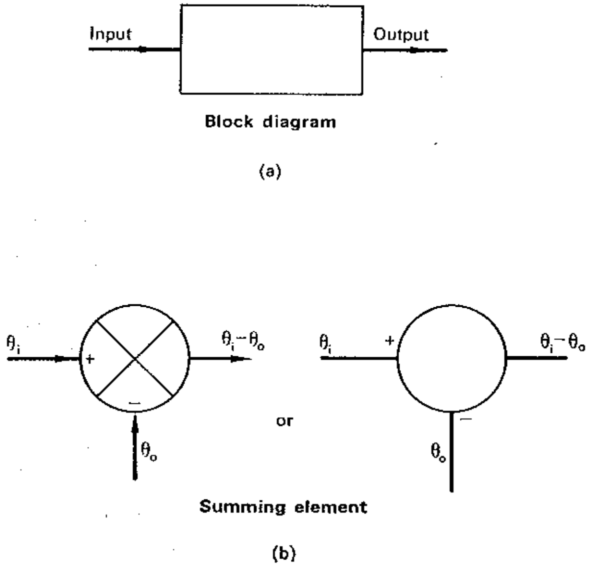


FIG. 1.1. Elements of block diagram.

The possibility of numerate treatment of transport systems arises from the fact that the behaviour of all continuous physical systems can be governed by a series of differential equations. Simple mechanical systems are governed by equations such as

$$M\frac{d^2x}{dt^2}+f(\dot{x})\frac{dx}{dt}+f(x)x = F(t), \qquad (1.1)$$

where x = linear displacement,

M = mass,

$f(\dot{x})$ = function governing variation of resistance to motion with velocity, e.g. aerodynamic drag,

$f(x)$ = function governing variation of force with displacement, e.g. the stiffness of a spring.

Thus the left-hand side of the equation describes the internal characteristics of the system.

$F(t)$, known as the forcing or driving function, describes the external force applied to the system expressed as a function of time.

Many practical systems comprise aggregations of masses and resistive elements, i.e. multiple degrees of freedom. The system will then be completely described by an array of

differential equations which can be solved simultaneously to define the response system to external stimulae.

Similarly all physical systems are governed by sets of differential equations of th

$$\mathcal{F}\left(a_0, \ldots a_n, x, \dot{x} \ldots \frac{d^n x}{dt^n}\right) = f(\text{u.v.}w..)$$

where the coefficients may take into account non-linear relationships as well as relationships and $f(\text{u.v.}w..)$ represents any physical disturbance including pl quantities such as temperature or voltage equally as well as mechanical force.

In the simpler systems, relationships may be governed by a few linear equatio solutions are easily obtained in analytical form. For more complex systems the equ may be non-linear, particularly including discontinuities, and the number of ele involved may be so large that only by the use of a computer can numerate val obtained. In a great many practical cases the precise mathematical solutions a obtainable without inordinate labour if at all, but nevertheless the study of a behaviour of elements and systems provides a powerful basis for the exercise of eng ing judgement.

1.2. Open-loop systems

Consider the simple system consisting of a man driving a motor car, Fig. 1.2(a) a road in direction "x" on which the speed limit is one of the constraints on the syste

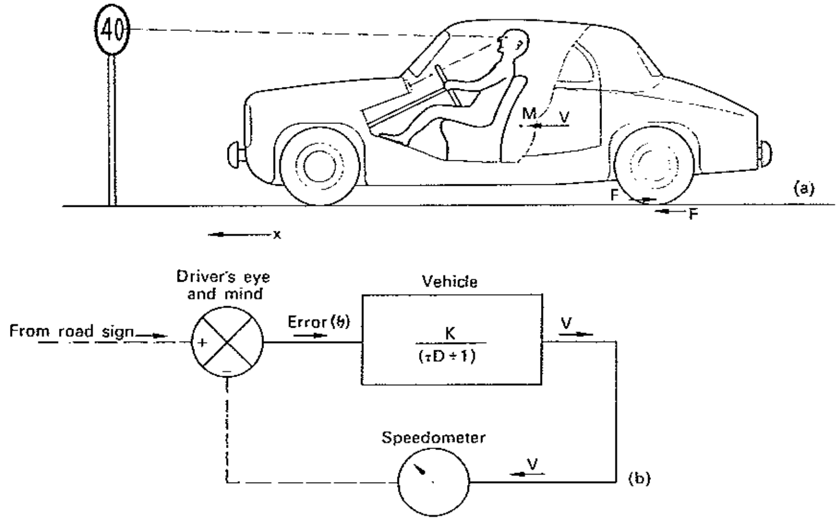


FIG. 1.2. "Feedback" in automobile driving.

passes a sign which indicates that the speed limit is increased from 30 to 40 m.p the next stretch of road. He proceeds to adjust his speed accordingly. From the d point of view the car is an *open-loop* system responding to the action of his foot accelerator. Let us determine the response of the system to this change in pedal pr thus regarding the accelerator as the input to the system. Assuming that the car is already in top gear and that the characteristics of the engine are such that the output torque is independent of the speed over the comparatively small variation imposed by these conditions, then the further depression of the accelerator pedal can be regarded as an immediate increase in torque from the pre-existing value to a new value. This torque acts on the mass of the car so as to produce an acceleration as well as to overcome any additional resistance to the motion of the car occasioned by the increasing speed. The increased pressure on the pistons arising from the action of the accelerator will be translated through the gear box and differential to a torque acting at the wheels, excepting that a certain amount of this torque will be required to accelerate engine components, shaft and transmission. For the present purpose, however, we will consider this to be lumped with the mass (M) of the car with which it is uniquely related for any gear ratio selected. We will consider the torque being represented by a force of tractive effort (F) acting at the wheel rim. η is taken as the coefficient governing the relationship between air plus rolling resistance and speed. It can be taken as being approximately linear over the range of interest. Then the equation becomes

$$M\frac{d^2x}{dt^2} + \eta\frac{dx}{dt} = F. \qquad (1.3)$$

We are particularly interested in speed so writing dx/dt as V and d^2x/dt^2 as DV where D denotes the operation of differentiation.

$$M\,DV + \eta V = F \qquad (1.4)$$

writing M/η as τ and dividing throughout by η

$$V(\tau D + 1) = F/\eta \qquad (1.5)$$

when $t \to \infty$, $V = F/\eta$ which we can describe as the "particular integral".

To describe the state of the system at any finite time this result must be combined with the "complementary function" which describes the rate of recovery of the system from any disturbance or its rate of adjustment to a new steady state.

Thus $$V(\tau D + 1) = 0. \qquad (1.6)$$

A solution taking the form $V = R\epsilon^{rt}$ may be selected so that $DV = rR\epsilon^{rt}$, substituting, in (1.6), $R\epsilon^{rt}(\tau r + 1) = 0$,

$$\therefore r = \frac{-1}{\tau}$$

and the solution is $V = R\epsilon^{-t/\tau}$.
Combining with (1.5) $V = F/\eta + R\epsilon^{-t/\tau}$.
If, when $t = 0$, $V = 0$, $0 = F/\eta + R$,

$$\therefore R = -F/\eta.$$

The expression for the speed of a vehicle accelerating from rest against a resistance varying linearly with speed is therefore

$$V = F/\eta\,(1 - \epsilon^{-t/\tau}). \qquad (1.7)$$

The expression $(1 - \epsilon^{-t/\tau})$ is known to control engineers as the *exponential lag* be of the presence of the term $\epsilon^{-t/\tau}$ and because it describes the lag of the response system to a sudden change in condition. The expression does not provide an acc description of the acceleration of a car from rest to full speed because η varies in a complicated manner than assumed and because engine torque does not remain co with change in engine speed to say nothing of the intervention of gear changes. Ho it will be sufficiently accurate to deal with the range from 30 to 40 m.p.h. (13·4 to 17· posed in our example.

Therefore at time $t < 0$
speed is constant at 13·4 and $DV = 0$,

$$\therefore \frac{F}{\eta} = 13{\cdot}4$$

and when $t > 0$

$$V(\tau D + 1) = \frac{F + \Delta F}{\eta}$$

or $$V = 13{\cdot}4 + 4{\cdot}5\,(1 - \epsilon^{-t/\tau}).$$

If the driver is skilful (and patient) enough to make an immediate adjustment whic enable his car to attain the newly demanded speed with a single adjustment of hi position, the new value of F will be $17{\cdot}9\eta$ and the speed–time curve will be as sho Fig. 1.3. Values of τ and η for a large motor car can be taken as 75 to 150 and 10 in S.I. units.

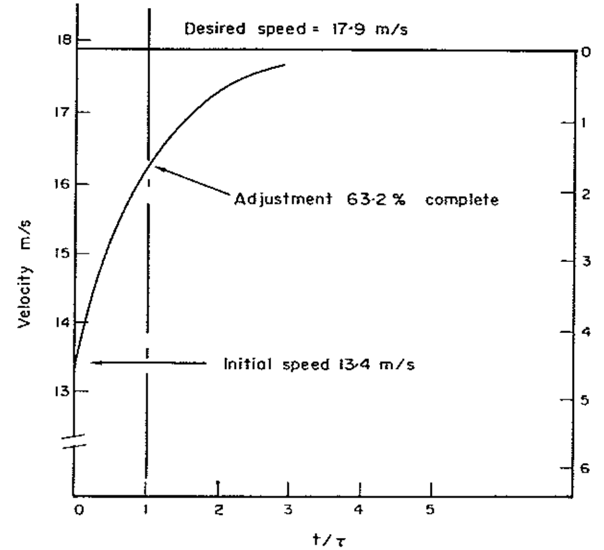


FIG. 1.3. Acceleration of car at constant torque.

The increase in speed will follow the exponential curve and will be 63·2% con when $t = \tau$. Although the curve is asymptotic to the desired value, the response c regarded for all practical purposes to be complete when t becomes equal to five ti

Thus the use of the term "time constant" to describe τ is justified. It can be regarded as a measure of the "sluggishness" of the system, being compounded in this case of mass and velocity dependent resistance to motion.

$$\frac{dV}{dt} = \frac{\Delta F}{\eta}\frac{1}{\tau}\epsilon^{-t/\tau}, \text{ when } t = 0, \epsilon^{-t/\tau} = 1,$$

$$\therefore \text{ initial slope } = \frac{1}{\tau} \times \frac{\Delta F.}{\eta} \qquad (1.9)$$

We can regard the accelerator pedal, engine, transmission, back axle, etc., as a second system relating angle of depression of the pedal with engine torque. We can endow further flexibility to the model by showing the region from the output of the engine shaft to the axle by a further system which enables such factors as the rotational inertia of the transmission components, action of gear changing and clutch slip to be incorporated into a more general study of the vehicle dynamics as in Fig. 1.4.

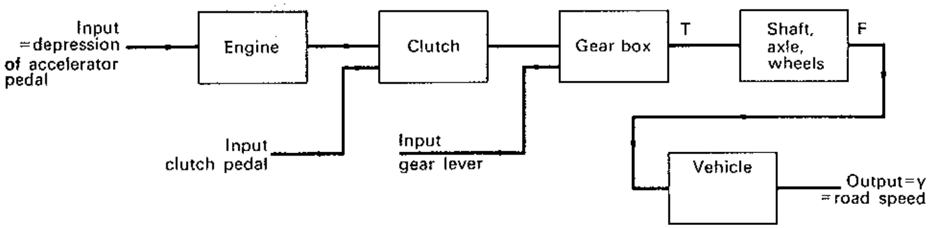


FIG. 1.4. Car control as open-loop system.

1.3. "Feedback"—Closed-loop systems

It will be seen that a single action on the controls will result in a definite and precise result which is determined by the characteristics of the system and environment. This is known as an open-loop system and an inexpert driver might provide either too much or too little torque to make the adjustment and would indeed soon become aware of his mistake and make a further adjustment. The means by which he becomes aware of his mistake must incorporate some means of measuring output, feeding an indicator of some description and a means of communication to the control system. In this case, of course, the speedometer provides the indication, and the eyes and brain of the driver interpret the information which is finally transmitted to the foot. This transmission of information about the controlled variable (in this case speed) to the controller (in this case the driver) is known as *feedback*. Figure 1.2(a) shows this diagrammatically and Fig. 1.2(b) is the corresponding block diagram. Information supplied to the driver comprises the desired speed from the roadside sign and the actual speed from the speedometer. The difference between these two is known as the "error" or "divergence" and this is what he sets out to minimise. It is represented diagrammatically by a "summing" element as in Fig. 1.1(b). The position of the accelerator pedal then becomes a function of the error. If we were designing an automatic control system such as is described in Chapter 2, the driver would be replaced by a system element which would convert the speed error into an accelerator displacement. This system could follow several laws, the simplest being proportional

control in which we can assume that the increase or decrease in engine torque is prc tional to the divergence between the desired and actual speeds.

Reverting to equation (1.5) and for simplicity combining drive and vehicle resist into transfer operator $\tau D + 1$

$$V(\tau D + 1) = \frac{F}{\eta}.$$

F/η now becomes a function of error so that writing

θ_i for input, in this case desired speed,
θ_o for output, in this case actual speed,
θ for error $= \theta_i - \theta_o$,

$$\frac{F}{\eta} = K\theta \quad \text{and} \quad V = \theta_o,$$

$$\theta_o = \frac{K\theta}{\tau D + 1} \quad \text{where} \quad \frac{K}{\tau D + 1}$$

is known as the "forward-path" transfer operator.

Writing $$\theta_o/\theta = \mathcal{N} \qquad ($$

$$\theta_o = (\theta_i - \theta_o)\mathcal{N},$$

$$\frac{1}{\mathcal{N}} = \frac{\theta_i - \theta_o}{\theta_o} = \frac{\theta_i}{\theta_o} - 1, \therefore \frac{\theta_i}{\theta_o} = \frac{1}{\mathcal{N}} + 1,$$

thus $$\theta_o/\theta_i = \frac{1}{1/\mathcal{N} + 1} \quad \text{or} \quad \frac{\mathcal{N}}{1 + \mathcal{N}}, \qquad ($$

thus $$\theta_o = \frac{1}{\frac{\tau D + 1}{K} + 1}\theta_i$$

or $$\frac{K}{\tau D + 1 + K}\theta_i.$$

If θ_i is a step input, as t tends to ∞

$$\theta_o \text{ tends to } \frac{K}{K + 1}\theta_i.$$

which is the particular integral.

For the complementary function $\theta_i = 0$

$$\therefore \theta_o(\tau D + 1 + K) = 0$$

and, by analogy with equation (1.7),

$$\theta_o = R\epsilon^{rt},$$
$$D\theta_o = rR\epsilon^{rt},$$
$$\therefore R\epsilon^{rt}(\tau r + 1 + K) = 0,$$

$$\therefore r = -\frac{1+K}{\tau}.$$

Thus
$$\theta_o = \theta_i \{1 - \epsilon^{-(t/\tau) \times (1+K)}\} \times \frac{K}{1+K}. \qquad (1.12)$$

The effect of simple feedback is therefore to increase the rate of response with some loss of accuracy. Without amplification, i.e. when K equals unity, the rate of response is doubled. Amplification produces further increases. In the example given, the effect of this will, of course, be restricted by the limitation in adhesion between the tyre and the road. The important feature of this equation is the value K.

Figure 1.5 shows the block diagram of the car system as a whole incorporating feedback. When control elements operate in series a little thought will show that the combined

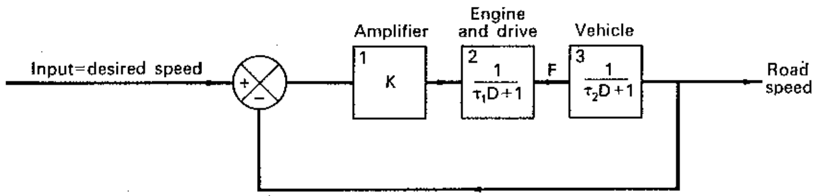


FIG. 1.5. Car control as closed-loop system.

output is the product of the two operations.

We can therefore write

Output of element (1) = input of element (2)
= Input of element (1) × K,
Output of element (2) = input of element (3)

$$= \text{Input of element (1)} \times K \times \frac{1}{(\tau_1 D + 1)},$$

$$\text{Output of element (3)} = \text{input of element (3)} \times \frac{1}{(\tau_2 D + 1)}$$

$$= \text{Input of element (1)} \times K \times \frac{1}{(\tau_1 D + 1) \times (\tau_2 D + 1)}.$$

We can therefore show the control and vehicle mass and drag as the elements of a system shown as a rectangle in Fig. 1.6 in which the product of the two elements is inserted. Such an expression is known as the *Transfer Operator*. From the point of view of the response of the system as a whole, variation of the amplification factor has precisely the same effect as variation of any other linear components such as mass or resistive drag coefficient in so far as they can be lumped together in an overall time constant τ. Comparing

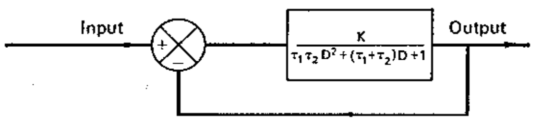


FIG. 1.6. Combination of transfer operators in series.

the feedback approach to the open-loop approach, it would be a very skilful driver in who could, with a single movement of his foot, produce a discrete increase in the t of the engine so as precisely to follow Fig. 1.3. It is more likely that a very skilful would have adapted himself to give an element of proportional control with a high of amplification so as to reduce the time constant of the combined vehicle–driver s to a minimum.

1.4. The phase-plane diagram

A convenient way of representing the response of the control system is to plot "r change" of error vertically against error horizontally. This is known as the "phase-p diagram.[3] Figure 1.7 reproduces the information of Fig. 1.3 in this form.

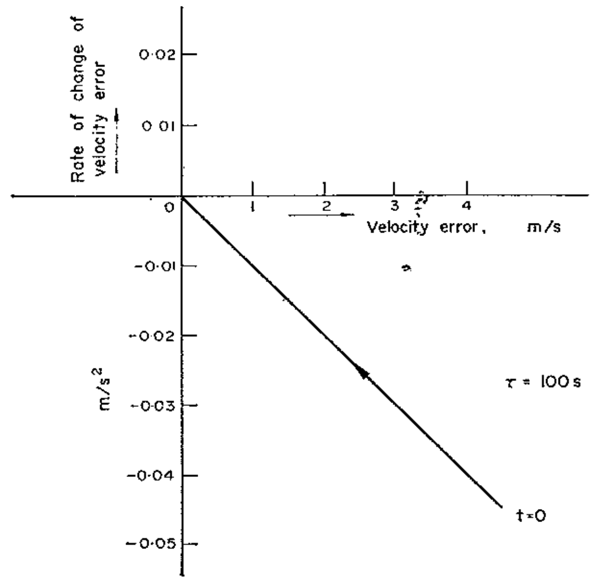


FIG. 1.7. Phase-plane representation of data of Fig. 1.3.

Reverting to the open-loop case, supposing a driver, perhaps a learner, faced wi desire to increase speed, depresses the accelerator until the desired speed is rea releases it completely until speed has fallen to a noticeable extent and then depre again. This would correspond to what is known as "*on/off*" *or* "*bang bang*" co commonly used in domestic heating installations and other thermostatic devices. T output would be limited only by the power of the engine and the time constant appl would be that applicable to the attainment of the maximum speed which the veh capable of obtaining rather than the desired speed. Figure 1.8 shows the phase- diagram for this case. Starting at A, speed will increase, thereby reducing the erro acceleration will diminish under the action of increasing drag until the desired sp reached. It would take some time, however, for the driver to appreciate that th happened so that there would be an overshoot to the point B then, on release accelerator, the car would slow down under the action of external and internal drag, notably that of the idling engine as indicated by line CD. When point D is reached, acceleration would be resumed. The desired condition would never be reached precisely but there would be a continuous oscillation of the actual condition around the origin of the phase-plane diagram.

If with the passage of time the extent of departure from this point diminishes, the system is said to be stable. Conditions sometimes apply where each oscillation involves a further

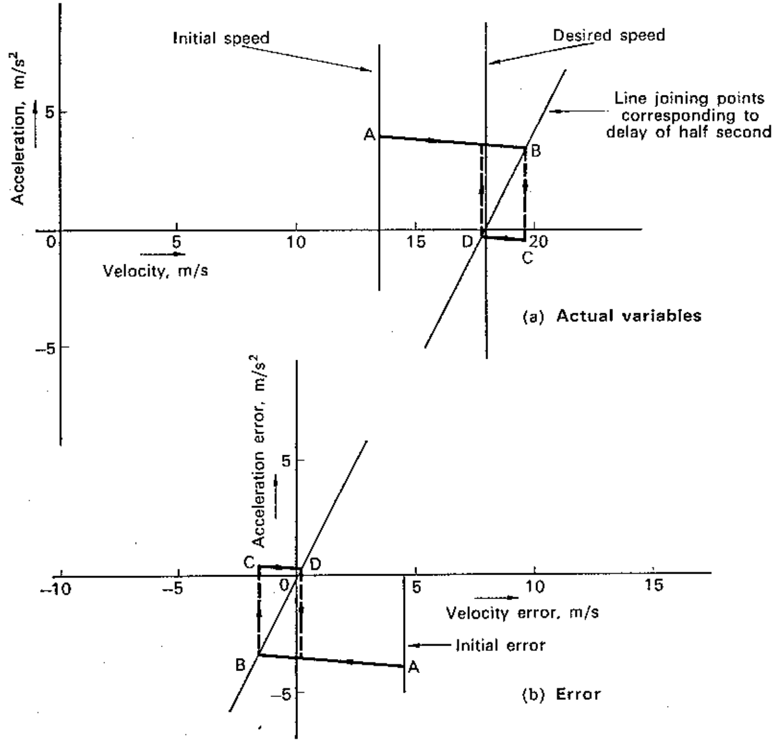


FIG. 1.8. Phase-plane diagram for "on–off" control showing limit cycle.

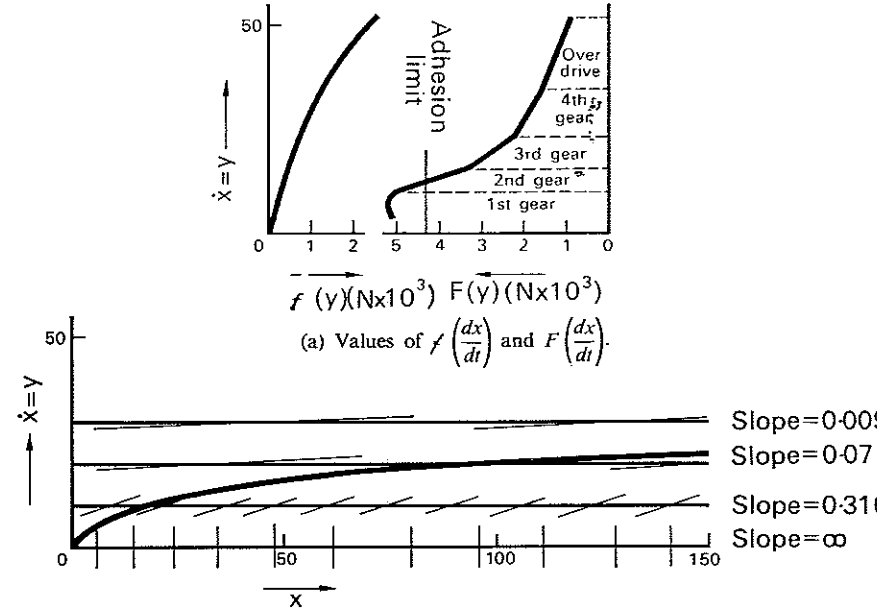
departure from the origin so that the error increases in magnitude with time. This is known as instability which is the first preoccupation of the designer and the main developments of control theory are intended to enable such conditions to be avoided.

In practice, non-linearities or limiting factors of some kind preclude the indefinite extension of the divergences so that the final situation is represented by a closed curve on the phase-plane diagram. This is known as a *limit cycle*. Limit cycles are further divided into two categories, soft and hard. In the case of all limit cycles, if a disturbance is applied to the system outside the bounds of the limit cycle the movements will in time tend to correspond to the limit cycle. In the case of a *soft* limit cycle this also applies to initial disturbances which are within its bounds, and which grow until the periodic oscillations attain the limiting value. In a *hard* limit cycle the initial disturbances represented by

co-ordinates of the phase-plane diagram which are bounded by the limit cycle will tend decay.

A driver would soon learn to moderate his action at least by reducing the intensity the increments or decrements of accelerator action and, possibly by continuous variati of foot pressure, introduce an element of proportional control. In this he would be act as an adaptive control system, i.e. one in which the transfer operators are themsel changed during the process in order to approach some optimum output.

The phase-plane diagram is of considerable use in introducing the effect of n linearities into a problem. These may take two forms. One is characterised by the dep dent variable not following a direct linear relationship with the independent variable; example, the variation of drag with velocity is more complex and cannot be represen over the complete speed range by the numerical coefficient so far used in this treatme The other case is where a certain discontinuous feature arises; for example, tractive eff can only be increased up to the point where the wheels begin to slip.



(a) Values of $f\left(\frac{dx}{dt}\right)$ and $F\left(\frac{dx}{dt}\right)$.

(b) Speed–distance curve showing isoclines.

FIG. 1.9. Phase-plane representation of speed–distance cycle.

Figure 1.9(b) shows a phase-plane diagram of the speed of a car during the accelerat period plotted against position. The equation governing motion is

$$M\frac{d^2x}{dt^2} + f\left(\frac{dx}{dt}\right) = F\left(\frac{dx}{dt}\right) \qquad (1.$$

where $f(dx/dt)$ and $F(dx/dt)$ are non-linear functions of velocity representing t vehicle drag and tractive effort at the wheel rims respectively as in Fig. 1.9(a).

Writing

$$\frac{dx}{dt} \text{ as } y \tag{1.14}$$

we have

$$M\dot{y} + f(y) = F(y),$$

$$\therefore \dot{y} = \frac{F(y) - f(y)}{M}, \tag{1.15}$$

$$\text{slope} = \frac{dy}{dx} = \frac{dy}{dt} \times \frac{dt}{dx} = \frac{\dot{y}}{y} = \lambda$$

$$= \frac{F(y) - f(y)}{My}. \tag{1.16}$$

Thus we have an expression for the slope of the speed–distance curve for every value of y and auxiliary curves can be drawn on the diagram connecting all points of equal slope. In this case, because acceleration is a function of y only, they represent horizontal straight lines. The actual trajectory constructed graphically from the isoclines will be approximate

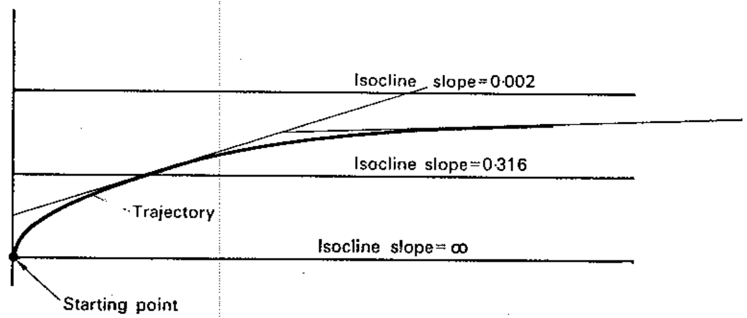


FIG. 1.10. Step-by-step construction of trajectories.

only but accuracy can be increased providing additional isoclines for intermediate values of λ. Figure 1.10 shows one method of step-by-step construction of a trajectory from isoclines as follows:

1. Draw a line through the starting point at slope corresponding to isocline.
2. Project line until it crosses the next isocline.
3. Bisect the segment between the two isoclines.
4. Draw a new line at the slope corresponding to second isocline through the mid point.
5. Project new line to third isocline and repeat the procedure.

An alternative construction, known as Liénard's construction, enables the slope at any point to be represented by the tangent to an arc of a circle constructed by compasses. The slope of the normal to the trajectory is represented by $-1/\lambda$ and in this case by

$$-\frac{My}{F(y) - f(y)}$$

or as line OB in Fig. 1.11(a).

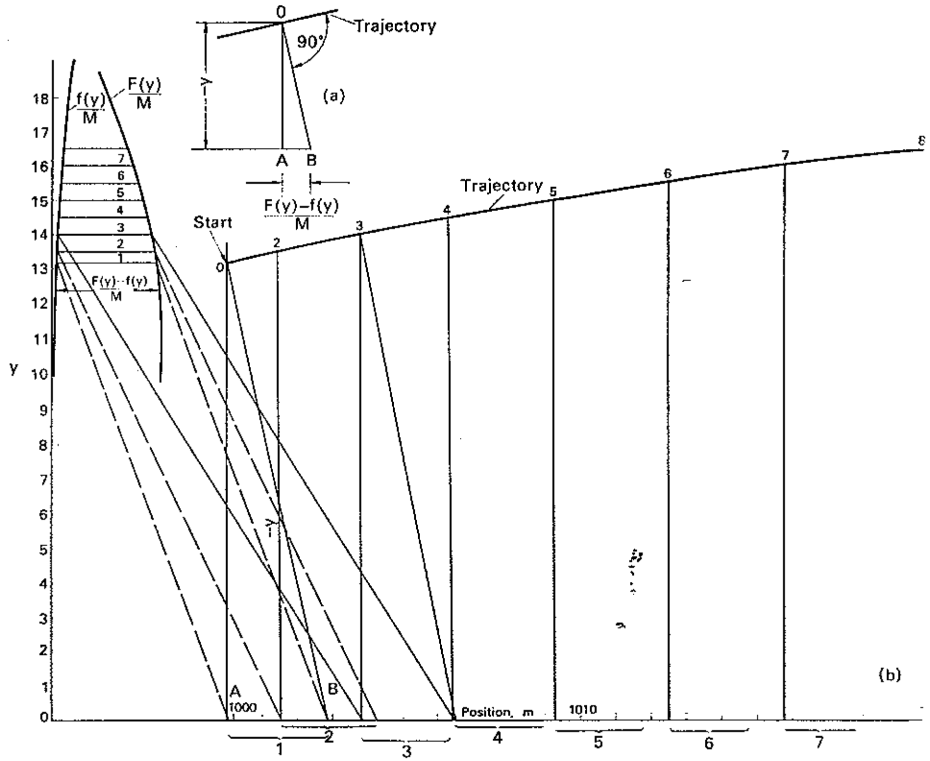


FIG. 1.11. Use of Liénard's construction.

(Acceleration of car from speed of 13·4 m/s)

Slope of normal to transectory $= \lambda = -\dfrac{M.y.}{F(y) - f(y)}$

where y = velocity $= \dfrac{dx.}{dt}$, $F(y)$ = tractive effort,

and $f(y)$ = tractive resistance.

Starting at O, construct triangle with vertical side AO equal to $-y$. Horizontal side AB is made equal to the v of $\dfrac{F(y) - f(y)}{M}$ for the particular value of y.

This can be obtained from curves $F(y)$ and $f(y)$ and transferred to the base line using dividers. An arc struck from cent and passing through O will form part of the trajectory.

Then select new point and repeat for new value of y. Repeat until trajectory complete.

NOTE: x does not appear in equation (1.13), therefore zero of x scale can be selected arbitrarily to indicate position of vehicle at any time.

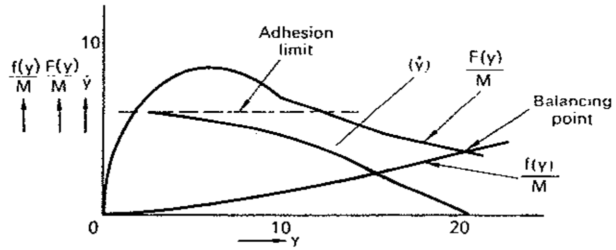


FIG. 1.12. Acceleration–speed curve.

An equally relevant plot shown in Fig. 1.12 represents equation (1.15), i.e. acceleration against velocity. The point where $F(y)$ and $f(y)$ are equal is known as the "balancing point" and corresponds to the maximum speed of the vehicle.

The effect of gradient is, of course, independent of speed and can be represented by an addition or subtraction of a constant quantity (positive for falling gradient and negative for rising gradient) to or from $F(y) - f(y)$. The point where gradient changes can be set out on the horizontal axis and the remainder of the trajectory can be constructed using the appropriate values of $F(y) - f(y) \pm Mg \sin \alpha$ where α is the inclination of the track to the horizontal. For practical purposes it is sufficiently accurate to take $\tan \alpha$, i.e. a gradient of one in thirty being defined as a vertical rise of one unit for every thirty units measured in a horizontal direction rather than along the inclined surface of the road itself.

Thus when $\alpha = 0{\cdot}0333$ (1·89 degrees), $\sin \alpha$ and $\tan \alpha$ are indistinguishable on four-figure tables.

Naturally if very steep gradients are involved more care is needed in definition of the method of measurement.

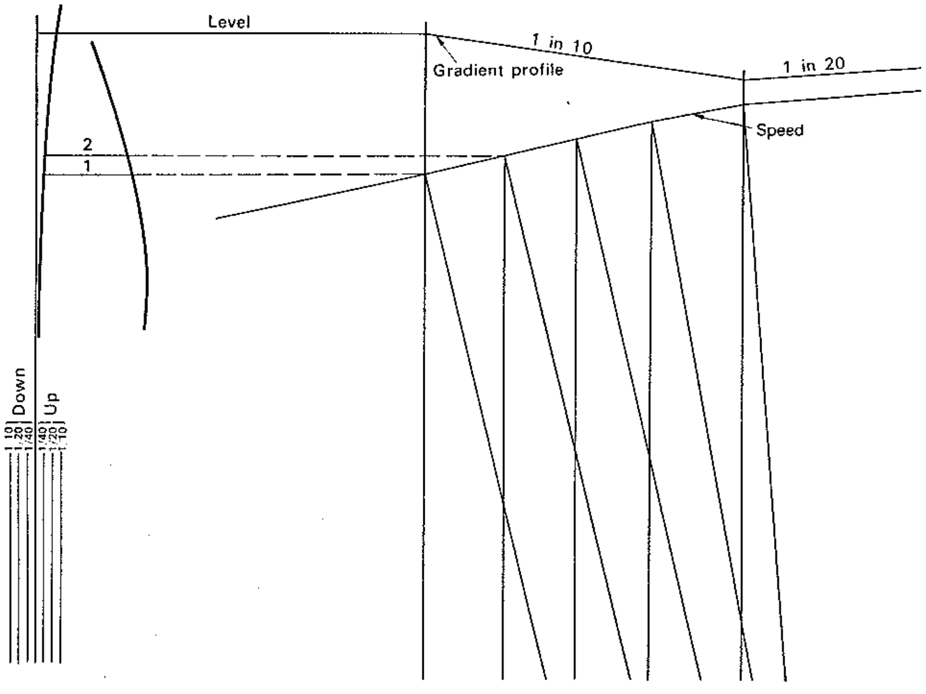


FIG. 1.13. Effect of gradients.
Diagram is divided into regions for different gradients. Scales for $g \sin \alpha$ are set up. Procedure is similar to that of Fig. 11 except that $g \sin \alpha$ is added to $\frac{F(y) - f(y)}{M}$ for falling gradient and subtracted for rising gradient.

Figure 1.13 shows the effect of gradients added to the conditions governing the construction of Fig. 1.11.

The phase-plane diagram does not reveal the effect of time implicitly. The trajectc however, represent the passage of time and, recalling $y = \Delta x/\Delta t$ for sufficiently s increments,

$$\frac{y}{\Delta x} = \frac{1}{\Delta t}. \qquad ($$

If therefore we assume a value for Δt, say 1 second, and construct an isosceles triang height equal to unity and base representing 1 second to a suitable scale, a similar tria will represent y vertically and Δx horizontally to a scale which is related to that w we have chosen to represent Δt. If, commencing at the value of x corresponding tc time at which we wish to commence this calculation, we draw a triangle between the sp distance curve and the abscissa then the height of the triangle will be y and the base represent the value of Δx corresponding to Δt. If we continue to divide up the whole in this manner as in Fig. 1.14 we can mark off the time occupied in accelerating to speed or to travel any given distance. Thus we have a speed–time plot against a non-li representation of t. This can be transformed to a speed–time curve by the simple const tion shown in Fig. 1.14(b) and to a distance–time curve as in Fig. 1.14(c).

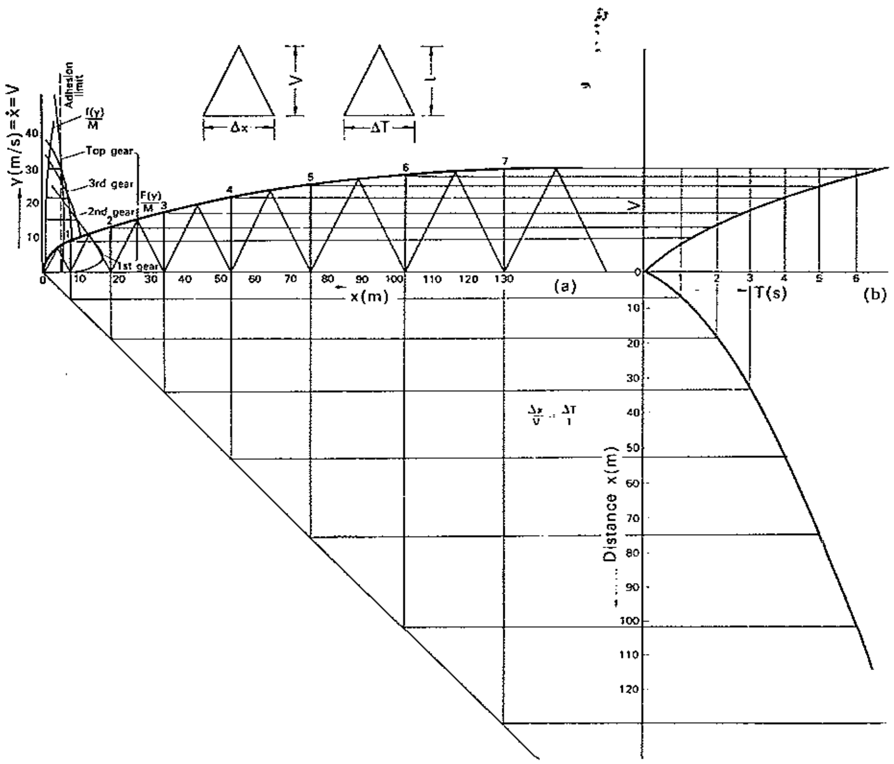


FIG. 1.14. Estimation of lapsed time from phase-plane diagram.

The concept of stability is crucial to any control problem. A system is "marginally stable" when an oscillatory variation, however caused, persists indefinitely without increasing or diminishing in magnitude. If it tends to decrease the system is stable and if amplitude increases with time the system is unstable. This is indicated on the phase-plane diagram by the shape of the trajectories and the direction in which they are traversed in time.

Some characteristic singularities met with in phase-plane diagrams are shown in Fig. 1.15. Figure 1.15(a) shows a "centre" which represents an oscillation of the simple harmonic motion type wherein amplitude neither increases nor decreases in successive cycles. The addition of damping provides a "stable node", Fig. 1.15(b), or a "stable focus", Fig. 1.15(c), according to whether the degree of damping is above or below the critical value. Thus a focus is characterised by overswing. Unstable conditions can also be represented by nodes and foci as in Fig. 1.15(d) and Fig. 1.15(e). A further phase

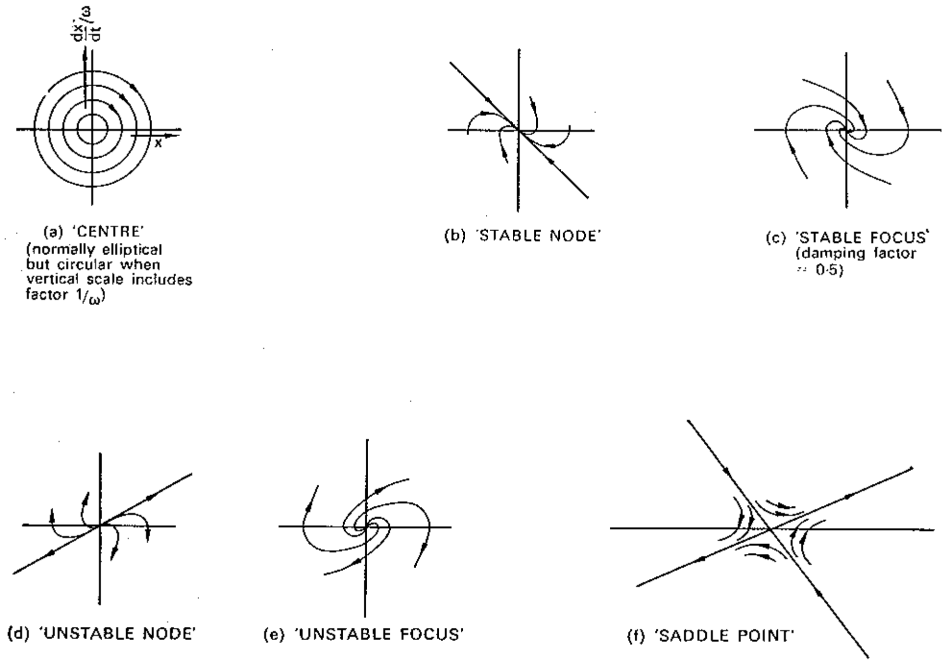


FIG. 1.15. Important singularities met with in phase-plane diagrams.

portrait known as the "saddle point", Fig. 1.15(f), is more difficult to describe. It represents the case where the force acting on a mass, for example, is directed away from the origin, increasing with distance therefrom. An illustrative example is provided by the phase portrait of a simple, rigid pendulum which is allowed to rotate above the support

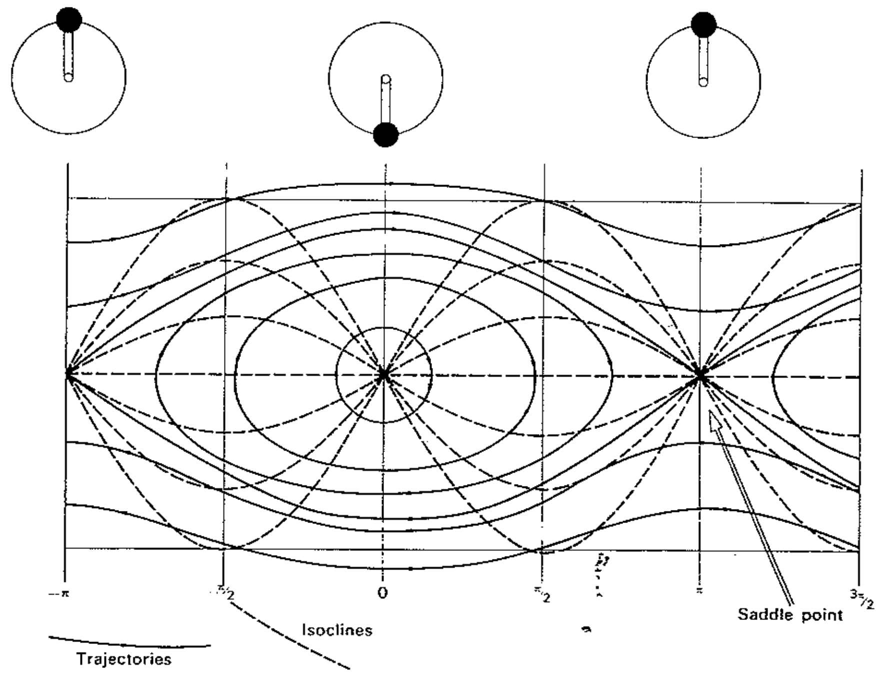


FIG. 1.16. Phase-plane diagram of rigid pendulum illustrating saddle point.

axis as well as below, Fig. 1.16. When the pendulum is hanging downwards and oscillatin with small amplitudes, its motion can be represented by a centre. The origin of the "sadd point" diagram represents the pendulum when it is motionless in a vertical positio directed upwards. The slightest motion in either direction will cause the pendulum to b subjected to forces tending to increase amplitude. Unless the velocity of approac diminishes to zero as the origin is reached, as represented by the straight line which slop downwards from left to right, the pendulum will continue to rotate in the same directio If it does, its further motion becomes indeterminate, being governed by either part of th straight line which slopes downwards from right to left. The trajectories which cross th abscissa represent the case where the energy of the pendulum is insufficient to carry it t the vertical so that it reverses direction before the point is reached. The trajectories whic cross the ordinates represent the higher energy cases where motion continues in the sam direction.

References

1. EDER, W. E. and GOSLING, W., *Mechanical System Design*, Pergamon, 1965.
2. NASLIN, P., *The Dynamics of Linear and Non-linear Systems*, Blackie, 1965.
3. WEST, J. C., *Servo Mechanisms*, English Universities Press, 1953.
For other texts on Control Theory, see Appendix.

CHAPTER 2

Route Capacity — Laws for Vehicle Following

2.1. Lane capacity

The capacity of a transport system may be defined as the product of the number of vehicles present in unit length of route and their average speed.

Thus

$$C = V \times d \tag{2.1}$$

where C = capacity of route in vehicles per unit time,

V = average speed,

and d = density of traffic in vehicles per unit distance.

Although in a practical system the possibility exists that faster vehicles may pass slower vehicles, this requires the occupation of previously unused space in the "fast lane". However, when we are concerned with maximum capacity, all lanes are assumed to be fully occupied and we can regard each vehicle constrained to a single lane. Its rate of progress is therefore determined by the vehicle in front and estimates of capacity may be made if we know the laws governing the average distance between vehicles as affected by their speed.

"*Headway*" is usually defined as the time interval between vehicles passing a given point, i.e. capacity of 1800 vehicles per hour (0·5 per second) corresponds to a headway of 2 seconds.

Thus

$$h = \frac{1}{V \times d} = \frac{1}{C} \tag{2.2}$$

where h = headway in seconds.

On the road, headway and therefore route capacity, is determined by the pattern of behaviour of drivers and by the mechanical characteristics (e.g. acceleration and braking rates) of their vehicles. The *Highway Code*, p. 6, states: "Never drive at such speed that you cannot pull up well within the distance you can see to be clear, particularly having regard to the weather and the state of the road." Diagrams are provided from which it can be calculated that the average deceleration has been assumed to be 21·5 ft/sec/sec (6.56 m/s/s) which corresponds to a coefficient of friction of 0·67. The maximum retarding force, neglecting aerodynamic resistance, is μMg, so that maximum deceleration attainable equals μg. Thus if

l = length of vehicle,

t = time of reaction of driver, sec (= 0·65, from the *Highway Code*),

a = rate of deceleration m/s² = μg when all axles are braked,

g = acceleration due to gravity (9·807 m/s²),

μ = coefficient of friction between wheel and track surfaces,

then braking distance = $V^2/2a$

thinking distance = Vt

and number of vehicles per unit length of highway

$$= \frac{1}{V^2/2a + Vt + l},$$

therefore route capacity in vehicles per unit/time

$$= \frac{V}{V^2/2a + Vt + l}.$$

Differentiating and equating to zero produces a value of V equal to $\sqrt{2la}$ for maxi capacity and by substitution that capacity is given by

$$\frac{1}{\sqrt{2l/a} + t} \quad \text{or} \quad \frac{\sqrt{\mu g}}{\sqrt{2l} + t\sqrt{\mu g}}$$

which shows the importance of "adhesion" or friction between wheel and track.

Assuming that the traffic is composed of private cars each 12 feet (3·66 m) long and d in accordance with the *Highway Code*, this works out at 2100 vehicles per hour per carr way. The optimum speed would be 15·5 m.p.h. (6·92 m/s). This relates to single veh If they are coupled into trains comprising "n" vehicles a general formula for capac

$$\frac{Vn}{V^2/2a + Vt + ln},$$

the optimum velocity is $\sqrt{2lan}$,

and the maximum capacity becomes $\dfrac{n}{\sqrt{2nl/a} + t}$.

This is the justification for the train and explains why it is more convenient technical achieve high rates of traffic flow at attractive speeds with captive vehicles marshalled trains rather than with individual vehicles freely driven.[1]

Comparing road and rail, it is usual (see the *Highway Code*) to assume that a coeffi of adhesion of 0·67 is applicable to pneumatic tyres on the roadway. British Railways curve for primitive freight trains is equivalent to a coefficient of adhesion of less one-tenth of this value, namely 0·05. Whilst this is not strictly comparing like with the former figure corresponding to good conditions and the latter to the worst condit for given vehicle assembly the road has the advantage of a factor of 13·4. To break at any given speed a train must have a length greater than thirteen road vehicles. It course very easy to provide this in any situation where route capacity is important.

2.2. Car-following theory

If we regard the driver's behaviour in greater detail we may suppose that he ado less cautious attitude than that implied by the *Highway Code*. For example, he may into account the fact that the vehicle in front will require some time (and distanc come to a stop. Therefore if he can brake as quickly as the driver in front, he will avoid a collision provided that the distance separating the two cars is sufficient to compensate for the reaction time required by the second driver to apply his brakes after the first driver has done so. This rather foolhardy approach ignores the possibility of other obstacles appearing suddenly, e.g. from other lanes. Thus multiple collisions occur with increasing frequency. Naturally different drivers will react differently to a given set of circumstances and it is not possible to provide a quantitative description of their action which can be generally applied. There are, however, certain constraints within which all drivers must operate and an attempt to quantify their behaviour in a manner which may have some significance on a statistical basis may be justified.

Thus, treating the driver as a control system, the input will be the position, speed and acceleration of the car which he is following and the output will be the behaviour of his own car. The variable which is at his disposal is the rate of acceleration or deceleration thereof.

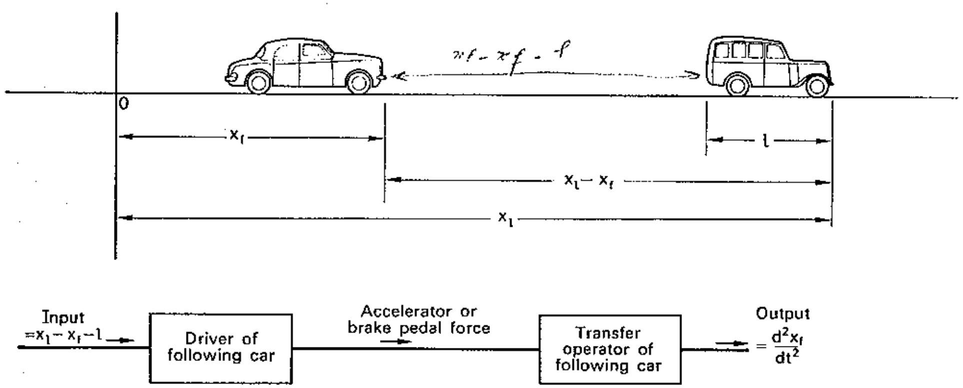


FIG. 2.1. Notation for car-following law.

Thus, as in Fig. 2.1, the input would be the driver's appreciation of the behaviour of the car in front. Suppose, for simplicity, he desires to catch up with the first car and to travel bumper to bumper with it at constant speed. This could be represented on a phase plane diagram as in Fig. 2.2 where displacement error, i.e. distance between the two cars, is plotted horizontally and velocity error, i.e. difference between their velocities, is plotted vertically.

Suppose that each car travelling within a single lane of roadway follows the one in front in accordance with a definite *stimulus response law*. Such a law could be expressed as follows

$$\ddot{x}_f(t+T) = \alpha\,[\dot{x}_l(t) - \dot{x}_f(t)] \tag{2.7}$$

where suffixes l and f refer to the leading and following cars respectively. Consider the acceleration of the following car as shown on the left-hand side of the equation. Its driver may be assumed to sense the difference in speed between his car and the previous one and to govern his acceleration accordingly. Thus, if he finds himself catching up with the

previous car he will slow down and if he falls behind he will speed up. He and to a les extent his vehicle, will embody a time lag T so that the response to a situation existing time t does not occur until $(t + T)$. The term α (having the dimension of time relates the intensity of the driver's action to other circumstances prevailing. Thus he v

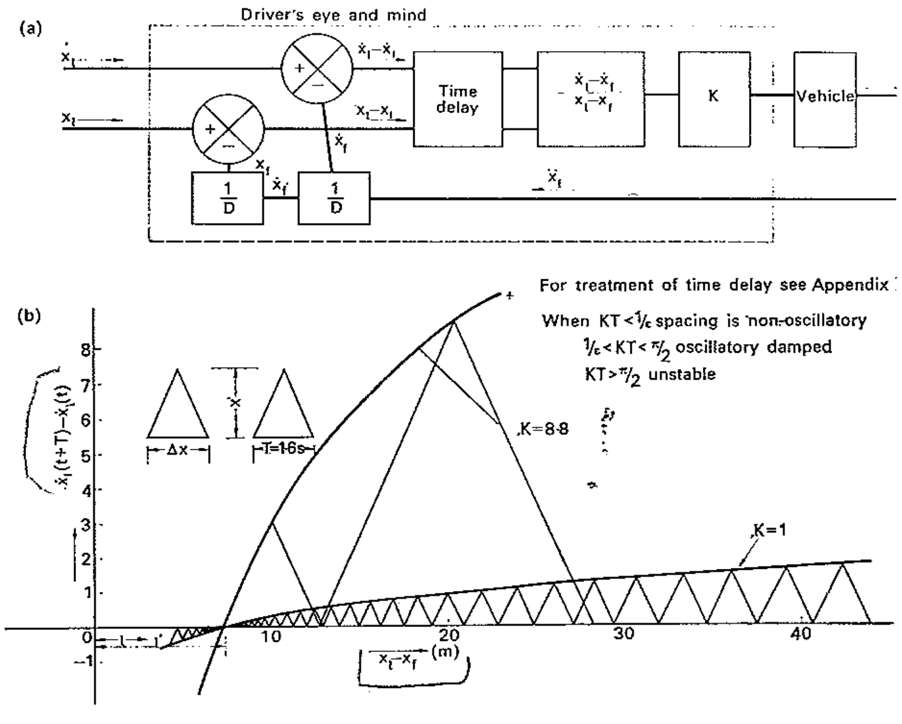


FIG. 2.2. Deceleration of following car according to Herman and Potts car-following law.

accelerate more readily if there is a long distance between cars than if they are cl together. Herman and Potts[2] suggest a reciprocal spacing law

$$\alpha(\rho) = \frac{K}{\rho} \qquad (\text{:}$$

where ρ is distance between cars and K is a characteristic speed.

Thus

$$\ddot{x}_f(t+T) = \frac{K\dot{x}_l(t) - \dot{x}_f(t)}{x_l(t) - x_f(t)}\,. \qquad (\text{:}$$

Integrating and suppressing (t) where meaning is obvious

$$\dot{x}_f(t+T) = K\int\frac{\dot{x}_l - \dot{x}_f}{x_l - x_f}\,dt$$
$$= K\,[\log_e(x_l - x_f) + C_1]. \qquad (2.$$

When the two cars are stationary and bumper to bumper

$$x_f - x_l = l \quad \text{and} \quad \dot{x}_f(t+T) = \dot{x}_l(t) = 0$$

$$\therefore C_1 = -\log_\epsilon l$$

$$\therefore \dot{x}_f(t+T) = K \log_\epsilon \frac{x_l - x_f}{l} \qquad (2.11)$$

where l is the effective length of each car.

They found that time lag T of 1·6 seconds and an average value for K of (19·6 m.p.h.) 8·8 m/s fitted their experimental results very well (correlation coefficient 0·97).

Of course, under conditions of sparse traffic, there will be gaps between cars sufficiently large for the second driver to act quite independently of his precursor. This has been estimated at a time interval of 6 seconds or a distance of over 60 metres.

One could well imagine a control system designed in accordance with this law but how does the individual driver learn to behave in this manner? Simulation of the situation on a computer at the General Motors Laboratory produced curves such as that shown in Fig. 2.3 for various values of K. The cars were initially spaced 100 feet (30 metres) apart. The leading car first accelerated at 16·1 ft/s/s (4·9 m/s/s) for 2 seconds and then decelerated for 2 seconds. It will be apparent that different values of K lead to different responses and that, below a certain value, a danger of instability exists. It is suggested that a driver learns by experience to select a value of K which leads to the most rapid response short of instability. Capacity in the steady state when $\dot{x}_f(t+T) = \dot{x}_l(t) = \dot{x}$ will be governed by

$$C = \dot{x}d = K \log_\epsilon \left[\frac{\text{spacing}}{\text{jam spacing}}\right] \times \frac{1}{\text{spacing}} \qquad (2.12)$$

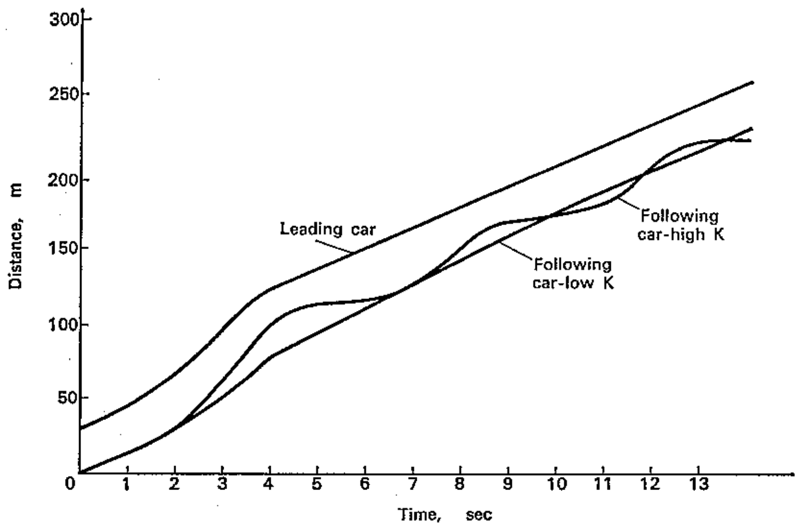


FIG. 2.3. Effect of value of K on stability.

where "jam spacing" represents the closest practical spacing of vehicles (l'). This naturally somewhat greater than l and Herman and Potts suggest a value of 7·3 met (220 vehicles per mile).

Let
$$N = \frac{\text{spacing}}{\text{jam spacing}}$$

then
$$C = K \log_\epsilon N \times \frac{1}{l'N}$$

and $\dfrac{dC}{dN} = \dfrac{K}{l'N}[1 - \log_\epsilon N] = 0$ for maximum and minimum values of C,

therefore
$$1 - \log_\epsilon N = 0,$$
$$\therefore \log_\epsilon N = 1,$$
$$\therefore N = \epsilon = 2{\cdot}7183$$

or
$$(x_l - x_f) = 19{\cdot}85 \text{ m.} \qquad (2.$$

Thus $C_{\max} = K/l'\epsilon = 0{\cdot}44$ vehicles per second (1580 vehicles/hour)

when
$$K = 8{\cdot}8 \text{ m/s.} \qquad (2.$$

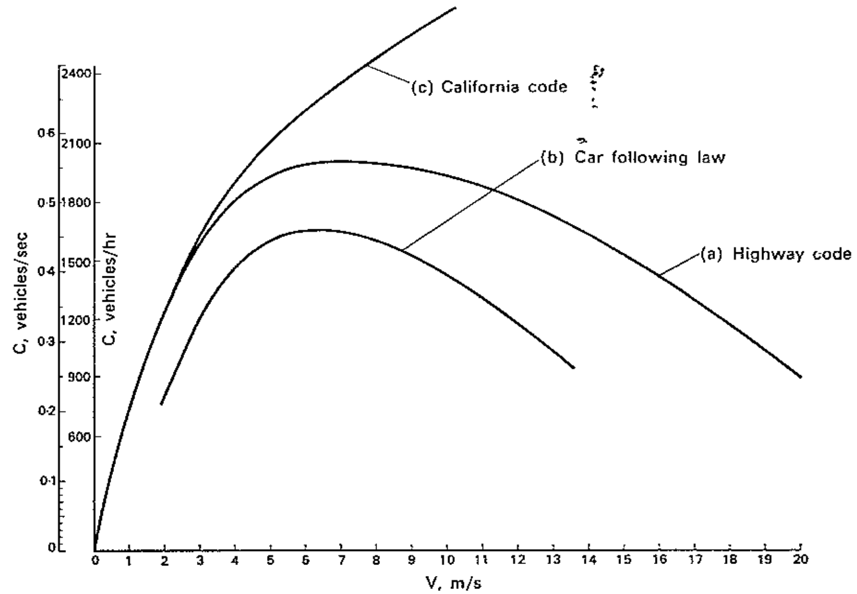


FIG. 2.4. Estimates of lane capacity.

The results of three different methods of estimating the relationship between capac and speed are plotted in Fig. 2.4. Curve (a) is based on the assumption that driv observe the *Highway Code*, curve (b) assumes that they follow the reciprocal spacing l using the value of the constant K determined experimentally by Herman and Potts, wh curve (c) assumes that they follow the "California Code" which ignores braking distai as such and which may be stated as follows: "A good rule for following another vehi at a safe distance is to allow yourself the length of the car for every ten mile/hour you travelling." It will be noted that there is good correspondence between the prediction of the *Highway Code* and the Car-following Law.

Some experimental results extracted from a report issued by the Road Research Laboratory[3] are shown in Fig. 2.5 which confirm that capacity passes through a maximum at a comparatively low speed and for practical purposes can be taken as not exceeding 1800 vehicles per hour (0·5 vehicles/second).

Whilst the agreement between theoretically estimated capacities and practical observation is good, nevertheless the success of drivers in following such laws remains to be explained, particularly as regards the nature of the input signal which they must receive to complete the feedback loop.

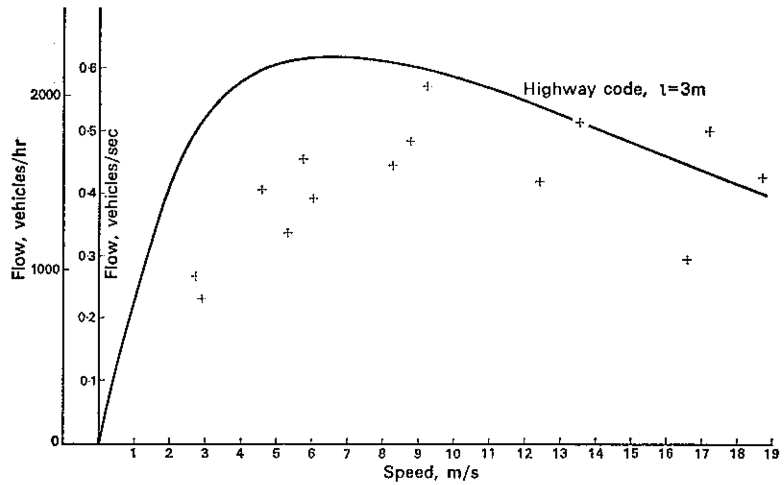


Fig. 2.5. Comparison of calculated with measured flows.

The driver is in charge of the vehicle moderating acceleration and deceleration by accelerator, brake, clutch and gearbox in response to information which he receives visually, aided in emergency by audio signals and supplemented by sensing the secondary aspects of the motion of the vehicle, i.e. pressure on his back as a correlate of acceleration rate. Clearly, hand signals or rear-light indicator signals on the leading vehicles provide important links.

Consider two vehicles travelling at different speeds, the leading vehicle being continuously visible to the following driver. How does the latter determine whether or not he is catching up on his precursor? Michaels[4] suggests that at some point in time the following driver can estimate the angle subtended by the lead vehicle. If at some later time this angle has increased sufficiently to be detected by him, he will infer that his distance from the vehicle in front has also decreased. Thus the change in size of the image of an object on the retina of the eye is the psychological correlate of distance perception. The contours of the leading vehicle expand or contract at a rate which is proportional to both relative

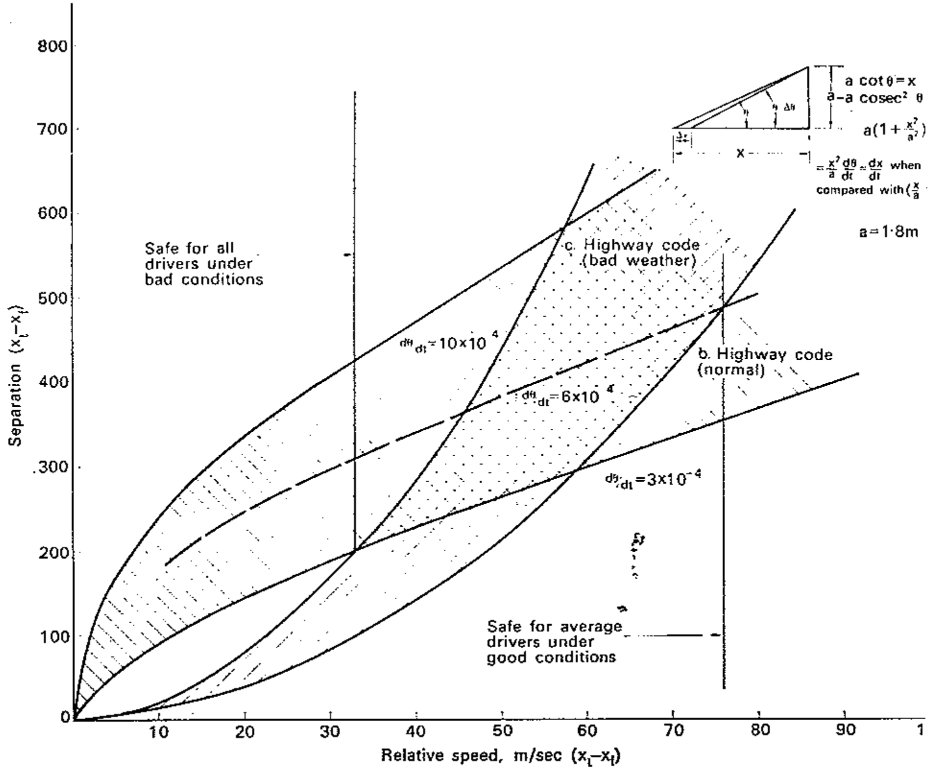


Fig. 2.6. Margin provided by visual estimation of relative velocity.

velocity and the distance apart so that the horizontal angle subtended changes at following rate:

$$\frac{d\theta}{dt} \propto \frac{\dot{x}_l - \dot{x}_f}{(x_l - x_f)^2}. \qquad (2.$$

Note the resemblance to equation (2.9). Experimental values for the minimum (thresho value of $d\theta/dt$ lay between 3 and 10×10^{-4} rad/sec.[5] Figure 2.6 shows the variation the minimum relative velocity which is detectable to a following driver with distance separation for various values of $d\theta/dt$.

Michaels[5] suggests that during the major part of the driving cycle a driver scales rate of overtaking on the basis of the angular velocity of the image of the lead vehi reducing his speed sufficiently to keep the angular velocity at or near the absolute thre old of detection. Taken to the limit this would lead to a smooth approach to bumper-bumper conditions as relative velocity approaches zero. In actual practice a driver v not approach closer than a certain minimum spacing necessary to ensure steering a speed control.

How adequate is this information to enable the driver to carry out his task? Curv

(Fig. 2.6) shows the thinking and braking distances derived from the *Highway Code* and extrapolated according to equation

$$x = 0{\cdot}682V + 0{\cdot}0762V^2 \text{ where } V \text{ is in m/s} \qquad (2.16)$$

from which it will be seen that a driver will receive adequate warning of a change in relative velocity of a leading vehicle to enable him to brake to standstill up to a speed of about 168 m.p.h. (75 m/s). This relates to good weather—broad daylight—good, dry roads. Curve c shows double the margin of safety as recommended for wet roads. This brings the maximum safe speed down to 100 m.p.h. (44·7 m/s) for the average driver. All drivers should be safe at 33 m/s (74 m.p.h.).

Thus it appears reasonable to operate roadways by means of direct vision although it is clear that poor visibility can render the system inoperative.

2.3. Automatic vehicle systems

It having been shown that the capacity of a road is determined by the characteristics of the average driver it is, therefore, very likely that the application of some automatic device would increase traffic capacity by expediting one or all of his three functions—namely, perception, decision and response. The controls of the conventional motor car have been evolved purely from mechanical and constructional considerations. Thus a joystick control which embodies speed and direction would almost certainly provoke a different driver reaction from the present unco-ordinated controls which operate by acceleration and yaw rate.

Several significant steps towards a practical automatic control system for automobiles have been made, particularly at the General Motors Technical Centre (Fig. 2.7). Attempts to devise a control in which all the necessary equipment was carried in the car and none on the road system were abandoned and co-operative equipment in the road has been found to be necessary.

Firstly a device for automatic steering is provided using the "electro lane" principle. This is a very simple device which can be used to warn a driver when he departs from a set path. This path is defined by a wire laid into the surface of the road and energised at a frequency of 2 kHz. This produces a circular magnetic field which may be detected by two search coils mounted on the vehicle symmetrically on each side of its centre. When the vehicle is evenly astride the wire, the signals in the two receivers are balanced but should its course depart from the predetermined path, an unbalance is created which gives a visual or audible warning to the driver. The "electro lane" can be regarded as a transitionary scheme which could be applied to any vehicle.

The next stage is to provide for automatic control of vehicle spacing. Firstly it is necessary to detect the presence of an obstruction. This can be achieved by dividing the road into blocks about 20 feet (6 metres) long. Within these blocks three coils are laid into the surface of the road, each extending to the full width of the traffic lane. Two coils, each embracing half the block length, are connected in opposition and fed with current at a frequency of from 15 to 30 kHz. The third coil embraces these two coils so that it normally experiences a balanced induction and therefore zero excitation. On any metallic object coming into the field of one of the coils, an unbalance is created which is detected in the third coil. This is used to operate the speed control which consists of a command

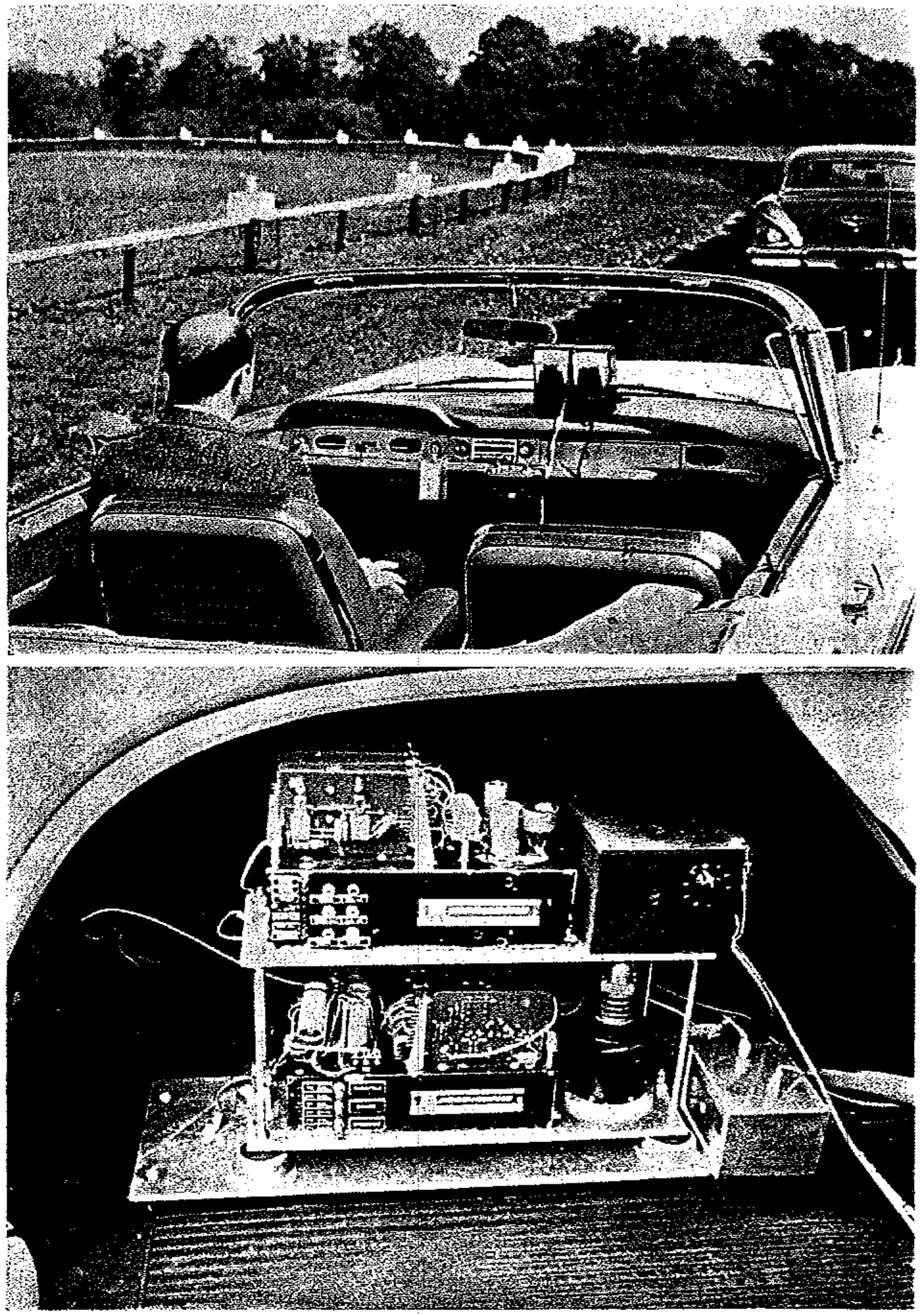


FIG. 2.7. (a) Car under automatic control. (b) Control apparatus in boot of car. Automatic Highway.

signal which is fed as a variable frequency pulse to a control cable laid into the road surface. This consists of two wires which are transposed laterally at fixed intervals so that the polarity of the magnetic pulses alternates as the car under control proceeds along the road. The frequency of this oscillation is naturally proportional to the speed of the vehicle and is compared with the imposed frequency representing the desired speed. Thus the elements necessary for error control are available.

On a car approaching a vehicle ahead, the speed control is modified to give spacing control. In one system under development, information about the speed and distance of the car ahead is supplied by road-based electronics into an analogue computer in the car itself. This immediately computes the safe spacing as a function of the speeds of the two vehicles and their accelerating or decelerating capacities. The resulting space-error signal, together with appropriate stabilising signals, provides the command signal to the accelerator or to the brake servo of the following car.

It is claimed that the automatic system, by eliminating lags and non-uniform behaviour of human drivers, can increase lane capacity to 3600 vehicles per hour at an optimum speed of 55 m.p.h. (24·6 m/s). It will have been gathered that the automatic road is technically feasible. Before it can be introduced, however, formidable problems of finance, insurance and legal responsibility for accidents must be overcome because of the large number of vehicle owners who would be responsible for maintenance of the control equipment. Here the railway has the advantage that track and vehicles are under the control of a single owner.

With automatic driving expression (2.4) simplifies to

$$C = \sqrt{\mu g/2l}$$

and if vehicles are coupled together, expression (2.6) becomes

$$n/\sqrt{2nl/a} \quad \text{or} \quad \sqrt{\frac{\mu g n}{2l}} \simeq 2{\cdot}2\sqrt{\mu n/l}. \qquad (2.17)$$

Thus coupling of vehicles can be important in road service.

Steering could be by the "electro lane" principle. the detector circuit providing three d.c. outputs, two steering signals and a monitor signal. This latter would be connected to the power and brake controls so that the vehicle would only continue in motion if a signal were present. The other two signals would initiate steering movements to the right or left by engaging clutches bringing into action a steering gear powered by a continuously running d.c. motor.

Any control system would be required to respond to different conditions of motion of the controlled vehicle. Consider that a route, perhaps under congested conditions, is being controlled at 7 m/s with a spacing of cars at 30 m. Thus the error would be ("actual spacing of cars" −30). Thus when spacing was 20 metres error would be −10. The vertical plot would be rate of change of error, i.e.

$$\frac{d}{dt}(x_l - x_f - 30) = \dot{x}_l - \dot{x}_f. \qquad (2.18)$$

Thus when the following car is faster than the leading car, velocity error is negative.

This can be represented as the phase-plane diagram in Fig. 2.8. The origin represents a point 30 metres behind the position of the leading car plotted horizontally and the speed of that car plotted vertically. The plane may then be divided into sections, each of which

represents the area of operation for a required control strategy. The control syste would be designed to bring the error to zero as represented by the origin. The norm mode of the controller would be to maintain the speed of the following car in order minimise error. To allow for minor fluctuations, an area represented by a circle wou contain a stable node or a stable focus, depending on the constants of the system.

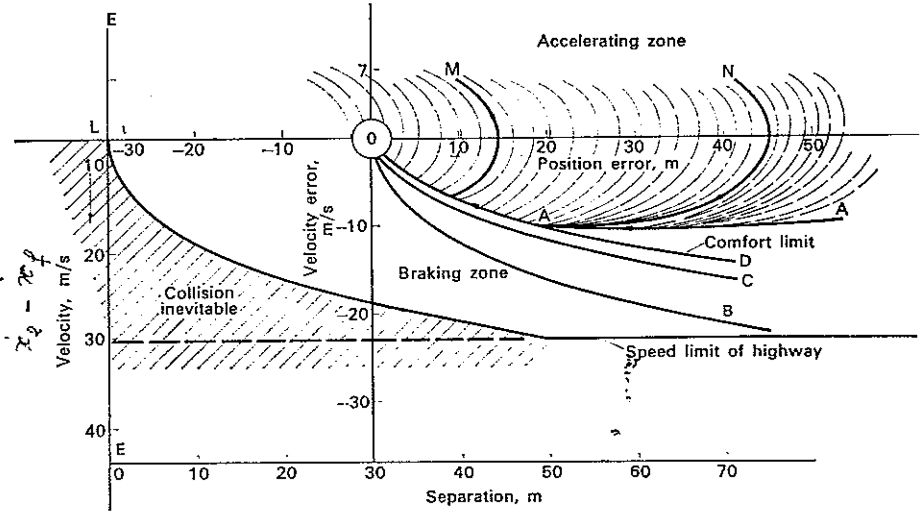


FIG. 2.8. Phase-plane diagram of automatic road control system.

There will, however, be transient conditions during which the following car has to brought into this region of control. Thus a car approaching at a higher speed will represented by AA. Line OB represents the maximum deceleration possible on a fine d OC represents the maximum braking which would be safe on slippery roads, OD can calculated to represent a rate of slowing down which would be quite comfortable for passengers. Thus a control region would be devised so that, when AA intersects C control would operate and the car would begin to slow down under the action of control system.

Whilst the case mentioned, i.e. a car coming from behind, is the only one which easily be imagined, there may be circumstances where the speed and position of the follo ing car can be represented by any point in the phase-plane diagram. A point to the left EE represents a collision, curves through L parallel to OB and OC define a zone wher collision becomes inevitable.

If a car were falling behind at a slow pace as represented by *M*, the control syst would impose an acceleration regime until the trajectory intersected OD when the braki regime would be entered. Starting from point N, the trajectory would continue until maximum designed speed, as represented by line AA, was reached. Speed would then held constant until the braking curve was reached.

The characteristics of the control system would be designed according to cert criteria so as to maximise economy, passenger comfort and to avoid wave action in following vehicles. The trajectories in the diagram were constructed on the assumption that maximum tractive effort is applied during acceleration so as to eliminate error as rapidly as possible. Thus from equation (1.13)

$$\ddot{x}_f = \frac{F(\dot{x}_f) - f(\dot{x}_f)}{M}, \qquad (2.19)$$

$$\theta = x_l - x_f - 30,$$

$$\frac{d\theta}{dt} = -\dot{x}_f \quad \text{when } \dot{x}_l = \text{system design speed which is constant,}$$

$$\lambda = \frac{\dot{y}}{y} = -\frac{F(-\dot{\theta}) - f(-\dot{\theta})}{M\dot{\theta}}. \qquad (2.20)$$

The isoclines are therefore horizontal straight lines as indicated in Fig. 2.8.

For braking $\ddot{x}_f = a$ where a is average rate of deceleration taken as 6·56 m/s^2 for emergency braking under good road conditions, 3·28 m/s^2 for bad roads and 2·65 m/s^2 for passenger comfort,

$$\therefore \lambda = -\frac{a}{\dot{\theta}}. \qquad (2.21)$$

References

1. BARWELL, F. T., Some speculation on the future of railway mechanical engineering. *Proc. Instn. Mech. Engrs.*, vol. 176, p. 61 (1962).
2. HERMAN, R. and POTTS, R. B., Single lane traffic theory and experiment. *Theory of Traffic Flow*, p. 120, Elsevier, 1961.
3. WARDROP, J. G., Experimental speed/flow relations in a single lane. *Proc. 2nd Int. Symposium on the Theory of Traffic Flow*, p. 194, O.E.C.D., Paris, 1965.
4. MICHAELS, R. M., Perceptual factors in car following. *Proc. 2nd Int. Symposium on the Theory of Traffic Flow*, p. 44, O.E.C.D., Paris, 1965.
5. MICHAELS, R. M. and COZAN, L. W., *Perceptual and Field Factors in Car Following*, Highway Research Board, 1963, Bureau of Public Roads, Washington, D.C.

CHAPTER 3

Control of Vehicle Spacing — Railway Signalling

3.1. Necessity for signalling on railways

As indicated in Fig. 2.6, the visual observation of the route ahead can provide suffici information for the preservation of a safe vehicle spacing under present-day road c ditions.

This is not the case when railways are operated in the traditional manner particul when freight is carried in wagons whereon the brakes are not under the control of driver.

Figure 3.1 applies the argument of Fig. 2.6 to railway conditions and takes into acco the fact that different drivers may have different threshold levels of perception of rat change of the size of the image of an object on the retina of their eyes. Curves a an represent the values of the minimum and maximum detectable value of $d\theta/dt$ which ran between 3 and 10 × 10^{-4} rad/sec. Curve c relates braking distance to speed for typ trains (the B.R. "S" curve) from which it will be noted that, above about 13 m. (6 m/s), the train would require a greater distance to pull up than that at which a dri having a threshold sensitivity of 3 × 10^{-4} rad/sec could form an opinion about the r tive motion of a preceding train. This then represents the maximum permissible speed "permissive" working, i.e. when trains are permitted to follow each other without intervention of fixed signals.

Certain urban electric railways and of course street railways operating with s contained electric vehicles whose wheel braking is often augmented with track bra (see Chapter 13) can operate at shorter braking distances. Where overall speed is limi by other considerations to, say, 40 m.p.h., permissive working can be adopted.

Curve d relates to a train consisting of sixty-five loaded 16-ton mineral wagons fi with simple vacuum brakes.[1] It will be apparent that it would be unsafe to operate train at anything faster than walking pace without some additional source of informat to the driver.

The successful evolution of railways has therefore depended on the provision additional information to the driver by means of signalling systems. Some considerat of the development of these systems may be justified, not only because of the continu viability of captive vehicle systems, but because such a study can serve to illust principles which must underlie the essential technology for automation of transpor general.

When railways were first operated there was no way of communicating informatio a driver regarding the condition of the track further ahead than he could see. Large

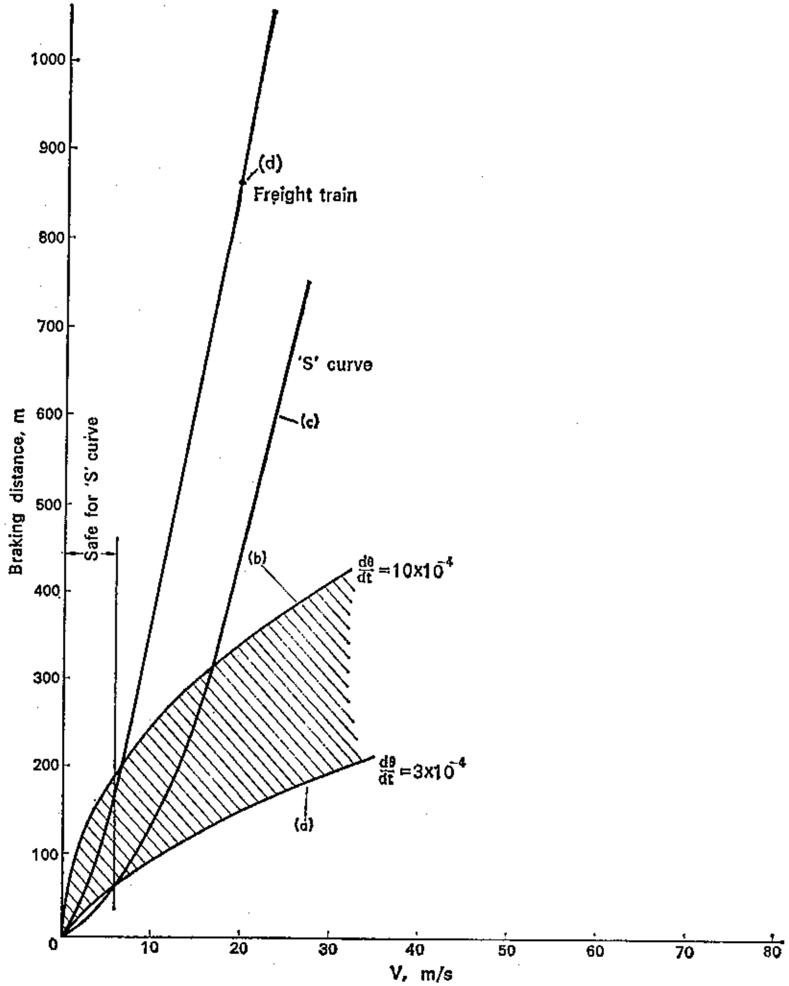


FIG. 3.1. Relation between visual perception of relative velocity and braking distance on rail.

and crossbar signals 8 feet wide and 40 to 60 feet high (Fig. 3.2) designed by I. K. Brunel, enabled station staff to indicate whether or not it was safe to pass a discrete manned point but beyond this it was a matter of inference. The only other information which could be given was the time which had elapsed since the passage of the previous train. The "Ball" was shown after 3 minutes and the "Board" after 10 minutes.(2)

Further progress had to await the invention of the electric telegraph and, due to the multiplicity of railway ownership in the nineteenth century, many trends were evident. It is possible, however, to distinguish between the different philosophies adopted in North America and Britain as represented by the "Train-order" and "Block" systems respectively.

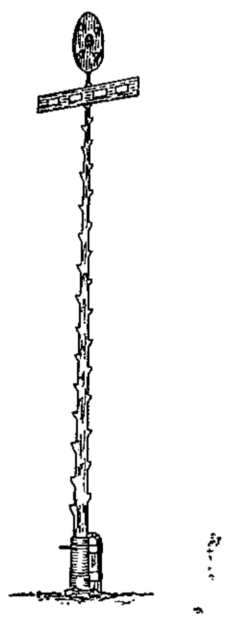


FIG. 3.2. Disc and crossbar signal.

3.2. The "train-order" system

The principle of the train-order system was that a centrally located "train dispatcl was responsible for ensuring the necessary separation of trains, i.e. no two trains sh be at one place at one time or, put another way, at any one time all trains should b different places. The orders were transmitted to local stations using the electrical graph. Instructions were sent to all concerned by written messages which had either t acknowledged by signature (a 31 order blank) or required no signature (a 19 order bla

The train dispatcher wrote the instructions in a book provided for the purpose, agent or agents responsible for onward transmission of the order at the local stat repeated it back to confirm its accuracy and then delivered it to the drivers responsibl carrying out the order. The fact that an order had been delivered was then sent back tc centre and in the case of a "31" order, when all recipients had been demonstrated to l received the order, a copy bearing the word "complete" was given to all concer Action on the basis of this complete copy should therefore involve no risk of conflic movement.

3.3. The "block" system

In the United Kingdom the telegraph was used in a different manner and on a m more localised basis to protect individual sections of the track or blocks from occupa by more than one train. Instead of telegraphic commands of a verbal character, sir special-purpose telegraphic instruments were developed for traffic control. Two ave of communication were usually provided; (i) a bell code which conveyed information mainly in the forward direction and (ii) an indicator code which transmitted information in the reverse direction. The system embodied a great deal of redundancy and thus was inherently safer than the train-order system. Given the telegraphic link between two ends of the section, all that was really necessary to operate the block system would have been a bi-stable relay operated from the exit point and giving an indication at the entry point. Thus on a train entering the section an impulse would be sent to the relay putting it into position which we shall refer to as the "on" position and on the train leaving the section the relay will be reversed into what we would call the "off" position. An indication from the relay at the point of entry would serve to show whether or not a train could proceed safely forward and indeed might well be connected directly to the signal without manual intervention. (Note. The writer has observed a system operating on this principle on the Chinese Railways.) To become automatic such a system would require some device at the exit to verify that all material which had passed through the entrance had also passed out. Thus an axle-counting device would be as effective as a continuous track circuit in providing an automated system.

The Glasgow Cable Subway, inaugurated in 1896, employed a simple automated block system. Trains were controlled by starting semaphore signals which were situated at the exit from the station platforms and which were worked by a dispatcher stationed on a rostrum placed centrally in each platform. Treadles were situated at the exits to the station which controlled an electric bolt interlock attached to the lever mechanism controlling the signal. The electric bolt was only released when a train operated a treadle as it departed from the station ahead. Thus, assuming that a signal was placed at danger when a train departed, it could not be placed at "clear" again until the previous train had left the section.

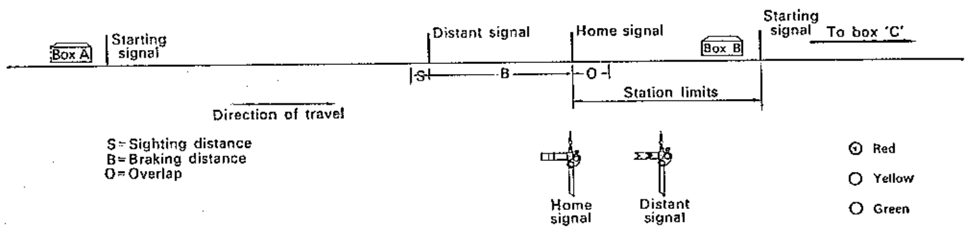


FIG. 3.3. Manual block system.

In practice, however, the operation of a block system on main lines in Great Britain is very much more complex than this. It serves to illustrate principles of redundancy which might well be employed in the modern idiom using electronic techniques in a computer-based system.

The entrance to a block is protected by a "starting" signal and the exit by a "home" signal (Fig. 3.3). Because, in main-line working, the braking distance often exceeds the sighting distance of a signal, a third signal—a "distant signal"—is provided which repeats the position of the home signal at such a distance in advance thereof that the driver can bring his train to rest there. (For interlocking of signals, see Chapter 10.)

Manned signal cabins were provided for each block and connected by a telegraph which

could be used to pass messages by means of a bell code. There were also instrur which gave three indications as follows:

LINE CLEAR
LINE BLOCKED (or CLOSED)
TRAIN ON LINE

Duplicate indicators were provided in each box but only one had an operating ke signalman at the entrance to a block having a keyless instrument.

A great deal of information could be passed using different sequences of beat pauses but the more important of some thirty codes are given below.

Message	*Bell code*
Call attention	1 beat
Is line clear for	
express passenger train?	4 beats consecutively
ordinary passenger train?	3 pause 1
light engine?	2 pause 3
slow freight train?	3 beats consecutively
Train approaching	1 pause 2 pause 1
Train entering section	2 consecutively
Obstruction danger	6 consecutively
Train out of section	2 pause 1
Train passed without tail lamp	9 consecutive to box in advance 4 pause 5 to box in rear

Consider three signal boxes, A, B, and C and suppose that the signalman at A req to dispatch a train towards B. It is assumed that he will already have received the " out of section" code signal relating to the previous train and that the Block indi communicating with box B is in the LINE CLOSED (or BLOCKED) position. He will s signal beat to B to "call attention" and must await acknowledgement. He will gi "is line clear?" signal for the appropriate class of train. He will repeat this proc until B is ready to accept the train; B will then acknowledge the bell-code signal and the Block indicator in the "Line Clear" position, that is, B's keyed instrument an keyless instrument will both show "Line Clear". B will not signal "Line Clear" to A u he had received "Train out of Section" for the previous train or unless he can s himself that this has passed at least ¼ mile beyond his home signal. A will then low starting signal to permit the train to proceed towards B and, on the train actually pa he will signal B with two consecutive beats. B will acknowledge and change the I Indicator from "Line Clear" to "Train on Line". A will verify that the train is com with tail lamps. If not he will immediately notify boxes both in advance and in the

B will meanwhile have repeated a similar procedure with C and, having obtained Clear", will pull off his "Distant", "Home", and "Starting" signals to allow the tr pass forward. He will observe the passage of the train, particularly satisfying himsel the tail lamp is in position. This proves to him that the section is now empty so h give the "Train out of Section" code to A and replace his keyed instrument int "Line Blocked" position thus completing the cycle. Figure 5.6 represents a "system chart" of a signalman's action.

Where the block signal was introduced in the United States the high degree of r dancy characteristic of the British system was not generally employed. The main consisted of a written register of trains known as the "Block Sheet" whereon the signalman recorded the entry of a train and also the fact of exit as telegraphed to him from the exit block. He would then refer to the block sheet to determine whether or not it was safe to admit a further train.

3.4. Axle-counting—radio link

An alternative both to the observation of tail lamps and to continuous track circuitry is the principle of axle-counting. If all the axles of a train are counted when it enters and when it leaves a section and if the results agree, then it is certain that the section is clear. Mechanical axle-counters, which depended on physical contact between the wheel and the detection apparatus, were difficult to maintain but it is now possible to detect the passage of an axle using entirely static apparatus. Fixed coils are attached to the rail so that the flux path is disturbed by the passage of a wheel flange and produces a distinct signal. The detectors are mounted at the entrance to a section of railway as in Fig. 3.4

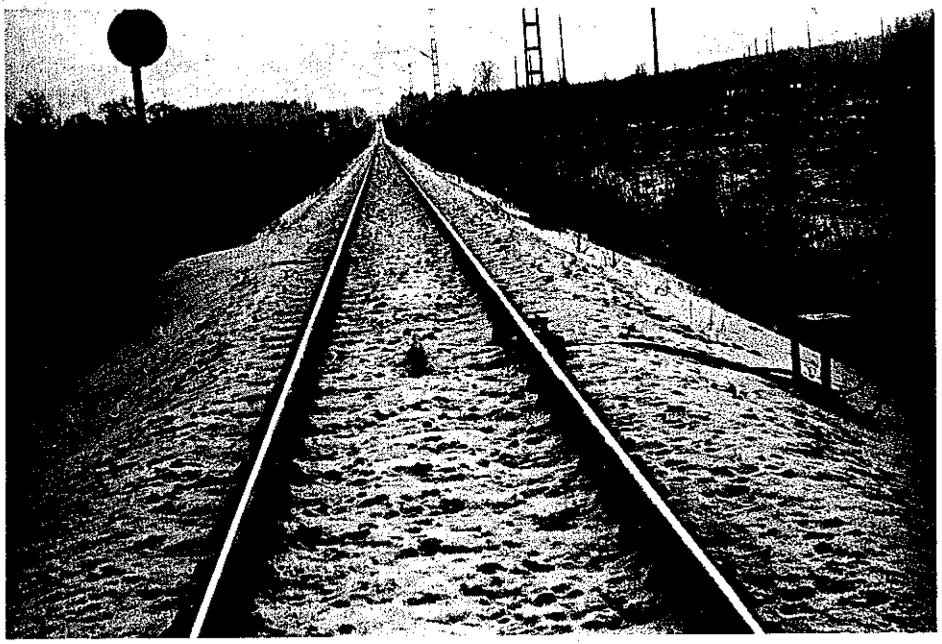


FIG. 3.4. Electronic axle-counter in use in Finland.

and, when a train passes, the detector sends information on the number of pulses corresponding to the number of axles to a register located at the next station, where the information is stored. When the train arrives at the next station the axles will again be counted and a comparator circuit will determine whether both detectors have counted the same number of axles.

The axle detectors are designed so that the direction of train movement is determ as well as the number of axles. Thus if a shunting movement takes place over a det and back again the total number of axles will be counted as zero. The actual informa about the number of axles is transmitted between the stations, using a security bi code. This is constantly repeated so that a disturbance in the transmission link is interpreted as a discrepancy in the number of axles. (Earlier systems of axle-cour transmitted actual pulses and external interference might produce pulses similar to t from the axle detectors thus leading to a mismatch which might be interpreted a obstruction upon the line.)

The railway between Tampere and Seinäjoki (157 km long) in Finland is automati signalled using axle-counters. No telecommunication wires or cables are used an formation and command information is passed by radio-link. A suggestion made b Finnish Railway for the future is that there should be no station-to-station commu tion links but that all axle-counting units should transmit their results to a cent situated computer.

3.5. Single-line working

In spite of the almost universal application of the block system this is not relied in Britain (or on many railways which follow British practice) to ensure the avoid of collisions on single-track railways. In addition to observing fixed signals the driv a train is required to carry a "token" (sometimes known as a "train-staff" or "tab which is specific to each particular section. It must be assured that only one tok available for a section at a given time and when only one token was provided trains c not follow each other and could only operate in alternate directions so that the t could be passed forwards and backwards along the section.

The problem was overcome using token instruments at each end of a section w were connected by telegraph wire (Fig. 3.5). These instruments contained magazine storing a number of tokens but only one can be out of the instruments at any time. token can be replaced and then the apparatus is in its normal state. Assuming intended to pass a train from A to B, the signalman at A sends "Call attention" t B responds. Then A sends "Is line clear?". B returns "Line clear" but when pressi the plunger for the last time keeps it in until a needle mounted in the block instru regains the upright position. This indicates that A's block instrument has been unloc "A" then withdraws a token which he hands to the driver. He sets a pointer on instrument to "up token out" or "down token out", according to the direction of train, and sends "train entering section" to B. B also sets his instrument to "token

On the arrival of the train at B, the driver hands the token to the signalman replaces it into the instrument, turns the pointer switch to "token in" and sends " out of section" to A. A repeats message in acknowledgement and sets his instrume "token in".

The actual retention or release of a token is controlled by rotary commutators w are situated at the central position on the token instruments where the four slots conv into one. A token can only be withdrawn from the slot by the insertion of the flat key rectangular piece at the end of the spindle of the token into a keyway in the commut

If the section is clear, the signalman at "A" lifts a token from the magazine of his instrument, inserting its end projection into the keyway of the commutator. He then turns the key through 90° in the anti-clockwise direction which sends the current from Box B to a polarised relay. Provided the instruments at both ends of the section are in phase, the polarised relay at A will close a local circuit which will release a lock on the commutator. This will allow the token to be rotated through a further 90° from which position it can then be withdrawn by the signalman for issue to the train driver.

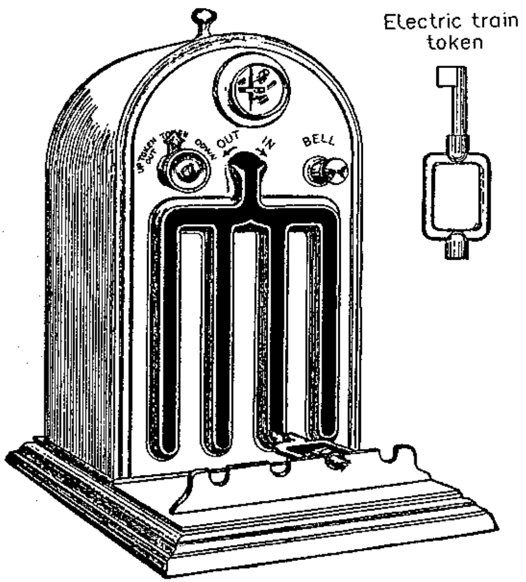


FIG. 3.5. Single-line block-token instrument.

The rotation of the commutator physically prevents the withdrawal of a second token at A and inhibits the release of the commutator at B through an electrical circuit. On the train reaching B, the token is inserted in the keyway of the instrument which is turned through 180° in a clockwise direction. This reverses the electrical connections to Box A so that a token can be obtained from either end.

Radio link

The instruments have to be connected electrically and overhead wires are subject to vandalism, theft and severe weather. For example, during the winter of 1977/78 a particularly severe blizzard destroyed about 65 km of overhead line between Inverness

and Wick. This provided an opportunity for the introduction of a radio system as alternative means of communication between signal boxes.[4]

At the heart of the control unit which interfaces with the block token instruments duplicated (i.e. 2 out of 2 redundant) microcomputer system designed to perform necessary logic and information processing tasks. These include sampling the input fı the token instruments, formatting and encoding data messages, decoding recei messages and outputting the information to the token instruments. The radio signa broadcast to all other block posts but each post has a unique address so that the eqı ment at the other end of the block section from the transmitter is the only one to respo

The electronic token

The combination of 19th and 20th-century technology represented by the connect of the manually-operated single-line token instruments to microcomputers does represent the optimum use of resources. A further development emanating from Research and Development Division of British Railways exploits the potential benefit the microcomputer more fully by eliminating the use of physical tokens and man signalboxes along the line. In a proposed installation between Dingwall and the Kyle Lochalsh in the Highlands of Scotland, the line is controlled by a despatcher at Dingv by radio. Each locomotive driving cab is provided with apparatus for transmitting receiving data messages through the radio network. An instrument is provided wh displays continuously the names of two places between which the driver has permiss to travel. This indication is the equivalent of a block token and, because the mic processor is programmed to ensure that not more than one locomotive receives d messages relating to any given section, the system is known as the "electronic" token

The dispatcher at Dingwall will be provided with a track diagram panel which display the occupancy of track from Dingwall to Kyle of Lochalsh. A number of phant token exchange points along the line will be designated and the train driver will responsible for requesting and returning tokens. Assuming that a train is at rest a token exchange point and waiting to enter the next section, the driver will call the patcher on radio, identify the train and its position, give the unique radio number of train token apparatus, and ask for the token of the section in advance. If a toker available, the despatcher will manually enter the unique radio number in the despa apparatus and inform the driver. On receipt of the token and with the verbal permiss of the despatcher, the driver will take his train forward until he reaches the next tol exchange point when he will call the despatch centre and offer to return the token. Wl the despatcher is ready, the driver will send the unique data message which returns token to the centre. The driver will inform the despatcher when the rear of the train cleared the station stop-board so that he can release the token for the section in rear t following train.

Whilst the system is "fail safe" provided the full procedure is operated, no provisio made for broken trains. It is believed, however, that with fully-fitted trains (vacuum air-braked) the brake application following rupture of the train pipe will alert the dri to the situation, who will in turn inform the despatcher who will take the necess precautions to protect the separate parts of the broken train.

3.6. Lock and block

Many operators employ an additional element of redundancy by interconnecting the block instruments with the signals controlling the entrance to a section. Should a signalman forget to replace signals at danger there is a risk that a following train might enter a block without the safety procedure being observed, therefore some railways use the "Lock and Block" system which ensures that, once a train has been accepted and signalled forward, it must pass through the section and all signals replaced to danger before the following train can be accepted. A treadle is provided to indicate the passage of the train. The system has many variants but in principle the communication circuits are provided with means for locking the signals and vice versa. Thus "Line Clear" cannot be given a second time until both the treadle has been operated and the home signal replaced at danger. One system ensures correct sequence of operations by arranging for the "Line Clear", "Train on Line" and "Line Blocked" indications to be given by a rotary switch that is constrained to turn in one direction only by a ratchet and pawl. The normal indication is "Line Blocked". Rotation of the switch by one-third of a revolution gives "Line Clear" and by another third "Train on Line". It cannot be turned further until the train passes over a treadle, after which the original position can be restored.

3.7. Multiple-aspect signalling

Given an automatic method of testing that a section of track is clear of obstruction, for example by track circuits described in Chapter 5, automatic multiple-aspect colour light signalling may be used.

In this system the normal aspect is clear. The passage of a train causes a relay to drop out which puts to red the signal immediately behind the train. This information is passed to the signal behind, which shows a "yellow". The next signal shows a "double yellow" and the last one of all shows "green". Even if a train were to break into two parts it would be protected because the track circuit would operate whenever any vehicle whatsoever was in the section. There is no need for observation of tail lamps by signalmen. Figure 3.6 shows the disposition of automatic signals for a speed of 100 m.p.h. (45 m/s). The braking distance is 2000 metres (6600 feet) assuming "S" curve conditions as represented by curve "c", Fig. 3.1. Assuming a coefficient of adhesion of one-tenth, the same signalling system would be appropriate for a maximum speed of 63 m/s (141 m.p.h.). The term "sighting distance" is self-explanatory and "overlap" represents an allowance to cover any failure of a train to come to a standstill at the precise location of the signal at danger.

If a train were stationary at A the following train would have to commence braking at C so as to be able to stop in time to avoid a collision. If the block was equal to the braking distance, point C would have both the starting signal for the box in the rear and the distant signal for the box in advance. Thus there would be three aspects as follows:

Position of signal		Indication
Home	Distant	
on	on	stop
off	on	caution
off	off	clear

With colour-light signals this is of course simplified to red, yellow and green indicati respectively.

The clear indication cannot be given until the first train has passed out of the n block. Therefore if two trains are to follow each other at uniform speed they must separated by two block lengths. However, if a fourth aspect, the double yellow, is int duced, a green may be given at C when the preceding train has passed only half a brak distance beyond the red. The separation of trains has now been reduced from 2× brak distance to $1\frac{1}{2}\times$ braking distance and track capacity has therefore been increa correspondingly.

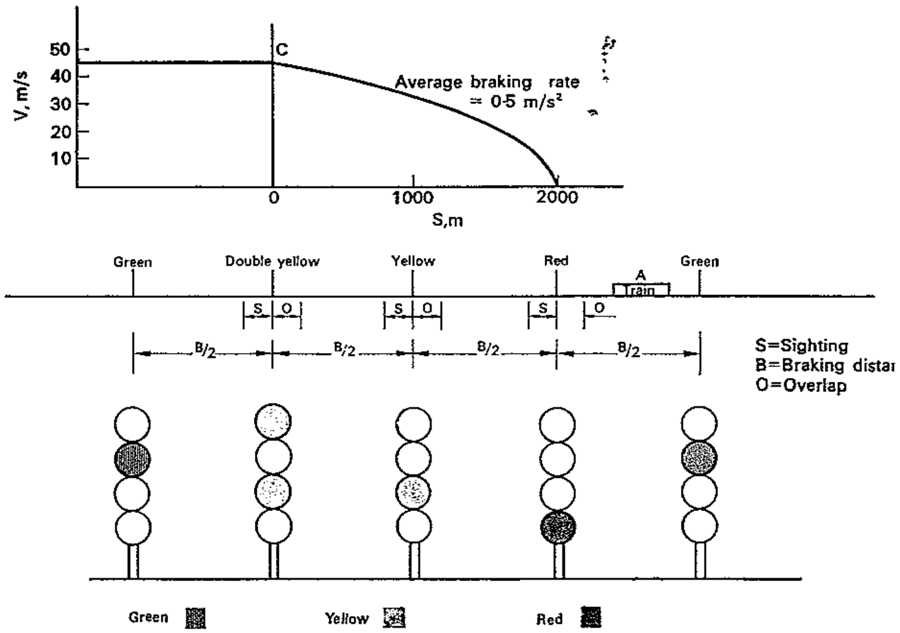


FIG. 3.6. Multiple-aspect colour-light system.

For very high-speed working as many as seven aspects have been suggested, but clea the return for increasing the number is a rapidly diminishing one whereas the cost fixed signals increases in direct proportion to the number of aspects.

Reverting to the consideration of maximum capacity and making the following assumptions:

1. braking rate $= 0{\cdot}5$ m/s^2,
2. maximum train length 420 metres (1400 feet), i.e. seventy wagons each containing 20 tons and being 6 metres (20 feet) long,
3. four aspect colour-light signals,
4. overlap 183 metres (600 feet),
5. sighting distance 183 metres (600 feet),

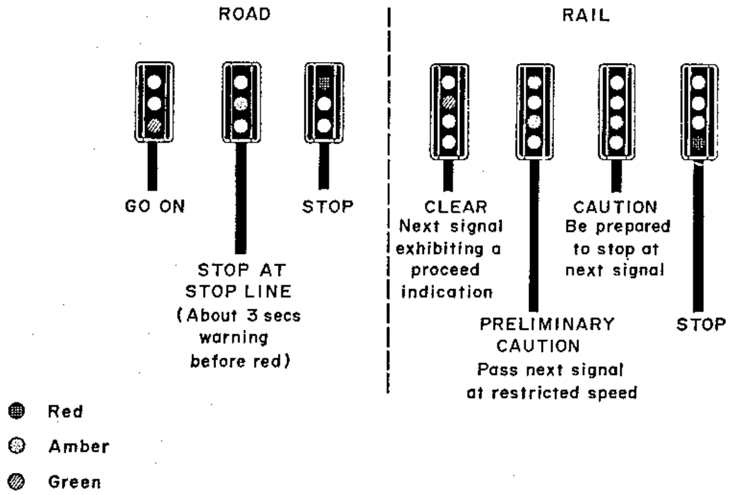


Fig. 3.7. Comparison of significance of colour-light signals for road and rail.

separation of trains
$$= l + s + o + \frac{3}{2}\cdot\frac{V^2}{2a} \text{ metres}$$
$$= 420 + 183 + 183 + \frac{3V^2}{2},$$

capacity in trains/unit time
$$= \frac{V}{786 + \frac{3}{2}V^2}.$$

Differentiating and equating to zero, $V = 23$ m/s, or 51 m.p.h.
Therefore capacity for maximum output equals fifty-two trains per hour with four aspect signals placed 262 metres (880 feet) apart.

Capacity equals $2{\cdot}0 \times 10^4$ kg/sec (72,800 tons/hour).

Let us consider ways and means whereby this capacity could be increased. Firstly there could be a reduction in braking distance but the value of the coefficient of adhesion presents a physical limitation to progress in this direction. Increase in train length would be possible and profitable subject to the limitations imposed by the strength of the draw-gear and the length of sidings and terminal facilities. Increase in speed without increase in train length would reduce capacity because of increase in braking distance. If accompanied

by change in train length, capacity can be increased without limit. An increase in t number of aspects would increase capacity in diminishing proportions. In the examj chosen there are already six track circuits and sectional points per mile and a furtl increase would be expensive. Automation would enable "overlap" and "sightin distances to be eliminated.

Comparison of the significance of colour-light signals as applied to rail and road illustrated in Fig. 3.7.

Although in simple sections of track, trains may be automatically signalled, manı intervention becomes necessary at junctions and other points where conflicting movemeı

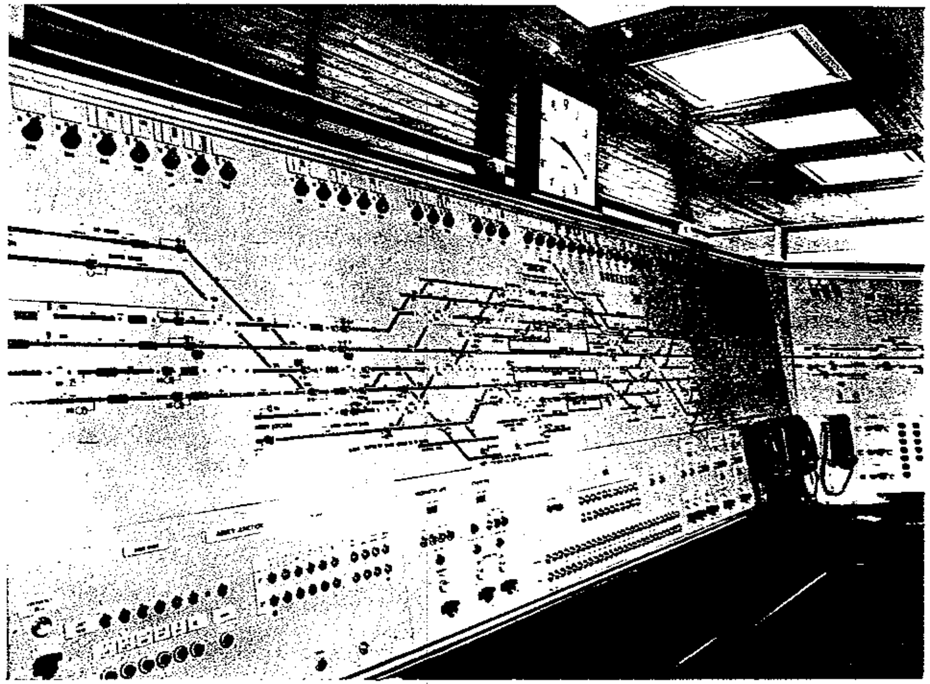


Fig. 3.8. Illuminated diagram and control panel. (Courtesy of British Rail.)

may occur. Control at such points has been progressively automated. Originally poiı and signals were operated by individual levers with mechanical interlocking (see Chap 10) to prevent dangerous situations from arising.

Progressive introduction of electromagnetic devices into an all electric signal b enables an operator to concern himself with overall movements rather than with indi dual points and signals. An illuminated diagram is provided on which the occupation each section of track is indicated. The actual train is identified by a number which is f forward automatically from box to box and from section to section on the diagram accordance with its actual movement.

A common system of route setting is known as the N–X system. In this method t route is set up by the signalman operating two control knobs, one at the entry and t other at the exit of the desired route. The proving of the freedom of the route and operation of the individual points and signals is then carried out automatically.

A high level of control is achieved by grouping the control of a number of local areas within a central box. Local circuits having a fully inbuilt "fail safe" capability are controlled by a multiplex system from the central box. Thus one team of men will now control as much as 60 miles of route with, of course, a track mileage much greater than this. A typical combined illuminated diagram and control panel is illustrated in Fig. 3.8.[3]

The foregoing account relates only to the basic principles of railway signalling which is, in practice, complicated by many geographical and operational requirements. For further information the reader is referred to a series of pamphlets published by the Institution of Railway Signal Engineers and listed below.

1. *Principles of the Layout of Signals.*
2. *Principles of Interlocking.*
3. *Mechanical and Electrical Interlocking.*
4. *Single Line Control* (out of print).
5. *Principles of Power Point Control and Detection.*
6. *Signalling Relays.*
7. *Typical Signal Control Circuits.*
8. *Typical Selection Circuits.*
9. *Track Circuits.*
10. *Mechanical Signalling.*
11. *Signalling Power Supplies.*
12. *Block Instruments* (out of print).
13. *Train Describers* (out of print).
14. *Multiple Aspect Signalling.*
15. *Circuits for Colour Light Signals.*
16. *Route Holding.*
17. *Track and Lineside Signalling Circuits in A.C. Electrified Areas.*
18. *Principles of Relay Interlocking and Control Panels.*
19. *Route Control Systems* (L.T. Practice).
20. *Route Control Systems* (W. B. & S., Co.).
21. *Route Control Systems* (A.E.I.–G.R.S. Co.).
22. *Route Control Systems* (S.G.E. Co.).
23. *Mechanical Control of Points and Signals.*
24. *Automatic Train Control (Warning and Trainstop Systems).*
25. *Level Crossing Protection* (out of print).
26. *Remote Control of Railway Signal Interlocking Equipment.*
27. *Signalling a Layout* [Replaces Booklet No. 1].
28. *Remote Control Systems* (L.T. Practice) [Replaces Booklet No. 19].

Booklets Nos. 1, 4, 5, 9, 11, 13, 16, 19, 20, 21, 22, 23, 24 and 25 are out of print but can be consulted in major engineering libraries.

The Hon. General Secretary,
The Institution of Railway Signal Engineers,
21 Avalon Road,
Earley, Reading
RG6 2NS.

References

1. Baldwin, T., Peacock, D. W. and Scales, B. T., Problems arising with continuously braked freight trains, *Proc. Instn. Mech. Engrs. Convention on Railway Braking*, p. 12 (1962).
2. Macdermot, E. T., *History of the Great Western Railway*, vol. 1, part 2, p. 593 (1927), Great Western Railway, Paddington Station, London.
3. Brentnall, E. G., Signalling and telecommunications works on the Euston main line electrification. *Proc. Instn. Mech. Engrs.*, vol. 181, pt. 3F, p. 65 (1966).
4. Cribbens, A. H. and Giles, L. J., The Inverness-Wick radio signalling scheme. *Railways in the Electronic Age*, I.E.E. Conference Publication No. 203, p. 11 (1981).

CHAPTER 4

Problems of Congestion — Traffic Regulation

4.1. Random events

Most transport situations embody an element of chance and an element of order. example, it is highly predictable that certain buses will run full during a peak perioc the actual number of people carried on any vehicle during a slack period is entir matter of chance.

Again, the ideal orderly progression of events exemplified by a railway timetable be contrasted with the random variations encountered in all real-life experience. important, however, to look for the underlying predictability of apparently ord systems.

If the time of occurrence of certain events is random but the total number of e occurring in a relatively long period of time is known, it is possible to assign a proba to the occurrence of a given number of events in a given (much shorter) time interval. is given by the Poisson distribution as follows. If the expected number of events aver over a long time is λ, the probability of an event occurring 0, 1, 2, 3, to n times in time interval is given by the successive terms of the expansion of the function ϵ That is $\epsilon^{-\lambda}$ multiplied by 1, λ, $\lambda^2/2$, $\lambda^3/3!$ to $\lambda^n/n!$

Thus $$P(n) = \frac{\epsilon^{-\lambda t}(\lambda t)^n}{n!}$$

where $P(n)$ = the probability of n events occurring in a given period of time, t
λ = the average rate of occurrence of events.

The derivation of the Poisson distribution as a particular case of the binomial dist tion will be found in many textbooks on statistics; for example, H. D. Young, *Stati. Treatment of Experimental Data*, McGraw-Hill.

4.2. Queues—Poisson arrivals—constant service times

Let us take as an example the formation of a queue at a booking office where passe may be assumed to arrive at various instants of time which are randomly distribute where they are served with tickets at a constant rate. We can estimate three quantit follows:

(a) number of persons waiting at any time t;

(b) length of time that a person arriving at time t would have to wait before being served; and

(c) the number of persons passing through a queue from a given state until it becomes empty.

Let $1/\mu$ = time occupied by service,

λ = rate of arrivals,

and ρ = traffic intensity = λ/μ.

It can be demonstrated that the following expressions apply in these circumstances:

$$\text{Mean queue length} = \rho^2/2(1-\rho) = \frac{\lambda^2}{2\mu(\mu-\lambda)},$$

$$\text{Mean number of units in system} = \rho + \frac{\rho^2}{2(1-\rho)} = \frac{\lambda}{2(\mu-\lambda)}(2-\rho). \quad (4.2)$$

It is obvious that if ρ is greater than unity the queue will grow longer and longer. If it is equal to unity there will be a constant variation in the length of the queue but it will never be empty. If ρ is less than unity there will be occasions when the queue will be empty.

$$\text{Mean waiting time} = \frac{\rho}{2\mu(1-\rho)} \quad \text{or} \quad \frac{\lambda}{2\mu(\mu-\lambda)}. \quad (4.3)$$

$$\text{Mean waiting time plus service time} = \frac{2-\rho}{2\mu(1-\rho)}$$

$$\text{or} \quad \frac{1}{2(\mu-\lambda)}(2-\rho). \quad (4.4)$$

Thus if booking takes 10 seconds and there are 300 per hour, $\mu = 6$, $\lambda = 5$ and therefore $\rho = 5/6$. The queue would then average three people and the mean waiting plus service time would be 35 seconds.

Suppose we have a station comprising several platforms and from which trains are scheduled to depart every 2 minutes over a single unidirectional track having a nominal capacity of forty trains per hour. Owing to delays caused by passengers, late arrivals due to fog, etc., the trains are actually ready for departure in accordance with a Poisson distribution.

Then μ = 40 trains per hour,
λ = 30 trains per hour.

$$\text{Average waiting time} = \frac{30}{80(40-30)} \text{ hours} = 2\tfrac{1}{4} \text{ minutes.}$$

In order to reduce average delay to below 1 minute we would have to reduce λ to 22·8. Alternatively we could increase μ by increasing the number of tracks. If this were doubled so that μ = 80 trains per hour, the average delay would be reduced to 14·5 seconds. The traffic intensity would be 0·375.

4.3. Exponential service times

Reverting to our example of the booking office, whilst the average time taken to b each passenger might remain unchanged at 10 seconds, some passengers might req special tickets, change, etc., so that the individual booking times might vary. It is usua take an exponential distribution of service times for this purpose. This distributio defined by the expression, density = $a\epsilon^{-ax}$ where "a" is a positive parameter. The m becomes $1/a$. The expressions governing queuing then become:

$$\text{mean queue length} = \frac{\lambda^2}{\mu(\mu-\lambda)}, \quad (\cdot$$

$$\text{mean number of units in the system} = \frac{\lambda}{(\mu-\lambda)}, \quad (\cdot$$

$$\text{mean waiting time} = \frac{\lambda}{\mu(\mu-\lambda)}, \quad (\cdot$$

$$\text{mean waiting time plus service time} = \frac{1}{(\mu-\lambda)}. \quad (\cdot$$

It will be noted that the effect of variation of service times is to double queue len and waiting time.

Other distribution functions are also used in operational research. Because transp problems generally reduce to consideration of headway at a bottleneck which is usu fixed by braking distance and is a constant, the concept of Poisson arrivals and Const Service times is the most convenient for traffic flow studies.

4.4. Effect of delays on headway of signalled systems

It is apparent that the effect of dispersion of headways in a traffic stream is to red capacity below the theoretical values calculated in accordance with the methods Chapter 2. Similarly in a signalled system complete regularity of flow is a requisite maximum through-put.

The factor limiting the capacity of any section of railway equipped with fixed signal of course, the length of the block. Where multiple-aspect signals are used this te applies to the smallest subsection. Clearly the shorter the blocks governed by a sig aspect the greater will be the line capacity, but the law of diminishing return operates v powerfully so that there would appear to be little justification for fixed equipment spa more closely than at intervals of 1000 metres for main-line operation.

The attraction of the moving-block system, however, may be relevant under conditions of congestion and irregular working and if the automatic system of communi tion selected is independent of fixed equipment any advantages may be secured with cost penalty. To clarify terminology the moving-block system referred to here is wherein the speed of a train is continuously varied so that in emergency it could come t stop immediately behind the previous train, should that stop instantaneously. Views h been expressed that this specification is unduly rigorous because the previous train mi in fact require some time to come to a stop. Nevertheless, it is possible for the equival of a sudden stop to occur as at a converging junction when the other entrance is suddenly occupied. Alternatively the track may be suddenly blocked by collision.

Figure 4.1 shows the effect of a delay which causes four trains to be brought to a standstill. They will thus be spaced out at a minimum spacing of 1000 metres. Assume that the signal governing the leading train goes to "clear". The first train must travel its own length plus 180 metres overlap before the second can commence to move.

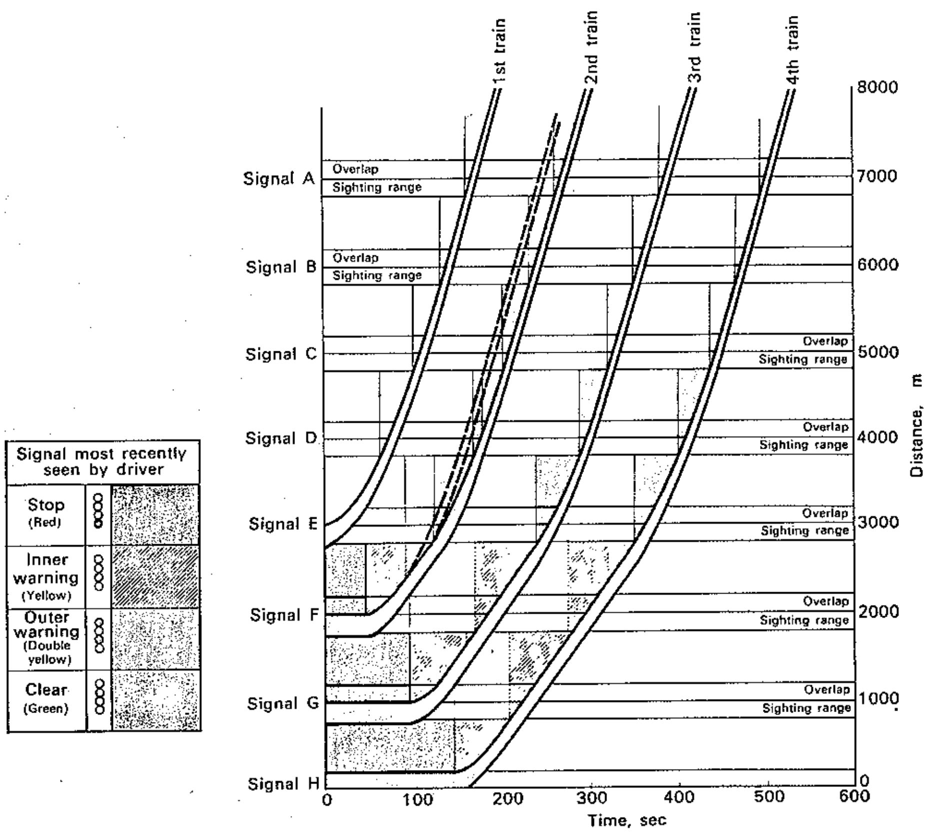


FIG. 4.1. Effect of congestion on multiple-aspect signalling.

Thus there is a delay of $\sqrt{2s/a}$ seconds before the second train can commence.

The nth train will not move until $n\sqrt{2s/a}$ seconds after the first one has moved,

therefore

$$\text{delay} = n\sqrt{\frac{2(l+o)}{a}} \tag{4.9}$$

where n = number of trains,

a = rate of acceleration (mean),

l = length of train,

o = length of overlap,

$s = l + o$.

Let us examine the effect of such a delay on line capacity. Consider that the distanc time relationship of the trains during acceleration from rest is as shown in Fig. 12.2 a that the line is signalled as in Fig. 3.4.

Trains are held at signals E, F, G, etc., spaced at 1000 metres (3300 feet). The fi signal goes to green and the first train accelerates in accordance with Fig. 12.2. T second train will wait at signal F until the tail of the first train has cleared the overlap signal E. The aspect of F will then, however, become yellow, and with the colour-lig system the driver will have no further information until he comes within sighting distan of the signal at E. He has no reason to assume that this will not be at red and therefo must not proceed at a speed greater than that which will allow him to bring his train to stop within sighting distance. This is about 13 m/s (30 m.p.h.) for a 180-metre (600 fe sighting distance and a braking rate of 0·5 m/s². Instead, therefore, of being able continue acceleration as indicated by the dotted curve in Fig. 4.1, once he has reached

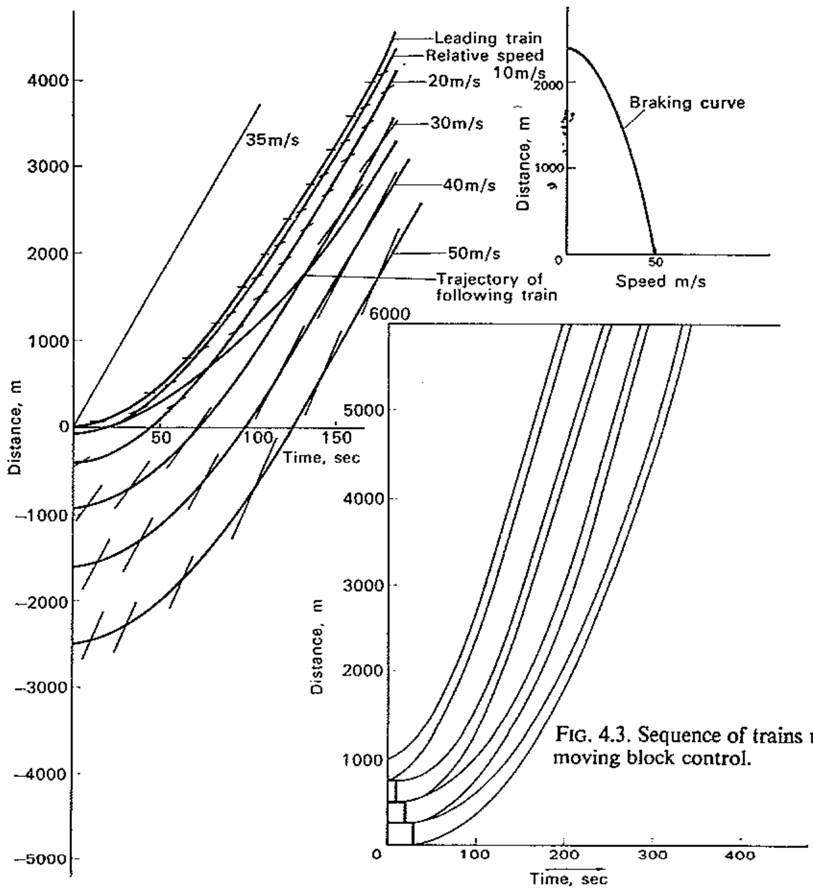


FIG. 4.3. Sequence of trains unde moving block control.

FIG. 4.2. Clearance for moving block system.

speed of 13·4 m/s (30 m.p.h.) he should not accelerate further until he has seen the aspect of the next signal (in theory he could continue to accelerate until he met the braking distance as given in curve 3.4 and then apply full braking but in practice this would place too much reliance upon estimation of distance by judgement only). His speed thus falls behind that of the previous train. However, on reaching E he receives a "double yellow" and can then resume acceleration and passes a "green" before his speed exceeds that permissible on a "double yellow". The following train, however, passes two single yellows before meeting a double yellow so that the period of running at restricted speed is prolonged. The fourth train will experience a similar set of aspects to the third train excepting that it will pass three yellows before receiving a double yellow. It is to be noted, however, that all the trains receive a double yellow at E and subsequent trains can therefore follow each other at uniform headway. This is 105 seconds rather than the 60 seconds based on uninterrupted flow at designed speed. Thus it is only possible to utilise the line to about half its theoretical capacity.

Were cab signalling installed, a driver would know immediately the signal ahead changed to a less restrictive aspect. Thus, in our example the signal at A goes from red to yellow at about the same time as the train reaches 13·4 m/s (30 m.p.h.). The driver does not know this, however, and continues to drive as though there were a "red" at A. With cab signalling he would become aware of the change and could continue to accelerate.

If we now consider the system to be automated and arranged with a speed-regulating system which maintained correct braking distance between any two trains as in Fig. 4.2. In the event of a delay they would draw up close together as in Fig. 4.3. It is assumed that the system for determining permissible speed operates at intervals of at least 10 seconds. Ten seconds after the first train moved off the second train would commence to move at a rate of acceleration so controlled as to leave a continually increasing gap between the two trains. This gap would increase monotonically until the speed and braking distance for optimum through-put was reached.

Thus, whilst the only contribution to increased track capacity made by the automatic moving block in the steady state condition may be the elimination of sighting distance and overlap, the advantage under transient conditions appears to be very great.

4.5. Statistical aspects of car-following behaviour

The treatment of car-following behaviour and route capacity presented in Chapter 2 depended on the assumption that individual drivers exhibited a sufficiently uniform behaviour for generalised laws to be valid. Some consideration of the probability of such an assumption is obviously warranted.

In the case of uncrowded roads each driver will act on his own and it can be shown that the number of cars passing any point in a given time interval "t" will be governed by a Poisson distribution as in equation (4.1),

thus
$$P(n) = \frac{\epsilon^{-\lambda t}(\lambda t)^n}{n!} \qquad (4.10)$$

where λ equals average number of vehicles in unit time averaged over a long period.

There will be some intervals when there is no car present, i.e. there will be a gap. Tl probability of this occurring is obtained by substituting o for n.

Thus
$$P(o) = \frac{\epsilon^{-\lambda t}\lambda^o}{o!} = \epsilon^{-\lambda t}. \qquad (4.1$$

If λ is 0·1 vehicles per second and T is in seconds, then the probability that the ro would be clear for 1 second would be

$$\epsilon^{-\cdot 1} = 0{\cdot}9048.$$

For more realistic times, i.e. for a gap of 10 seconds, we can substitute λT for λ, i. λT is the average number passing in T seconds.

Thus $P(o)$ for a 10-second gap is $\epsilon^{-1} = 0{\cdot}3678$.

Such an expression can form the basis for prediction of the behaviour of interacti traffic streams. For example, a motorist wishing to emerge from a side road and to mer with the traffic stream on a main road will require a gap of minimum time in order to c so. His mean waiting time will be governed by equation (4.10) and becomes

$$W = \frac{1}{\lambda\epsilon^{-\lambda t}} - \frac{t}{1 - \epsilon^{-\lambda t}} \qquad (4.1$$

where W = average waiting time and t = acceptable gap for manoeuvre.[1]

A theory can be constructed describing the existence of any queue of vehicles whi may form behind him. Similar considerations apply to pedestrians requiring to cro traffic streams at uncontrolled crossings. Numerous cases have been studied and report in the literature.[2–6]

When a road junction is controlled by fixed-time traffic signals the length of sign settings can be adjusted in accordance with the estimated flow of traffic on the conflicti routes.[7] For each approach to the junction the signal cycle is made up of an effecti red period during which no traffic departs, and an effective green period during whi traffic is either undelayed or, if there is a queue, will depart at a steady rate equal to tl saturation flow. In setting up the time sequence, certain assumptions have to be ma regarding such factors as lost time between phases, arrival rates and saturation flow Allsop[8] has made a large number of calculations to demonstrate the result of erro made when assigning numerical values to these factors. He shows that the effect of sma errors in the lost times can be large. The introduction of a traffic-responsive automat system (see Chapter 5) can therefore improve the capacity of a junction.

An important point to note in circumstances where speeds and headways vary alo a section of route concerns estimation of the average speed of traffic flow. If the speed vehicles passing any given point is averaged, this, as was shown by Wardrop,[9] gives a erroneously high result. The correct average is the one taken over space, i.e. the avera of the speeds of a number of vehicles on a section of highway taken at one instant of tim

As the amount of traffic on a road increases, drivers no longer act entirely indepe dently but become subject to constraints occasioned by the presence of other vehicles. I particular, a tendency exists for successive vehicles to form themselves into "platoons Prigogine *et al.*[10] have based a theory of traffic flow on the concept that the interacti of vehicles may possess features in common with the behaviour of molecules as describ in the fundamental Boltzmann equation of the kinetic theory of gases.

It is assumed that each individual driver desires to travel at a particular speed and that these desires are distributed in accordance with a particular distribution function $f(x, v)$ when all interactions between vehicles can be neglected

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + V\frac{\partial f}{\partial x}. \tag{4.13}$$

When this is not the case two terms exist

$$\frac{df}{dt} = \left(\frac{\partial f}{\partial t}\right)_{\text{relaxation}} + \left(\frac{\partial f}{\partial t}\right)_{\text{interaction}}$$

where the term "relaxation" represents the tendency of a driver to resume his desired speed and "interaction" represents the presence of other vehicles.

$$\left(\frac{\partial f}{\partial t}\right)_{\text{interaction}} \text{ is given as } C(\bar{V} - V)(1 - P)f$$

where C = concentration of vehicles,
$\bar{V}$ = average speed of concentrated cars,
and P is the probability of passing.

$$\left(\frac{\partial f}{\partial t}\right)_{\text{relaxation}} = -\frac{(f - f^0)}{T}$$

where f^0 is desired velocity distribution and the expression quantifies the desire of the driver to achieve this condition within time T.

$$\frac{\partial f}{\partial t} + V\frac{\partial f}{\partial x} = -\frac{(f - f^0)}{T} + C(\bar{V} - V)(1 - P)f. \tag{4.14}$$

Prigogine[10] points out a particularly interesting singularity where $1 - CT(1 - P)\bar{V} = 0$ which represents a boundary between two flow regimes which he describes as an individual flow regime and a collective flow regime.

Because we are interested primarily in congested conditions where passing is seldom possible, we may neglect P when the velocity of transition becomes

$$\frac{1}{\int_0^\infty \left(\frac{f^0}{C}\right)\frac{dV}{V}}. \tag{4.15}$$

This leads to a flow/concentration curve of the form shown in Fig. 4.4.

The abscissa represents normalised concentration, i.e. concentration divided by the jam concentration when flow is zero. Because concentration diminishes as speed is increased, the portion of the curve at the right of Fig. 4.4 corresponds to the left of those of Fig. 2.4 so that at speeds below the approach to the optimum collective conditions operate and individual factors are only introduced in the region of the optimum and beyond. The collective approach of Chapter 2 is therefore justified.

The existence of an optimum flow rate in the vicinity of a normalised concentration of 0·5 also provides confirmation of the validity of the generalised approach.

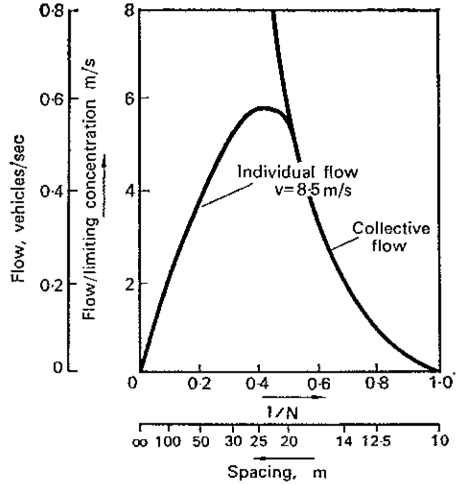


FIG. 4.4. Individual and collective flows.

4.6. Traffic waves

In 1955 Lighthill and Whitham[11] postulated an analogy between the flow of traffi and the flow of a fluid, presenting a theoretical model based on classical hydrodynamics Whilst this approach may be considered to be superseded by the car-following models o subsequent workers, the results present many similar features. A most important con clusion was that there was a characteristic velocity of propagation of disturbances in a traffic stream and that a sudden increase or decrease in velocity of one vehicle wa

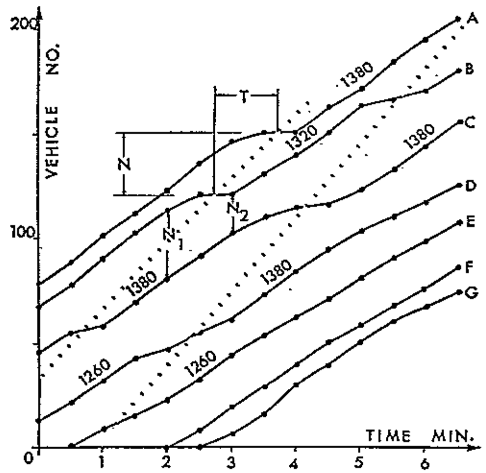


FIG. 4.5. Time-sequence data showing the arrival times of various vehicles at observers A through G. The slopes of the curves give the rates of flow and the figures following the slowdown waves (indicated by dotted lines) give flow rates in vehicles per hour. (After Edie and Baverez.) (Courtesy of Elsevier.)

reproduced in the behaviour of following vehicles being either amplified or attenuated in the process.

Thus if in Fig. 2.3 additional lines were added to represent additional vehicles following behind the two involved therein, the perturbation would be reflected in the trajectories of the following vehicles with a delay in each case. The amount of the perturbation would depend on the value of K, the characteristic speed defined in (2.8).

Such traffic waves are encountered in practice under congested conditions and may even involve following vehicles coming to a standstill from time to time. In extreme conditions multiple collisions may result.

Observations of actual traffic behaviour, particularly by Edie and Baverez[12] in the Holland tunnels of New York, have confirmed the existence of these waves. Figure 4.5 taken from their report shows results of counting of vehicles by seven observers (indicated by letters A to G) spaced about 500 metres apart. From these curves they read the following quantities:

N = number of vehicles passing through a wave as it moves between two observers,
T = travel time of a wave between two observers,
N_1 = number of vehicles in a section just before a wave departs,
N_2 = number of vehicles in a section just after a wave departs,

and from these they were able to calculate the following quantities:

C_j = jam concentration—vehicles/metre $= N/X$ where X = length of section in metres,
C_1 = concentration in section when a stoppage wave is present $= N_1/X$,
C_2 = concentration in section after a stoppage wave $= N_2/X$,
R = reacceleration response time following a stoppage wave $= N/T$,
V = speed of propagation of the stoppage wave,
a = wave amplitude $N_1 - N_2$ (vehicles).

Their findings, converted to S.I. units, are given in Table 4.1.

TABLE 4.1. WAVE BEHAVIOUR
(Reported by Edie and Baverez)

	Section			
	A to B	B to C	C to D	D to E
Length of section (metres)	250	524	572	445
C_j (vehicles/metre)	0·128	0·116	0·122	0·115
C_1	0·086	0·06	0·062	0·066
C_2	0·044	0·038	0·045	0·052
R (sec)	2·04	1·80	1·87	1·83
V (m/s)	3·80	4·79	4·38	4·74
a (vehicles)	11	12	9	5

The agreement between the quantities calculated for sections B to C, C to D and D to E

is good. There is some doubt about the position of observer A. Mean values for B t are as follows:

$$R = 1{\cdot}83 \text{ sec}, \quad V = 4{\cdot}52 \text{ m/s (10·4 m.p.h.)}.$$

The mean value of the ratio $C_1/C_j = 0{\cdot}517$ which corresponds closely with the valu be expected for maximum through-put. It is suggested that, if the rate of flow of tr entering the tunnel could be regulated to something slightly below this value, traffic f would be optimised. Computer control has now been applied to the Holland and Linc tunnels in New York. Vehicle detectors working on the photocell and induction l principles are located in the tunnel roadway. The output from these detectors is fed to computer which is programmed to monitor traffic volume and to predict danger conditions of instability. Flow of traffic into the tunnel is then adjusted by the timin conventional traffic lights so as to avoid unstable car-following behaviour.[13]

Results of the control system operating in the New York tunnels are summarised Table 4.2.

TABLE 4.2. RESULTS OF COMPUTER CONTROL OF TUNNEL TRAFFIC

	Uncontrolled	Controlled
Maximum through-put over half-hour period (cars per lane per hour)	1260	1430
Average speed (miles/hour)	19·3	27·5
(m/s)	8·18	12·3
Average density (cars/mile)	75·8	47·5
(cars/km)	47·0	29·5

Although the increase in through-put is not large, the benefit of increased operat speed will be apparent and it will be noted that this benefit arises not from increasing value of d in equation (2.1) but rather reducing it, that is, by increasing separation of c so that the value of V can be increased.

Similarly, in the Rheinallee tunnels in Düsseldorf, vehicle movements are detected inductive loops installed successively every 100 metres under the surface of each lane. T output is fed to a computer which, in the event of an obstruction, arranges for l diversion to be indicated to minimise congestion.[14]

References

1. ADAMS, W. F., Road traffic as a random series. *J. Instn. Civil Engrs.*, vol. 4, p. 121 (1936).
2. HELLY, W., Simulation of bottlenecks in single-lane traffic flow, *Theory of Traffic Flow*, p. 2 Amsterdam, Elsevier, 1961.
3. WEISS, L. H., The intersection delay problem with correlated gap acceptance. *Operations Res.*, vol. p. 614 (1966).
4. OLIVER, R. M. and BISBEC, E. F., Queueing for gaps in high flow traffic streams. *Operations R* vol. 10.
5. TANNER, J. C., The delay in pedestrians crossing a road. *Biometrika*, vol. 38, p. 383 (1953).
6. GARWOOD, L., The application of the theory of probability to the operation of vehicle-control traffic signals. *J. Roy. Stat. Soc.*, suppl. 7, p. 65 (1960).
7. WEBSTER, V. F., Traffic signal setting. *Road Research Technical Paper* No. 39. H.M.S.O., London
8. ALLSOP, R. E., Effects of errors in lost times on the delay to traffic at an isolated road junction c trolled by signals. *Transport Research*, vol. 7, p. 145, Pergamon Press, 1973.
9. WARDROP, J. G., Some theoretical aspects of road traffic research, *Proc. Instn. Civil Engrs.*, part vol. 1, no. 2, p. 325 (1952).

10. Prigogine, I., Herman, R. and Anderson, R., Further developments of the Boltzmann-like theory of traffic flow. *Proc. 2nd Int. Symposium on the Theory of Road Traffic Flow*, Paris, Office of Economic Cooperation and Development, p. 129, 1965.
11. Lighthill, M. J. and Whitham, G. B., On kinematic waves. II. A theory of traffic flow on long crowded roads. *Proc. Roy. Soc. London*, series A, vol. 22, p. 317 (1955).
12. Edie, L. C. and Baverez, E., Generation and propagation of start-stop traffic waves. Vehicular Traffic Science, *Proc. 3rd Int. Symposium on the Theory of Traffic Flow*, p. 26, New York, Elsevier, 1967.
13. Hauslen, R. A., A computerised tunnel traffic control system. *Proc. 1er Symposium Int. sur la Régulation du Trafic, Versailles*, 1970. Preprint no. 1, p. 69.
14. Freiberg, S., Automatic supervisory and control system at the Rheinallee Tunnel in Düsseldorf. *Proc. 1er Symposium Int. sur la Régulation du Trafic, Versailles*, 1970, p. 61.

CHAPTER 5

Computer Aids to Operation — Traffic Surveillance and Control

5.1. Application of digital computers

Digital computers, which can memorise and manipulate large quantities of numerica data, enable objectively based operating decisions to be made regarding much larger area than was previously possible. When used in association with the principles of contr engineering, a measure of automation can be applied to the overall transport system o say, a town or a group of railway routes.

Some form of computer is to be found in every transport operation at the present tin although the place of the machine in the operational system varies widely.

Computers were first employed to ease the handling of "historical data"—for exampl accounting and the manipulation of transport statistics. Other clerical functions whic were made more efficient were the construction of timetables, duty rosters and se reservations.[1-3]

A more creative use was in prediction. Thus surveys of transport usage and custom preference could be made on a very wide basis.[4] Computer methods can now hand very much more data than was previously possible enabling predictions to be made of t best means for meeting demands of the future.

5.2. Continuous progress control (C.P.C.), Dynamic programming

Any transport system must involve the movement of a great many vehicles. In the ca of passenger services, movements may repeat themselves with great regularity but ofte special circumstances intervene to break up the movement pattern. Freight, however, made up of a predominance of individual items consigned at random to a wide variety destinations. Keeping some form of control over the movement of wagons, for exampl requires the collection, transmission and appreciation of a great deal of data.

Bennett[5] describes the development of a continuous progress control system f informing a wagon controller continuously about the numbers of wagons of each class i marshalling yards, trains and terminals. It also shows whether they are loaded or empt and if empty, whether they are allocated for reloading.

The principle is as follows. "If a full description of the system is available stating t exact state and location of every wagon at a single given instant, then the state of affairs any subsequent time is determinable by modifying the description of the initial system t incorporate the effect of all changes or movements which have occurred in the interi period." This is simply a matter of addition and subtraction but one involving the manipulation of a great deal of data.

Figure 5.1 shows a basic control table for C.P.C. The upper part shows the number of wagons of different types and in different conditions at a number of depots in South Wales and the lower chart shows the number of wagons in transit. The problem is to keep

STATIC																
	Lowfit								Medfit							
	1	2	3	4	5	8	8A	9	1	2	3	4	5	8	8A	9
Tondu																
Blaengarw																
Pontycymmer																
Wyndham																
Ogmore Vale																
Repair shop																
Total																

TRANSIT												
	Lowfit				Medfit				Hyfit			
Train No.	L	E	GC	R	L	E	GC	R	L	E	GC	R
Total												

Code:

'Lowfit' 'Medfit' etc. denote type of wagon

1. Inwards loaded
2. Empty
3. Outwards loaded
4. Crippled
5. Total 1 to 4
8. Net No. wanted tomorrow
8A. No. arranged to meet 9
9. Surplus
L. Loaded
E. Empty
GC. Defective
R. Runner

FIG. 5.1. C.P.C. system of wagon control.

the upper part up to date as a result of arrival and departure of trains and of other changes such as loading of wagons.

Data-processing on the required scale by manual methods would be virtually impossible. Use of a computer can permit some 1000 messages daily to be handled, the computer instantly and continuously recording the current position and giving out answers on demand. Shelley and Holmes[6] claim that as a direct result of the application of C.P.C. the number of wagons required to maintain the important freight business in the Cardiff area has been reduced by 4000, notwithstanding a 10% increase in traffic handled. As many as 40,000 wagons of seventy-five different types might be present in the district at any one time and during periods of peak movement activity, as many as seventy trains might be on the move simultaneously.

The flow of information in a C.P.C. system is shown diagrammatically in Fig. 5.2. The guard or shunter makes up a "Train Consist Record" which is a form on which he records the composition of every starting train. This information is passed by telephone to the computing centre where it is fed into the computer by keyboard 1. The train is given a four-digit number. Subsequent movements of the train, and particularly its arrival, are communicated to the computer but it is no longer necessary to repeat the composition of the train because this information is already stored in the computer, having been recorded from the starting message.

When a train departs, the wagons are deducted from the stock of the originating y and when it arrives its composition is added to that of the arrival station. Thus, once distribution of the wagons of different types is known, the computer will always give distribution of wagons at any other time when interrogated by the interrogating keybo:

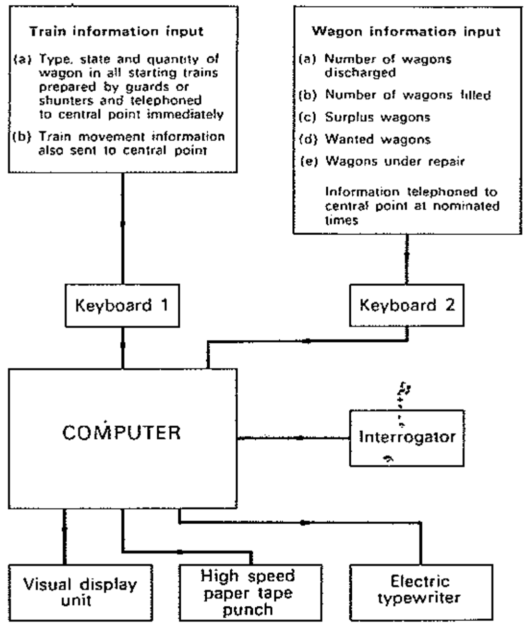


FIG. 5.2. Basic control board for C.P.C.

The output from the computer may be displayed on a visual unit showing partic aspects of the situation. It may also be presented in a typed form or on a punched tape onward transmission or further processing.

The state of the wagons actually in the yards, i.e. whether loaded, unloaded or awai repair, will not be known from the "train consist record" and a second channel of in information is required. This is represented by the rectangle to the right of Fig. 5.2. H at certain predetermined intervals throughout the day, information regarding the stat all stock in hand, the estimates of forward requirements and surpluses and also detail all loadings and discharges since the previous return, are telephoned to the control rc and fed into the computer through keyboard 2.

The computer can now provide completely up-to-date information whereby the of responsible for empty-wagon distribution can order wagon movements to meet estimated demand with great precision. It is to be noted that the computer in this appl tion is used essentially as a monitoring unit rather than as a controller. Action is take a human who makes his own decision regarding future procedure but who uses fac information from the computer in order to do so.

The useful information provided by C.R.C. is the distribution of empty wagons. The ultimate destination of loaded wagons is not known, neither can the location of any particular consignment be traced.

Use of the wagon identification system described in Chapter 7 would enable the guard's duties in making up the "train consist record" to be eliminated. The line-side detectors could be strategically placed and connected direct to the computer by land line so that the area in which every wagon was situated could always be known.

If the yard staff notified the computer at intervals of the loading of individual wagons, as well as their destination, the information would be complete and could form the basis of a positive control.

Alternatively, a wagon could be coded to give its destination and marshalling yards made completely automatic.

Such an extended form of C.P.C. could form a very useful component of an automatic railway system which was controlled by a master computer. A C.P.C. computer could interrogate the master computer at any time to determine the position of any train. Thus the position of any wagon or even consignment could be traced.

5.3. Total operations processing system (TOPS)

Computer-based planning and control systems have been applied on a number of railway operations. One development, initiated on the Canadian National Railway and developed by the Southern Pacific Railroad of the U.S.A., has been adopted by British Rail to help in revitalizing its freight business. The system resembles C.P.C. but operates on a much extended scale. The information which is transmitted and manipulated by TOPS is essentially the status and loading of each wagon. One hundred and fifty Area Freight Centres were established at marshalling yards and major freight depots which were connected with a computer in London. A system of wagon identification applied to trains entering marshalling yards based on television was tried but not adopted, nor were any of the other systems described in Chapter 8.

5.4. Optimum train sequence

Coates and Hawkes[7] have suggested that the computer process of solving problems, known as "critical path analysis", may also be developed for finding the minimum route through a series of time-connected operations. Taking the Borough Market area of the Southern Region of British Railways for example, there are eighteen possible routes through the junction to and from the six platforms of London Bridge Station shown at the right of Fig. 5.3. A matrix can be formed as follows:

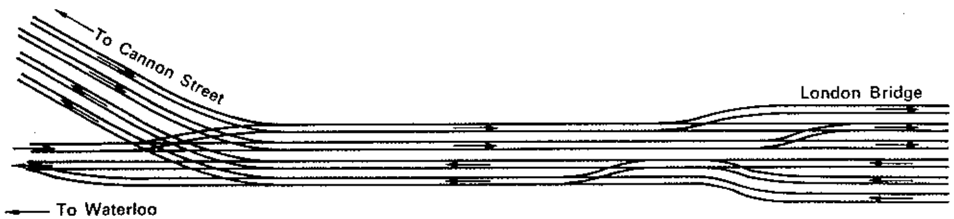


FIG. 5.3. Borough Market junction.

$$T = \left\{\begin{array}{cccccc} & 1 & 2 & 3 & n & 18 \\ 1 & t_{11} & t_{12} & t_{13} & t_{1n} & t_{1\ 18} \\ 2 & t_{21} & t_{22} & t_{23} & t_{2n} & t_{2\ 18} \\ 3 & t_{31} & t_{32} & t_{33} & t_{3n} & t_{3\ 18} \\ & & & & & \\ m & t_{m1} & t_{m2} & t_{m3} & t_{mn} & t_{m18} \\ 18 & t_{18\cdot 1} & t_{18\cdot 2} & t_{18\cdot 3} & t_{18\cdot n} & t_{18\cdot 18} \end{array}\right\} \quad ($$

The elements of the matrix will be the times which must necessarily elapse between a t passing on route m and another on route n. Some routes will not conflict with each o so that their time elements will be numerically equal to zero.

If the proportion of routes to be followed in a given time is known, the problem i determine the order which makes for maximum through-put. A sequence of event ti may be arrived at as follows:

$$\begin{aligned} T_0 &= 0 \text{ (train } a \text{ passes)}, \\ T_1 &= T_{ab} \text{ (train } b \text{ passes)}, \\ T_2 &= \max\,[(T_1 + T_{bc}), (T_0 + T_{ac})], \\ T_3 &= \max\,[(T_2 + T_{cd}), (T_1 + T_{bd}), (T_0 + T_{ad})], \\ T_4 &= \ldots \quad \ldots \quad \ldots \quad \ldots \\ T_r &= \max\,\{(T_{r-1} + T_{r,\,r+1}), (T_{r-2} + T_{r-1,\,r+1}) \\ &\quad \times (T_{r-3} + T_{r-2,\,r+1}) \ldots (T_0 + T_{1,\,r+1})\}, \end{aligned}$$

where T_r is the time at which the rth train passes and the time elapsing between passage of two trains a and b is T_{ab}. This of course, is numerically equal to the figure for the corresponding routes used by trains a and b.

Such techniques are still in their infancy. They depend on the infinite patience of computer in trying combination after combination in its search for the optimum, in case for T_r to be a minimum. There is little doubt that, as the control of vehicles in physical sense is improved, these operations will be linked more closely to the comp until, ultimately, human intervention will be confined to dealing with the abnorma other words, to management by exception.

5.5. Junction optimisation technique (JOT)

Optimisation of train sequence and platform occupation in complex junction areas be achieved in real time by incorporating the working timetable, track layout featu and other relevant information into the train describer computer which accesses location of all trains approaching a junction or station. This computer then proce and continually updates the information and, if actual operations depart from the requirements of the timetable, a revised pattern of working is worked out to minin delay to traffic. Such a system, known as "Junction Optimisation Technique" has b used at Glasgow Central Signal Box to advise signalmen on the best choice of platf and train sequence in the event of delay and a more advanced version, known as "T Regulation Advisory Control", has been developed to deal with multiplicity of probl over a dispersed geographical area such as the approaches to Borough Market Juncti

5.6. Traffic surveillance and control

It has been shown in Chapter 4 that the behaviour of vehicles in traffic streams is governed by two distinct sets of laws, those relating to the desires and performance of individual drivers and those determined by group behaviour of aggregations of vehicles. As congestion increases, the group behaviour patterns are strengthened so as to enable predictions to be made regarding traffic behaviour and, as a consequence, the possibility of optimising that behaviour by control systems is introduced.

On the most elementary plane, the introduction of traffic signals or even policemen on point duty at intersections can increase capacity by replacing the Poisson distribution of gaps by regular interruptions of flow in each direction to allow conflicting movements to take place with safety. The benefit of divisions of traffic on a major road into discrete platoons can be experienced even at uncontrolled inlets situated down stream of the controlled crossing.

Optimisation of the capacity of an intersection relative to the demands of the various routes can be achieved by adjustment of the duration of green phases and a further development is so to time the occurrence of green phases of successive intersections as to allow uninterrupted passage of platoons of vehicles in the appropriate direction.

Figure 5.4 shows a graphical timetable covering three intersections. This particular setting allows co-ordination in both northbound and southbound directions but with a shorter time interval for southbound traffic.

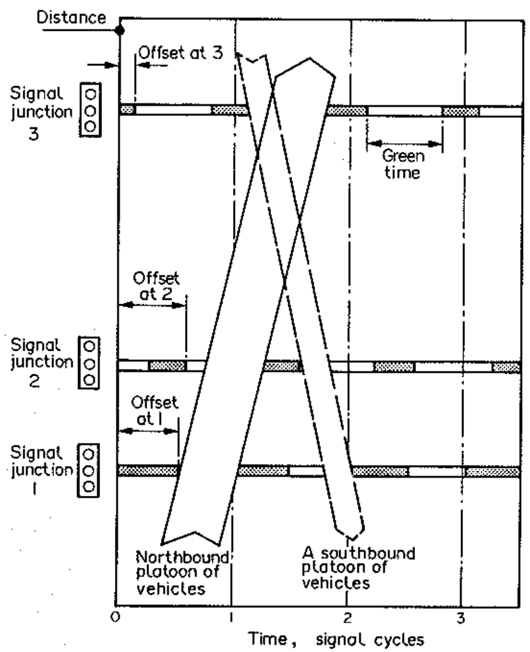


Fig. 5.4. A time–distance diagram that shows signal co-ordination in a fixed time plan.

The next step in evolution is to adjust the duration of green phases in accordance with the actual traffic demand at the time in question and to interrelate the action of signals over a wide area. Such a system only becomes possible when a computer is available which will rapidly assimilate a great deal of input data and manipulate it in sufficient time for a decision to be effective.

In addition to the computer, means are required for the detection of vehicles and lines of communication from the detectors to the computer and from the computer to the traffic signals. An installation to serve a large city is, therefore, expensive and requires a great deal of preliminary study of local conditions. Nevertheless, a number of cities throughout the world are introducing central surveillance and control to a greater or lesser degree and pioneer installations in this country have included Glasgow, Liverpool and West London.

One of the earliest schemes to be introduced was that at Toronto which is shown schematically in Fig. 5.5. Detectors which are sensitive to the presence of a metal mass are placed in selected streets and transmit binary information, i.e. 0 signifies "no vehicle" and 1 indicates its presence. These signals are transmitted through telephone lines in "Multiplex" and held in a scanner which is sampled several times per second by the computer. Thus a 1 followed by a 0 is stored in the computer memory as a one-car count.

The traffic pattern so established is fed to a master computer which optimises the control pattern and then sends back control signals to operate individual traffic lights at intersections. The forms of traffic control so far developed may be summarised as follows:

(a) *Fixed progression*

This permits vehicles to proceed at normal speed whilst experiencing minimum delay in passing a series of traffic lights.

(b) *Volume density*

This is a direct application of queueing theory. A green phase is provided sufficiently long to allow all the vehicles which had arrived during the previous red to clear the signal. This is prolonged only as long as the number of vehicles approaching green is greater than the number waiting on the other limb of the intersection. Traffic is therefore divided up into compact "platoons" so that otherwise-idle periods can be used.

(c) *Variable progression*

The computer calculates average flow of traffic over a grid in each of two opposing directions, inbound and outbound, and introduces appropriate diversions.

(d) *Traffic responsive—local*

The computer balances accumulated vehicle-seconds of delay for waiting traffic with volume of traffic proceeding on green.

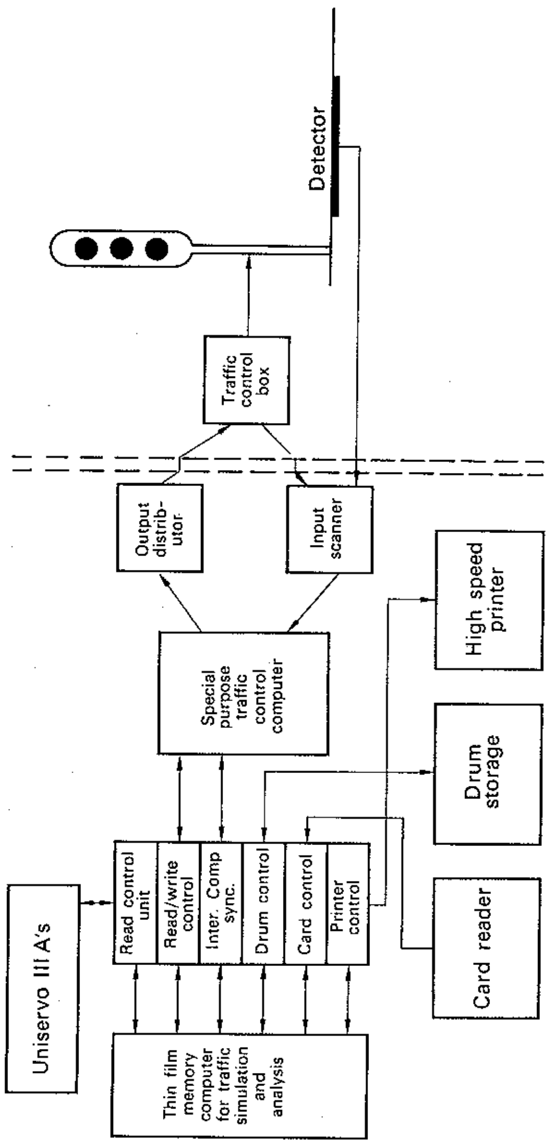


FIG. 5.5. Toronto automatic traffic signal system.

(e) *Traffic responsive—co-ordinated*

A common cycle length is established for all intersections of an artery or grid based o maximum flow in any direction.

A great advantage of computer control is that, in addition to the control action, dat regarding the results of that control action can be automatically recorded so that im proved control strategies can be evolved. An account of the experience gained with th Toronto scheme has been presented by Hewton.[8]

As a result of the successful operation of the Toronto system, numerous advances hav been made partly as the result of the experience of operation and partly arising fron developments in computer technology. During 1975 a computerised traffic signal contrc program was developed which was capable of optimising the cycle events. An evaluatioı of this system, known as the Real Time Optimization Program (RTOP), has beeı published by Rach.[9] The main components are indicated in Fig. 5.6 and consist of (1) a traffic prediction routine to estimate traffic flows on each network link, (2) a com putational routine for individual intersections based on Webster's method[10] anc (3) a network optimisation process for grouping and evaluation of possible cycle length

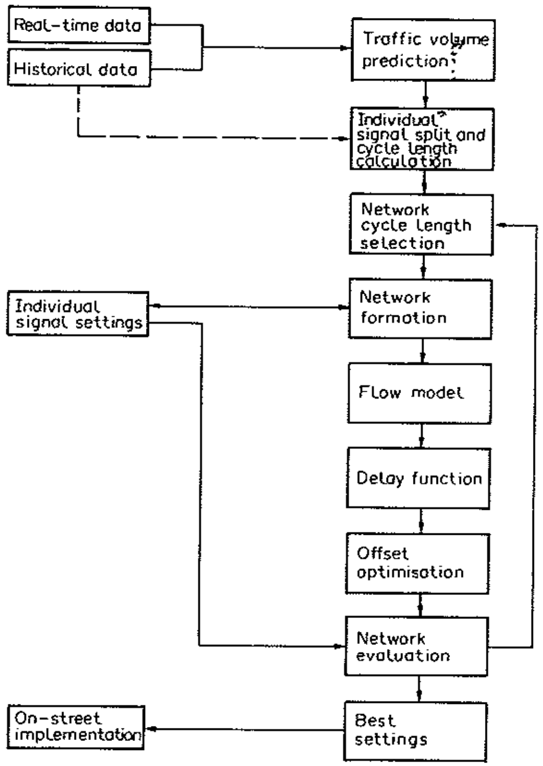


FIG. 5.6. Real-time optimisation program package.

in a co-ordinated network including consideration of those elements which may be excluded from area co-ordination by reason of local constraints.

A comparison of the effectiveness of the RTOP and the fixed time/time of day settings was made on two separate areas selected from Toronto's Metropolitan system which covers an area of 240 square miles and includes 1200 traffic signals. The test areas, one located in the city centre and the other in the suburbs, contained 24 traffic signals and 51 respectively The results of the central area test was a 3·6% improvement and in the suburban area a 6·3% improvement as the result of application of RTOP.

5.7. Split cycle and offset optimisation technique (SCOOT)

A new method of real time optimisation of traffic signal timings has been developed by the British Transport and Road Research Laboratory in collaboration with the Ferranti, G.E.C., and Plessey Companies. Figure 5.7 shows the basic principle of operation. Induction loop vehicle detectors are located on the approaches of all signalised junctions as far upstream as possible. Ideally, if just downstream of the previous signal

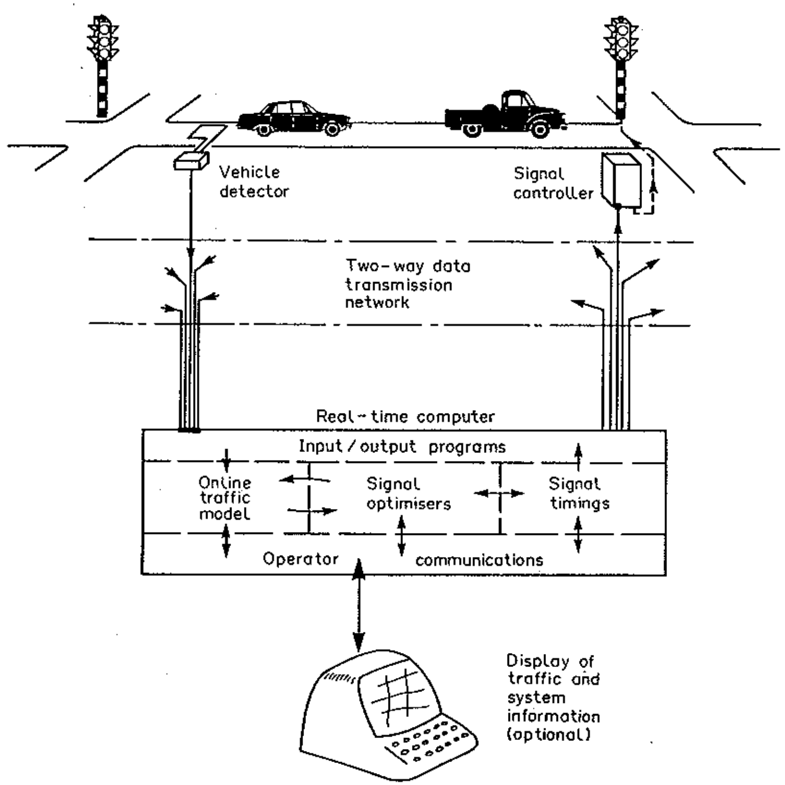


FIG. 5.7. The flow of information in a SCOOT urban traffic control system.

the whole of the traffic stream is detected. A concept known as a traffic-flow profile c be regarded as a histogram of the number of vehicles passing in each instant of time d ing a "green period". This is continuously updated at each cycle of operation and us to predict the duration of each green period. The function of the signal optimiser is use the information provided by the profiles to determine the best overall comprom governing all the streets in the SCOOT area. All the signals in that area operate on common cycle time but the period within that cycle during which each signal shows gre is determined by the computer. Each signal is governed by the "offset" time and t duration. The "offset" is defined as the time within the cycle that a light turns gre relative to a common datum. The "duration" is self-explanatory. A pattern of control carried on from cycle to cycle unless, a few seconds before each event, a signal optimi known as the "split" optimiser estimates that a change is necessary. This part of t programme relates to individual junctions and the duration of the green signal is det mined from the current estimates of queue lengths. Relation between conflicting rou is governed by "offset" changes. The "offset optimiser" uses the information stored the cyclic flow profile to estimate whether or not an alteration to the offset will impro the overall traffic progressions passing through that particular junction. The criterion the performance index (PI) which is defined by Webster and Cobbe[10] as "the ratio the average flow to the maximum flow which can be passed through that intersect from that particular approach".

The traffic capacity of junctions controlled by signals increases as cycle time is i creased. On the other hand, individual vehicles may be delayed for longer times wh stopped at lights. It is usual to set the computer so that the cycle time is used which w just cause 90% saturation at the current value of traffic flow on the most heavily load junction. Other junctions may be lightly loaded so as to permit "double cycling" on cycle time which is one-half of that for the sub-area.

The SCOOT system has been tested in Glasgow and Coventry and the results of traf surveys indicate a reduction of average delays at traffic signals of about 12%.[11]

The system can be adapted to give priority to special classes of vehicles, notably bus by restricting the flow of other traffic onto bus routes which become congested as well providing means for altering signal timings for individual buses fitted with passi responders without undue delays to other traffic.

5.8. Control of cascaded vehicles

The traffic surveillance and control systems can only instruct groups of vehicles discrete points in a system. Where a direct line of communication is open to each vehicl as in an automatic railway or road, corrective action may be applied to each vehicle.

As will be developed in Chapter 15, it is an open question as to how much of the contr system for a locomotive should be located therein and how much control should b exerted by a central computer. Thus, if in a fully loaded system a number of trains a required to follow each other at minimum headway, that headway may be monitored b the central computer and the controls of individual units adjusted so as to eliminate a deviation. Alternatively, the central computer might simply indicate to each vehicle un the desired speed and position. The unit would itself monitor its true position and spee deriving an "error" function by difference from the desired conditions. The acceleration and braking controls would then be acted upon by the "error" signal so as to eliminate the deviation.

Depending on the characteristics of the controls, the latter system might well re-introduce an element of instability similar to that described in the Holland tunnel and referred to in Chapter 4. A conflict may exist, therefore, between the selection of control characteristics for optimum capacity and for stability. However, the system makes minimum demands on the communication circuits and on the time of the central computer.

If sufficient communications exist for the central control computer to be in virtually continuous contact with each train unit, then the former system may be applied. A strategy for this has been published by Powner *et al.*[12] The equation of motion of the vehicles is linearised and accelerations and velocities written in the form

$$M_r\{\dot{V}_r(t)+\Delta\dot{V}_r(t)\}=\left\{\begin{array}{l}F_r(t)+\Delta F_r(t)-\eta_r[V_r(t)]\\ -\dfrac{\partial\eta}{\partial V_r}\Delta V_r(t)\text{—higher}\\ \text{order terms}\end{array}\right\}, \quad (5.2)$$

$$\dot{x}_r(t)+\Delta\dot{x}_r(t)=V_r(t)+\Delta V_r(t)$$

where the suffix r relates to the "rth" vehicle in a sequence, $F(t)$ represents the driving or braking force applied to that vehicle, η represents the drag coefficient of the vehicle and M_r is its mass. V_r is the velocity of vehicle r as initially specified.

For small perturbations of speed the higher order terms may be neglected leaving

$$\Delta\dot{V}_r(t)=[\Delta F_r(t)-\frac{\partial\eta_r}{\partial V_r}\Delta V_r t]/M_r$$

$$\Delta\dot{x}_r(t)=\Delta V_r(t).$$

Spacing between vehicles $= x_r - x_{r+1}$ so that

$$\Delta\dot{V}_r(t)=(\Delta F_r(t)-\frac{\partial\eta_r}{\partial V_r}\Delta V_r(t))/M_r \quad r=1,2,\ldots m, \quad (5.3)$$

$$\Delta\dot{x}_r-\Delta\dot{x}_{(r+1)}=\Delta V_r(t)-\Delta V_{r+1}(t) \qquad r=1,2,3,\ldots(n-1).$$

If there are n vehicles under consideration, a set of $n-1$ differential equations of the form of equation (5.3) can be formulated. The problem then becomes the selection of a combination of values for $\Delta F_1, \Delta F_2, \ldots, \Delta F_n$ so as to provide a satisfactory separation and progression of the appropriate vehicles. The computer would sample the physical system at intervals for x_1, x_2 to x_r and $V_1, V_2, V_3, V_r, \ldots, V_n$ and compute the appropriate values of tractive and braking effort for each vehicle.

Levine and Athans illustrate the method to derive the spacing between three vehicles to enable comparison to be made with analogue computer results[13] (see Chapter 6) and for twelve vehicles with position and velocity errors at the front of the series only. Simulation of the control action showed that full correction could be made in 4 seconds and that there was no appreciable effect on the vehicles subsequent to the seventh.

5.9. System flow charts

The bi-stable elements which go to make up a computer can be regarded as "Yes" or "No" elements and the logic underlying the operation of a system may be displayed on a flow chart comprising three symbols, oblongs to represent the starting and finishing events, diamonds indicating interrogation and rectangles to indicate necessary operations.

Hix[14] has published diagrams indicating the sequence of actions to be performed by the driver of a train and Fig. 5.8 shows the action of a signalman in box A operating the block system as described in Section 3.3. The regulations require that the "call attention" signal must always be given before any other signal and must be acknowledged immediately on receipt. It must always be repeated before the "is line clear?" signal no matter how often the latter has to be repeated.

The first requirement is shown on the diagram by the line running from "No" on the first diamond-shaped box to a point above the rectangle marked CALL ATTENTION. The second requirement is shown by the fact that the line from "No" on the second diamond-shaped box is directed to join the main sequence before the CALL ATTENTION operation is indicated rather than at a point immediately before the IS LINE CLEAR? rectangle. The remainder of the diagram is self explanatory.

In drawing up this flow chart certain simplifications were necessary and, because it relates to one line only, it represents only part of a signalman's activity. Such flow charts, however, demonstrate capacity of a trained man to carry out rapidly and reliably a great many individual operations in succession. Any attempt to automate a system by reproducing electronically each and every operation at present performed manually would present complex problems. The only way forward is to disregard existing procedures and to redesign the system of control, simplifying it as much as possible within the limits imposed by the physical restraints on the system. Thus the direct use of track circuits to prove that sections are unoccupied enables the elaborate system of communication between block posts to be eliminated.

5.10. Train describers

The train describer, originally a device for feeding forward a coded description of a train from block to block, can be associated with a digital computer so that the information may more readily be stored and handled in a complex installation[15] and ultimately used to evaluate the optimum train sequence matrices referred to in Section 5.3.

5.11. Microprocessors

It is just over 10 years since the first microprocessor, the Intel microprogrammed mini-computer, was placed on the market. This minute device embodied 2250 transistors and was followed by a long line of microelectronic devices which have revolutionised electronic design by combining many functions within miniature devices which are readily available at acceptable cost. For example, a microprocessor unit consisting of two printed-circuit boards could provide the same logical functions as 3000 relays in a signal

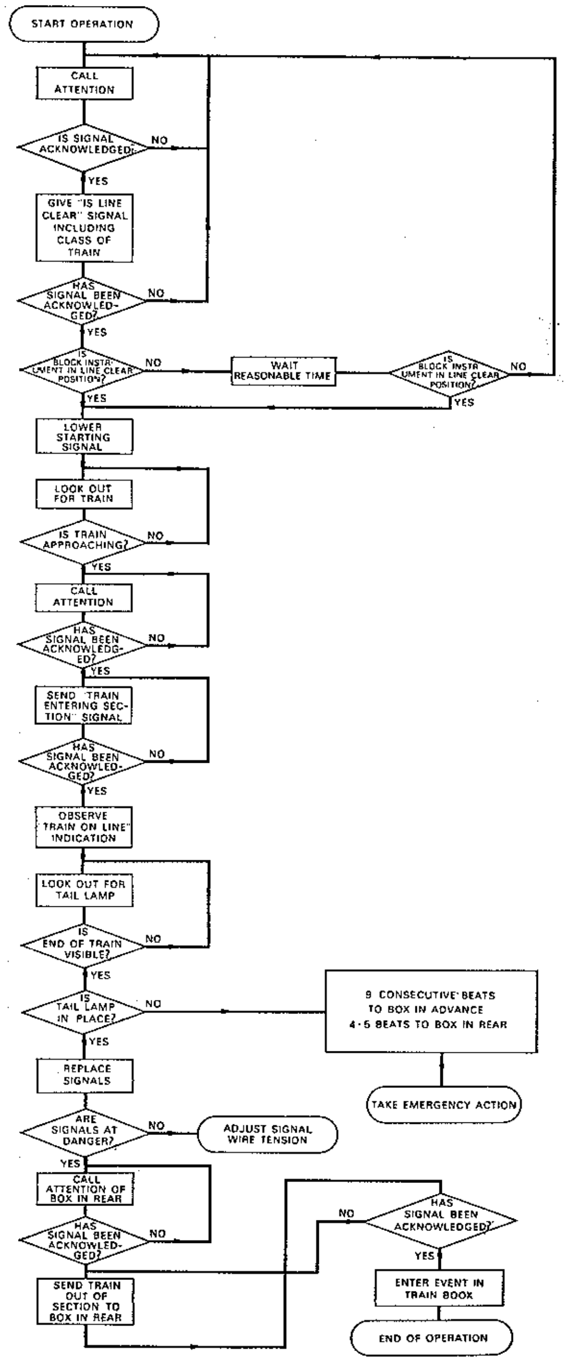


FIG. 5.8. Action of signalman analysed on system flow chart.

interlocking frame.[16] This implies a convenience and economy which should revol tionise transport-control technology. Already there are numerous transport applicatio and it is somewhat strange that these include the interfacing of these very mod devices with archaic technology such as the token instruments on the Inverness–W line (page 38) and in the control system of the Hong Kong Mass Transit System association with power control by rheostats and camshaft (page 246).

Whilst microprocessor elements are inherently reliable, when they do fail there is basis for the presumption of a "right side" failure (that is a failure which cannot lead a "clear" signal indication involving the risk of collision between trains). Indeed "wrong side" failure (which could cause an accident) is equally probable.

To minimise the risk of "wrong side" failures, a degree of "redundancy" has to built into the system. In developments by British Railways Board two levels of redundan are employed—either "duplication" or "Triple Modular Redundancy".[17]

Duplication involves performing the desired tasks in two parallel sub-systems a subjecting the two sets of output information to a process of comparison which indica a failure when the two sub-systems are not in agreement. A "wrong side" failure wou only be possible if both sub-systems failed simultaneously and produced identical outp

A higher level of availability is provided by "Triple Modular Redundancy" (T.M.R which involves performing the desired task in three parallel sub-systems and subjecti the three sets of output information to a voting process. The output will be determin by either three identical results or those of two out of three of the sub-systems. T.M. can provide greater system availability if an organisation exists for prompt repair of t defective sub-system whilst the other two sub-systems continue to function. Each modu contains a redundancy management device accessible to all three processes. This embodi the following features:

—a redundant testable means for enforcing the isolation of a module in the event of majority vote against it;

—means for ensuring that this action is irreversible (a further faulty circuit must n be allowed to reverse the vote);

—automatic reconfiguration of the system when one module is isolated, so that t system becomes a "duplicated" one in which no further failure can be tolerated;

—in the event of failure of the surviving duplicated system means are provided f enforcing irreversible system shut-down to give a "right side" failure.

5.12. Simulation of train-following behaviour

Train schedules and estimates of track capacity are usually calculated on a determini tic basis, whereas all the quantities which determine the motion of vehicles are subject variations of a stochastic kind. The results of the calculations do not therefore correspon with practical experience.

The queueing theory discussed in Sections 4.2 and 4.3 provides a useful first approxim tion but may itself be unrealistic when applied to an actual situation because its predi tions are based on the eventual attainment of a stable situation, whereas in practice th system may require an unrealistically long time to settle down and much interest attach to the early stages of an operating sequence.

The advent of the computer with its ability to make thousands of draws from realistic probability distributions enables the behaviour of real systems to be modelled taking into account such factors as the variation in the operational characteristics of individual vehicles. Thus in a rapid-transit system running at full capacity one might expect trains to attempt to run at a constant headway but to diverge from this due to variation in the duration of station stops. The times required to reach a given speed and stop from that speed would also be probablistic. Repeating a simulation on this basis for a number of following trains is similar to a number of queues in series and as such is much better studied by simulation than analysis.

The first requirement is to provide representation data which must be characterised by the appropriate range and distribution function. In a model which simulated a sequence of inter-city trains[18] these were assumed to arrive at the first station at constant time headway between trains, to wait at the station for a constant period plus a negative exponentially distributed lateness and then to accelerate to cruise speed in a normally distributed time. Minimum headway was prescribed in front of a train as it accelerated and ran at cruising speed. Figure 5.9 shows the result of a simulation of 201 trains with the following parameters: headway for arrivals at first station—150 seconds, deceleration time—mean 100 seconds, standard deviation 10 seconds. Stopping time—minimum 60 seconds plus lateness negative exponential with $\lambda = 4$. The mean acceleration time was 140 seconds with 10 seconds standard deviation and the minimum headway was 111 seconds.

It can be concluded from Fig. 5.9 that, although the system was not stable, the distribution of headways is acceptable. After 201 trains had followed each other through five stations the majority were operating at headways which did not differ seriously from the initial value of 150 seconds. Some sixty trains operated at a reduced headway, i.e. lower than 120 seconds, and only six were seriously delayed (headways greater than 276 seconds).

The computer is required to assign a time duration for each event using pseudo-random number. This is for all practical purposes a random number but is generated by

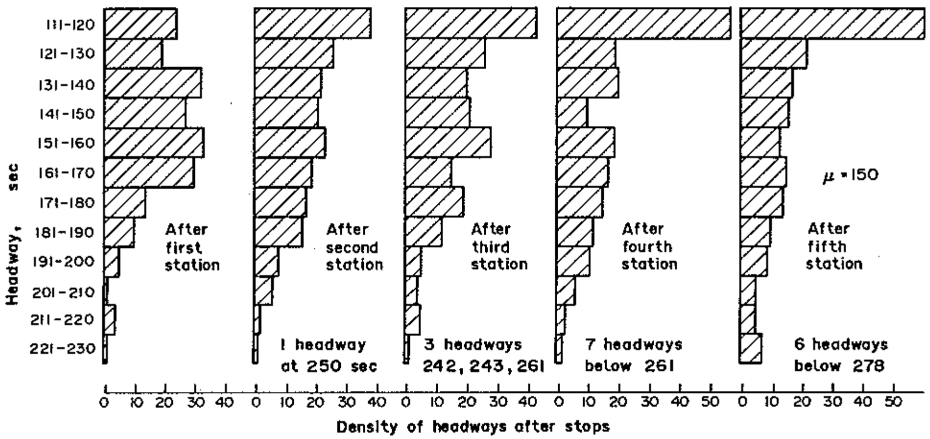


FIG. 5.9. Simulation of train-following behaviour.

the selection of residual numbers produced by the execution of deterministic formulae t long series of places of decimals to obtain the "draw" from an arbitrary single valu probability distribution.

Writing $F(X)$ as the probability that the random phenomenon has a value less than c equal to x. Let this be an exponential density function for example; then

$$f(t) = \lambda e^{-\lambda t} \; t \geqslant 0 \qquad (5.$$

and

$$F(t) = \int_0^x \lambda e^{-\partial t} \, dt = 1 - e^{-\lambda t}. \qquad (5.$$

Given a value of V where $0 < u < 1$ there is a unique value for x such that $F(x) =$ Let V_n denote a uniform random decimal number, then from equation (5.5)

$$V_n = 1 - e^{-\lambda \, xn} \qquad (5.$$

so that

$$x_n = \frac{-\log_e(1 - V_n)}{\lambda} = \frac{-\log_e U_n}{\lambda} \qquad (5.$$

where $U_n = 1 - V_n$, and hence is itself a uniform random decimal number. Thus, in simulation exercise, a series of random numbers U_1, V_2 and V_3 are generated and equa tion (5.7) is used to obtain a randomly exponentially distributed variable.

Of course where a phenomenon is purely random (rectangular), the generated numbe is used direct.

The most common case is that of the Gaussian Distribution. Unfortunately the di tribution function for the normal density with mean 0 and variance 1

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \, dt \qquad (5.$$

does not yield an analytic formula for the inverse function $F^{-1}(a)$. The method of cor volution[19] employs the sum of K independently and identically distributed unifor random variable. Let U_i for $i = 1, 2 \ldots 12$, be independent draws of a uniform rando decimal number; then the value x is computed by equation (5.9):

$$x = \sum_{i=1}^{12} U_i - 6 \qquad (5.9$$

which is approximately normal having a mean of 0 and variance 1. The approximation poor for values beyond three standard deviations from the mean.

References

1. MILLARD, S., The production of working time-tables, passenger station platform workings, locomotive and train crew schedules by computer. *Proc. Symposium on the Use of Cybernetics on the Railways*, p. 133, The International Union of Railways, Paris, 1963.
2. SAISE, K., Automation of J.N.R. seat reservation business. *Proc. Symposium on the Use of Cybernetics on the Railways*, p. 274, The International Union of Railways, Paris, 1963.
3. ROSCHER, H. G., Electronic seat reservation system used by the Deutsche Bundesbahn. *Proc. Symposium on the Use of Cybernetics on the Railways*, p. 279, The International Union of Railways, Paris, 1963.
4. TRESIDDER, J. O., MEYERS, D. A., BURRELL, J. F. and POWELL, T. J., The London transportation study: Methods and techniques. *Proc. Instn. Civil Engrs.*, vol. 39, p. 433 (1968).
5. BENNETT, M. G., Operational research in British transport. *Brit. Trans. Rev.*, vol. VI, p. 3 (1960).
6. SHELLEY, N. and HOLMES, R. G., Continuous progress control. *Brit. Trans. Rev.*, vol. 7, p. 264 (1964).
7. COATES and HAWKES. Private communication.
8. HEWTON, J. T., The Metropolitan Toronto traffic control signal system, *Proc. Instn. Elect. Electron. Engrs.*, vol. 56, p. 577 (1968).
9. RACH, L. S., The development and evaluation of Metropolitan Toronto's real-time programme for computerised traffic control devices. *Proc. IFAC/IFIP/IFORS. Third International Symposium on Control in Transportation Systems*, p. 349 (1976).
10. WEBSTER, F. V. and COBBE, B. M., Traffic signals, *British Road Research Laboratory Technical Paper* No. 56, H.M.S.O.
11. HUNT, P. B., ROBERTSON, D. I., BRETHERTON, R. D. and WINTON, R. I., Scoot—A traffic responsive method of coordinating signals. *Transport and Road Research Laboratory Report* No. 104. Crowthorne, Bucks, England.
12. POWNER, F. T., ANDERSON, J. H. and QUALTROUGH, G. H., Optimal digital computer control of cascaded vehicles in high speed transportation systems. *Transportation Res.*, vol. 3, p. 101 (1969).
13. LEVINE, W. S. and ATHANS, M., On the optimal error regulation of a string of moving vehicles. *Trans. Instn. Elect. Electron. Engrs.*, AC-11, p. 355 (1966).
14. HIX, L., Automatic control and systems aspects of train and railway operation. *Proc. Instn. Mech. Engrs.*, vol. 179, pt. 3A, p. 80 (1964).
15. BOURA, J., SAVAGE, M. J., ALLINSON, J. S. and WILLISON, W. E., The role of computers in train regulation. *Proc. Ier Symposium Int. sur la Régulation du Trafic, Versailles*, p. 59, Preprint no. 4, 1970.
16. SHORT, R. C., The impact of micro-electronics on railway signalling. *I.E.E. Conference Publication* No. 203. *Railways in the Electronic Age*, p. 6 (1981).
17. CRIBBENS, A. H., FURNISS, M. J. and RYLAND, H. A., The solid state interlocking project. *Ibid.*, p. 1 (1981).
18. BARWELL, F. T. and LEECH, D. J., Simulation of train following in HSGT systems, *I.E.E. Conference Publication* No. 117, *Control Aspects of New Forms of Guided Land Transport*, p. 96 (1974).
19. WAGNER, H. M., *Principles of Operational Research*, Prentice Hill, London (1969).

CHAPTER 6

Measurement of Power — Analogue Computing

6.1. Mechanical manipulation of data

A sensible proportion of the power output of a vehicle is absorbed in air resistan[ce] which is difficult to measure. A useful method of testing a vehicle for its total power is [to] use a chassis dynamometer wherein the vehicle remains stationary in space but with i[ts] wheels resting on rollers which are driven thereby. A resisting torque applied to the rolle[rs] represents the resistances due to air or gradients. Flywheels or electronic controls, whic[h] apply a torque proportional to acceleration, represent the inertia of the vehicle. Mo[st] testing is, however, carried out at constant speed. The power output can be derived eith[er] from the resistive torque of the dynamometer or from the pull of the vehicle on t[he] drawbar or other member necessary to restrain its horizontal movement.

Chassis dynamometers for railway locomotives require to be very large and are know[n] as locomotive testing plants. Notable examples exist at Swindon, Rugby and Vitry su[r] Seine.[1,2]

Where the purpose of the vehicle is to act as a tractor, the output of significance is t[he] drawbar pull and tests may be carried out under road conditions using a trailer equippe[d] with means for measuring drawbar pull and distance travelled. Then power equals

$$F\frac{dx}{dt} \quad \text{and work done equals} \quad \int_0^x F\,dx.$$

Such trailers, when used on the railway, are known as "dynamometer cars" and consi[st] essentially of means for measuring drawbar pull and distance travelled. Drawbar pull ma[y] be measured from deflection of a spring or by pressure in a hydraulic system and [a] special wheel, independent of the normal suspension, is used to measure forward motio[n].

The output from these devices is manipulated to provide an output as shown in Fig. 6.1. Here curve (a) represents speed of motion, (b) tractive effort, (c) power developed an[d] (d) total work done. These curves are plotted by stationary pens on to rolls of paper whic[h] are propelled in proportion to the actual motion of the vehicle. The horizontal scal[e] therefore, represents distance. Some form of chronometer completes the input.

Given the inputs F, x and t we have for curve (a)

$$V = \frac{dx}{dt} \qquad (6.1)$$

for curve (b) F recorded directly,

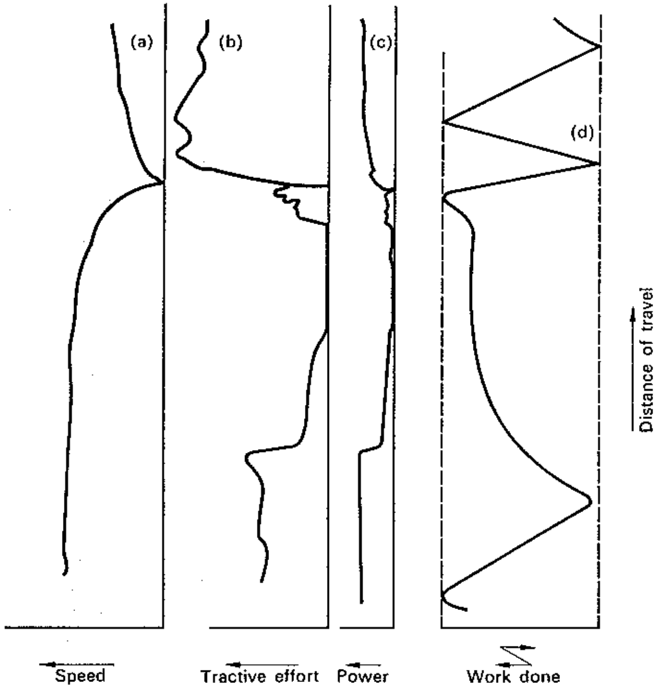


FIG. 6.1. Traces recorded in dynamometer car.

for curve (c) $$H = F\frac{dx}{dt} = \text{(a)} \times \text{(b)} \quad \text{or} \quad \frac{dW}{dt}, \tag{6.2}$$

for curve (d) $$W = \int F dx. \tag{6.3}$$

The operations of differentiation and integration were formerly carried out using the mechanical integrator shown in Fig. 6.2. The edges of two discs placed at right angles ran in contact with a sphere. If the sphere were unconstrained and the discs rotated at different speeds, one for example rotating at constant speed and the other corresponding to the motion of the vehicle, then the motion of the sphere would take place about an axis determined by the vector sum of the two rotations. The inclination of this axis then represented the ratio between the two inputs $\Delta x/\Delta t$ and thus represented the velocity.

Similarly, if the axis of rotation of the sphere were constrained at a certain value to represent an input quantity, say F, and one of the two discs represented a second input, say distance, the output of the other disc would be a function of the two inputs, e.g. $F \times S$ or W.

Torque amplification can be achieved by using a cord wrapped around a drum which is rotated at a constant high speed using the well known relationship between the two tensions, i.e. $T_1/T_2 = \epsilon^{\mu\theta}$.

Thus devices exist which will carry out the actions of integration mechanically and which can be arranged to provide the solutions of differential or integral equations such as (6.3) which can be regarded as a very simple example. Quite complex differential analysers [3] were constructed using these principles but these have now been superseded by the electrical methods described later.

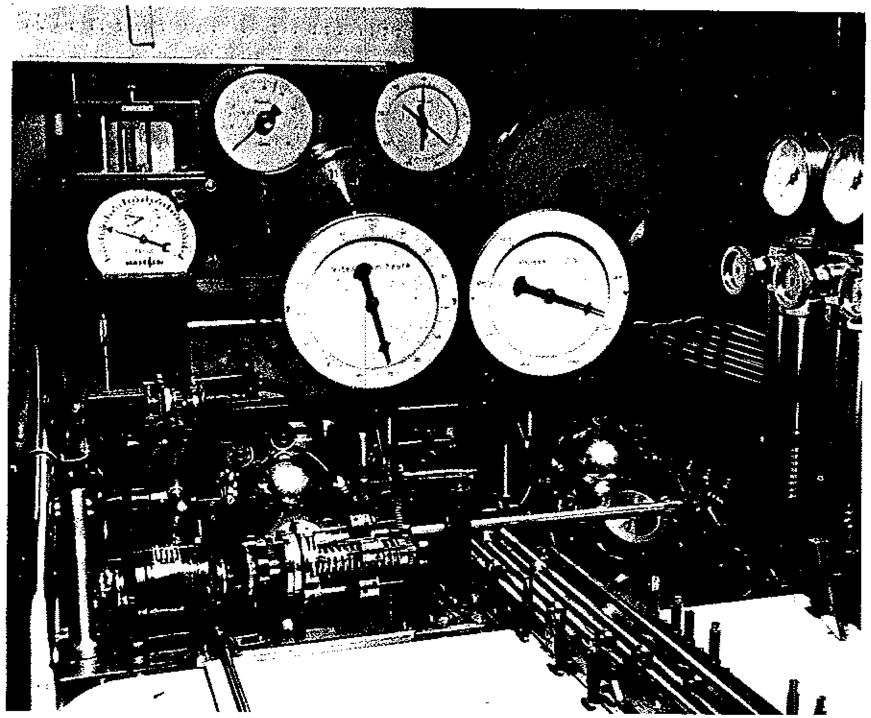
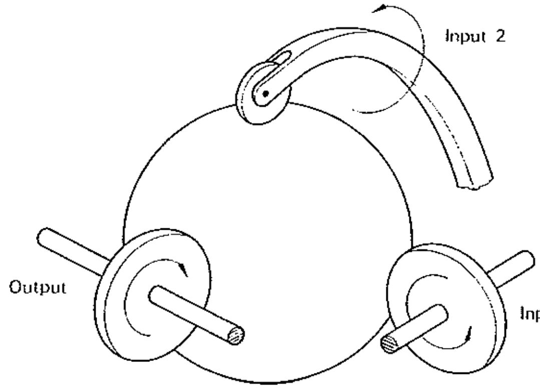


FIG. 6.2. (a) Interior of dynamometer car. (Courtesy of S.N.C.F.)



(b) Mechanical analogue computer. Mechanical manipulation of data.

Incidentally, it is often desirable that the trials of a locomotive should be carried c at constant speed and power irrespective of the configuration of the route. This used be achieved on the Continent by using several locomotives coupled behind the dyr mometer car which assisted or resisted the locomotive under test to maintain consta speed and drawbar pull. An electrical method for achieving this represents one of t first examples of the use of feedback on a large scale.[4] Three vehicles fitted with axle-driven alternators were coupled behind the dynamometer car as illustrated in the Frontispiece. On this car a tachometer measured the actual speed which was compared with the desired speed. This gave an error signal which was amplified and communicated to the trailing vehicles. After amplification the signal was fed to thyratrons on the vehicles which controlled the excitation of the generators so that the torque was increased or diminished in order to restore the desired speed. Because of the time delay of the generators, additional stabilising circuits were provided as described in Appendix I.[5]

6.2. Equivalence of mechanical and electrical quantities

Modern electronic computers represent all quantities by voltages and can be used in two ways, either to simulate the behaviour of the actual components of a system or as a means for solution of differential equations.

Usually more than one consistent system of quantities may be selected to satisfy equations. Table 6.1 shows the equivalence of some electrical and mechanical quantities.

TABLE 6.1

Feature	Electrical	Mechanical
Component characteristics	Capacitance Conductance Inductance	Mass Viscous friction Compliance $= \frac{1}{\text{stiffness}}$
Potential	Voltage	Velocity or displacement
Flow variable	Current	Force

6.3. Potential and flow

To prepare an analysis of a system for treatment by analogue computer it is necessary to construct an equivalent network consisting of circuit elements, i.e. the smallest sections to which a particular impedance can be attributed. Three concepts necessary to complete this process are potential, flow and nodes as defined below.

Potential (across) is a quantity which must be specified with respect to some other point in the system.

Flow (through) relates to the flow of a quantity through a node without reference to the rest of the system.

Node. Whatever the potential, a node is a point in the system where the flow variable is balanced. Thus inward flow equals outward flow as in Kirchhoff's Law.

6.4. Operational amplifiers

Figure 6.3(a) to (f) shows how amplifiers may be used to carry out a variety of operations. Let the point at the centre of the string shown in Fig. 6.3(a) be at the potential v_x. Resistances on the input and output sides are R_i and R_o respectively and let v_i and v_o be the input and output potentials.

Then current $i = \dfrac{v_i - v_o}{R_i + R_o} = \dfrac{v_i - v_x}{R_i} = \dfrac{v_x - v_o}{R_o}$

and $v_x = v_i - R_i i = v_o + R_o i = \dfrac{v_i R_o + v_o R_i}{R_i + R_o}$. (

If now an amplifier is inserted in parallel with R_o so that the output maintains v ; value équal to $-Av_x$ where A is the amplification factor (Fig. 6.3(b)) then $R_o i = v_x -$

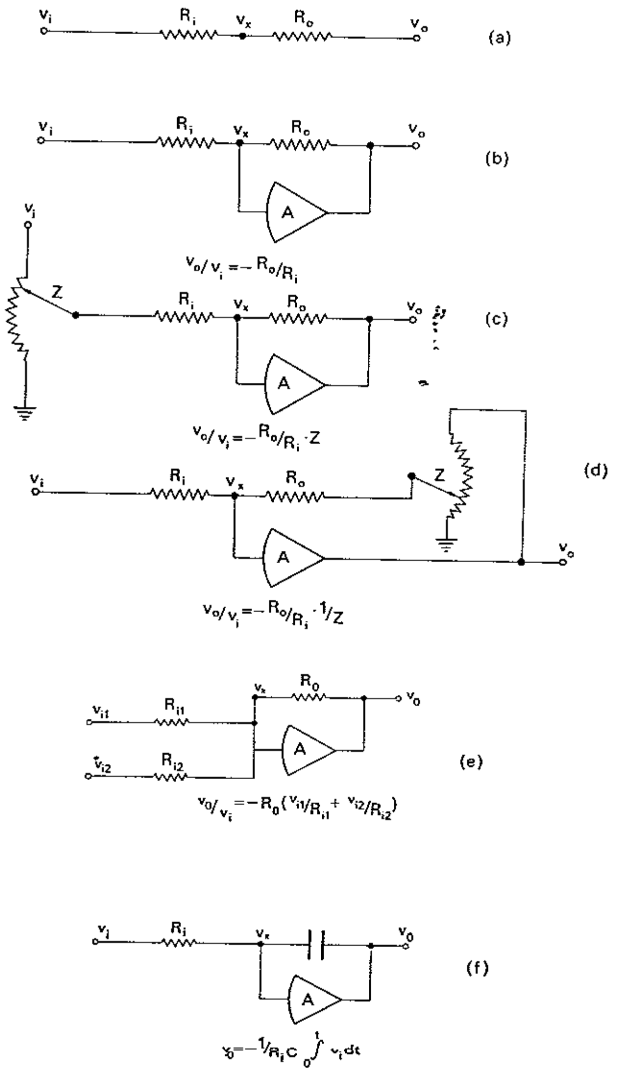


FIG. 6.3. Function of operational amplifiers.

The amplifier is purely a voltage device deriving its energy from an external source so that $R_i i = v_i - v_x$,

therefore
$$\frac{v_i - v_x}{R_i} + \frac{v_o - v_x}{R_o} = 0 \tag{6.5}$$

substituting for v_x,
$$\frac{v_i}{R_i} + \frac{v_i}{AR_i} + \frac{v_o}{R_o} + \frac{v_o}{AR_o} = 0. \tag{6.6}$$

If A is now made indefinitely large

$$\frac{v_i}{R_i} + \frac{v_o}{R_o} = 0 \quad \text{or} \quad \frac{v_o}{v_i} = -\frac{R_o}{R_i}. \tag{6.7}$$

Thus, provided that amplification is adequate, we can adjust the values of R_o and R_i to give a desired value of the ratio of input to output.

Multiplication can be arranged as shown in Fig. 6.3(c) where a variable potentiometer divides v_i into ratio Z which can be varied from 0 to 1,

i.e.
$$\frac{v_o}{v_i} = -\frac{R_o}{R_i} Z. \tag{6.8}$$

Division as in Fig. 6.3(d),

i.e.
$$\frac{v_o}{v_i} = -\frac{R_o}{R_i} \times \frac{1}{Z}. \tag{6.9}$$

Summation as in Fig. 6.3(e),

i.e.
$$v_o = -R_o\left\{\frac{v_{i1}}{R_{i1}} + \frac{v_{i2}}{R_{i2}}\right\}. \tag{6.10}$$

If a capacitor is used in the feedback loop as in Fig. 6.3(f) then

$$i = dQ/dt,$$
$$v_x - v_o = Q/C,$$
$$(v_1 - v_x)/R_i = dQ/dt,$$

$$\therefore \frac{v_i - v_x}{CR_i} - \frac{d}{dt}(v_x - v_o) = 0, \tag{6.11}$$

then
$$v_o = -\frac{1}{R_i C}\int_0^t v_i dt. \tag{6.12}$$

This circuit can then be used for integration.

6.5. Vehicle suspension analogy

Paul[6] illustrates this system by analogy with the vehicle-suspension problem. He considers the situation at one wheel only and ignores the interchange of energy which would occur in the presence of "pitching" or "rolling" motion.

He takes three nodes. Node 0 represents the track, node 2 represents the axle centre an node 1 the vehicle body. Nodes 0 and 2 are connected by the tyre whose stiffness denoted by K_{02}. Node 2 is connected to node 1 by spring stiffness K_{12} and dashpot who: viscous damping coefficient is n. The mass of the vehicle quarter section is M_{01} and th: of the wheel and axle assembly M_{02}. Then at node 1 forces must balance and can be take as the flow variable, thus $F_{21} - F_{10} = 0$. In physical terms the forces exerted on the ax by the vehicle must equal those it exerts on the ground less inertia forces caused by i own acceleration, subject of course to the wheel being in contact with the ground. Thı

$$n(V_2 - V_1) + K_{12}\int(V_2 - V_1)dt - M_{01}\frac{dV_1}{dt} = 0 \tag{6.1.}$$

and for node 2

$$k_{02}\int(V_0 - V_2)dt - M_{02}\frac{dV_2}{dt} - n_{12}(V_2 - V_1) - K_{12}\int(V_2 - V_1)dt = 0. \tag{6.1·}$$

Figure 6.4 shows the equivalent electrical and mechanical networks.

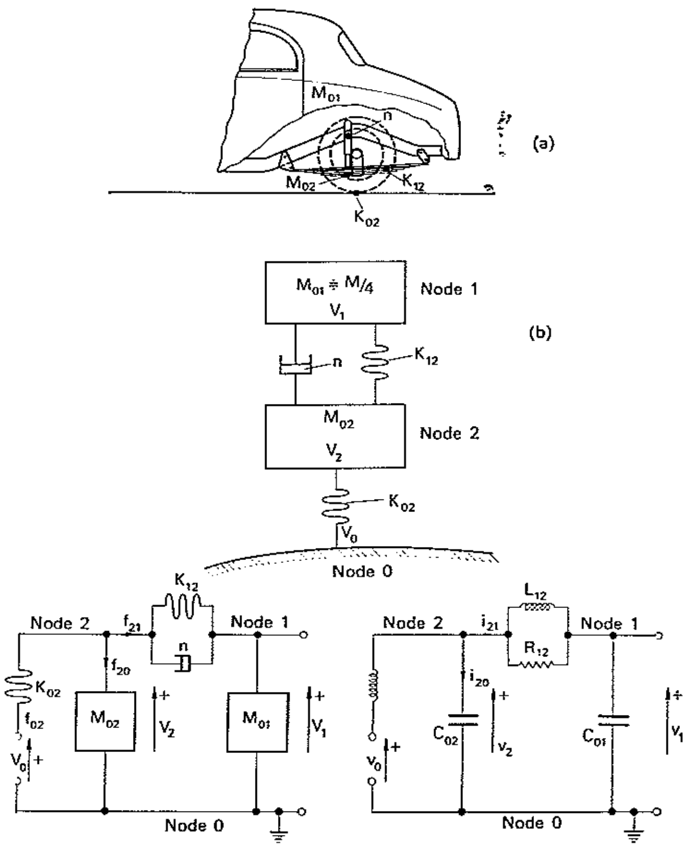


FIG. 6.4. Electrical analogy of vehicle suspension.

So far the physical characteristics of the system under study can easily be recognised from the electrical analogy but further manipulation may cause this correspondence to be lost sight of. The suspension problem may be simplified by neglecting the mass of the wheel-set together with any damping.

Then
$$M\frac{dV}{dt} + K\int V dt = F$$

where F is the force causing displacement from the point of equilibrium. Where functions have to be integrated twice it is convenient to do this separately so that the operation may be broken down into the sequence of Fig. 6.5(a).

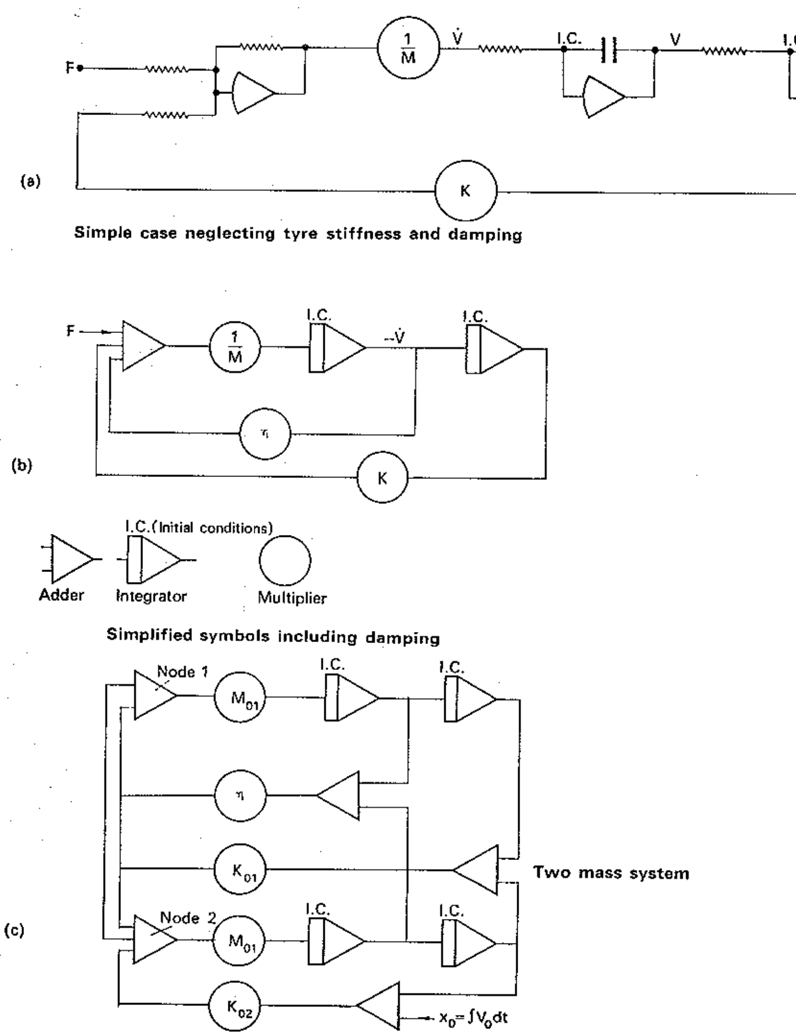


FIG. 6.5. Analogue computer representation of vehicle suspension.

Figure 6.5(b) introduces the conventional symbols and includes the effect of dampin
Figure 6.5(c) represents the two mass systems of Fig. 6.4.

Analogue computers can be used to assess the stability of control systems by setting u the forward path, including any non-linearities of the components, so as to determine tl "open-loop" response to an applied stimulus. This enables the "Nyquist" or "Whitel criteria to be applied[5, 6] (see Appendix I).

6.6. Application to strings of vehicles

The application of digital computers to secure optimum spacing of vehicles wi stability has been mentioned in the previous chapter. An analogue computer simulatic has also been presented.[8] Consider some vehicles moving in a string as in Fig. 6.6. L

$x_k(t)$ be the position of the kth vehicle at time t,

$V_k(t)$ be the velocity,

and M_k be its mass.

$\eta_k(t)$ as in Chapter 1 can be taken as the tractive resistance at time (t) assumed to be function of velocity only and $F_k(t)$ representing the quantity at the right-hand side equation (1.3).

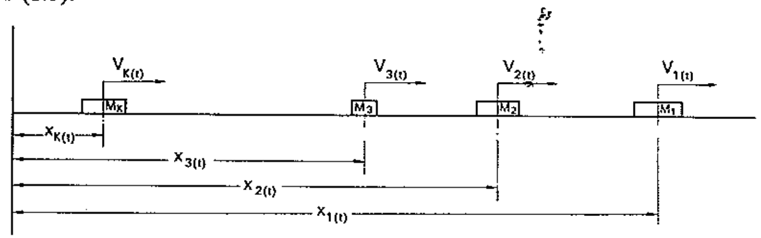


FIG. 6.6. Notation for spaced vehicles.

Then
$$\frac{d}{dt}x_k(t) = V_k(t) \qquad (6.1$$

and from (1.3)
$$M_k\frac{d}{dt}V_k(t) = -\eta_k(t)\,V_k(t) + F_k(t) \qquad (6.1$$

where k equals 1, 2, . . . , n.

The vehicles are required to travel at system velocity v with a spacing of S. The velocit error $\epsilon_{1k}(t)$ at any time t will be

$$\epsilon_{1k}(t) = V_k(t) - V_{k+1}(t) \qquad (6.1$$

and the spacing error ϵ_2 will be

$$\epsilon_{2k}(t) = x_k(t) - x_{k+1}(t) - S. \qquad (6.1$$

The problem is to minimise ϵ_1 and ϵ_2 in the most advantageous manner.

$$\frac{d}{dt}\epsilon_{2k}(t) = V_k(t) - V_{k+1}(t)$$

and
$$M_k\dot{\epsilon}_{1k}(t) = -\eta_k(t)(\epsilon_{1k}(t) + v) + F_k(t) \qquad (6.1$$

$$\eta_k(t)\,(\epsilon_{1k}(t) + v) = \eta v + \sigma_k \epsilon_{2k}(t)$$

plus higher-order terms where

$$\sigma_k = \frac{d\eta_k(t) V_k(t)}{dV_k(t)}$$

where $V_k(t) = v$.
Neglecting higher-order terms

$$\frac{d}{dt}\epsilon_{1k}(t) = -\frac{\sigma_k}{M_k}\epsilon_{1k}(t) + \frac{1}{M_k}\epsilon_{3k}(t) \qquad (6.20)$$

where $\epsilon_{3k} = F_k(t) - F_{vk}$, where F_{vk} is the steady state tractive effort required to propel vehicle k at the system speed v. Then the two errors can be described by

$$\frac{d}{dt}\epsilon_{2k}(t) = \epsilon_{1k}(t) - \epsilon_{1k+1}(t),$$

$$\frac{d}{dt}\epsilon_{1k}(t) = -\frac{\alpha_k}{M_k}\epsilon_{1k}(t) + \frac{1}{M_k}\epsilon_{3k}(t). \qquad (6.21)$$

It is now necessary to define the most advantageous manner of regulating the system. Levine and Athans[8] suggest a function J so that

$$J = \tfrac{1}{2}\int_0^\infty \left\{ \sum_{K=1}^{N-1} q_k\epsilon^2_{2k}(t) + \sum_{K=1}^{N-1} p\epsilon^2_{1k}(t) + \sum_{K=1}^{N} r_k\epsilon^2_{3k}(t) \right\} dt \qquad (6.22)$$

where $q_k \geqq 0$, $p_k \geqq 0$ and $r_k > 0$ for all values of k.

The first term governs the separation of vehicles and ensures that this is always positive. The second term is concerned with velocity variations and the third limits the value of accelerations. Selection of the coefficients q, p and r determines the relative weight to be assigned to each aspect of performance.

Particulars of a string of three vehicles are set out below in matrix form.

$$\frac{d}{dt}\begin{bmatrix}\epsilon_{11}(t)\\ \epsilon_{21}(t)\\ \epsilon_{12}(t)\\ \epsilon_{22}(t)\\ \epsilon_{13}(t)\end{bmatrix} = \begin{bmatrix}-1 & 0 & 0 & 0 & 0\\ +1 & 0 & -1 & 0 & 0\\ 0 & 0 & -1 & 0 & 0\\ 0 & 0 & 1 & 0 & -1\\ 0 & 0 & 0 & 0 & -1\end{bmatrix}\begin{bmatrix}\epsilon_{11}(t)\\ \epsilon_{21}(t)\\ \epsilon_{12}(t)\\ \epsilon_{22}(t)\\ \epsilon_{13}(t)\end{bmatrix} + \begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\epsilon_{31}(t)\\ \epsilon_{32}(t)\\ \epsilon_{33}(t)\end{bmatrix} \qquad (6.23)$$

where $k = 1, 2$ and 3,
thus $\epsilon_{3k} = \epsilon_{31}$, ϵ_{32} and ϵ_{33}.

This can be manipulated as an analogue network: Figs. 6.7 and 6.8 show some results.

The authors comment that such a system would make great demands on communications (4950 inputs for a string of fifty vehicles). A possible compromise would be for each vehicle to be controlled relative to the one in front and the one behind only.

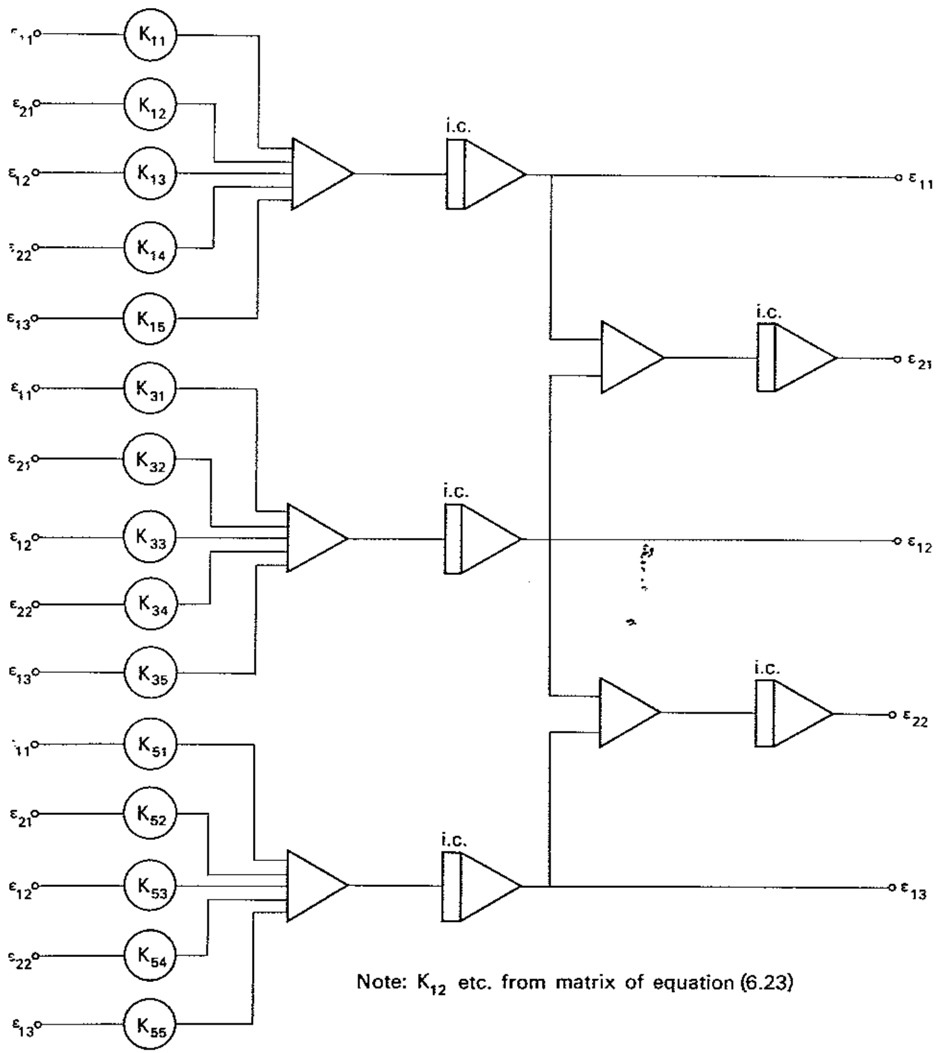


FIG. 6.7. Control of spacing of three vehicles.

6.7. Hybrid computers

It will be apparent that both digital and analogue computers have a part to play in transport. However, situations exist where the best solution is provided by a Hybrid Computer wherein both digital and analogue features are used.[9]

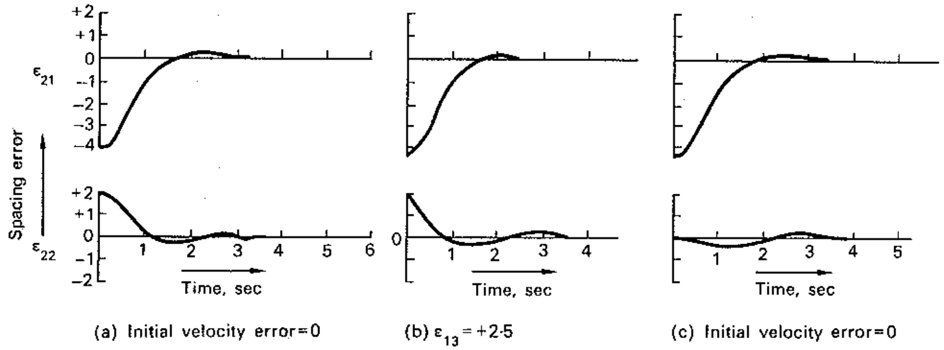


FIG. 6.8. Levine and Athans' results for various initial conditions.

6.8. Simulation of service environment

During testing and development of elements and subsystems it is often necessary to study their performance under the action of control signals which correspond to those which would be applied in a complete system in service. For example, the control mechanism for an aircraft may be tested on the ground, the effect of aerodynamic forces and their alteration in response to control action being simulated by computer.

An example in the ground transport field is provided by a system used by the S.N.C.F. for simulating in the laboratory the action of the braking system in bringing a train to rest (see Fig. 6.9). The braking equipment of a train consisting of, say, twenty-five bogie vehicles was assembled in the laboratory and connected by an array of pipes in a configuration devised to match the time characteristics of the train pipe. The force exerted by the cylinders can be regarded as the output of the rig but, converted into retarding torque, also as the input to the hypothetical train system.

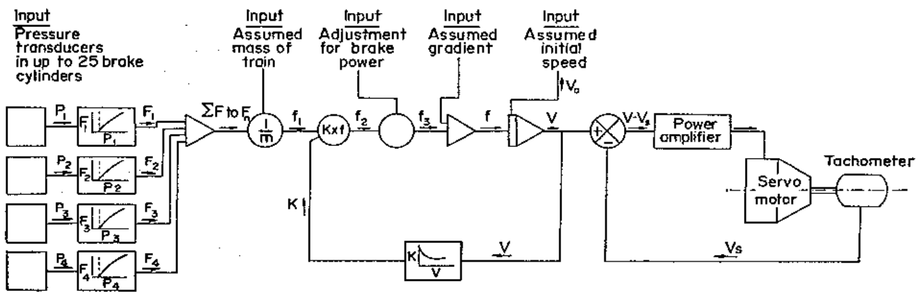


FIG. 6.9.

The output from the pressure transducers in the brake cylinders is fed to functi generators which allow for the lost motion in the brake gear and for the coefficient friction between brake block and tyre so that outputs F_1 to F_n are each proportional to tangential force exerted on a wheel set. These outputs, when added together, represent total retarding force which is then divided by the mass of the train to give a crude value deceleration equal to f_1. The coefficient of friction is, however, a function of speed. feedback loop is therefore provided which embodies a function generator giving an o put K representing the value of the coefficient of friction for any input speed. T correction gives f_2 which is the deceleration of the train on level track. A facility adjusting brake power is added at this stage followed by an input which adds or subtra from the deceleration an amount determined by gradient which can be preset or modif during a test. The final deceleration f is then integrated to give velocity of the train at a time after the instant of application of the brakes. This is represented by the init conditions V_0 fed into the integrator. A servo motor is driven at a speed proportional that of the hypothetical train through the power amplifier feedback circuit. The to angular rotation of the servo motor is therefore proportional to the distance travelled the hypothetical train after a brake application.

References

1. CARLING, D. R., Locomotive testing stations. *Trans. Newcomen Society*, vol. XLV, p. 105 (1972-
2. PLACE, P., Locomotive testing plants (with special reference to the testing plant at Vitry). *J. Inst. Lo Engrs.*, vol. XXV, p. 380 (1935).
3. ROSE, H. E., The mechanical differential analyser, its principles, development and applicatio *Proc. Instn. Mech. Engrs.*, vol. 159, p. 46 (1948).
4. ANDREWS, H. I., Locomotive test plant of the London, Midland and Scottish Railway. *Proc. Ins Mech. Engrs.*, vol. 158, p. 450 (1948).
5. WHITELY, A. L., Theory of servo systems with particular reference to stabilisation. *J. Instn. El Engrs.*, vol. 93, pt. II, p. 353 (1946).
6. PAUL, R. J. A., Review of analogue computing techniques. *The Use of Computers in Mechani Engineering*, p. 1, Instn. Mech. Engrs., 1963.
7. NYQUIST, H., Regeneration theory. *Bell System Technical Journal* (1952).
8. LEVINE, W. S. and ATHANS, M., On the optimal error regulation of a string of moving vehicles. *Ins Elect. Electron. Engrs. Trans. Automatic Control*, Vol. AC-11, p. 355 (1966).
9. PAUL, R. J. A. and ROWLEY, G. C., Hybrid computing techniques. *The Use of Computers in Mechani Engineering, Proc. Instn. Mech. Engrs.*, p. 16 (1965).

CHAPTER 7

Vehicle Detection

7.1. Presence detectors

Any system for remote control of vehicles requires some means of detecting their presence and, if possible, their velocity.

One form, consisting of three co-operating inductive loops, has been mentioned in Chapter 2 in connection with automatic road operation. Detectors can be made to work using any physical characteristics of the vehicle. Thus the weight of railway vehicles has been used to detect their presence by arranging for the flanges of the wheels to depress treadles placed adjacent to the rails. Gravity is also used in the familiar pneumatic tubes laid into the surface of a road at vehicle-actuated traffic lights.

Conductance and capacitance have been employed and magnetic nuclear resonance has been proposed for detection of vehicles.

Where inductive loops are used in road systems an approximate estimate of the speed may be obtained by noting the duration of the pulse. This will vary, of course, with the length and electromagnetic properties of the vehicle, but an approximate expression based on Toronto experience (see Chapter 5) is of the following form:

$$V = \frac{(A - Bp)}{p} \tag{7.1}$$

where V = speed in m/s and p = duration of pulse in seconds. Values of A and B of 5 and 1·5, respectively, provide a correlation coefficient of 0·98 between true and calculated speeds. For practical control purposes speed is best estimated by summing the total detector pulse output for a finite time during which a number of vehicles will have passed.

The longest experience of vehicle-presence detection has been that of track circuits used on railways and tramways.

7.2. Track circuits

As a safeguard against the risk of a signalman forgetting a train which had been held for a long period at a signal, the railways adopted a rule—Rule 55—which stated that when a train had been brought to a stand owing to a stop signal being at danger, the driver must sound the engine whistle. If the train was still detained (after 3 minutes in clear weather, immediately during fog or falling snow) the guard or fireman was required to go to the signal box to remind the signalman of the position of the train and to remain there until permission was obtained for the train to proceed.

This provision could be very irksome and time wasting when the signal in question some distance from the signal box and various devices were adopted for indicating in signal box the fact that a train was held at certain stop signals. The most successful the track circuit invented in 1872. The section of line adjacent to the stop signal insulated electrically from abutting track at each end and a battery was connected ac the two rails at one end and a d.c. relay at the other (Fig. 7.1). Current flowed from

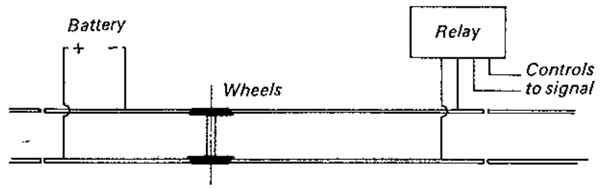


FIG. 7.1. The track circuit.

battery, along one rail, through the relay and back to the battery via the other rail. Wl a train was present on the section, the wheels and axles completed the circuit ther by-passing or shunting the relay which then "dropped out" and transmitted a " occupied" indication to the signal box. It will be noted that the device has a "fail safety" character because, should current supply fail, the relay will drop out and give danger signal. A measure of broken-rail protection is afforded because an open cir indicates danger although all broken rails do not necessarily lead to open circuits.

A certain amount of leakage of current will occur through the sleepers and ball particularly in wet weather, so that the relay must discriminate between the "train shu and the leakage. It will be required to "pick up" at a certain value and to "drop out' another value which is above zero. The conductivity of the wheel and axle circuit mus regarded as finite to allow for such factors as rusty or dirty rails.

The relay used for simple track circuits consists of soft iron cores which are wound a soft iron armature. The material used has low retentivity. The difference between "pick-up" and "drop-away" current should be as small as possible so that the chang the value of the air gap as the relay operates must be restricted. This is achieved by fitt stops which ensure the existence of a definite air gap in the energised position. Tyle describes various types of signalling relay.

Where track circuits are used to indicate the presence of a train at a stop signal, t signal is marked with a white diamond-shaped sign on the signal mast. This indicate the train crew that they need not comply with rule 55.

From this simple beginning the track circuit has been applied to facing point locking indicate the state of occupation of the track at points not visible from the signal box finally to form the basis of complete automatic signalling.

Where electric traction by d.c. is operated, interaction between traction return curr and signalling current is avoided by using a.c. in the track circuits.

There are two systems used, known respectively as "single-rail" and "double-r track circuits. In the former, one rail is bonded throughout and provides the path for traction return current. The other rail is provided with insulated joints at the extremi of the track circuit sections. Train-detection voltage is applied across the track so that unbonded rail provides a common return and the presence of a wheel set shunts the signalling current. The relay, however, is unaffected by traction return currents.

This system has the advantage of simplicity but the disadvantage that only one track rail is available for the return of traction current.

The double-rail track circuit is designed to overcome this disadvantage. Joints at sectioning points on both rails are insulated and cables are provided which are of sufficient cross-section to carry the traction return current. These are led to "impedance bonds" laid between the track which, whilst permitting the easy passage of d.c. current, offer a very high impedance to restrict the passage of a.c. signalling current. The relay used is also unaffected by d.c. and is usually of the two-element type, having two separate magnetic circuits, one an exciting circuit fed directly from the supply and another which embodies the control function. Of the various types of two-element relay, the induction motor type will be most familiar to electrical engineers. This is similar in principle to a squirrel-cage induction motor but usually has a laminated iron core which is fixed in space, the rotor proper taking the form of an aluminium cylinder. This is carried in bearings and drives the control mechanism by means of a pinion. The stator carries two windings, the exciting winding and the control winding. These windings are displaced so that torque on the rotor depends on the relative magnitude of the two currents and the phase angle between them. The presence in the control circuit of stray currents, d.c. or any other frequency than the exciting current, does not affect the relay provided they are absent from the exciting current.

For modern electrification at 50 c/s for traction, other frequencies, e.g. $83\frac{2}{3}$, are adopted. It is important to ensure that false operation of relays, due to the interaction of harmonies of the two frequencies, should not occur. Using $83\frac{2}{3}$ and 50 c/s the lowest harmonies to coincide are the 5th harmonic of 50 c/s and the 3rd harmonic of $83\frac{2}{3}$ c/s.

Direct current can be used provided the relays are immunised against operation by a.c. by fitting copper rings known as "slugs" on to the cores between the coils and the armature air gap. These massive copper rings act as short-circuited secondary transformer windings so as drastically to attenuate any a.c. flux component. An auxiliary magnetic path is provided for a.c. flux, avoiding the air gap.

Simple d.c. track circuits may be used up to lengths of about 500 metres whereas a.c. double-rail circuits are suitable for distances up to 1500 metres.

7.3. Jointless track circuits

The use of long, welded rails is rapidly increasing and the necessity for providing joints therein simply for signalling purposes is to be avoided. Two methods for jointless track circuiting are now available, one based on the use of different frequencies, e.g. 1500 c/s and 2500 c/s in adjacent sections, and the other one based on pulsed sweeps. Figure 7.2 shows the section terminated by bonds FC and JO shorting the two rails. These bonds have an appreciable impedance. Taking the subcircuit BEDC, power at frequency f_1 is applied across condenser C_{10} which, together with the rail BC and part of the "Z" bond CDE, form a resonant circuit. A similar resonant circuit is formed by C_{12}, MJ, and JKL which serves for reception. A train on rail AF or OH would cause detuning and between F and O de-energisation of the receiver by shunting in the normal way.

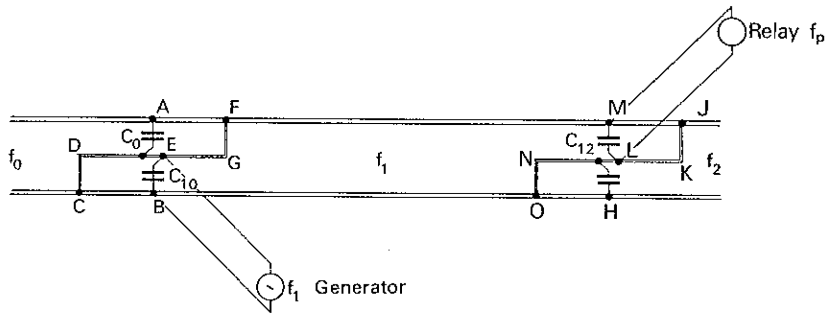


FIG. 7.2. "Aster" jointless track circuit.

In an alternative proposed by Ogilvy (Fig. 7.3), short circuits AB and CD terminate th track circuit section. A pulse is induced into the loop so created by loop E. A simila loop F detects the pulse if the track is unoccupied and after a predetermined tim interval, transmits a similar pulse. Thus the arrival of pulses indicates at both ends of th

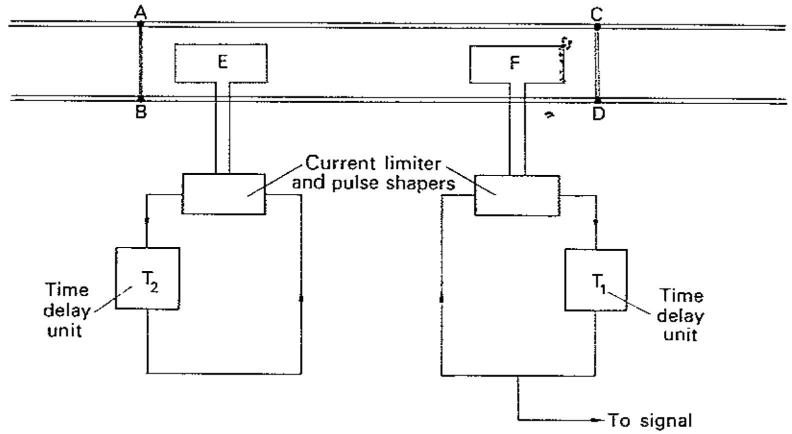


FIG. 7.3. Ogilvy's proposed electronic track circuit.

section that the line is clear. Presence of a train shunt will destroy the pulse and cause th oscillation to cease. It will be noted that for a clear indication to persist, pulses have t traverse the section in both directions. This adds to the sensitivity of the system because should an axle provide a poor shunt to a pulse coming from one direction, it will probabl respond to a pulse from the other direction.

7.4. Guided radar

The track circuit suffers from considerable attenuation and is very limited in the amoun of information which it can convey.

The possible use of pneumatic tyres or air cushions will, in any event, preclude its us and a search for alternative means for train and obstacle detection is taking place i

Europe, Japan and the U.S.A. Some form of surface waveguide would appear to provide an attractive solution.

The simplest type of surface waveguide, the Goubau line (familiarly known as the "G" line), has been proposed for obstacle detection in railway operation.[2] It consists simply of a single conductor covered with a uniform thickness of low loss dielectric. The theoretical description of the mechanism of wave propagation is complex but comparison with the field of coaxial conductors provides a useful, if simplified, explanation. Figure 7.4 shows the electric field of the coaxial two-conductor transmission line, firstly with the outer conductor close to the inner, secondly with the outer conductor remote and finally, absent. In the first case all lines radiate from the inner to the outer; in the second case, where the separation is comparable with a half wavelength, the lines tend to originate and terminate on the inner as well as the outer conductor (because the potential differences are similar). Thus, smaller currents flow in the outer conductor and finally become vanishingly small as the separation is increased. The energy is now propagated entirely as a wave motion on the surface of the inner conductor, with a radial evanescent field—the so-called Sommerfeld wave. The effect may be enhanced by increasing the surface reactance (using, for example, a dielectric). The thickness of the dielectric will control the radial decay of field, which is also a function of frequency. The radio wave is, therefore, caused to adhere to the surface of the "G" line and consequently may be guided along the curves of a railway track clear of the usual trackside obstacles. Conventional transmission-line theory is relevant in terms of effects of discontinuities so that obstacles lying within the field will create a change of line impedance, producing the usual reflections.

Experimental observation has indeed shown that trains or obstacles may be detected by radar techniques using a "G" line although the circular field introduces problems of support and, more seriously, causes severe losses when the line is mounted within a few inches of ground level. Figure 7.5 shows an experimental installation erected near Edinburgh, by Ferranti Limited, under contract to British Railways. Operating at 900 MHz the attenuation of a typical line has increased to over 35 dB/mile compared with a theoretical value of 8 dB/mile (in free space) when mounted 12 inches above ground. Clearly this is unacceptable and points to the need for a surface wave guide supporting an asymmetric field, with the energy concentrated away from ground and supports.

A technique suggested by Barlow[3] resembling a combination of "G" line and coaxial line, presents a possible solution. In this arrangement (Fig. 7.6) the "G" line is partially encircled by a concentric conductor coated internally with a low loss dielectric.

Barlow has shown that if the surface reactance modified by the dielectric is suitable and surface resistance is slightly negative, the concentric screen will appear matched to the evanescent field of the "G" line permitting the wave to propagate along the guide in "free space" conditions, the field extending through the slot for detection of obstacles. The concentric conductor, being entirely passive, cannot of course be given the property of negative resistance; if the positive losses are maintained at a low value, however, the mismatch introduced will be insignificant, the vector impedance being governed by the very much greater magnitude of the reactive component.

Granted a satisfactory conclusion to the current research, many engineering problems will need to be solved to make a practical installation possible. The prize for success, however, is attractive. Apart from detection of obstacles for safety in high-speed operation, pure moving-block control becomes feasible without undue complexity; many

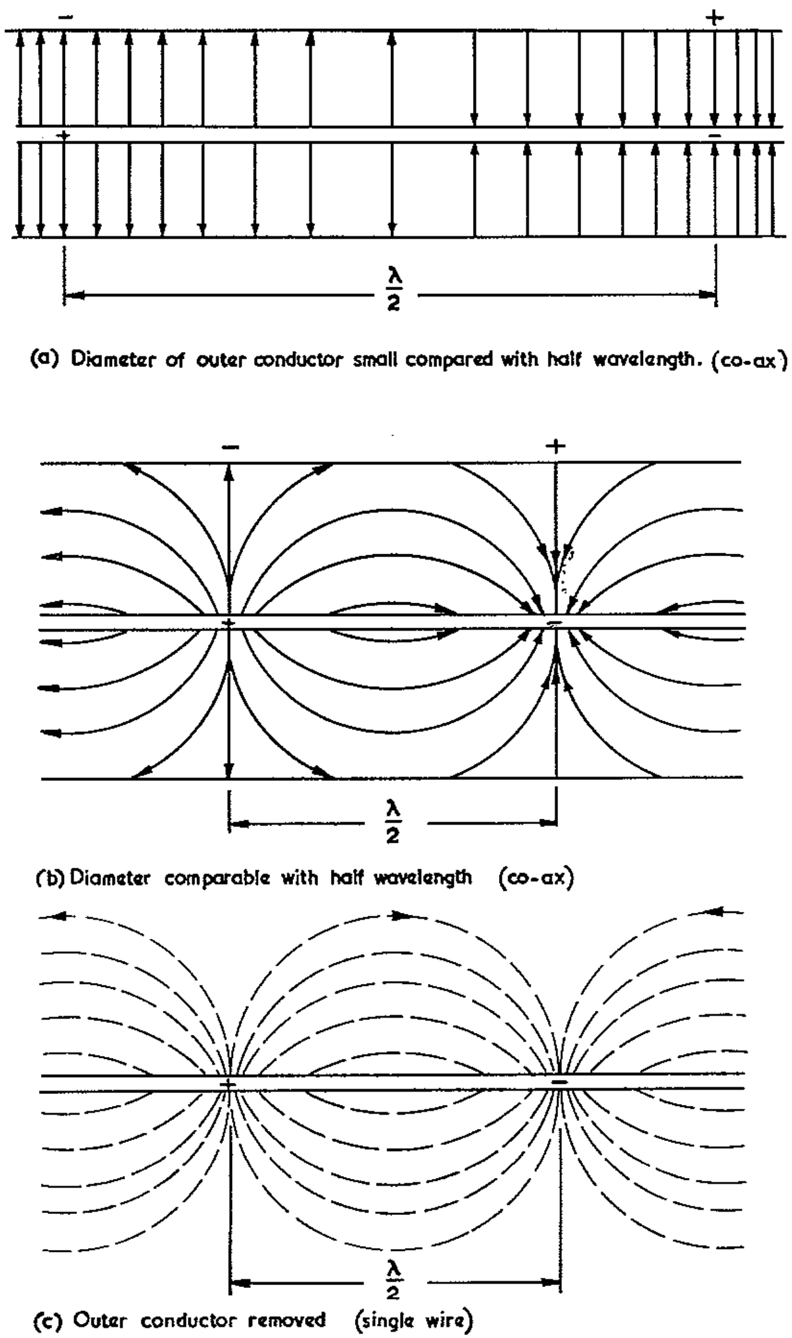


FIG. 7.4. Stages in surface wave evolution.

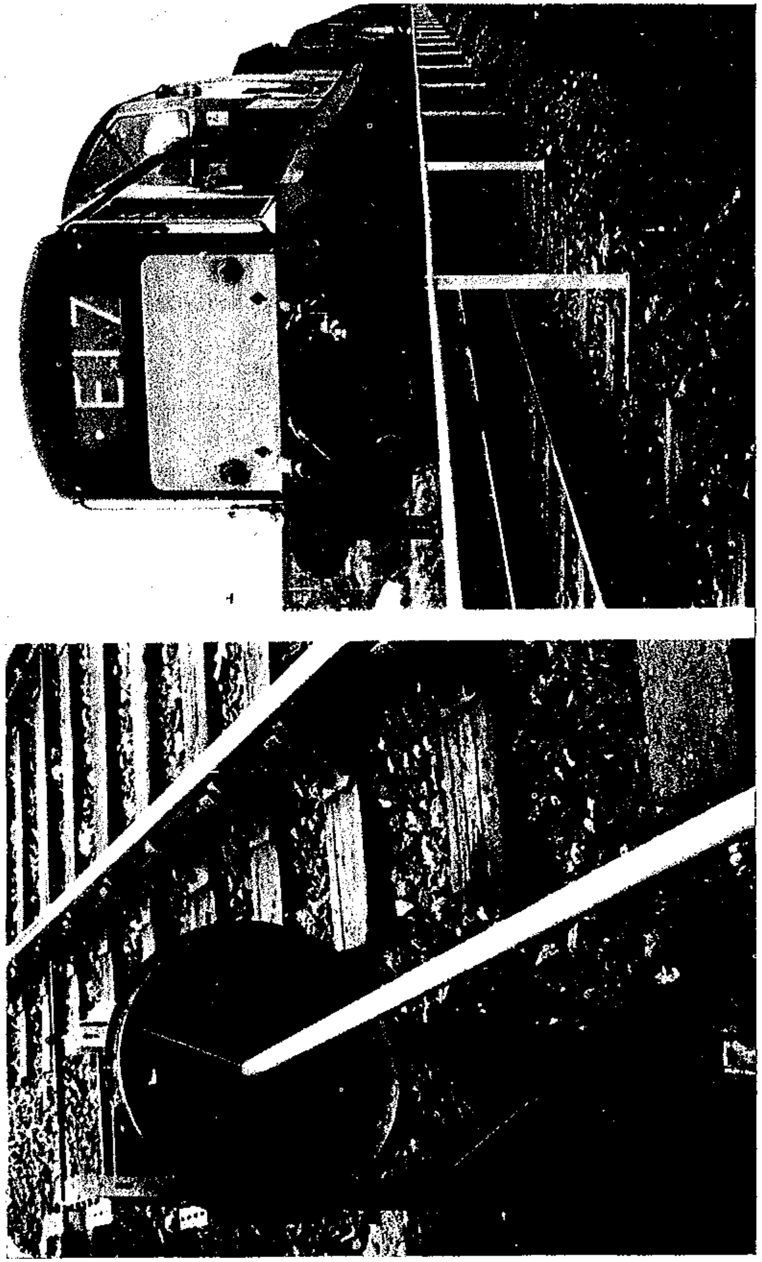


FIG. 7.5. Experimental "G" line.

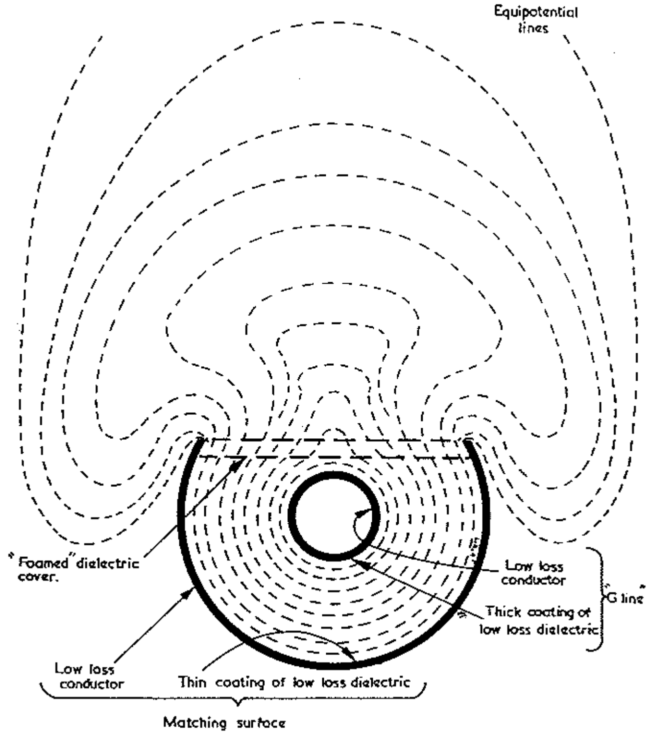


FIG. 7.6. Possible field distribution of modified "G" line.

channels are available for control and communication between fixed points and fixed aı mobile stations simultaneously over the same wave guide. Laboratory and full-sca demonstrations have already proved that communication from a mobile transmitter field modulation of the wave guide is possible, using very simple transmitting equipmeı For example, if the "transmitting" aerial is made approximately resonant to one of t carrier frequencies and placed adjacent to the wave guide, echoes will be produced wi a magnitude related to the resonant frequency of the aerial. If the resonance is vari sinusoidally at some convenient frequency, for example 50 KHz, a subcarrier signal w be generated and received at the wave guide transmitting point, appearing as a modul tion of the main signal, the latter being of course in the region of 1000 MHz. Many su carriers may be generated simultaneously by this method.[4]

An advanced concept of wave-guide communication link is described in reference 5.

7.5. Obstruction detection by radar

An experiment was successfully concluded in the U.S.A.[6] in which a radar transmitt receiver continually scanned a reflector placed on the opposite side of the track. It w envisaged that, on a double-track railway, the instrument would be located at the centre of the right-of-way between the two tracks with a longitudinal spacing of 100 metres. On the outside of both tracks a reflecting "fence" was mounted on posts. The height of the fence was 60 mm and it consisted of plastic, the back surface of which was embossed with many "corner cubes" which possessed the property that, when suitably orientated, they reflected a light beam in a direction parallel to the incident beam. Thus each transmitter/receiver unit scanned 100 metres of the fence on each side and any substantial difference between the reflected and incident beam strengths indicated the presence of an obstacle.

7.6. Ultra-sonic train detection

Rubber-tyred vehicles cannot of course operate track circuits and ultra-sonic waves have been used on the Lille Metro (described in Chapter 16) as a means for train detection.[7] This is referred to as a negative detection device because the presence of a vehicle results in the absence of a signal although, of course, the track circuit itself could be similarly described. An ultra-sonic wave generator on one side of the track normally feeds a receiver on the other, and vice versa. Thus a train passing causes the absence of a signal in two receivers for a short interval of time. Sets of transmitter/receiver instruments are placed at the entrance and exit of each block section so that the track-occupied disposition of the control apparatus is set up on a train entering a section and is cancelled when it passes the second set of instruments located at the exit.

7.7. Two-way communications at radio frequency

The presence of a train can be detected at a central or local control centre by a process of interrogation rather than by physical identification. Thus a radio signal can be broadcast and the train equipment can respond, reporting its position and other operational data which may be required. Such a system is operational in British Columbia. However, in the U.K. the frequencies available for mobile radio are limited to the VHF and UHF bands and an insufficient number of channels can be made available for railway use and some form of guided transmission is required. Various forms of radiating cable are described by Hutchings and Cree.[8]

A very high-frequency range (up to 9 GHz) can be obtained using a leaky wave-guide in the track linked with a suitable "reflector-coupler" system on the train. The wave-guides may take the form of metal tubes protected with a polymer sheath. The amount of radiation is controlled by providing rows of small holes on the side of the tube facing the vehicle. Performance characteristics are given in reference 9.

References

1. Tyler, J. F. H., *Signalling Relays (British Practice)*, Institution of Railway Signal Engineers, 1958.
2. Barlow, H. M. and Brown, J., *Radio Surface Waves*, Oxford University Press, 1963.
3. Barlow, H. M., Screened surface waves and some possible applications, *Proc. Instn. Elect. Engrs.*, vol. 112, p. 477 (1965).

4. Galawa, R. L., Beery, W. M., Chu, T. M., Cook, K. R., Fitz Gerrel, R. G., Haidle, L. L., Part J. E. and Rosner, K., *Use of Surface Waves in Communicating with High-speed Vehicles*, Institute Telecommunications Sciences, Boulder, Colorado, U.S.A. (1968).
5. Abele, M. and Medechi, H., A GHz dielectric waveguide communication line. *U.S. Federal Railr Administration Report No. FRA-RT-72-29* (1971).
6. Obstruction Detection Program Final Report, RCA/New Business Program, Princeton, N.J., U.S (1963).
7. Ferbech, D., Lille Metro—The VAL–ATO system, *Rail Engineers Forum, 11th March 1982*, I.E. London.
8. Hutchings, B. W. and Cree, D. J., *Longitudinal Track to Vehicle Communication*, I.E.E. Conferer Publication No. 117, p. 153 (1974).
9. Nakahara, T., Yoshida, K. T., Kuroda, M., Kurachi, M., Takemura, K., Kitani, H. a Migamoto, Y., *Verification of a Feasability Study on a Coupled Leaky Waveguide System for Co munication in High-speed Ground Transportation*, U.S. Department of Transportation, Washingt D.C. (1970).

CHAPTER 8

Vehicle Identification

THE previous chapter described means for detecting the presence of a vehicle. For control, however, it is often necessary to know more than this. In the case of railways it is necessary for a signalman to know the destination of a train in order to set the route correctly. Locomotives usually carry head codes, i.e. lamps or discs arranged in various patterns or illuminated number and letter codes which serve to identify the train by visual observation. The number of a vehicle has frequently to be read in all forms of transport for control and routing purposes and means are now being developed for this to be done automatically.

8.1. The "Identra" system

This is used to enable trains to describe themselves as belonging to one of nine categories. The train is equipped with a coil which is connected to a capacitance. The value of the capacitance can be changed by a nine-position rotary switch so that the coil and capacitance form tuned circuits which will resonate if placed within an electromagnetic field. According to the switch position, the resonating frequency will be one of nine values.

The wayside equipment consists of two coils placed one above the other in a vertical plane and 0·18 metre (7 inches) from the path of the train-carried coil. The two fixed coils are arranged so as to avoid mutual magnetic coupling so that when one is coupled to an input, no output is detected at the other coil which is connected to a two-stage resistance coupled amplifier.

When the train-carried coil enters the magnetic field of the track-side coils, the mutual coupling causes the amplifiers to oscillate. The frequency of this oscillation is determined by the setting of the resonant frequency selected for the train-carried unit. The oscillation signal from the amplifier is then applied to frequency selective networks which consist of band pass filters, each sensitive to a different band of frequencies. The network which happens to coincide with the frequency of the input signal from the amplifier accepts this signal, rectifies it and applies the output to a relay which operates an external control circuit. The networks tuned to other circuits of course reject the input signal. This output can be used to actuate the points at junctions so that nine different routes can be selected by varying the position of the rotary switch.

8.2. Bus electronic scanning indicator (BESI)

The regulation of buses operating on long routes in large cities poses many difficulties which would be eased if information of their location could be continuously available at a

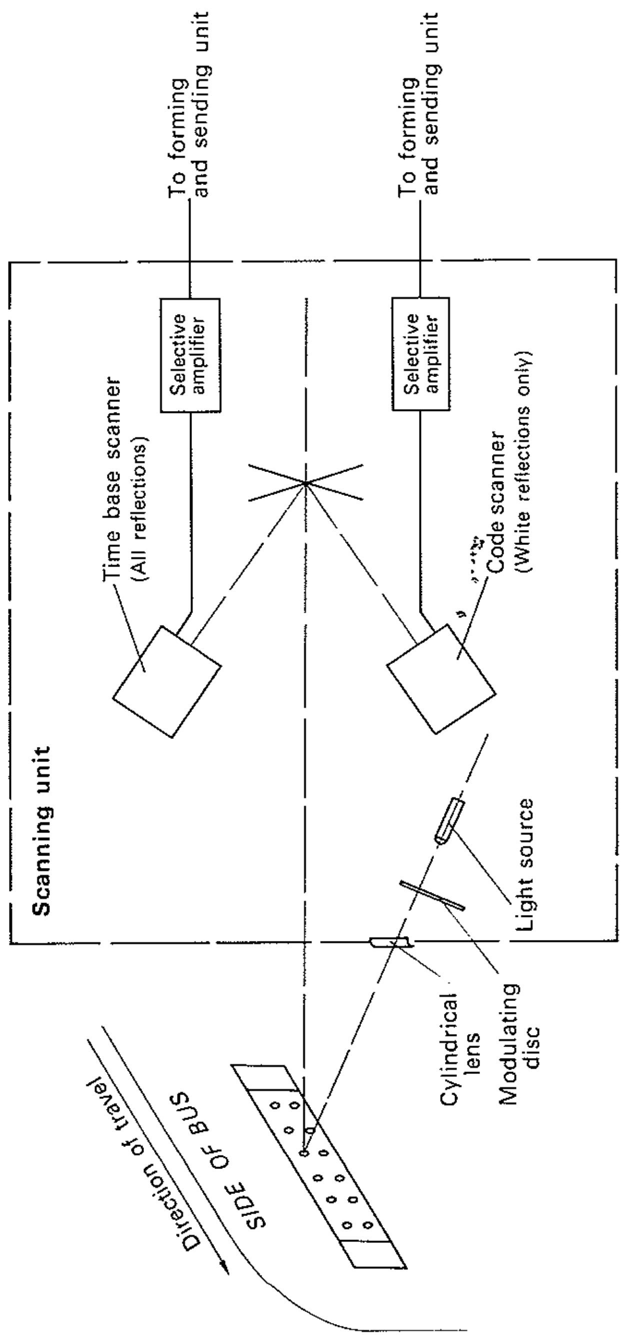


FIG. 8.1. Bus electronic scanning indicator.

central control point. London, for example, has some 7000 buses operating on 500 routes so that continuous verbal reporting would impose an intolerable load on control staff.

A system has been developed whereby the position of every bus on a route can be displayed on an indicator at the control room.[1] Each bus is fitted with a plate measuring 50 × 350 mm to which twelve dome-headed cast-glass reflectors, 15 mm in diameter and arranged in two rows, are attached. Reflectors are of two types, clear and red, and are arranged to indicate the running number of the bus in binary code. The "clear" reflectors represent the units in the code and the "red" reflectors the ciphers.

The scanning apparatus, which is mounted at the roadside (Fig. 8.1), comprises a light source which is modulated by passing through slots in a spinning disc. The beam is interrupted 5000 times a second. This is done so that reflections of the beam from scanner can be distinguished from reflections due to sunlight or other spurious sources. After modulation the beam passes through a cylindrical lens so as to shine on to the coded plate in a beam which is elongated in the vertical direction. The typical distance of the scanning apparatus from the side of the bus is 4 metres. Light reflected from the code plate is received in an object lens and then directed by two plane mirrors inclined at 150° on to

FIG. 8.2. B.E.S.I. in London street.

two photocells, one of which responds to all light and the other to red only. The out[put] from both cells are fed to selective amplifiers which are tuned to respond only to [the] modulation frequency of the incident light beam and thus intercept spurious reflecti[ons].

The output from the amplifiers is converted into suitable pulses for transmission o[f] bus number in binary code over telephone lines to the central control room where [it is] decoded for display on a board which shows the most recently indicated position of e[ach] bus on the route.

Figure 8.2 shows the appearance of the scanner unit as installed in a London stree[t].

8.3. Automatic wagon-recording system (A.W.R.S.)

Railway wagons, particularly those engaged in mineral traffic, are subject to d[...] conditions which could rapidly render mirrors ineffective.

Code bars, as shown in Fig. 8.3, are therefore used and receptor photocells are arran[ged] to respond to the spectral scattered light therefrom[2] in a system developed by S.P.[...] co-operation with the British Transport Commission. The binary decimal system con[sists] of a series of groups of four bars disposed in the upper and lower sections of the c[ode] plate so as to represent a single decimal digit in binary form. A comparison with [the] binary system is shown in Table 8.1.

TABLE 8.1.
COMPARISON OF BINARY DECIMAL SYSTEM WITH BINARY SYSTEM

	Decimal system	Binary system	A.W.R.S.
one	1	1	
two	2	10	
three	3	11	
four	4	100	
five	5	101	
ten	10	1010	
twenty	20	10100	
twenty-five	25	11001	

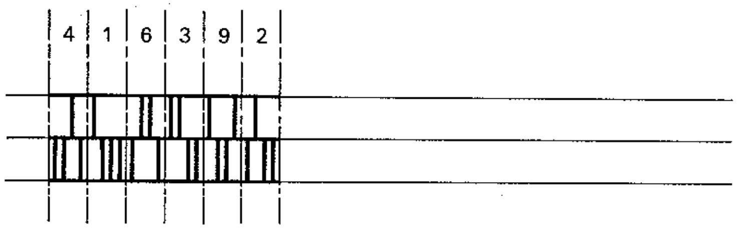


FIG. 8.3. Combined binary and decimal code for wagon identification, wagon No. 416392.

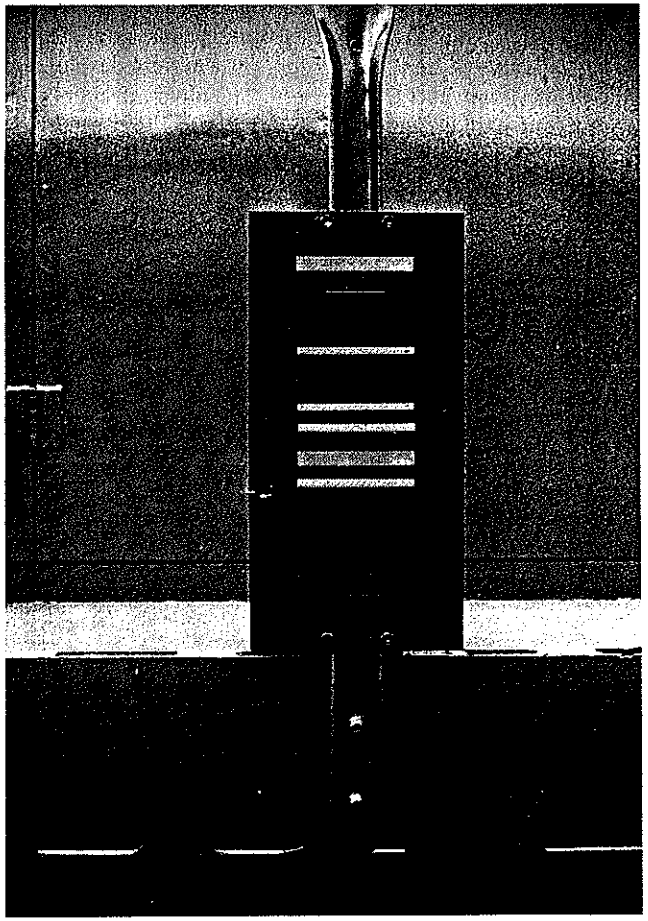


FIG. 8.4. Car identification tag fitted to locomotives and cars in North America.

Thus the plates of Fig. 8.3 can report six digits, that is, numbers up to 999,999. estimated that a code plate giving up to 10 million possible identities and capabl being detected at vehicle speeds up to 44·7 m/s (100 m.p.h.) would require to be metre long (23 inches) and 0·178 metre (7 inches) high.

Light modulated at 14,000 c/s is directed at the code bars and that which is scatt from them in a direction normal to their surface is collected and focused on to two ph cells, one of which views the upper row and the other the lower. The output of these is decoded and processed for transmission by land line.

Figure 8.4 shows a target indicator of a type which is fitted to every rail-borne veh in the U.S.A. and Canada. The code is so arranged that all the permanent informa about a car, i.e. ownership, registration number, etc., can be read by a lineside opt colour-sensitive scanner.

Although vehicles have been fitted, comparatively few railroad operators have insta scanning equipment.

Other methods of identification

Optical methods have been criticised because of the possible effect of obscuratio fog, ice, flying dust, snow, etc., and a number of alternatives have been studied.[3] Tl include Isotropic Radiation, Microwave or Ultrasonic Reflection, Static Magnetic Fie Electromagnetic Induction Fields and Infra-red Light.

All these devices are static in so far as the apparatus carried on the vehicle sin reflects energy supplied by the line-side equipment. Active vehicle equipment ma supplied with energy by a beam of a certain frequency which supplies transistor circ which energise circuits specific to each digit which can in turn emit radiation for detec by the line-side equipment.[4]

References

1. PICK, T. S. and READMAN, A., The recognition of moving vehicles by electric means, *Proc. Instn. l Engrs.*, vol. 106, pt. B, p. 186 (1958).
2. HIX, L., Automatic vehicle identification. *Symposium on the Use of Cybernetics on the Railways*, p. International Union of Railways, Paris, 1963.
3. KELLER, W. M., A method of developing hardware for automatic car identification. *Symposium o Use of Cybernetics on the Railways*, p. 404.
4. ROSCHER, H. G., Present developments in the mechanical reading of vehicle numbers. *Symposiu the Use of Cybernetics on the Railways*, p. 424.

CHAPTER 9

Communication of Control Signals to Moving Vehicles

9.1. Automatic warning systems—cab signalling

Before any question of automatic trains arose, railway operators in many countries were becoming aware of the hazards associated with the transmission of control instructions to train drivers by visual signals only and were seeking safeguards of a more positive nature. These, it must be stressed, were intended to assist the driver who retained full responsibility for the safety of his train in contrast to a fully automatic system where full reliance must be placed on the apparatus itself.

In 1905 the Great Western Railway commenced installation of what was then known as "Automatic Train Control".[1] This consisted of a system which gave an audible indication of the position of a distant signal and, in the event of the driver failing to acknowledge an adverse or "caution" signal, the brakes were automatically applied.

The signal was given by means of a ramp placed between the rails (Fig. 9.1). This ramp consisted of a baulk of timber surmounted by an inverted steel tee section. The length of the ramp was about 12 metres (40 feet) and its maximum height was 0·08 metre (3½ inches) above rail level. When the signal was at caution, the ramp was unenergised—a "fail to safety" feature. A plunger, carried on the locomotive, engaged with the ramp and was forced upwards. This interrupted a circuit normally energised by a battery carried on the locomotive allowing an electropneumatic valve to open, admitting air to the vacuum brake system and gradually applying the brakes. The passage of this air energised a siren which gave audible warning to the engineman. He could acknowledge this signal by lifting a small handle which restored current to the circuit which by now would have been remade as a result of the plunger, having left the ramp, resuming its normal position.

When a clear indication was to be given, the ramp was made live with d.c. so that current was collected by the plunger. This current was arranged to replace that interrupted by the raising of the plunger so that the electropneumatic valve remained closed. The circuit was so arranged that this current operated a relay governing the ringing of a bell. Thus an audible "clear" warning was given.

This system, therefore, gave information about position as well as the state of the road ahead and so was particularly helpful in fog. Although fitted to some 2000 miles of line it suffered from the fault that it was discontinuous, i.e. a driver could cancel a danger indication by acknowledgement and then forget that he had done so. Also, the physical contact with the ramp was open to objection.

Fig. 9.1. G.W.R. automatic train-control system.

A modified system, now known as A.W.S. (Automatic Warning System), has been developed by British Railways. The fixed part of this system consists of two magnets in the centre of the track placed about 200 metres in advance of each distant signal. The first magnet is a permanent magnet and the second, an electromagnet, is only energised when the signal concerned shows clear. The basic circuit is shown in Fig. 9.2.

Considering first the "caution" sequence, the flux from the permanent magnet is picked up by a receiver mounted below the locomotive and attracts an armature which cuts current from an electropneumatic valve after a delay of one second. This delay is achieved by the use of a condenser. After 1 second the air begins to enter the brake system and a horn is sounded separately. At this stage the driver may acknowledge and restore

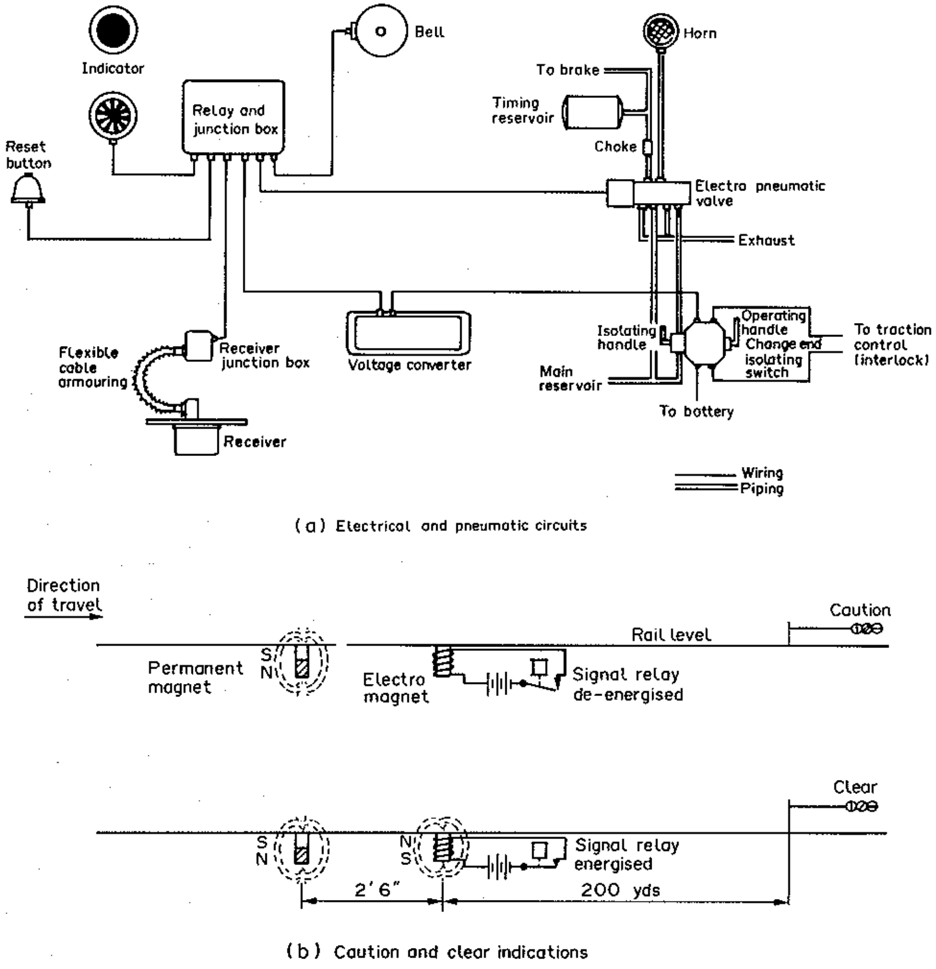


FIG. 9.2. British Railways automatic warning system.

the system to normal except for an indicator which is made to show black and yellow segments. This indication remains until it is cancelled by the permanent magnet at i next distant signal.

When a "clear" indication is to be given the electromagnet on the track is energis which restores the armature to its normal position. This occurs well within the 1-seco delay mentioned earlier. At this stage a bell is arranged to ring for 2 seconds only.

Possible disadvantages of the above system are that it requires an electromagnet to fed with power at the track-side and that it can only give two aspects, "clear" a "caution".

9.2. The train stop

A simple form of reinforcement of visual signals which is frequently used on urb railways consists of the train stop. A simple cock is connected to the train pipe of t continuous brake (air or vacuum) as described in Chapter 13. When closed, the han points in a downward direction. An arm, usually electro-pneumatically operated. carried in an assembly which is mounted on the sleepers at the side of the track near the stop signals. The arm is normally raised so that it would be struck by the handle the trip cock which would be rotated, connecting the train pipe to the atmosphere w consequent full application of the train brakes. When the signal is in the clear positi the arm is lowered and the trip cock remains closed.

In the event of operation of the trip cock, the driver must leave his cab and actua restore the trip cock to normal position. Of course all such incidents should be report

9.3. The "Indusi" system

A system used on the Continent, known as the "Indusi" system, employs three reso tors or "beacons" in the track. One, at the warning signal, being tuned to 1000 Hz, o 150 metres in advance of the stop signal, tuned at 500 Hz, and one, at the stop sign tuned at 2000 Hz. These coils are open-circuited when signals are at clear so as not resonate. When passing a warning signal at caution, the driver must operate a vigila control within 4 seconds otherwise a full application of the brakes will occur. Twe seconds after having passed the signal he must have reduced his speed to below 25 n (56 m.p.h.). At 150 metres in advance of the stop signal, he must have reduced speed below 18 m/s (40 m.p.h.) and of course full braking is applied if the stop signal is pass at danger.

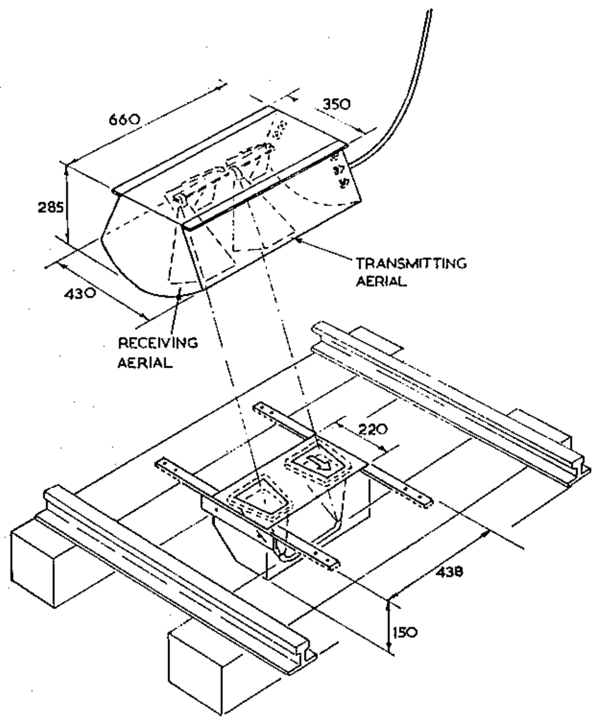
The apparatus on the train consists of shoes containing search coils which are tuned the above frequencies and fed from a three-frequency alternator. When these search cc come within range of the appropriate track-mounted coils in the switched-on conditi the energy passing increases and is used to operate relays.

The system has the advantage that no power has to be supplied to the track-moun coils and is, in fact, the functional converse of the "Identra" system where, it will recalled, the inert coils were mounted on the vehicle, power being supplied to the trac mounted equipment.

The system has the advantage over the British system in that it applies to the "stop" as well as the "distant" signal and that no power is required to be supplied to the track equipment. It has the disadvantage, however, that the destruction of the fixed coils would lead to a clear indication, i.e. the system does not "fail to safety".

9.4. The "Signum" system

This system is used in Switzerland and is based on induction between magnets placed on the track and in the vehicle. In the track are two coils which are connected together if the signal is at danger and are separated if the route is clear. The locomotive carries two coils, one which is permanently magnetised and one which is energised from the track magnets. In the danger position, the permanent magnet on the locomotive induces a flux in the first track-mounted coil which links with a current which is passed to the second coil. This in turn induces flux in the second coil on the locomotive which generates a current which can be used to operate controls.



Dimensions in mm.

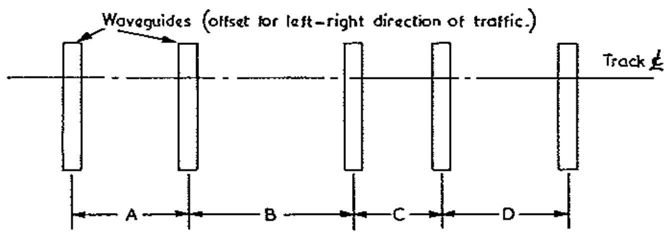
FIG. 9.3. Wave-guide type of passive beacon.

9.5. Use of beacon devices

All the aforementioned devices can be compared to beacons in so far as they communicate information from a fixed point. It is the writer's view that a successful control system, as distinct from one intended to reinforce the action of a manual controller, will require continuous information as described in Sections 9.6 to 9.8.

Nevertheless, the use of beacon devices for the transmission of information relating to specific locations may reduce the complexity of a continuous system as well as enhancing its reliability.

Another common feature of the devices described so far is that they rely on conduction or magnetic induction rather than on electromagnetic radiation. An experimental system, developed in Sweden and known as the "PHAR" system, employs a Doppler radio transmitter mounted on the underside of a locomotive and directed towards the track at an angle of 45°. A passive beacon installed in the track consists of a wave guide which accepts high-frequency energy derived from the vertical components of the radar beam, this energy being almost instantaneously emitted vertically upwards from the second aperture (Fig. 9.3). Thus a locomotive is able to detect the instant of transit over a wave guide. In a practical installation the frequency of operation selected is 9 Gc/s and a group of five such wave guides is laid a standard distance "a" in advance of an information point such as the location of a change of permissible line speed or of gradient. The spacing of the successive wave guides may be varied so as to convey quantitative information. Figure 9.4 shows distances A to D separating the wave guides in a particular group.



Typical arrangement of waveguide group

A — Gauge distance for checking accuracy

B — Distance giving gradient value

C — Distance giving line speed

D — Distance giving position of next information change

FIG. 9.4. Transmission of fixed information using passive beacons.

Space A is a standard or gauge length for checking distance measuring equipment which uses the horizontal component of the emission from the Doppler radar transmitter. Distance travelled is obtained from a count of the different pulses and is claimed to be accurate to 3 parts in 1000. Correct functioning is checked against the time required to traverse distance A at any speed and any lack of correspondence results in a "failure" indication. Distance B can be used to convey information about the route, etc. In practice, due to the fact that the spacing of the wave guides is determined by sleeper spacing, up to eight units of gradient information, twenty-five units of speed information and 100 units of distance information in 50-metre steps are provided by each group of wave guides. The final spacing D is used to convey the distance to the next information point. This introduces a "fail safe" element to a system which does not otherwise comply with this requirement. Thus if a wave guide becomes damaged or removed but its position has been notified by the previous wave guide, the absence of the expected response can be monitored to give a "failure" indication. Figure 9.5 shows how three such groups can provide information about changes in gradient and permitted speed.

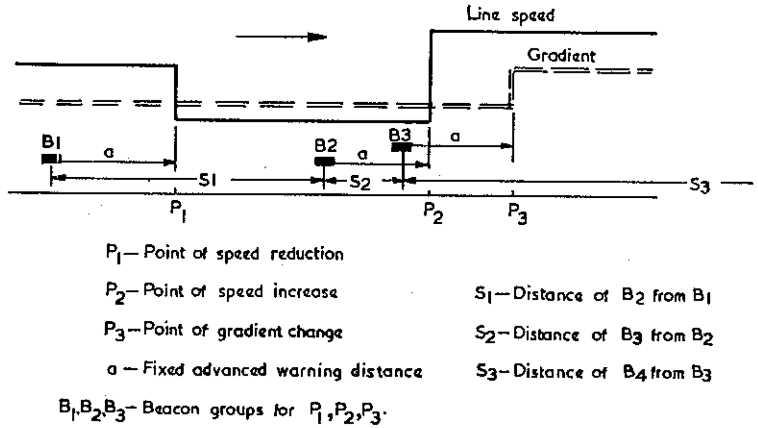


FIG. 9.5. Arrangement of beacon groups for typical line diagram.

A modern system developed by L. M. Ericsson, known as the J.Z.G.700 Automatic Train Control, uses pairs of beacons located between the track at fixed signals and at other points where a train needs new information. These transmit signals to a mini-computer through an antenna on board the train. Figure 9.6(a) shows the general configuration.

Beacons are duplicated for safety reasons: to provide automatic registration of the direction of motion of a train and to obtain high information capacity. In the event of failure of one beacon, the vehicle equipment will detect the other and actuate a "beacon fault" alarm on the driver's panel.

The first beacon of a pair gives information on the permissible speed and track conditions ahead and the second gives the distance to the next information point.

The message transmitted by a beacon consists of five eight-bit words made up of binary "one" or "zero" states. One of the states consists of a 4–5-MHz signal, the other state is the absence of that signal. The transmission rate is 5·0 K bit. Each message consists of five eight-bit words of which the first and the fifth serve to identify the beginning and end. The fifth word of one message is the first of the next, as indicated in Fig. 9.6(b).

The vehicle antenna is mounted underneath the floor and has separate transmitting and receiving loops. Even at a speed of 300 km/h (190 m.p.h.) the complete information message will be received at least eight times by the receiving loop as it passes the beacon.

The message from the beacon is processed by the micro-computer which delivers visua cab-signal displays, audible alarms or brake actuation, as appropriate. The mini-compute is controlled by two independent and completely different programs with data stored i separate memories. The two programs together supply the train with the signal "Inhibi alarm". If one fails the alarm and the brake will be activated. This will also happen i the event of a disparity between the results of the two programs.

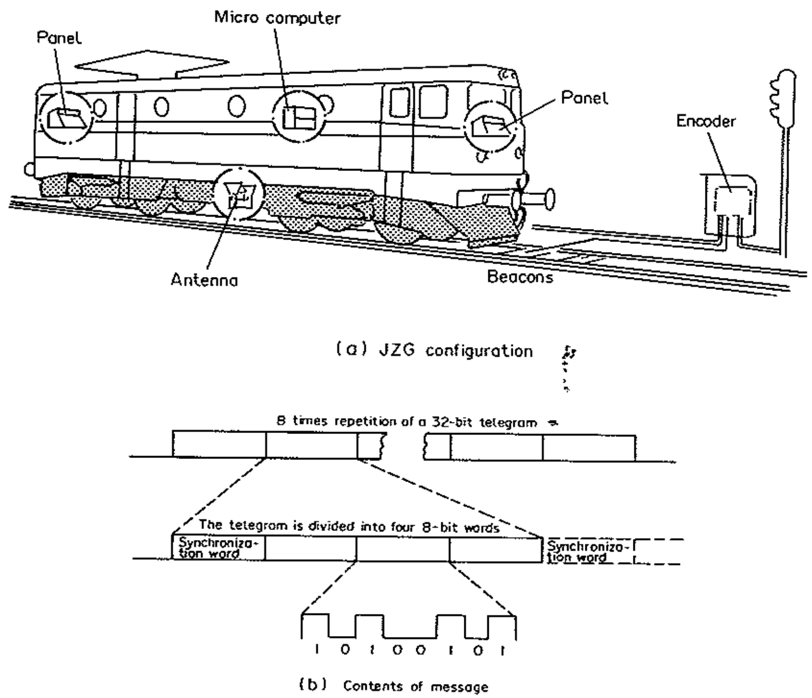


FIG. 9.6. Beacon-type automatic train control.

9.6. Coded track circuit

When a track circuit is fed with a.c. this can form the basis for communication o information to a train if search coils are placed between the point at which current is fe to the track and the nearest axle which would, of course, short circuit the track. The firs train would therefore cut off any signal to a following train so that absence of a signa indicates danger, a valuable "fail to safety" feature.

Further information can, of course, be transmitted by modifying the current used t operate the track circuit. A cab-signalling system based on this principle has been used b the Pennsylvania and certain other railways since 1933. The a.c. fed to the track circuit is pulsed at a rate of 180 c/m for a clear (green) indication, 120 for an "approach medium"

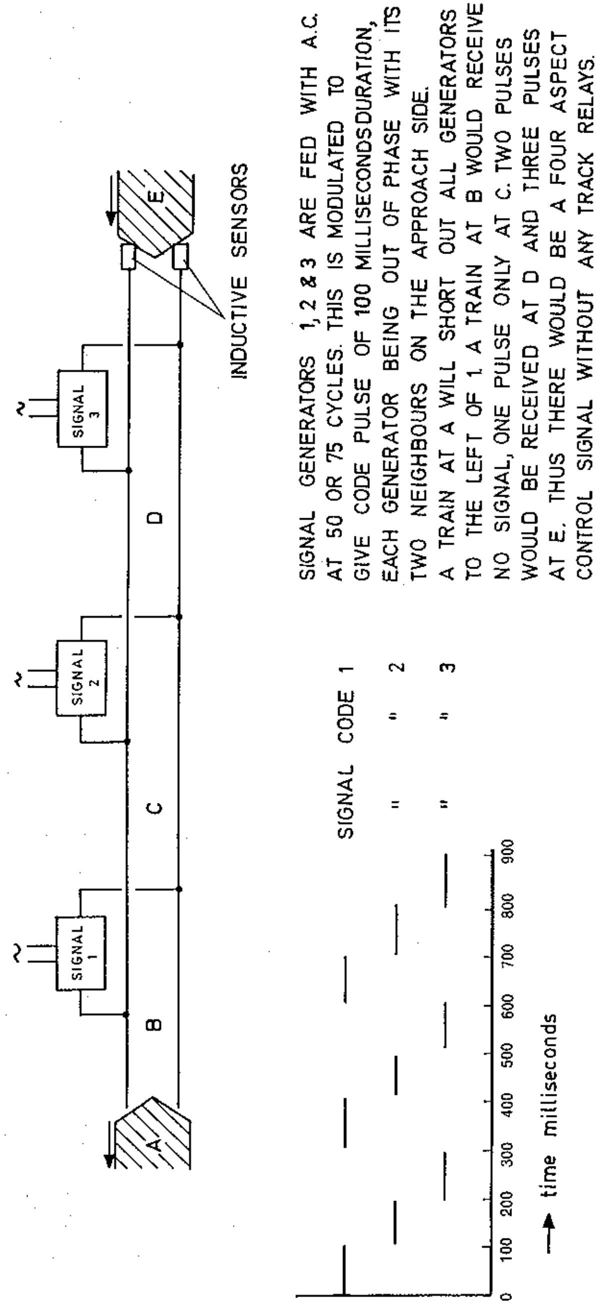


FIG. 9.7. Professor Poupé's control system.

(or double yellow), 75 for an "approach" (yellow) indication and, of course, no curren for "stop" (red). Consider a series of blocks. The first block would be shunted by a trai and would have no pulse to the rear thereof. The next block would have 75 pulses. If th line were clear these pulses would be fed to a relay at the end of the block which woul translate the signal to be fed to the next section at a frequency of 120. Then the nex section would be fed at 180 and all the way down the line as far as this was clear. Coil mounted on the locomotive sense these pulsations and provide continuous cab signalling

9.7. Professor Poupé's system of coded track circuits

The relays used in the coded track circuit system are somewhat complicated an expensive. This has led Professor Poupé of the Transport University of Zilina in Czecho slovakia to devise a system for communicating information about the state of the trac ahead without the necessity for track-side relays. His system also avoids the necessity fo insulated joints at the extremities of each block which are difficult to install when con tinuous welded rails are used.

Signal generators are connected across the rails at the beginning of each section an supplied with alternating current at 50 or 75 Hz. This is modulated in the signal generato to give pulses of 100 milliseconds duration interspersed with gaps of twice that duration Each generator is out of phase with the two generators on the approach side (Fig. 9.7)

The presence of a train will short out all impulses between its position and the firs feeding point in the rear. In the next section there would be only one pulse of 100 m duration. In the following section there would be a combined pulse of 200 ms duratio and in the third and subsequent sections there would be a continuous signal.

Locomotives are fitted with detectors which are sensitive to the duration of pulse Thus when a continuous signal was received this would signify a clear run or "green" aspect. Two pulses would correspond with a "double yellow" and one pulse to a "yellow" The absence of any pulse would signify a "red". The system is "fail safe" in so far as th inadvertent absence of any pulse leads to a more restrictive aspect.

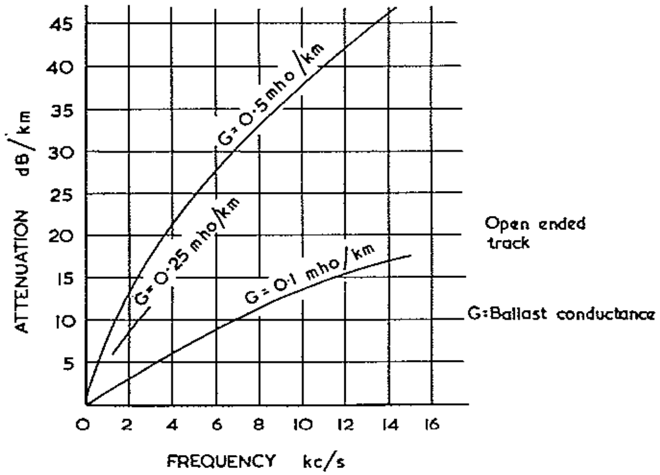


FIG. 9.8. Attenuation characteristics of railway track.

9.8. Use of continuous conductors in the track

Figure 9.8 shows the attenuation characteristics of rail as affected by variation in the electrical resistance of the ballast from which it will be gathered that the coded track circuit is limited in the range and variety of information it can transmit at quite moderate frequencies. There is no real possibility of using it for any reverse transmission of information other than the crude indication of track occupancy previously mentioned.

The alternative, which has been developed by British Railways[4] and brought to a considerable state of development before similar activities on the Deutsche Bundesbahn became known,[5–6] consisted of using an insulated conductor fed at a frequency somewhere between 20 kHz and 150 kHz so as to permit simultaneous transmission of information in both directions over considerable distances.

In the present experiments, lightly insulated conductors are laid upon the centre part of the track and provided with a high-frequency carrier, representing a safety signal and frequency modulated to provide variable information. Various configurations have been tried with the intention of providing a direct measurement and transmission of both vehicle position and train speed.

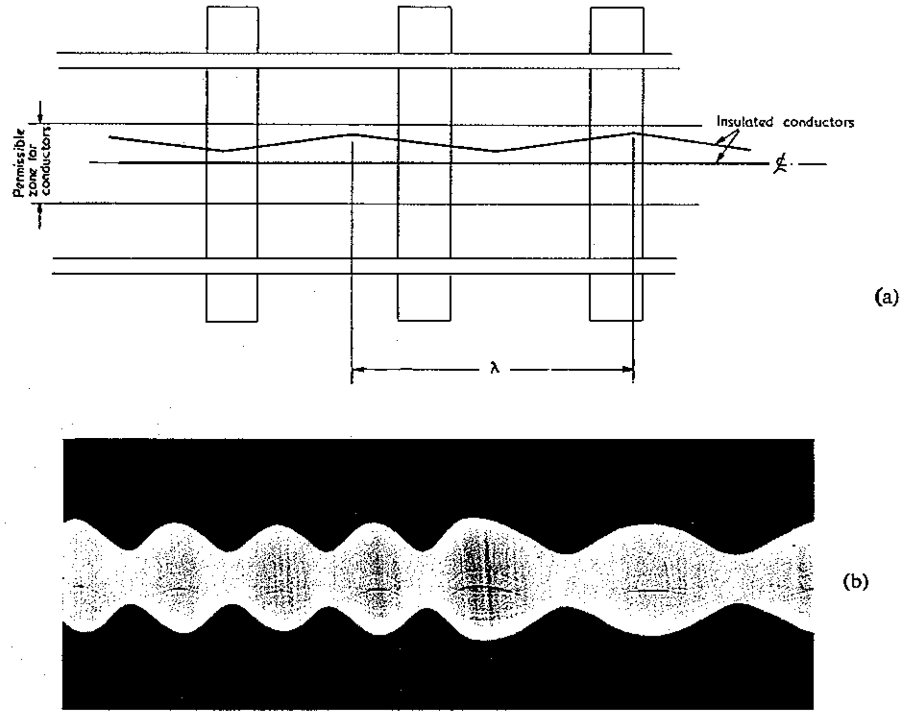


Fig. 9.9. (a) Conductor arrangement used to indicate line speed. (b) Example of signal produced by motion of train. Speed measurement from line conductors.

If, of the two conductors laid in the track, one is disposed in a triangular configuration, an approximately sinusoidal wave is received in a detector placed in the locomotive as shown in Fig. 9.9. If the wavelength of the conductor (λ) was chosen as 4 metres, the

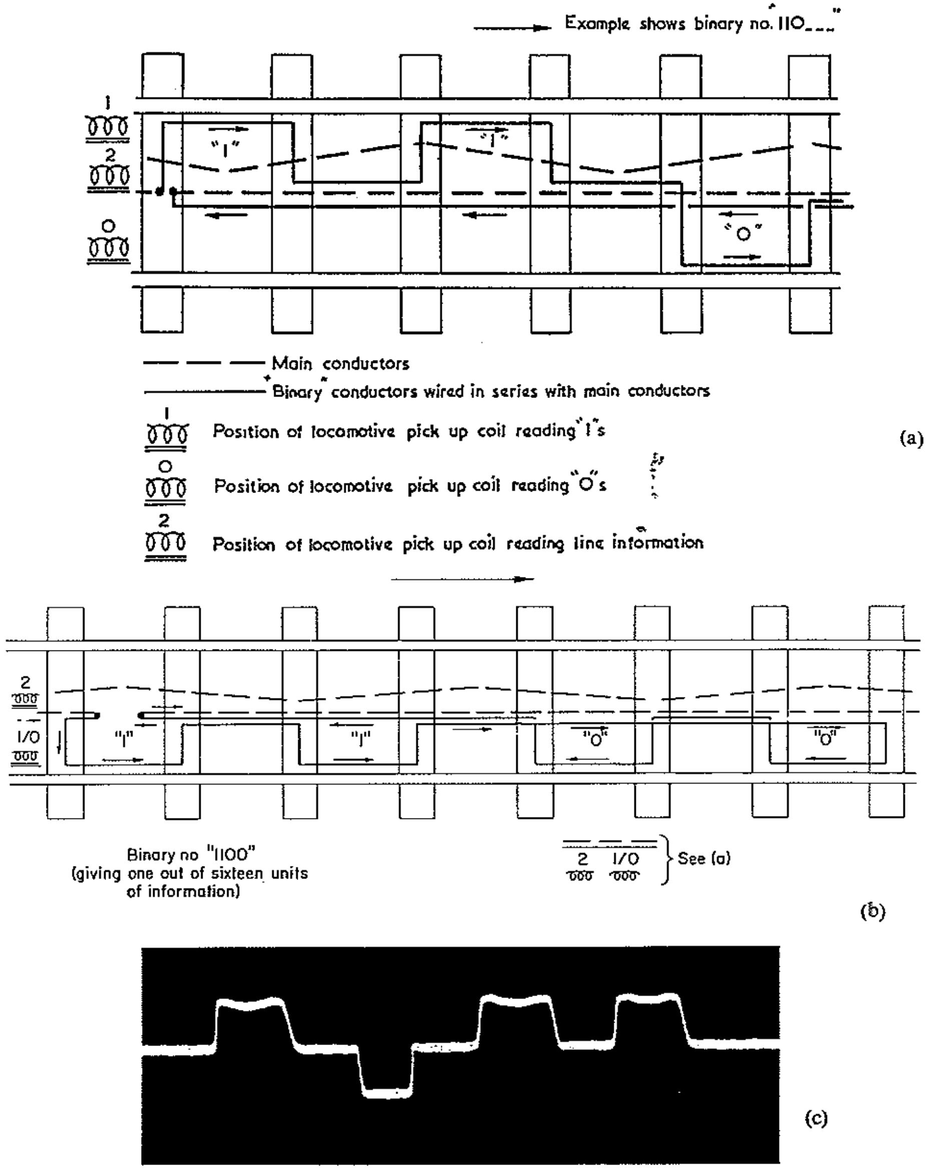


Fig. 9.10. (a) Method of writing binary no. into track. (b) Binary no. produced by phase comparison. (c) Oscillogram of binary no. derived from track conductors. Binary information from track conductors.

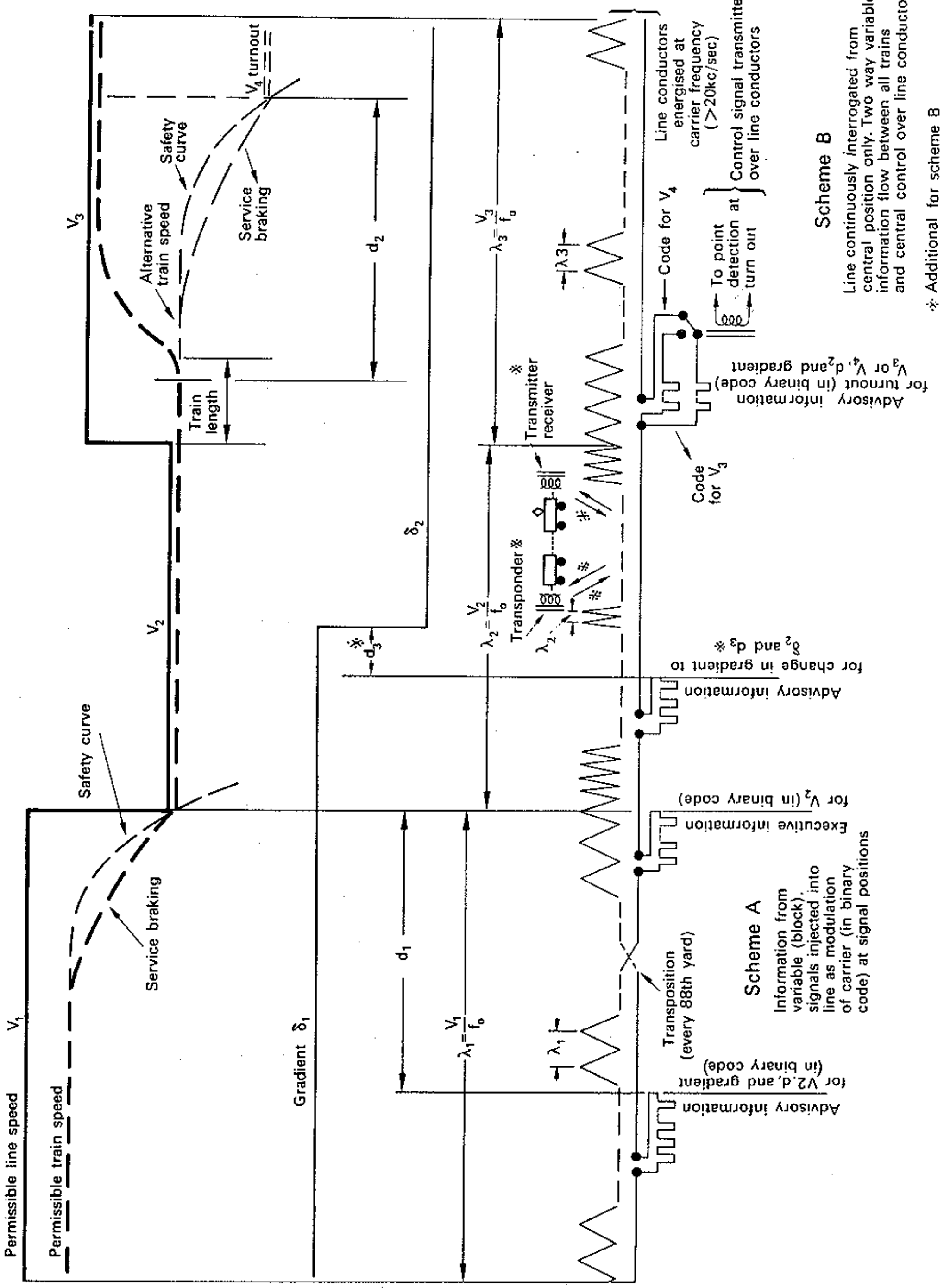


FIG. 9.11. Basic communication system for continuous-speed control.

frequency of modulation (f) would be 12·5 Hz for a speed of 50 m/s (112 m.p.h.). I general

$$\text{line speed} = f \times 4 \text{ m/s}.$$

In addition, physical modulation can be embodied in the track to represent such factors a speed restriction due to curvature or other operating features. Whether or not suc programming features should be embodied in the track must depend on local circum stances but the facility is likely to be useful on occasions. Additional loops can b incorporated as shown in Fig. 9.10(a) and (b) to provide binary information. Figu 9.10(c) shows an oscillogram which reads 1011 and was taken during an early experimen

Figure 9.11 shows, in a simplified form, a system combining fixed and variable informa tion. It also shows that the presence of a speed restriction can be indicated by reductio of the wavelength of the zigzag conductors. Variation in gradient is indicated by tran posing the conductors to generate digital information. Where there is a change in rout the conductors can be duplicated, the appropriate set being selected by switching.

Figure 9.12 shows an experimental installation on British Rail at West Drayton whic was used for checking field patterns and general line characteristics.

9.9. Combination of magnetic and inductive-loop systems

Most transport developments are concerned with systems which are already oper tional and improvements must usually respect, to a greater or lesser degree, the necessi for compatibility with existing equipment.

FIG. 9.12. Experimental installation of line conductors.

The proposals illustrated in Fig. 9.13[7, 8] employ the permanent magnet of the conventional A.W.S. equipment, described in Section 9.1, to trigger off reception from inductive loops some 180 metres long. This provides a "fail safe" feature because with discontinuous inductive loops, absence of a signal due to failure of the loop cannot be distinguished from the normal absence of signals between loops. Assuming that the track magnet imposes the most restrictive aspect, the inductive loops can cancel this and impose a less restrictive aspect, either yellow, double yellow or green. In the proposed system these are displayed in the driver's cab and he is required to acknowledge them specifically, failure to acknowledge a restrictive aspect entailing a brake application. The signal having been acknowledged it continues to be displayed in the cab until the next magnet is passed. Thus a continuous indication is given on the basis of a discontinuous input.

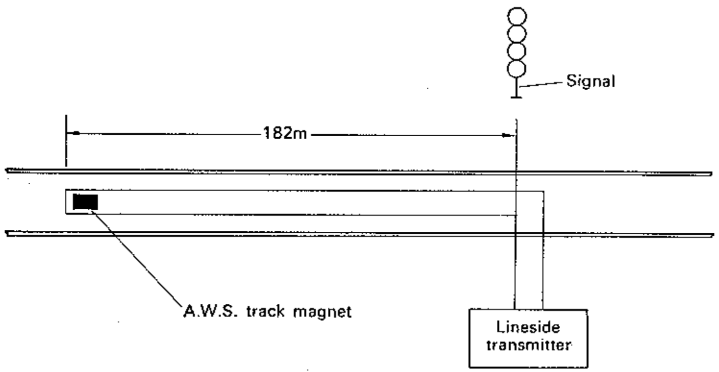


FIG. 9.13. Combination of inductive loops and track magnets.

The British Railways Automatic Warning System was considered to be adequate for speeds up to 100 m.p.h. (45 m/s) but not for proposed operations at 125 m.p.h. (55 m/s) without so increasing the warning distance that the operation of the majority of conventional trains travelling at 100 m.p.h. (44·7 m/s) or less would be impaired. It was accordingly proposed that continuous cab signalling should be provided for the 125 m.p.h. (55 m/s) trains only, the remainder continuing to operate on the discontinuous equipment but improvements in train braking, notably the introduction of the two pipe air brakes and disc brakes, has rendered this unnecessary.

References

1. DYMOND, A. W. J., Forty years of automatic train control—the Great Western System. *Proc. Instn. Loco. Engrs.*, 1948.
2. Professor POUPÉ, Private communication.
3. BARWELL, F. T. and OGILVY, H. H., Communications and their effect on railway operation. *Proc. Instn. Railway Signal Engrs.*, p. 135 (1966).
4. HUTCHINGS, B. W. and CREE, D. G., Longitudinal track to vehicle communications. *I.E.E. Conference Publication*, No. 17 (1974).

5. SCHMITZ, W., Möglichkeit der selbsttätigen Steuerung eines Zuges in Verbindung mit einer modernen Signaltechnik. *Neue Technik*, No. A.2.70 (1964).
6. KILB, E., Grundsatzlicche Planung der selbsttätigen Fahr und Bremssteuerung für Eisenbahnzuge. *Neue Technik*, p. 1, Ap. 19 (1964).
7. TYLER, J. F. H., Signalling for high speed trains. *Proc. Instn. Railway Signal Engrs.*, preprint (1970).
8. ALSTON, L. L., Electrical research in British Railways. *Electronics and Power*, vol. 16, p. 3 (1970).

CHAPTER 10

Interlocking — Sequence Control

10.1. Mechanical interlocking

A railway signalman has a great deal to remember and devices to prevent mistakes are provided wherever possible. These devices, although somewhat primitive in construction, illustrate principles which are important in modern control engineering.

Take as a simple example the "distant" signal. This must only be lowered when both the "home" and "starting" signals are clear because it indicates to a driver that he has a clear run right the way through the station and into the next section.

Connected to each lever used by the signalman to operate the signal is a rectangular rod known as a "tappet" and usually mounted in a vertical plane. These rods have parts cut away on one side to allow for penetration of pins known as "locks" carried on other members which run at right angles thereto. Consider the situation shown in Fig. 10.1(a).

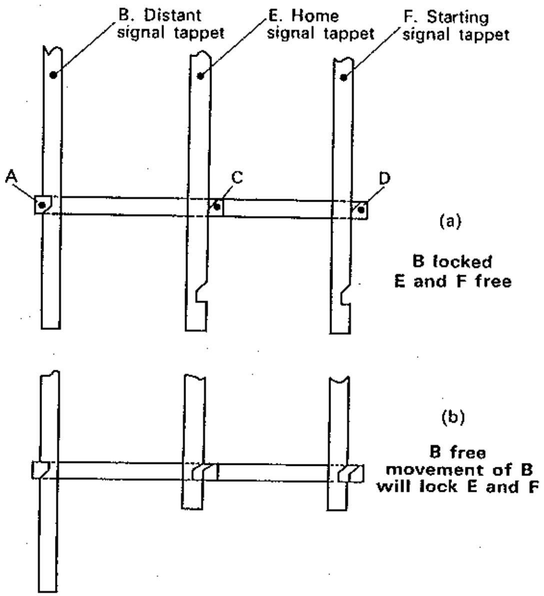


FIG. 10.1. Principles of interlocking.

The pin A fixed in the cross member cannot move to the left to allow tappet B to move because pins C and D are prevented from moving by tappets E and F which are connected to the Home and Starting signal levers respectively. Assuming that these signals have been pulled off and that their associated tappets have taken up the position shown in Fig. 10.1(b), then there is now no resistance to the horizontal movement of pins C, D and A and the distant signal can now be pulled off. This will cause pins C and D to move to the left, locking bars E and F and consequently the home and starter signals in the "off" position. They cannot be restored to danger until the distant signal has been pulled "on". This is a form of "sequence" control and enables the driver to know that once the "distant" is "off" he can be certain of a clear run.

The most important function of interlocking is to prevent conflicting movements at junctions. Thus in the layout shown in Fig. 10.2, once permission has been given for a train to leave A for D, no train must leave C for B until the first train has passed clear of the junction.

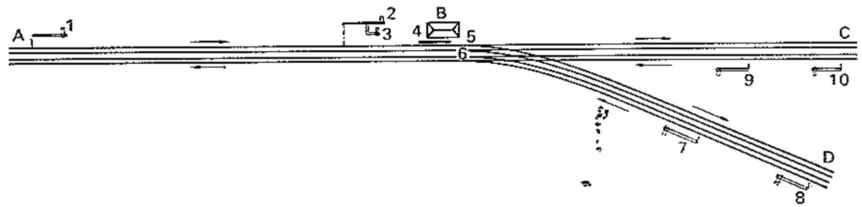


FIG. 10.2. Layout of signals at junction.

Similarly, the signalman at B must not accept trains from D and C simultaneously and must be protected from so doing by interlocking of points and signals.

The design of an interlocking system which will permit all legitimate movements and yet eliminate any possible movement involving risk of collision is a complex task. The first stage in the design process is to construct a control table as shown in Table 10.1. The lever under consideration is denoted by a number in the central column. By looking at the left-hand column the numbers of the levers which have to be pulled before the chosen lever can be released are determined. The right-hand column shows the levers which will be locked once the chosen lever is pulled. Table 10.1 is taken from reference 1

TABLE 10.1. CONTROL TABLE FOR JUNCTION

Released by	Lever number	Locks
2	1	
4	2	5
4 5	3	
	4	5 (in either position)
6	5	2
	6	9
6	7	
7	8	
	9	6
9	10	

Signal levers are "normal" for signals at danger and point levers are "normal" when the route is set for the main line. Facing point locks secure points either in the normal or in the reversed position. The lever performs its locking whenever it is pulled. Thus in setting up a route from A to C, a signalman would first ensure that points (5) were in the normal position, that is, set for the main line. He would then operate (4) the facing point lock which would lock (5) in either normal or reverse position. He could only pull off signal (2) if (5) were in the normal position and he had proved this by operating (4). Distant signal (1) could only be pulled off if (2) had been and, by inference, if the points had been correctly set for the main line and locked in this position.

Were a train required to proceed from A to D, points (5) could only be operated after points (6). These cannot be operated unless signal (9) is at danger. Therefore a train cannot be sent from A to D if the route is set up for a train to run from C to A. Thus the protection required is built into the interlocking system. If (6) has been reversed, points (5) can now be reversed which locks signal (2) so that a driver bound for the branch cannot be given the signal for the main line. Because signal (2) is locked, signal (1) cannot be pulled off so that a caution indication is ensured for any train destined for the branch. After operating the facing point lock (4), signal (3) can be pulled off, thereby completing the setting of the route for the branch. A train from D to A could be accepted simultaneously with one from A to D because (6) would already have been reversed so releasing (7) which in turn would release (8).

A train from C to A can only receive a clear signal at (9) if (6) is in the normal position.

Although with modern systems the interlocking is carried out electrically, the same principles apply.

10.2. Boolean algebra

The layout shown in Fig. 10.2 is extremely simple. A modern electric signal box may control a large and complex area and many industrial processes are equally complex. The use of Boolean algebra[2] enables the logical processes required to set up an interlocking system to be set down and manipulated, using a simple notation.

The notation and concepts of Boolean algebra are as follows:

$\in$ signifies *inclusion*, thus $a \in A$ indicates that "a" belongs to "class A",

$\notin$ signifies *exclusion*, i.e. $a \notin A$ indicates that "a" is not a member of "A",

$\subset$ signifies the *inclusion* of one class within another class,

$\odot$ signifies a fictional *empty* or *null* class,

1 signifies a *universal* class containing everything under consideration,

$\cup$ signifies the *union* of two classes, i.e. $A \cup B$ contains all members of classes A and B but once only,

$\cap$ signifies the *intersection* of two classes, i.e. contains all members common to both A and B,

A' indicates the *complement* of A, i.e. everything that is not A,

$\cap$ and $\cup$ are commutative, associative and distributive, each over the other so that manipulation by processes similar to those of ordinary algebra is possible. In particular there is the "algebra of sentences".

Following reference 2, Table 10.2 can be constructed.

TABLE 10.2

A	B	$A \cup B$	$(A \cup B)'$	$A \cap B$	$(A \cap B) \cap (A \cup B)'$
0	0	0	1	0	0
0	1	1	0	0	0
1	0	1	0	0	0
1	1	1	0	1	0

This is known as a truth table and this particular example is set up to test whether

$$A \cap B \cap (A \cup B)' = 0$$

which can be stated in words as follows: "Nothing exists which is both A and B bu is not contained in A and B taken together."

In the first two columns values of 0 and 1 are assigned to A and B so that all possibl combinations are displayed on four lines. The final column shows that the sentenc $(A \cap B) \cup (A \cup B)$ is valid.

Now reverting to our junction we can write (3) is released by (4) and (5) but (5) is onl released by (6). Suppose we had arranged for (3) to release (6) we would have had complete impasse and the system would be unworkable. Supposing we had not arrange for (5) to be released by (6) there would have been no connection between the control c the up and the down lines and collision between an "up" branch train and a "down main-line train would have been possible.

Use of sentence logic can therefore assist in the design of interlocking systems.

Let class A include all levers in normal position and class B all levers which have bee pulled. For an up main-line train

$$B = 1 \;\&\; 2 \quad \text{and} \quad A = B',$$

but now (5) is locked and in turn locks (3).

If two trains are offered and if the lever movements required by one are included in th complement of those required by the other, the movement is possible. If there is a intersection then only one of the trains can be accepted.

In signalling terms, the logic functions may be simplified to

And
Or
Not

A device may be arranged to give an output when two or more input circuits a energised. This is an "And" function. Simple switching circuits are shown in Fig. 10.

The "Or" function is achieved by providing an output if one of a number of inputs energised.

The "Not" function is achieved by arranging for an overriding signal to suppress t output of units of either of the above types.

Heald and Gore[3] describe the application of logic elements to a signal box Henley-on-Thames. Although technically successful this application was not extend because of the high cost of the special-purpose wound components. Progress in future is likely to be based on the introduction of micro-electronic units.[4]

A number of manufacturers provide solid state logic units which can be easily assembled into control systems.

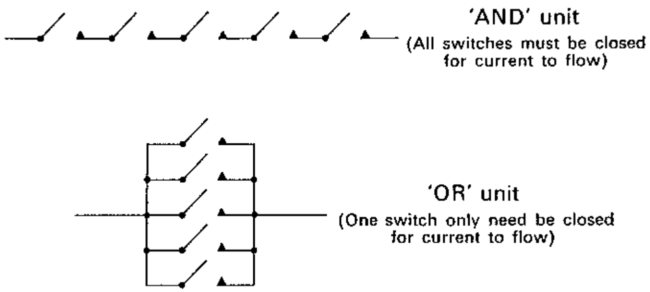


FIG. 10.3. Switching equivalent of logic unit.

A number of manufacturers provide solid state logic units which can be easily assembled into control systems.

FIG. 10.4. Intersection of railway and airport runway.

10.3. Control of intersections—automatic half barriers

Where two or more routes intersect on the level, collision of controlled vehicles can avoided by some form of interlocking but problems arise where more than one mode involved. Figure 10.4 shows an unusual intersection at Le Touquet where a passeng carrying railway crosses the runway of the airport.

The intersection of railways and roads (grade-crossings) is governed by regulatio which are often based on the history and state of development of the district concerne Thus when railways were introduced into Great Britain there was already a reasonal well-developed highway system and the legislation introduced placed full responsibil for the safe operation of the crossing on to the railway company which was required fully fence the railway, to provide gates at the level-crossings which were normally clos across the road, and to employ "good and proper persons" to open and shut them. Th the public was fully protected provided the crossing keeper had accurate knowledge the movement of trains and acted accordingly.

Many crossings were protected by signals, either specially provided or forming part the block system. These signals could only be pulled "off" if the gates were closed acr the road and proved to be so by a gate stop which was mechanically interlocked to t signals.

In the U.S.A. railways were developed in completely different circumstances and w generally the means whereby districts were opened up to agriculture and commerce. Th were therefore unfenced and it was the responsibility of the public to protect themsel against impact with the train. The railway usually marked crossings by two boards for ing a "St. Andrew's Cross" and locomotive drivers were expected to use steam whist as audible warnings of approach. One legal requirement was that drivers of such vehic as school-buses or petrol tankers were required to bring their vehicles to a halt bef crossing the railway.

The British Transport Commission Acts of 1954 and 1957 gave power to the railw authorities in Britain to substitute lifting barriers for gates at level-crossings and relax previous requirements relating thereto. For example, the traditional crossing gates swu across the railways so that, when open for the passage of road vehicles, the railw right-of-way was completely fenced off. (It should be mentioned that the original legis tion was relaxed in many cases so that the normal position of the gates become "ope to road, "closed" to rail.) Following a study of the practice in Holland a number level-crossings (over 200) were converted to the "Automatic Half Barrier" type.

Instead of massive wooden gates capable of completely excluding a horse-dra vehicle (initially) or a private motor car (latterly), lifting barriers were provided wh had little more than token value. The single barrier was placed across the left-s carriageway (in Britain, or the right-hand carriageway in Europe and America) so as bring oncoming traffic to a stop before the crossing was reached. The other carriagew was unobstructed so as to avoid the risk of vehicles being trapped within the cross area. This arrangement was thought to be a hazard because careless drivers mi "zig-zag" through the crossing even when the barriers were down.

When a train approached a level-crossing it would actuate a track circuit which wo de-energise relays at the crossing, setting warning bells ringing and warning lights fla ing. After a certain interval the barriers are lowered and the timing of this event is arranged in relation to the arrival of the train that zig-zagging is discouraged.

The position of the controlling track-circuit and the timing of the barrier movement must be such that as a train approaches the audible and visual warnings begin to act at 6 to 8 seconds before the barriers commence to fall. The actual lowering of the barriers should take another 6 to 8 seconds and should be complete 5 seconds before arrival at the crossing of the fastest train. The audible warning ceases as soon as the barrier is lowered but the lights continue to flash. As soon as the train has passed, the barriers must rise and the lights extinguish except if, on a double-line section, another train has passed the appropriate track circuit. An indicator will then display the message "second train coming".

If a second train "strikes in" after the first train has passed, the barriers must rise and remain upright for the full warning period before they start to fall again and the track circuits must be located accordingly.

It will be realised that this practice sets aside the basic principle of railway signalling, namely "fail-safe" operation. If the crossing is obstructed for any reason by a road vehicle there is no way of alerting the train driver to this effect. Although the signalman may have received "line clear" from the next box and lowered distant, home, and starting signals informing the driver of this fact, disaster may be lurking at the next level-crossing in the form of a vehicle stalled across the track. That this was possible was tragically demonstrated on 6th January 1968 when a transformer weighing 120 tons was being carried on a transporter belonging to a well-established haulage contractor who was experienced in handling exceptional loads, came to a halt on a level-crossing at Hixon in Staffordshire. It was struck by an electrically-hauled express travelling from Manchester to London at a speed of 75 m.p.h. (120 km/h), resulting in the death of eleven passengers.

Such an event had been anticipated. Even in the days of gated level-crossings a rule (No. 107) required the operators of traction engines, etc., to notify the station-master in advance of their intention to cross any level-crossing not protected by signals so that appropriate safety precautions could be taken. In the case of the modernised crossings notices were exhibited at the entrance requesting the drivers of exceptional loads to telephone the signalman before attempting to cross the track. Had this been done the disaster would have been avoided. The signalman would have waited for a suitable interval between trains, placed all his signals at "danger" and then given permission for the driver to cross, requesting him to telephone again when he was safely across. On receipt of this message he would re-open the line to rail traffic.

The enquiry into the disaster took over 40 days. Much attention was devoted to the sighting of the warning notice and to the responsibility of the team of policemen who accompanied the load. The author, who was present as an expert witness, proposed an automatic gated crossing developed in co-operation with Mr. Trevor Davies (now Reader in Mechanisms at Loughborough University) which consisted of gates which were each in two leaves, pivotted in a special manner, so that as they were moved from being closed across the railway to closed across the road they swept the full area of roadway within the crossing. Any obstruction likely to cause danger to a train would prevent the gate from closing. On attaining the closed position the gates would operate a relay which would enable the trains to be signalled through the crossing. The device is sketched in Fig. 10.5.

The current practice for control of level-crossings on important lines in Britain is to

use closed-circuit television to prove that the crossing is clear of traffic before a train i allowed to pass. The crossing may be controlled from a signal-box which may be as fa as 30 kilometres away. The crossing is illuminated at night. After a signalman has close the gates by remote control he examines the television monitor and, when satisfied tha it is safe to do so, he presses a button which releases the signals.

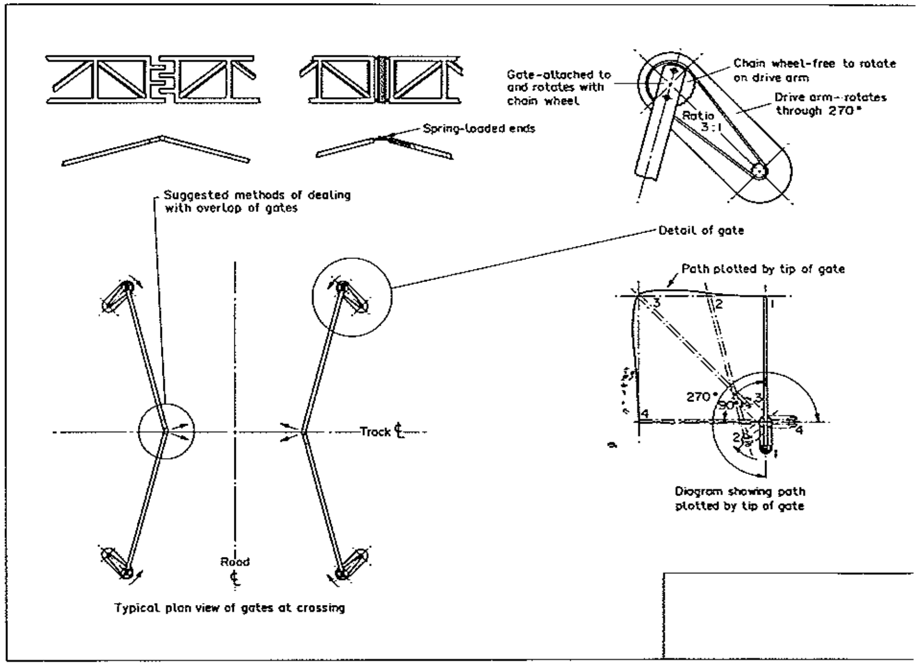


FIG. 10.5. Special linkwork for level-crossing gates devised by Mr. Trevor Davies.

References

1. SUCH, V. H., *Principles of Interlocking*, Institution of Railway Signal Engineers, London, 1956.
2. GOODSTEIN, R. L., *Boolean Algebra*, Pergamon, 1963.
3. HEALD, J. A. and GORE, G. W., Contactless switching with particular reference to square loop ferrite *Proc. Inst. Railway Signal Engrs.* (1962).
4. SHORT, R. C., The impact of micro-electronics on railway signalling. *I.E.E. Conference Publicati* No. 203, *Railways in the Electronic Age*, p. 6 (1981).

CHAPTER 11

Sorting and Marshalling

A NATIONAL system of freight transport must comprise a number of agencies of varying capacity. Whilst certain flows of traffic such as the transport of coal from mine to power station may be arranged in merry-go-round fashion so that the load is carried in complete trains which are never uncoupled, the majority of traffic items originate from widely distributed points and are consigned to an equally wide distribution of destinations.

Economy demands that the various individual consignments be grouped together so as to take advantage of trunk haulage arrangements making minimum demands on man-power, rolling stock and route capacity. This grouping takes the form of collecting individual parcels into van loads and vans into trains. The van may be a road vehicle which makes the complete journey without the economy of further aggregation; it might be a railway vehicle operating between private sidings or, most probably, a container which would commence and terminate its journey by road but which could cover large distances on rail or sea.

In modern industry much transport is concerned with transfer of semi-finished products from one manufacturing plant to another; standardisation of containers, cribs and stillages is therefore necessary so that they may be taken from the point of production, stored where necessary, sorted, transported and distributed with a minimum of man-handling. The unit consignment may, therefore, be a single parcel, a crib suitable for handling by a fork-lift truck, a container of capacity varying from 1 to 20 tons, a road trailer capable of being rail-hauled "pick-a-back" fashion or a railway wagon.

The high capacity and productivity of the railway system depends on the controlled passage of a sequence of trains consisting of large numbers of vehicles coupled together, thereby operating in closer average headways than the road system in which each vehicle must allow a braking distance between itself and the preceding vehicle. Where the traffic is offered in train loads (e.g. power-station coal), the advantage of the railway system is undisputed. However, most consignments to particular destinations are on a much smaller scale and the cost of assembly of a multitudinous number of items to make up a profitable train represents a serious economic drawback to the railway as a competitor for the freight business.

11.1. The hump yard

Initially small wagons were staged from point to point, being marshalled into trains at a great many small yards. At one time the sorting was performed by loose-shunting of wagons into a fan of sidings by means of a shunting engine. This practice is still current

in some yards. The next stage in development was the provision of a hump over which train was pushed slowly by a locomotive. The wagons would accelerate down the hum to provide sufficient spacing for the switches to be altered between cuts of wagons. As wagon passed into the appropriate sorting siding men known as "runners" would appl the handbrake often while riding to the target point. At other times slipper brakes wei used.

Typically the track falls away from the crest of the hump at a gradient of 1 in 16 fc about 20 metres to provide adequate spacing between adjacent cuts of wagons.

11.2. Automatic retarders

Operation was so hazardous that a degree of automation was obviously required an partial solutions eliminating the employment of runners have been devised. These ai based on electrically controlled, hydraulically operated retarders which grip the whee of the vehicles, slowing them down by friction. These are usually arranged in two group: primary and secondary retarders. As the wagon comes down the road from the hump it speed is measured and also its rate of deceleration. The section where this is done i

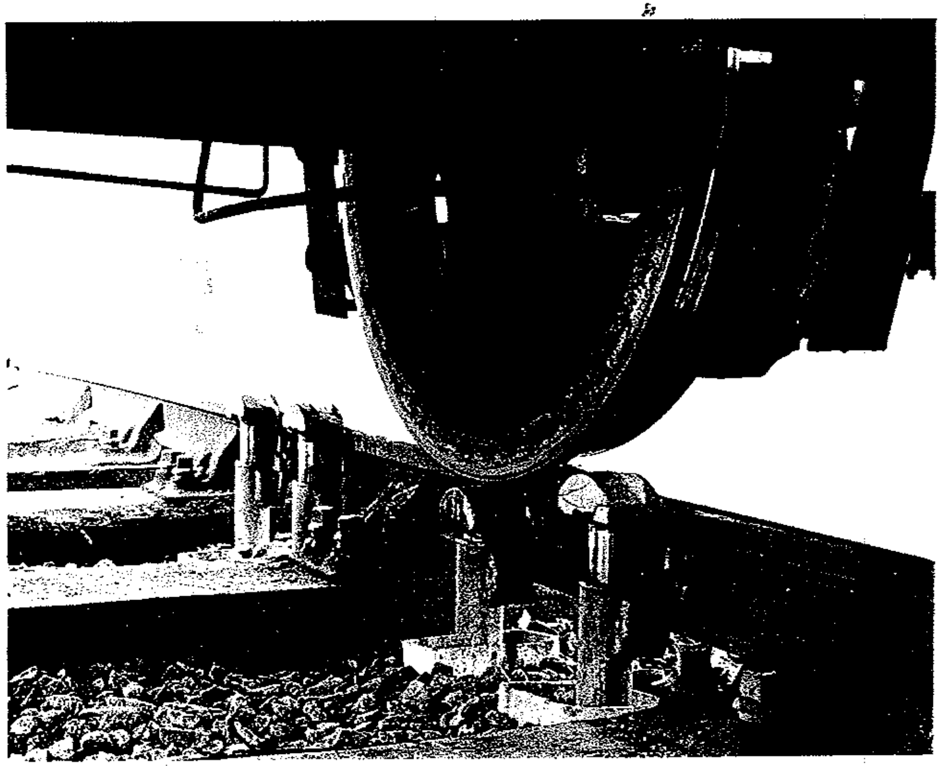


FIG. 11.1. Dowty retarder.

usually inclined with a gradient of 1 in 50. The gradient through the retarder itself is 4% (1 in 25) so that, if a wagon is inadvertently brought to rest within the retarder, it can be restarted by gravity. The main sorting sidings are laid out with a gradient of about 1 in 1200 to compensate for the rolling resistance of the wagon. This accounts for the term "gravity yard" which is sometimes used in place of "hump yard".

The measured characteristics of the wagons making up each cut is fed into a computer together with information about wind speed, destination siding (including whether full, half full or empty) and the retarders are so controlled that the wagon leaves the second retarder at just that speed which will carry it as far as the wagons already in the siding without undesirable impact.

Unfortunately, this system does not conform to the dictates of good control engineering because, once the wagon has left the secondary retarder, there is no "feedback" to compare the accuracy of the prediction with actuality and no means for making the correction. For example, a wagon may be a "bad runner" at low speed on curves so that it would not travel down the sidings as far as appeared to be likely from its motion on leaving the hump. It would, therefore, come to a stop at a point near the entrance to the road and the following wagon would strike it with excessive impact.

11.3. Hydraulic retarders[1]

The Dowty system

The hydraulic retarder unit shown in cross-section in Fig. 11.2 enables the speed of the faster wagon to be limited to 1·6 m/s (3½ m.p.h.) without affecting the slower wagons. The operation of the unit is as follows:

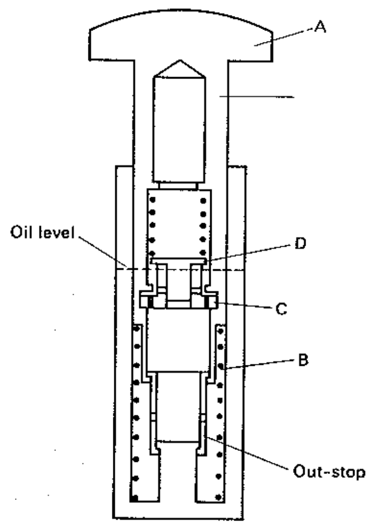


FIG. 11.2. Cross-section of retarder.

(a) The flange of the wagon wheel engages with piston A forcing it downwards aga spring B.
(b) The increase in internal pressure causes speed-control valve C to close.
(c) Further increase in pressure causes relief valve D to open. This is set so that furt depression of the piston is met by the hydraulic pressure necessary to force through to the relief valve thereby extracting energy from the moving wagon.
(d) After the passage of the wheel, the action of spring B causes the piston to m upwards, opening the speed-control valve so that oil may flow freely past the pist If the wagon is travelling more slowly than 1·6 m/s the piston moves down slowly to lift speed control valve C so that there is no hydraulic resistance to pis movement.

In certain parts of a yard it may be necessary to propel wagons as well as to retard th A hydraulic booster-retarder (Fig. 11.3) may be used for this purpose. This is similar the aforementioned retarder but is connected to a central source of hydraulic energy that the wagon may be speeded up if necessary. When acting as a retarder the actio similar to the preceding, except that when the speed-control valve closes, the consequ pressure acts on the back of "cut-off" valve E via the ports of slide valve F. Furt movement traps fluid behind the cut-off valve thus locking it in position. The press continues to rise until it operates a relief valve as before.

When the piston is depressed by a slow wagon, oil escapes from the main cylinder return via the orifice in the speed control valve. Further movement carries the sleeve va E to a point where its ports coincide with openings in valve stem G which allows ma pressure to act on the piston where it lifts the speed-control valve and forces the h sharply upwards forcing the wagon forwards.

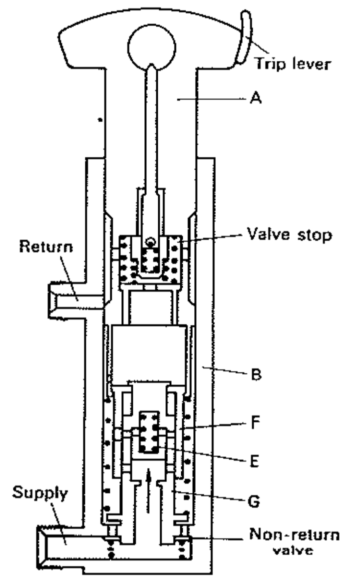


FIG. 11.3. Booster-retarder.

Retarders can also be provided in the form of arresters. These may be remotely controlled to retard wagons irrespective of their speed. Figure 11.4 shows a possible sequence for bringing a wagon to a halt from 10 m.p.h. (4·47 m/s).

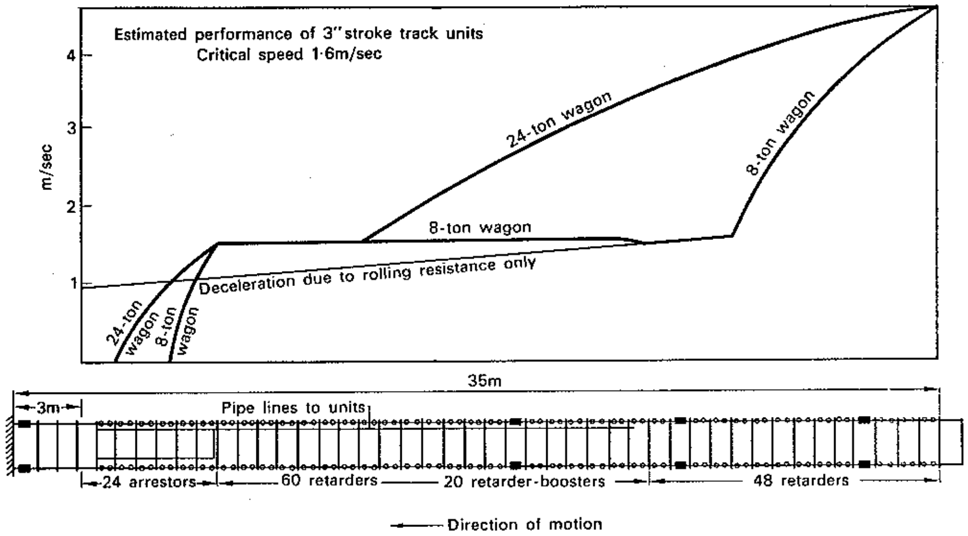


FIG. 11.4. Sequence for bringing wagon to rest.

The piston-type oleo-pneumatic retarders are somewhat noisy and are prone to leakage of oil.

Spiral-type retarders

The Dowty type retarder is not considered suitable for use in the extreme North, partly because of the effect of extreme cold on their operation and partly because of the vulnerability of small isolated units to damage by snow ploughs. Spiral-type retarders have therefore been developed by A.S.E.A. in Sweden to operate over a wide range of ambient temperatures, although a special hydraulic oil is recommended for temperatures below −25°C.

The general arrangement of a spiral retarder is shown in Fig. 11.5(a) and a cross-section in Fig. 11.5(b). The principle of operation is as follows: referring to Fig. 11.5(a), the wheel of the wagon strikes the helical fin "A" which is attached to the periphery of cylinder "B" causing this to rotate. This rotation is resisted by the action of pump plungers acting on the spiral cam within the cylinder 2 (Fig. 11.5(b)).

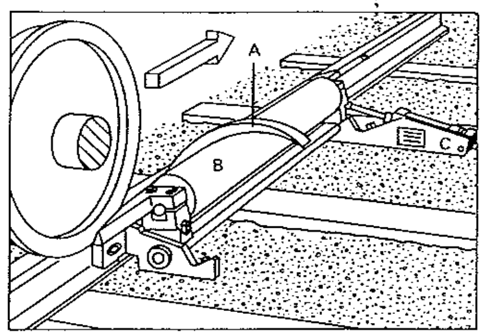
Control valves 3 are adjusted so as to impose hydraulic resistance to suction if the speed is higher than a predetermined value. If the speed is slower than this, only minimal resistance is offered. The speed is set when the unit is manufactured, to between the limits of 1 m/s and 4 m/s ($2\frac{1}{4}$ to 9 m.p.h.). The maximum speed for wagons passing over the retarders is 6 m/s ($13\frac{1}{2}$ m.p.h.).

The complete retarder may be lowered to permit the passage of locomotives over retarder by the action of a compressed air cylinder, shown as "C" in Fig. 11.5(a). application of this system to the marshalling yard at Helsinborg has been described Carlson and Stenow.[4]

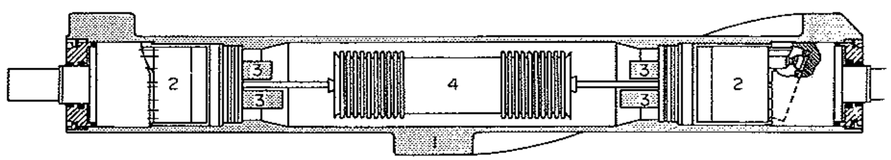
11.4. Control of wagon movement using rope-hauled mules

A number of yards are fitted with systems of continuous wagon handling which ens that, after separation, each wagon is caused to maintain a constant predetermined sp until it reaches the target point. The actual propulsion of the wagons is usually by mı which run on separate narrow gauge tracks between the main rails or on the lower fla of these rails. Propulsion of the mules is usually by wirerope but in some cases low-volt d.c. powered rubber-tyred mules are used.

A representative rope-operated yard is that at Basel Muttenz II opened in May 197(is fully automated with computer control and follows the general pattern of a hump y although, because the receiving sidings are inclined, there are no hump locomotives. computer is arranged to control the retarders which hold the approaching trains on slope, primary and secondary electro-magnetic retarders, switches for distribution of



(a) General view



(b) Section of rotating cylinder
1 Cylinder with spiral cam
2 Oil pump
3 Control valves
4 Volume compensator

FIG. 11.5. Spiral-type hydraulic retarder.

cuts into the classification tracks and rope-hauled mules which propel the wagons along those tracks at a speed of 1·5 m/s (3·3 m.p.h.).

A general view across the yard (Fig. 11.6) shows a number of mules lined up at the entrance to the classification tracks. These mules are arranged to run on the inside of the running rails, being guided and supported by rollers which interact with the rail where the foot joins the vertical flange. They are so designed that a locomotive or wagon may pass over them without interference, but when propulsion or retardation of a wagon is required, a roller (Fig. 11.7) can be raised so as to engage with the flange of the wagon wheel. Each classification track has its own mule system for propulsion and the driving motors are arranged in a tunnel which runs across the yard at right angles to the tracks, as shown in Fig. 11.8. The general arrangement of the system is shown in Fig. 11.9. The capacity of the Basle-Muttenz yard is 4500 cars per day.

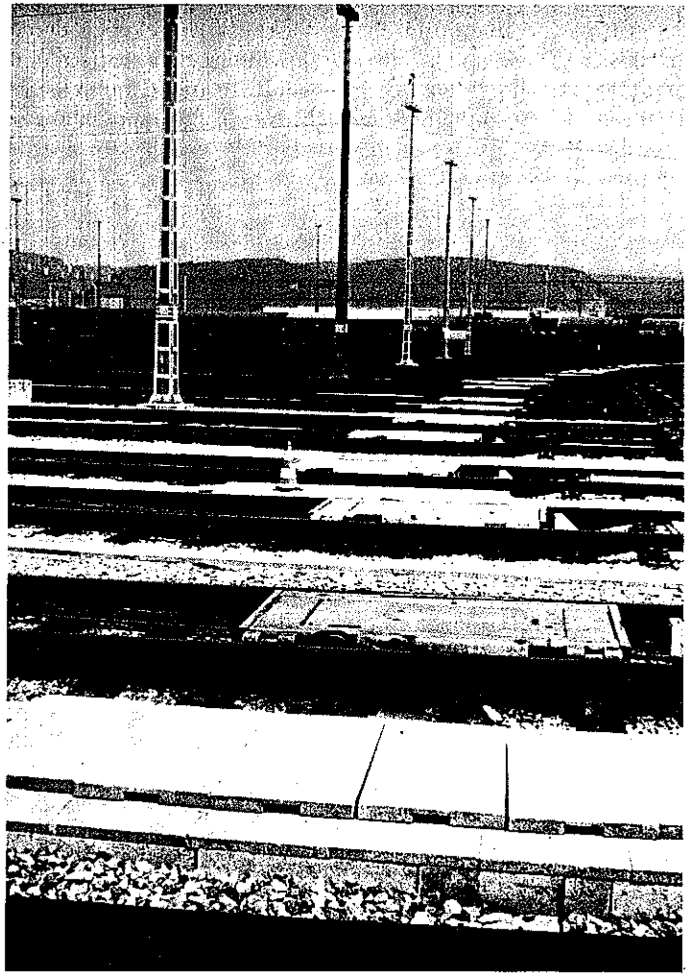


FIG. 11.6. General view across Basel Muttenz II classification yard.

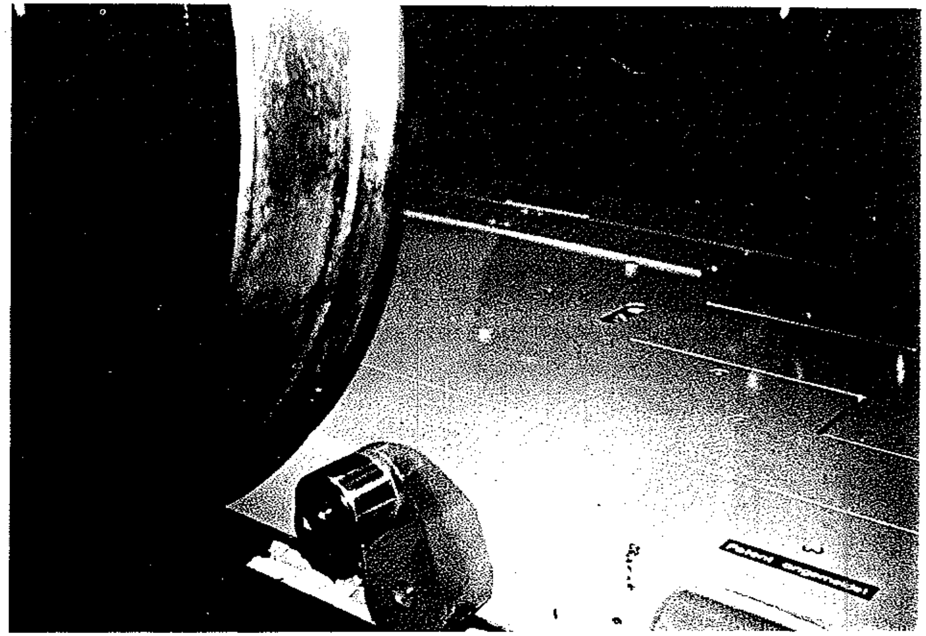


FIG. 11.7. Propulsion system.

FIG. 11.8. Rope haulage system for marshalling yard.

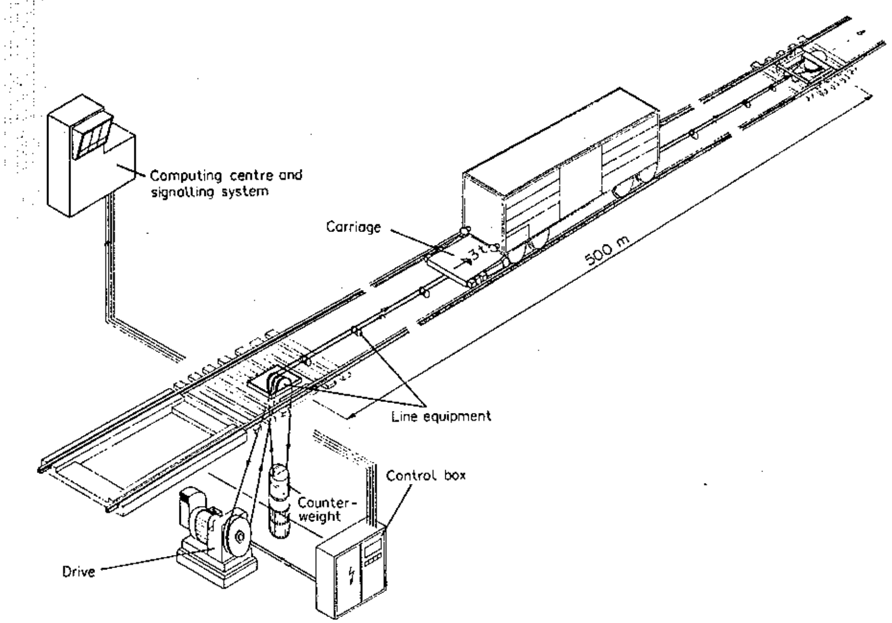


FIG. 11.9. Assembly of rope-hauled mule system.

11.5. Linear motor propulsion along classification tracks

In Japan the Shiohama marshalling yard uses linear motors. Although the yard is basically a hump yard fitted with conventional primary and secondary retarders, linear motor propelling and retarding units provide extra controls necessitated by high winds and the handling of freight wagons containing explosive fluid and requiring lower coupling speeds.

Inside the main tracks along the whole length there are separate narrow-gauge tracks for the linear motor propelling and braking units. These units are normally positioned at the entrance to each classification track. As a wagon approaches it is detected by the distance unit which actuates the pusher unit which starts in the same direction, engages with the wagon wheels and positively propels it down the yard at 13–15 km/h (8–9½ m.p.h.) until a point is reached at a predetermined distance from the waiting group of wagons. The speed of the wagon is then reduced to a pre-set level when it is released to coast until it bumps into a group of wagons with an impact limited by the controlled speed. Meanwhile the linear motor unit reverses its direction and returns towards the entrance of the classification track at a higher speed. If, on its way back to the initial waiting position, another wagon is encountered this wagon is sensed and the linear motor unit stops, reverses direction and controls the wagon in the normal manner. At the end of the classification tracks, remote from the hump, there are spring-loaded end stoppers which bring the first wagon to stop at a standard position. The propulsion units are

provided with counting devices which record the distance travelled from the last enga ment point so that the wagons are always released at an acceptable distance from last wagon in the siding.

Figure 11.10 shows a pusher unit. The driving rollers are retractable horizontally a are mounted on a sub-frame which is connected to the main frame by damping un which minimise the shock when the pusher unit engages with a wagon which happens be stationary or moving at a different speed from the pusher unit.

FIG. 11.10. Linear motor propulsion unit for Shiohama yard.

The apparatus for controlling the speed of the wagons is distributed between fi separately supported units as follows:

(a) A distance unit which detects the presence of a wagon and which embodies t main sensor-controlled functions.
(b) A motor unit which carries two linear induction motors which react with t reaction plates fixed to the track.

(c) A control unit which incorporates the automatic operation programs and is capable of making many decisions independently of the main computer.
(d) A brake unit which incorporates a pneumatic brake which supplements when necessary the regeneration braking action of the linear motor.
(e) A pusher unit which incorporates two sets of rollers which engage with the wheels of the wagon under control.

The speed of impact is normally limited to 4 km/h ($2\frac{1}{2}$ m.p.h.) but is halved for dangerous goods.

Among the many advantages claimed for the system are the following:

(a) the effect of disturbing factors such as wind or unfavourable wagon rolling characteristics are minimised;
(b) wagon speed is controlled at a comparatively high value;
(c) the system is independent of the adhesion limit;
(d) wagons requiring safe bumping can be catered for;
(e) the track can be easily maintained;
(f) the system contributes little noise and no other pollution to the environment.

11.6. Simulation of throughput of an automated flat yard

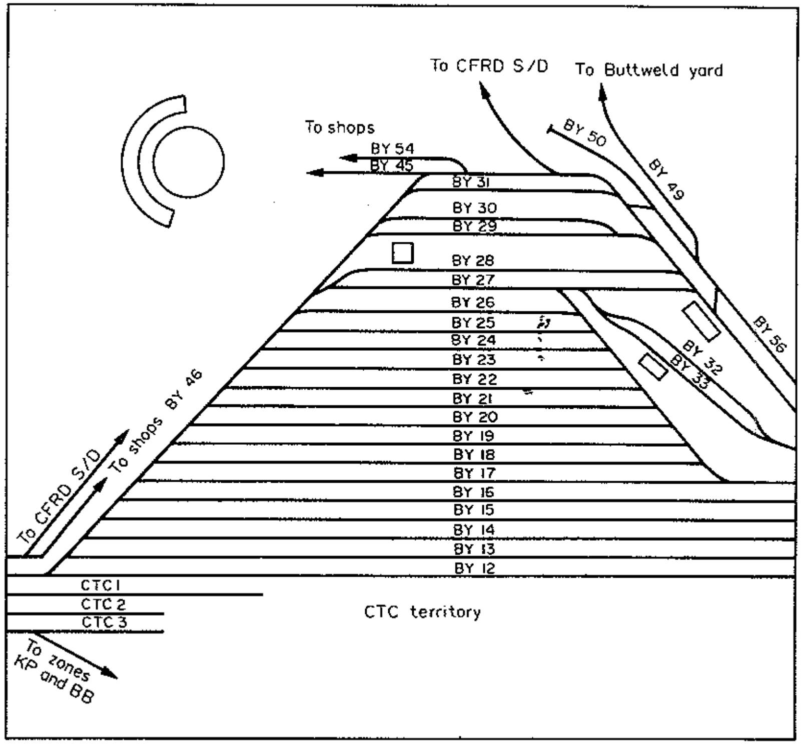
A number of yards exist throughout the world which carry a volume of traffic which is important from the point of view of the locality served but which is insufficient in volume to justify investment in a full-scale hump yard. However, the switching crews in classification yards carry out a task which can be dangerous and which is very unpleasant during inclement weather. Moreover, much of this work has to be done at night. It is therefore to be expected that it will become increasingly difficult to recruit staff and an economic means for automating these yards is desirable.

A recent study[5] was therefore undertaken to simulate the traffic flow in a moderately sized flat yard, assuming that it were equipped with a simple form of linear induction motor arranged for moving wagons on the level. The question at issue was whether or not the proposed system would produce an adequate throughput.

A yard in Canada, the Belleville Yard situated between Toronto and Montreal, was selected for study. Figure 11.11 shows the layout of the yard, and Fig. 11.12 shows the method of shunting envisaged. The first stage is simply a feeder system which takes a cut from the rear of the train and positions it at a fixed point called the "feeder point". (This movement is not necessary when starting on a train which is drawn up with its leading vehicle to the left of the feeder point.) The linear motor in stage 2 picks up any cut positioned at the feeder point and pulls it along the main track until it is past the clearance point. It then returns to collect the next cut whilst the linear motor in stage 3 picks up the cut and moves it back along the main track and then past the points and, whilst it does not itself enter the classification track, the cars do so covering a short distance over the points under their own momentum. At a fixed point just inside the classification track, the final linear motor will engage the cut and propel it down the track to a point some distance from the rear of the train under assembly, known as the "release point". (It is suggested that this can be most economically determined by radar but a counting device might be used, as in Shiohama.) Once the release distance has been

reached the cut driven by the linear motor would be decelerated down to a low sp after which it would be released to roll on and gently couple with the stationary wa; ahead.

The simulation was based on realistic data of traffic flow. The most important fea was the representation of the amount and dispersal with time of the traffic to be hanc A computer printout showing the number of wagons arriving and departing during



Track assignment BY zone no. 3	Marshalling yard	Track assignment BY zone no. 3	Marshalling yard
12	Receiving and departure	26	Bad order cars
13	Receiving and departure	27	Run around track
14	Receiving and departure	28	Cleaning track
15	Montreal and East traffic	29	Cleaning track (preference to hoppers)
16	Toronto hump traffic	30	"Hold" traffic
17	Marshalling track	31	Storage track
18	Marshalling track	32	Scale track
19	Marshalling track	33	Van sliding
20	Short haul cars	45	Shop run around
21	Short haul cars and supplies	46	West End shop
22	Cars to be cleaned	49	Cut off to Butt-Weld yard and Rail yard
23	Rail yard cars		
24	Ottawa and Smith Falls S/D traffic	50	Stores track
25	Local commercial traffic (for "BA" and "BB" zones)	54	To repair tracks—East End shops
		56	Departure track for northbound trains

FIG. 11.11. Layout of Belleville yard.

month was analysed on the basis of the destination of each cut of wagons. The distribution of wagons per train was found to be negative exponential and the distribution of wagon lengths and wagon weights was gaussian. The probability of any wagon being required to be sent to a particular destination was based on the actual proportions recorded on the computer printout.

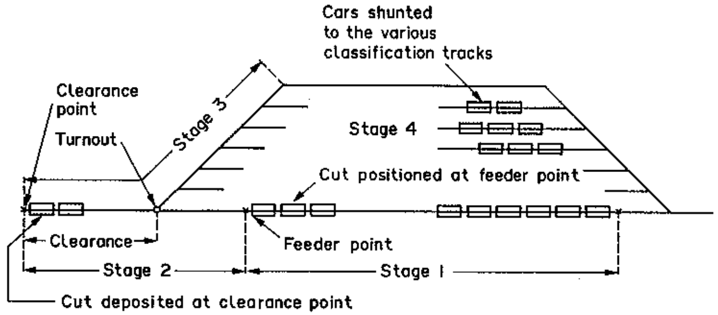


FIG. 11.12. Shunting method.

A flow chart for the logic is set out in Fig. 11.13. The value of the time interval required for the train to be correctly positioned is obtained from a normal distribution using a pseudo-random generator. Similarly, the values of the other variables such as length and weight of wagons which go to make up a train are determined. The destinations are then

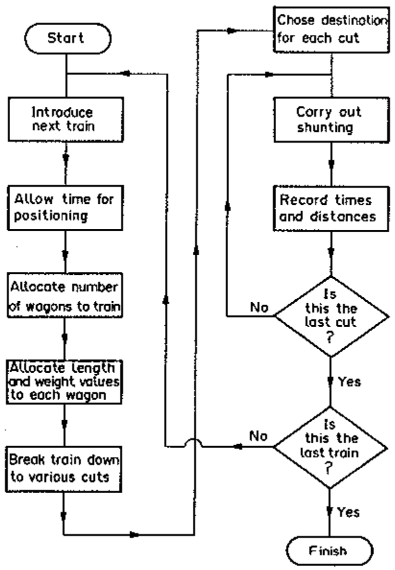


FIG. 11.13. Flowchart for the program logic.

generated and the actual time required to propel the individual cuts into the yards is t calculated and recorded.

An example of the make-up of three trains and the manner in which the wagons w distributed within the classification tracks is given in Table 11.1.

After a few initial simulations it was discovered that the linear motor in stag (Fig. 11.12) would remain idle for appreciable amounts of time while waiting for sta to supply a new cut. Accordingly two modifications were carried out. Firstly, the sp of the linear motor when returning empty was allowed to become greater than w propelling a cut of cars, and secondly, the feeder point was allowed to move progressi toward the train being marshalled.

TABLE 11.1. TRAIN MAKE-UP BEFORE AND AFTER SHUNTING

	Number of cars	Number of cuts	Length of train, m
Before shunting			
Train 1	40	14	615·5
Train 2	43	14	624·6
Train 3	17	6	246·8
After shunting			
Track 1	22	6	306·6
Track 2	11	4	178·4
Track 3	6	2	84·4
Track 4	13	4	210·7
Track 5	5	5	79·1
Track 6	2	1	30·0
Track 7	13	4	185·9
Track 8	14	3	207·5
Track 9	14	5	204·3

A typical result indicates that the time required to sort 100 wagons was 1 hour minutes so that the system would very adequately cope with the amount of traffic to expected at a yard such as Belleville which would be something less than 200 wag per day.

A total daily capacity of 1800 wagons would, however, be quite inadequate for a l central yard which might be required to handle up to eight wagons per minute.

The principle of rope or linear motor propulsion on the classification tracks themsel is of course applicable to a high-capacity yard but some form of gravity separatio necessary for high throughputs. This requires a falling gradient (typically 1 in 16 20 metres) and some form of retarder to regulate the velocity of the cars as they enter classification tracks. If a proper system of speed control exists in the classificat tracks there would appear to be no need to vary the speed of wagons as they leave retarders. Therefore all retarders could be set at constant speed and could be passive action such as those types described in Section 11.3. The speed of the wagons being c trolled at all times, the role of the computer would be confined to setting the routes the individual cuts.

References

1. SADLER, J., The design and equipment of modern marshalling yards, *Proc. Inst. Civil Engrs.*, Part 2, vol. 2, p. 755 (1953).
2. *The Railway Gazette*, 13th May 1960.
3. BUYANOV, V. A., Technical and technological realization of system "Man and Machine" as applied to the control of railway yards. *Proc. 1er Symposium Int. sur la Régulation du Trafic, Versailles*, 1970, Preprint no. 4, p. 37.
4. CARLSON, H. and STENOW, A., Neues system für die geschwindigkeits beeinflussung von ablauferiden güterwagen. *E.T.R.—Eisenbahntechnische Rundschau*, Heft 5, p. 279 (1976).
5. BARWELL, F. T. and LEECH, D. J., Automation of marshalling yards with particular reference to linear motors. *Proc. Inst. Civil Engrs.*, Part 2, vol. 6, p. 627 (1979).

CHAPTER 12

Control of Acceleration and Power

12.1. Limitations

The tractive effort or torque applied to the axles of a powered vehicle must fall with the limits shown in Fig. 12.1. The curve on the right-hand side of the figure implies th the power available is limited to a constant value whatever the speed. It therefore tak the form of a hyperbola. In many propulsion systems, power developed is not indepe dent of speed due to limitations in the transmission or torque amplification systems that the curve represents an envelope bounding the area occupied by the real values. other cases, particularly where electrical equipment is involved, high levels of power m be developed for short periods only, i.e. a certain amount of overloading may be permitt for a short time. This could be represented on Fig. 12.1 by a series of hyperbolae corr ponding to various degrees of overloading.

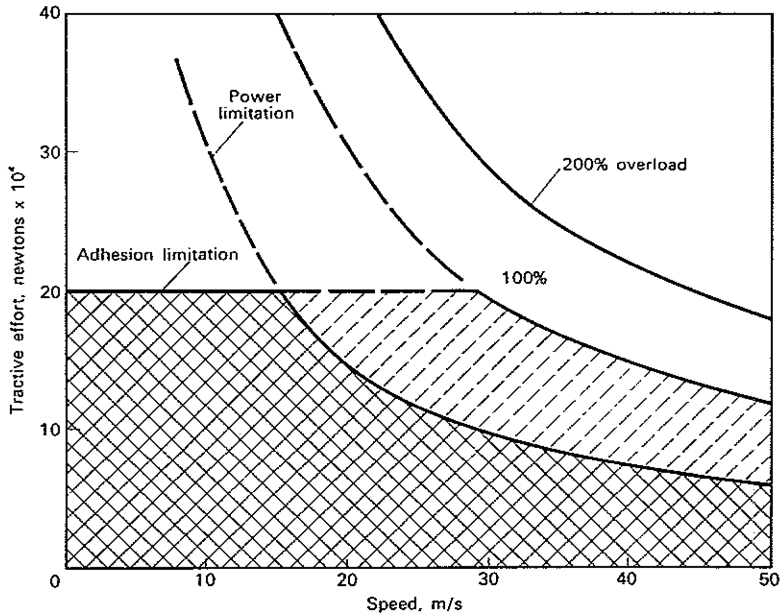


FIG. 12.1. Limits of tractive effort.

The horizontal line at the left-hand side of the diagram represents the adhesion limit. This is exceedingly variable and means must be provided to prevent it being exceeded. In the absence of automatic slip control, this makes one of the most critical demands for skill on the part of the driver.

Serious consequences likely to result from exceeding the adhesion limit are, in the case of road vehicles, loss of steering control (skidding) and wear of tyres. In the case of the railway, severe damage may be done to the rails and motors may be damaged by overspeeding.

12.2. Equations of motion

As indicated in previous chapters, tractive effort (F) is the quantity which has to be controlled in order that the motion of a vehicle shall be constrained in conformity with the requirements of the system as a whole. The equation of motion may be written

$$M\frac{d^2x}{dt^2} + Mg\sin\alpha = F - R \qquad (12.1)$$

where R equals tractive resistance. This is usually a non-linear function of dx/dt and generally expressed as a function of mass, i.e. so many newtons per kilogramme and traditionally, as so many pounds force per ton. This is inconvenient because at high speeds air resistance, which is determined by the shape and volume of a vehicle rather than its mass, predominates. It will be recalled from previous chapters that where small deviations of velocity were involved resistance was linearised, the symbol η representing force in newtons per metre per second.

Thus
$$\frac{d^2x}{dt^2} + \frac{\eta}{M}\frac{dx}{dt} + g\alpha = \frac{F}{M} \leqslant \mu g \qquad (12.2)$$

writing α for $\sin\alpha$ for small values of incline, or in terms of velocity

$$\frac{dV}{dt} + \eta\frac{V}{M} + g\alpha = \frac{F}{M} \leqslant \mu g, \qquad (12.3)$$

or, introducing power (equals H),

$$\frac{dV}{dt} + \eta\frac{V}{M} + g\alpha = \frac{H}{MV} \leqslant \mu g. \qquad (12.4)$$

The maximum possible acceleration will occur as $V \to 0$ when it will be limited to μg whatever the power installed.

As speed is increased, the margin of power available for acceleration will diminish until a point is reached when the vehicle will move forward at constant speed. This is known as the "balancing speed".

Thus
$$H = \eta V^2 + \alpha g VM \text{ watts} \qquad (12.5)$$

or, if R is a function of speed denoted by R_v and expressed in newtons/kg,

$$H = (R_v + \alpha g)VM \text{ watts.} \qquad (12.6)$$

Balancing or maximum speed becomes

$$V_m = \frac{H}{(R_v + \alpha g)\,M} \text{ m/s.} \qquad (12.$$

12.3. Values of resistance coefficient

The value of R_v is therefore of vital importance in traction calculations but reliat data, particularly relating to road vehicles, is very difficult to obtain.[1]

Some values quoted from the literature are given below in the units reported by tl investigators.

For road vehicles, Steeds[2] separates air resistance from rolling resistance as follow

$$\text{Air resistance} = KAV^2 \text{ lb-f,} \qquad (12.$$

where K is a constant having values ranging from 0·0005 to 0·003 depending on the sha of the vehicle,

A = frontal area in square feet,
and V = speed relative to air in miles per hour.

$$\text{Rolling resistance} = (a + bV)W \text{ lb-f,} \qquad (12.$$

where a = coefficient ranging from 15 to 500,
b = coefficient ranging from 0·1 to 3·5,
W = weight of vehicle in tons.

For railway work Dover[3] quotes

$$R = \text{total resistance} = (4{\cdot}1 + 0{\cdot}055V)W + AV^2\,(0{\cdot}0028S + 0{\cdot}0000122nl) \qquad (12.$$

where S is a shape factor equal to unity for a flat end but reduced for improved aer dynamic shapes, n equals the number of coaches and l equals the length of each coach.

Andrews[4] quotes a specific resistance of $2{\cdot}193 + 0{\cdot}041V + 0{\cdot}001029V^2$ for eig coaches hauled in still air. The coaches were 57 feet (17·4 metres) long and each weigh 31 tons (31,500 kg). They were tested using the constant-speed equipment referred to Chapter 6. This author does not distinguish between rolling resistance and air resistan but, separating terms, it may be inferred that the first two terms of Dover's expression a over-estimates.

Dover's method of allowing for shape by calculating volume is not convincing and tl method of Lipetz[5] using peripheral area may be more realistic. Thus Andrews' resu may be written in the form

$$R = (2{\cdot}193 + 0{\cdot}041V)W + 1{\cdot}62 \times V^2 \times pnl \times 10^{-5}$$

where p equals the length of the periphery of the cross-section of the vehicles in fe measured from the level of the top of the rail.

The resistance to motion or drag of a vehicle is made up with the following componen

1. rolling resistance,
2. bearing friction,
3. surface friction drag,
4. normal pressure drag.

At low speeds items (1) and (2) predominate. Rolling resistance represents the energy loss occurring between wheel and track. This is very low for rail on straight track in the absence of side forces. Where lateral forces exist due to side winds, centrifugal force or vehicle misalignment, drag may be of considerable significance.

For pneumatic tyres, the drag is affected by type of tread and tyre construction and degree of inflation. A typical value may be taken as 0·2 newtons/kg.

Tests on the Transit Expressway Vehicle (see Chapter 17) reported by Westinghouse give results in accordance with the following expression over the range from 0 to 50 m.p.h. (0 to 22·4 m/s).

$$\text{Tractive resistance in newtons} = 0{\cdot}077W + 137{\cdot}6 + 54{\cdot}5V + 0{\cdot}37AV \qquad (12.11)$$

where W = mass of vehicle in kg,
V = velocity in m/s,
and A = vehicle frontal area in square metres.

The majority of modern vehicles are fitted with rolling contact bearings giving low and fairly constant friction. Plain bearings lubricated by solid grease are traditional on railway freight wagons and present a high coefficient of friction when starting from rest. This falls to a low value when speed increases sufficiently for hydrodynamic lubrication to be established. Refined modern designs of oil-lubricated bearings also operate on the hydrodynamic principle, often attaining lower operating temperatures at high speeds than the corresponding roller bearings.

The aerodynamic contribution to drag becomes of predominant importance at quite moderate speeds. It arises basically from the viscosity of the air. The difference between surface friction drag and form drag may best be explained by considering the resistance to motion of a thin plate. When the plate is aligned parallel to the direction of motion it will be subjected to a tangential force due to the shearing action on the air with which it is in contact. This is surface friction drag.

When it is aligned with its plane normal to the direction of motion, drag will arise from any difference in pressure acting on the surfaces at front and rear.

The total drag acting on a moving vehicle will therefore be dependent on its size and shape. This can only be determined experimentally and is expressed as a drag coefficient C_D where

$$C_D = \frac{\text{Drag force,}}{A\dfrac{\rho V^2}{2}} \qquad (12.12)$$

$\rho V^2/2$ is the dynamic pressure of the airstream.

For any given body C_D will be dependent on Reynolds number, $\rho VD/\eta$, and Mach number. The Mach number relates to the effect of compressibility of the fluid and is not important for the speeds likely to be attained in land transport. Reynolds number determines the ratio of normal pressure drag to skin friction drag, systems having the same Reynolds number giving rise to the same value of drag coefficient. The significance of the terms are

ρ = density of air,
V = velocity of vehicle relative to air stream,
D = characteristic dimension of vehicle,
η = viscosity of air.

The effect of natural wind must be taken into account, particularly during the slowe parts of an operating cycle. The effect of relative direction or "yaw" may be particularl important. Andrews[4] reports that for freight trains, the worst condition will occur at a angle of yaw of 50°.

The existing data both for road and rail tractive resistances is totally inadequate for th control engineer who wishes to know not only the mean values on the occasion of test bu how these depend on the design features of the vehicle tested. Most of all he desires t know the "dispersion" of the data, that is, the frequency of occurrence of variations whic he will have to take into account in order to ensure that his control system will hav adequate range and speed of response.

The results of wind-tunnel tests on models of railway vehicles have not correlated we with track tests, probably because of the difficulty of reproducing the effect of the groun particularly as regards the space between the floor of the vehicle and the track. Rece tests on full-size automobiles and quarter-scale models at the Motor Industry Researc Association's Laboratories[1] have, however, provided good correlation (within 6%) wit measurements on the road after the latter have been corrected for rolling resistance. Th was determined by testing the vehicle under study within a shroud which was towed b another vehicle. The drag link from which measurement of towing force was derived wa also situated within the shroud so that aerodynamic forces acting on the car body we excluded. Drag due to the proximity of the ground to the underside of the vehicle was c course included.

The mechanical drag forces so estimated are reproduced as curves in ref. 1 from whic values of a and b in the expression $R_v = (a + bV)$ newtons/kg can be estimated over range of V from 9 to 35 m/s. An approximate value of b is $4{\cdot}9 \times 10^{-3}$ when V is given i metres per second and "a" ranges from 0·0438 to about twice this value.

Regarding aerodynamic drag coefficient, measurable variations are attributable t whether or not the radiator is open and to the action of the fan. However, far great variations are caused by the action of side winds.

Drag is quoted in the form

$$D = C_D A \frac{\rho V^2}{2} \qquad (12.1$$

where A is the frontal area,
ρ is the density of the air,
and V is wind speed component parallel to the longitudinal axis of the car.

For cars of recent vintage, C_D may be taken in these authors' units as about 0·4 for wind straight ahead and 0·59 for 20° of yaw. Units are not stated explicitly bu assuming that

D is in pounds force,
A is in square feet,
ρ is in slugs/ft^3 (taken as 0·00238 at sea level)
and V is in feet per second,
we have in the S.I. system

$$F = \begin{pmatrix} 0{\cdot}0438 \\ \text{to} \\ 0{\cdot}0876 \end{pmatrix} + 0{\cdot}0049V_g \Bigg) M + 0{\cdot}43\,\frac{\rho V_a^2}{2}\,A \text{ newtons} \qquad (12.14)$$

where V_g is velocity along the ground in m/s and V_a is the component of wind speed measured relative to the longitudinal axis of the car. ρ can be taken as 1·249 kg/m³.

A general treatment of aerodynamic factors is available in reference 6.

12.4. Estimation of distance–time relationships

It has been shown in Chapter 1 how the speed–distance curve may be constructed on the basis of the phase-plane diagram and how this may be further used to determine the speed–time curve.

The traditional method for evaluating these quantities is by means of the Dalby diagram.[7] This method is illustrated in Fig. 12.2 and is based on four quadrants, speed

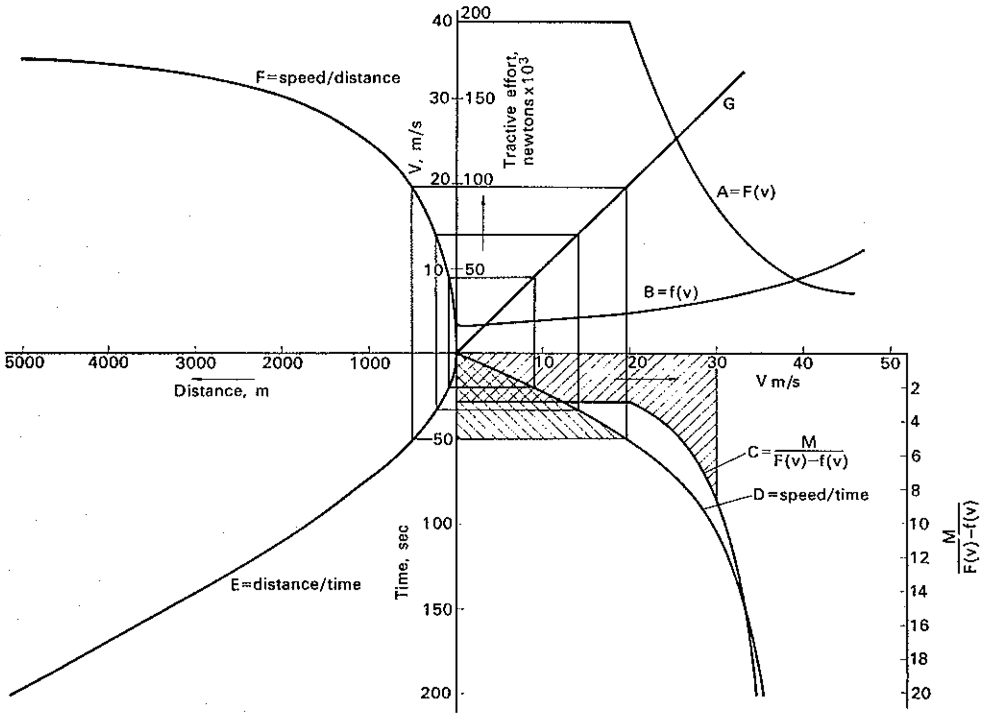


FIG. 12.2. Dalby diagram.

being plotted horizontally to the right of the origin and distance to the left. Tractive and resistive forces are plotted vertically in the upper quadrant and time is plotted downwards from the origin. Tractive effort is plotted in the upper right-hand quadrant against speed

(curve A), as is also tractive resistance (curve B). The effect of gradient, not being function of speed, may be introduced by adjustment of the zero.

The speed–time curve is governed by the equation

$$V = \frac{1}{M}\int (F - R)\,dt \qquad (12.1$$

but $(F - R)$ is a function of $V = f(V)$, (12.1‹

therefore

$$M\frac{dV}{dt} = f(V), \qquad (12.1$$

$$\int_{V_1}^{V_2} \frac{M}{f(V)}\,dV = (t_2 - t_1). \qquad (12.1$$

This integral is evaluated by plotting the reciprocal of net tractive effort against speed i the lower right-hand quadrant (curve C). Then the time corresponding to a particula speed is determined by drawing the perpendicular from the origin corresponding to th speed and evaluating the area enclosed between this and the ordinate. The result is plotte as curve D. Having determined V in terms of t, it is possible to determine distance, i. $S = \int V dt$. Thus taking the area bounded by D and the horizontal corresponding to an time t a distance–time curve, curve E can be constructed in the lower left-hand quadran The speed–distance curve F in the top left-hand quadrant is obtained by taking distanc and projecting downwards to E, across to D and then to G which relates the horizont and vertical speed scales.

In the case where incremental information is required, e.g. where a gradient is encou tered, the beginning of the slope is set off from the distance scale on to E and transferre across to D. A new line C_1 is set out taking into account the additional resistance of th gradient. The new speed–time curve is drawn breaking away from the original curve the point corresponding to the speed at which the gradient was entered. Curve E_1 will b continued until it intercepts a vertical corresponding to the next change in gradient. Th procedure would then be repeated until the braking cycle intervened.

The same procedure can be adopted for braking although it must be recalled th resistance has now to be added to braking effort. It is better to use a separate constructio commencing from the stopping point and to work backwards in time until operating spee has been reached. This will give braking distance, which, when subtracted from the tot distance, will show how far to the left the horizontal axis of the driving diagram should b extended.

There may also be speed restrictions. These can be shown in the top left-hand quadra and, if curve F tends to go above, then it must be broken off at the point of intersectio Curve E would continue from the corresponding point as a straight line of slope corre ponding to the speed limit and curve D would become a vertical straight line. Pow would be reduced to correspond to the net resistance leaving no margin for acceleratio With downward gradient some braking might be necessary.

12.5. Coasting

The assumption has been made that the maximum available tractive effort has been employed. This may not be the case for several reasons. Firstly, there may be no economic justification for making the journey in the minimum time or, secondly, traffic conditions may not permit it or, thirdly, the driver may have insufficient skill to achieve the best results.

Provided that the comfort and adhesion limits can be observed during the initial period, full power may be applied unless speed limits are encountered or until it is necessary to commence braking so as to terminate the journey at the desired point. The highest power is not necessarily the optimum from the point of energy consumption and, unless regeneration is used, braking represents an unmitigated waste of energy. If, however, power is cut off before braking becomes necessary so that the vehicle "coasts", i.e. uses part of its kinetic energy to overcome tractive resistance during the later stages of the journey, energy which otherwise would have been dissipated in braking is employed. In some suburban railways "coast" boards are placed sufficiently in advance of stations to assist drivers to save power. The curve for a journey on level track would look something like Fig. 12.3(a). Figure 12.3(b) shows power consumption.

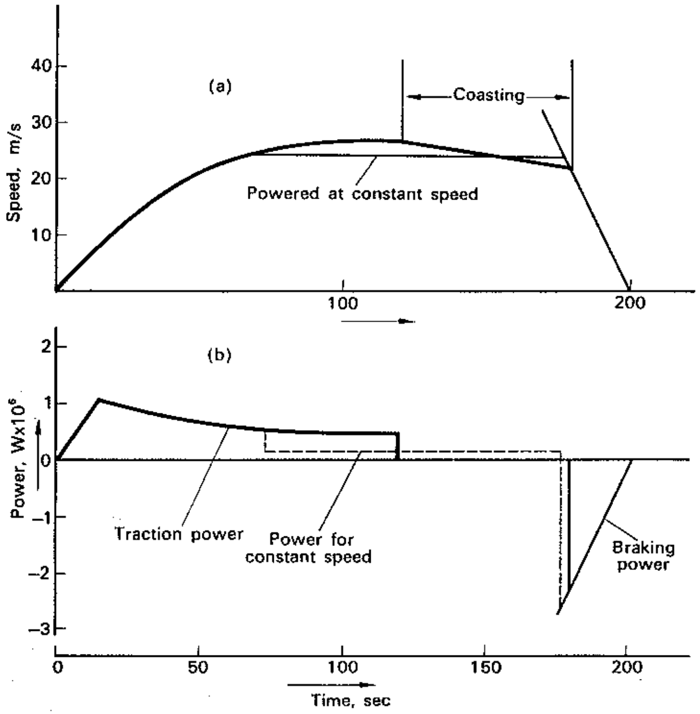


Fig. 12.3. Effect of coasting.

For extensive coasting to be worth while the normal running condition must be well removed from the balancing speed. Switching off power at the balancing speed would

result in a very rapid reduction in speed which would only be appropriate immediately before braking. Thus coasting is important for suburban rolling stock where power is installed mainly for acceleration and where speed is limited. In high-speed working a substantial proportion of power may be required to overcome tractive resistance so that coasting is less important.

12.6. Control of engine speed

Where the source of power required for propulsion of a vehicle is carried on the vehicle itself in the form of an engine or turbine, the means by which the power output is controlled must be integrated in some way with the control of the vehicle as a whole. Power output may be varied by change in output torque at constant engine speed or control of speed at constant torque or by combinations of both methods.

The control of engine speed by use of a centrifugal governor in the hands of James Watt in the year 1775 represents the birth of control engineering. This was recognised by Maxwell[8] as the first application of a feedback system to a device for control of power, although similar devices had been used previously as "regulation" in timepieces. Some doubt exists as to whether the use of the device in wind or water mills precedes or follows Watt's application to the steam engine.

Consider the system illustrated in Fig. 12.4. The main crankshaft is driven by piston and connecting rod which apply an average torque directly proportional to the steam

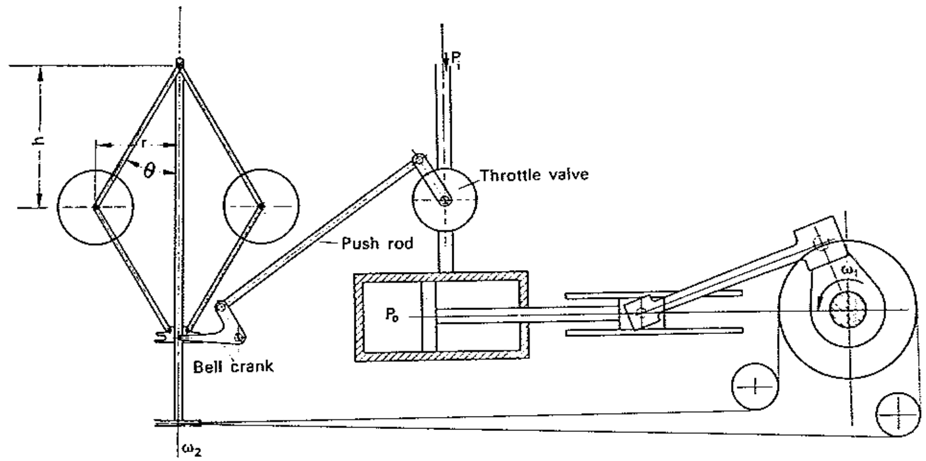


Fig. 12.4.

pressure P_o fed to the cylinder. This pressure may be varied as a proportion of the feed pressure by manipulation of the throttle valve. The governor is driven directly from the crankshaft and is so arranged that, as the balls fly outward under the action of centrifugal force, the sleeve will be raised and, acting through bell crank and push rod, will close the throttle valve.

The relation between the speed of rotation and the shaft and that taken up by the governor will be governed by the following expressions:

let ω_1 = speed of shaft,
ω_2 = speed of governor spindle,
k = gear ratio ω_1/ω_2.

Each ball of the governor experiences centripetal acceleration equal to $\omega_2^2 r$ and, neglecting friction and the weight of the linkage mechanism, this has to be balanced by the gravitational force acting downwards thereon. Therefore

$$\tan\theta = \omega_2^2 r/g \qquad (12.19)$$

where θ represents the angle made between the link and the vertical. If h equals the vertical height between the c.g. of the ball and the point where the axis of the arm intersects the axis of rotation then

$$r/h = \tan\theta, \quad \therefore\ h = g/\omega_2^2. \qquad (12.20)$$

This value of h is inconvenient for most machines but the construction will be assumed herein for the sake of simplicity. Supposing the arms form a simple diamond pattern, then the movement of the sleeve will be twice the change in h.

Let ω_t be the designed speed of the engine. Then $\omega_t - \omega_o$ represents the "error". When the error equals zero

$$\omega_2 = k\omega, \quad \therefore\ h = \frac{g}{k\omega^2} = h_0. \qquad (12.21)$$

Thus when the machine is running too fast

$$h - h_0 = \frac{g}{k}\left(\frac{1}{\omega_t^2} - \frac{1}{\omega_o^2}\right). \qquad (12.22)$$

Now when $h = h_0$ the throttle will be in its mid position so that $P_o = P_{t/2}$. When $h = h_{\min}$ the valve will be arranged to be fully closed and actual speed equal to twice the designed speed, i.e. $2\omega_o$. Then pressure in the engine cylinder can be represented by

$$P_t\left\{1 - 4\left(\frac{\omega}{\omega_t}\right)^2\right\} \qquad (12.23)$$

where P_t is inlet pressure. Cylinder pressure can be taken as a measure of the driving torque applied to the system. Any difference between this torque and the resistance to motion occasioned by the load will lead to a change in speed of the engine. The torque difference will be $I(d\omega/dt)$ where I is the inertia of the engine and its driven system calculated as the polar moment of inertia about the axis of the crankshaft. Thus

$$I\frac{d\omega}{dt} - K_2\left\{1 - 4\left(\frac{\omega}{\omega_t}\right)^2\right\} = T \qquad (12.24)$$

where K_2 relates cylinder pressure to driving torque and T equals load torque, or, grouping constant terms into a and b,

$$I\frac{d\omega}{dt} = a\omega^2 - b = 0. \qquad (12.2$$

Thus the action of the governor is described by a differential equation. The inertia term i obviously of great importance and the possibility of overshooting the mark is very rea In fact, a system as above would hunt very badly. A modern governor, for a diesel engin for example, might take the form shown in Fig. 12.5. Masses A, subject to centrifuga

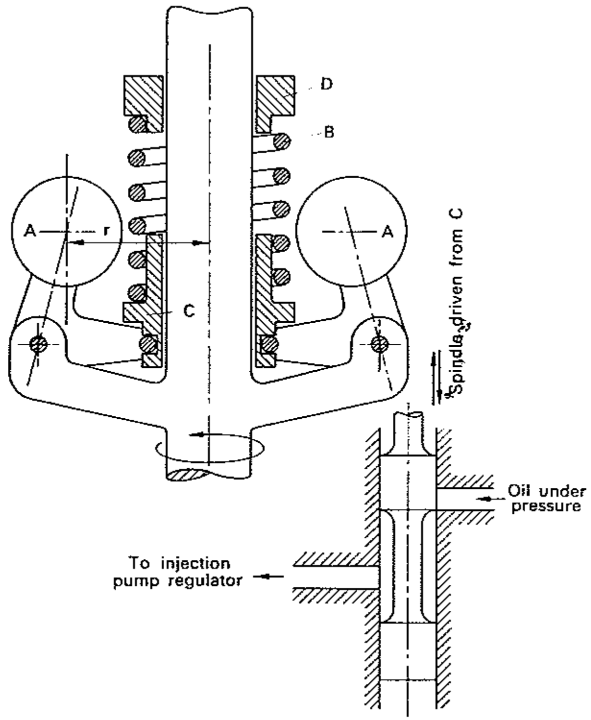


FIG. 12.5. Governor and oil servo.

force, act against spring B forcing up sleeve C. The force on each mass equals $M\omega^2(r + x$ assuming r = normal radius,
x = increase in radius,
M = mass.

Assuming bell-crank dimensions are such that sleeve and spindle C move vertically by a amount equal to x and assuming that this is resisted by the inertia of the masses whic will be neglected here and the spring and friction forces acting axially, then

$$M_1\omega^2(r + x) = n\frac{dx}{dt} + qx \qquad (12.2$$

where n represents viscous friction in the system and q equals the spring constant.

By selecting suitable values of M, n and q a sensitive, stable system may be obtained. This will operate an amplifying element so that, as soon as speed exceeds a predetermined value, the slide valve opens and oil flows to the injection pump to increase the speed of the engine. It will continue to do this until the governor returns to the normal position. Speed could be varied from outside the engine by moving headstock D and thus changing the value of x required to open the oil valve.

If a governor system were based on simple proportional control an increase in power demand would be compensated by input in proportion to the error. Thus in the steady state the error would attain a definite non-zero value. The speed of the engine would be slightly below its designed speed. This is known as "droop". The introduction of an additional term into the control whereby torque was increased in proportion to the integral of the error would enable "droop" to be eliminated. Integral control is destabilising but stability can be enhanced by the introduction of a function proportional to the differential coefficient of error. We thus have a "three-term controller" as shown schematically in Fig. 12.6. Table 12.1 displays some standardised definitions.

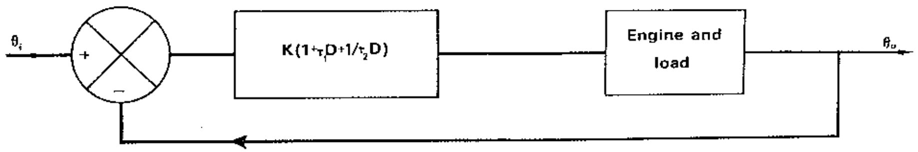


FIG. 12.6. Three-term controller.

TABLE 12.1. DERIVATIVE AND INTEGRAL CONTROLS

Type of control	Nature of feedback	Proportionality constant from B.S. 1523:1954	Contribution to servo-motor torque	Effect on performance of simple system
Simple proportional	Deviation	S_0	$-S_0\theta$	—
Phase advance stabilisation	Negative first derivative of deviation	S_1	$-S_1 d\theta/dt$	Decreases tendency for free oscillations
First integral	Negative first integral of deviation	C_1	$-C_1\int\theta dt$	Gives delayed compensation for velocity misalignment
Second integral	Negative second integral of deviation	C_2	$-C_2\int\int\theta dt dt$	Gives delayed compensation for acceleration misalignment

A full treatment of the application of control theory to governors has been published by Welbourne.[9]

12.7. Automatic transmissions

Referring to Fig. 1.14(b), it will be noted that for maximum acceleration there exist distinct values of vehicle speed at which change in gear ratio should take place. A Watt

governor connected to the engine output shaft would provide a measure of speed a initiate gear changing when appropriate.

A system whereby this can be effected is shown in Fig. 12.7 and represents the arran ment used in an automatic British Leylands mini-car.[10, 11] The governor is arranged control the admission of oil to a series of cylinders which engage the clutch or the bra drums which restrain the appropriate member of the epicyclic gear train. Under the f throttle conditions of Fig. 1.14(b) the gear-change action takes place at an engine spe of 5000 rev/min (524 rad/sec).

Circumstances frequently occur when it is required to constrain a vehicle to operate less than full power. These may result from traffic conditions, speed limitations or the ne to economise in energy consumption.

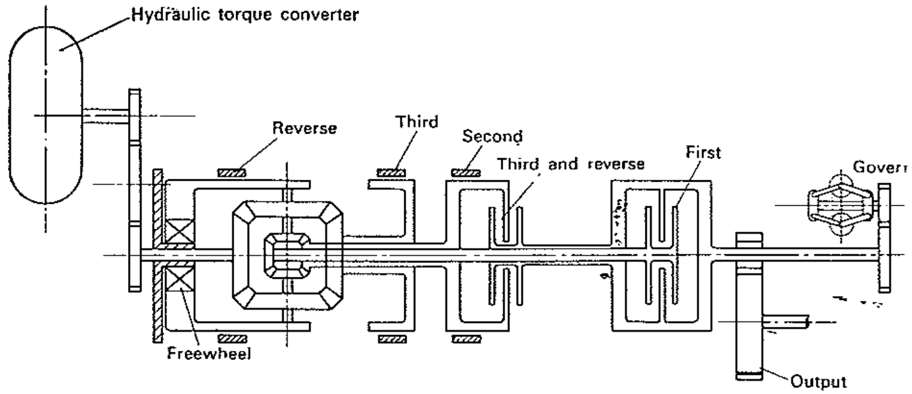


FIG. 12.7. Automatic gear change—British Leylands.

In a "two-pedal" control system the input of power control is the accelerator pe which is fully depressed for maximum power. Thus under conditions of partial loa control of gear ratio is insufficient and must be related to engine output. In the "mi system a connection is made from the accelerator pedal to the governor so that the engi speed at which the gear change takes place may be related to engine output. Thus un light throttle conditions, gear change takes place at 2000 rev/min.

The behaviour of gasolene engines at full and partial throttle openings is shown Figs. 12.8 to 12.10 which are based on tests by Davies.[12]

For a given road speed, part load power can be supplied by any combination of engi speed (gear ratio) and engine torque (throttle opening). Some other factor must considered in order to select the strategy for gear changes. This is provided by conside tion of fuel economy as indicated by the curves for specific fuel consumption

$$\left(\frac{\text{Fuel consumed}}{\text{work done}}\right).$$

The selection of the optimum engine speed for each value of power output and ro speed requires an infinitely variable or "stepless" gear of the type used on the "Daf" c In this design a belt is arranged to transmit power between two pulleys consisting double cones. By adjusting the axial spacing of these cones, their effective diameter m be changed and thus the gear ratio.

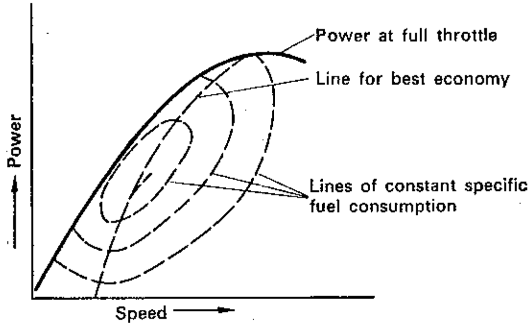


FIG. 12.8. Power characteristic of automobile engine.

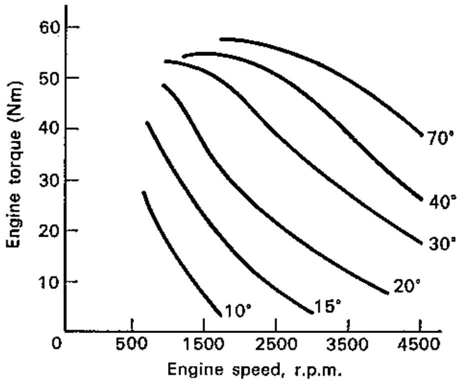


FIG. 12.9. Torque–speed curves for gasoline engine for various throttle positions.

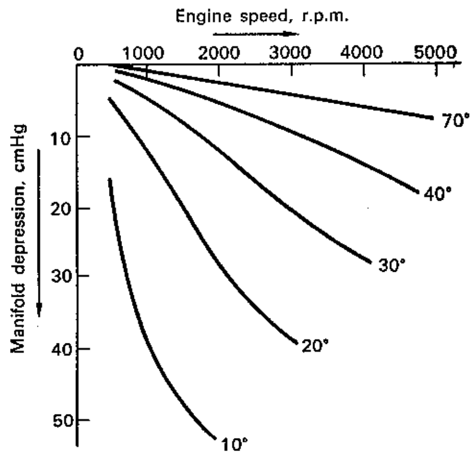


FIG. 12.10. Manifold depression for various speeds and throttle positions.

As demonstrated in Fig. 12.10, manifold depression is related to engine speed ar throttle position. It can be employed as a measure of engine load. In the Daf systeı shown diagrammatically in Fig. 12.11, gear ratio selection is effected by a combinatic of engine speed and manifold depression.

Possible modes of control for stepless transmissions and proposals for regenerati systems are discussed by Macmillan.[13]

12.8. Control of electric motive power

The renewal of importance of electric trains for urban, suburban and intercity tran port, the predominant use of electric transmission for diesel locomotives, and possib developments such as the electric car, all justify serious attention to the associated contı problems.

In contrast to the practice in most other fields where electric power is used, the mot most commonly used in traction is the series motor. When supplied at constant volta this has the characteristic that as speed increases, torque falls off. In fact the falling off torque is so marked as to bring about a major reduction of power below that correspon ing to the constant power hyperbola (Fig. 12.12). Correspondingly, on a train encount ing a gradient, for example, the effective resistance exceeds the torque supplied by t motor so that the train loses speed and in consequence the motor torque increases unti new equilibrium at a lower speed but higher power results. The series motor thus has inbuilt or inherent control feature which leads to stable operation. A note of warniı however, must be sounded here. If load is removed entirely from the motor with volta

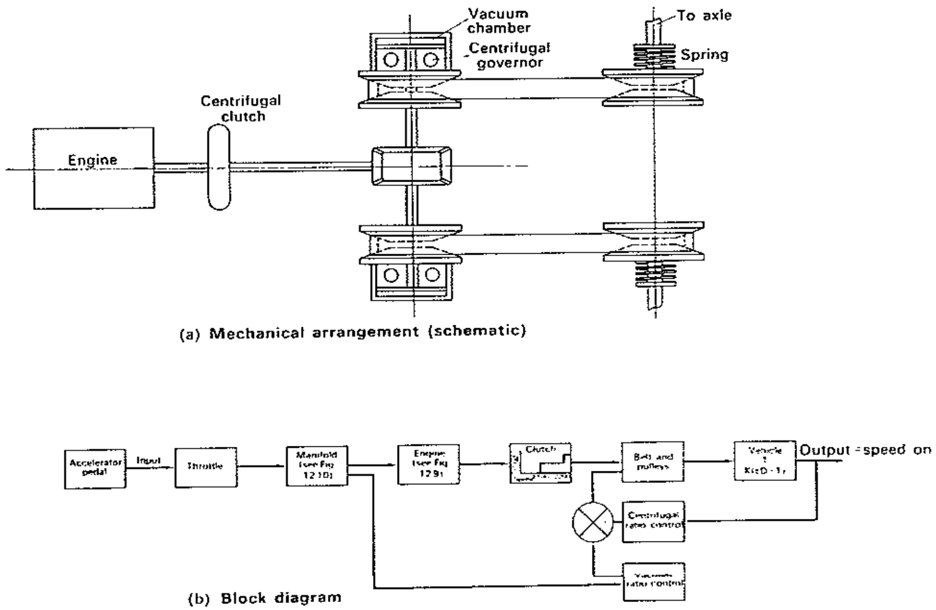


FIG. 12.11. Automatic control of Daf car.

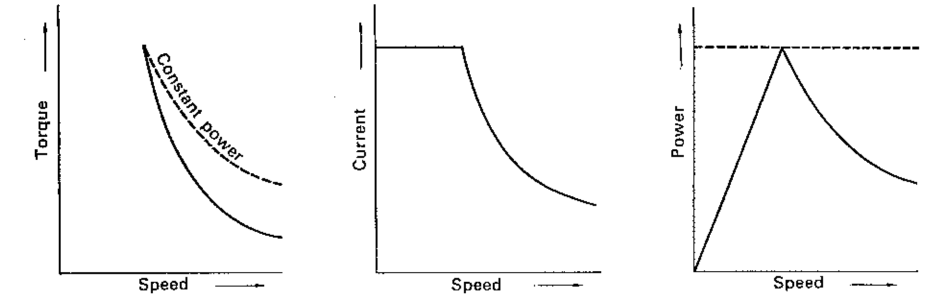


FIG. 12.12. Series motor characteristics.

remaining constant, acceleration will continue until mechanical failure intervenes. Series motors are, therefore, virtually confined to traction (railway, trolley-bus and crane) applications and even there, safeguards must be provided against the danger of over-speeding under conditions of low adhesion.

The fairly steep falling off in current with increasing speed facilitates the control of series motors by "notching". Each value of applied voltage produces a specific speed–voltage curve. Where the power supply is a.c. the different voltage steps may be obtained by selection of different tapping points on the transformer. Figure 12.13 shows a system whereby forty speed torque loci may be selected. Another form of control used to counteract the "throttling" effect of the series characteristic at high speed consists of "field

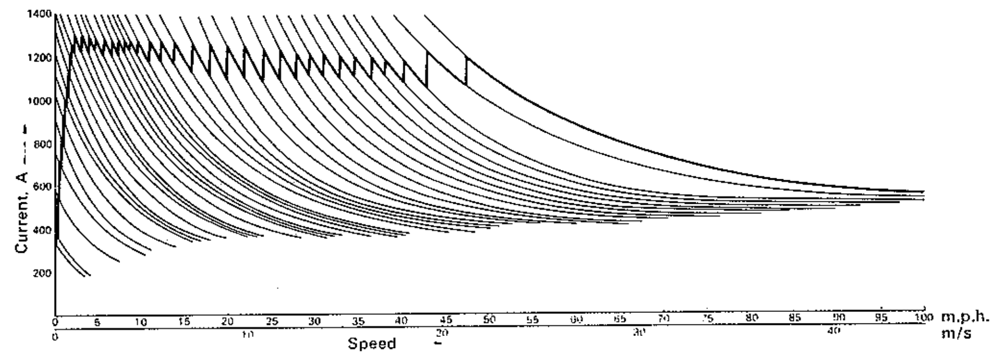


FIG. 12.13. Control of a.c. locomotive.

weakening". Part of the field coil is shunted out so as to reduce the flux for a given current. Thus the "back e.m.f." of the motor is reduced so that more current flows for a given applied voltage, giving rise to a different torque speed locus.

Manual control of locomotives usually consists of providing the driver with a controller enabling him to select any notch and an ammeter on which is shown the current taken by each motor and which bears markings indicating the maximum permissible current. He advances his controller to a notch which allows the desired current to flow. As the train

gathers speed, the current will fall off. The driver will then advance the controller notc by notch in an attempt to obtain maximum permissible amperes continuously. The torqu will be proportional to the amperes and the driver must ensure that the amount permitte by adhesion will not be exceeded.

Automation of the driver's action during the acceleration period is readily achieved b the introduction of a current relay which, on current falling to a predetermined value, cut out resistance on a d.c. system or varies the tap changing on an a.c. system.[14] Th maximum value of tractive effort must be set below the adhesion limit or automati wheel-slip protection (see Section 12.11) must be installed.

12.9. Use of transductors in power control

A transductor resembles a transformer but performs a control function by virtue of th pronounced saturation characteristic of its cores which are made of material having th square-loop characteristic of Fig. 12.14. When unsaturated, the impedance is very hig but very low when saturated by the passing of direct current through control winding embodied in the device. Therefore a.c. voltage and current can be controlled by variatio in the d.c. current supplied to the transductor.

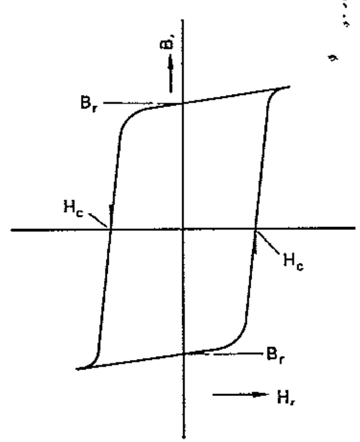


FIG. 12.14. "Square loop" magnetic material characteristics.

A transductor could be used to control the voltage applied to a locomotive motc acting over the full range of voltage required. Because the bulk and weight of such device would be at least one and a half times that of a conventional transformer. a alternative solution has been adopted.[15] Figure 12.15 shows a form of control for an a. locomotive wherein transductors are employed to give a smooth change of volta between notches and to ensure that the tapping switches are only operated when there is n current flowing through them.

A transformer is provided with a number of taps so as to break the total voltage ran into stages which are each within the capacity of a transductor of reasonable bulk ar weight. Four transductors are incorporated in two of the arms of a conventional rectifier bridge. Silicon diodes are used in this case. The transductors are so arranged that one pair can be connected in sequences to switches S_1, S_3, S_n and the other to S_2 and S_4 to $S(n+1)$ where n is any odd number.

Consider the situation at the commencement of a run. Contact S_1 is closed so that a definite voltage is applied to transductors T_1 and T_2. During one-half cycle, current will pass through T_1 through the traction motor and back to the transformer via rectifier R_1. During the next half cycle it will pass through R_2 then through the traction motor and then through transductor T_3. Thus an alternating current applied to the bridge circuit becomes a unidirectional one as applied to the traction motors.

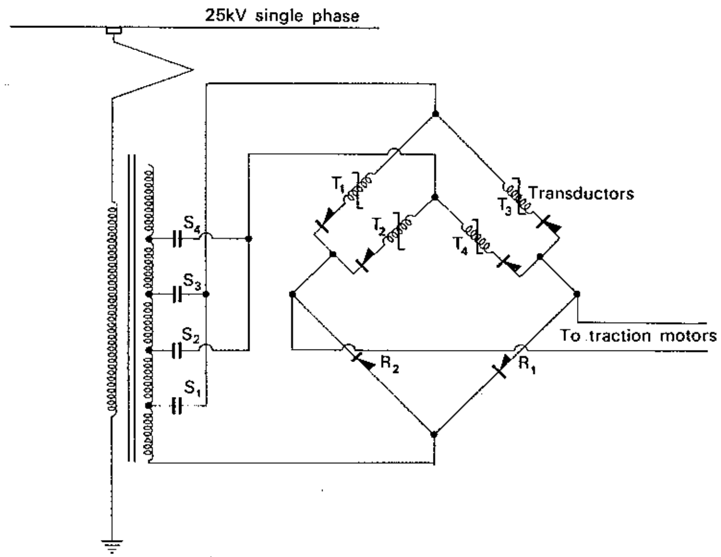


FIG. 12.15. Transductor control for electric locomotive.

Initially the control current (d.c.) applied to T_1 and T_3 will be low so that they are in the high impedance unsaturated condition. As the motor accelerates the impedance of the transductors will be reduced in consonance with increased back e.m.f. of the motors so that current and therefore tractive effort remain constant. When transductor impedance has been reduced to its minimum, that is, when its cores are saturated, the voltage will have risen to that represented by tapping S_1. We are of course ignoring the resistive volt drop of the components and the cable connecting them. Transductors T_2 and T_4 are now brought into consideration. If these are in the high impedance state, contactor S_2 can be closed without affecting the flow of current in the circuit.

Now the impedance of T_2 and T_4 is gradually diminished so that current begins to flow through S_2. When this has been established, flow through S_1 can be blocked by raising the impedance of T_1 and T_3. Thus contactor S_1 may be opened without interrupting current. Further reduction in impedance of T_2 and T_4 will enable the acceleration of the motors to be continued. Contactor S_3 may now be closed in readiness for the next voltage range which will be traversed under the control of transductors T_1 and T_3 as before.

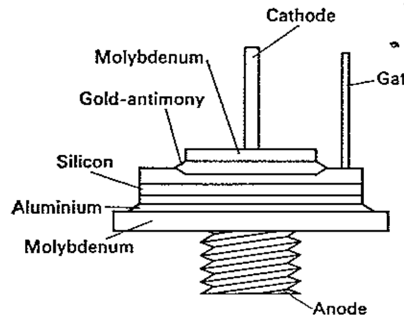
In addition to the removal of arc rupturing duty from the tap changing gear wi consequential savings in first cost and maintenance the system enables relatively wea currents applied to the transductors to exert a continuously variable control over tl main power currents. It is therefore possible to improve the system of control of tl locomotive thereby rendering it more adaptable to automation. Thus a driver or automat controller can determine the tractive effort required by setting a current relay whi adjusts transductor-control current and tap-changer position to maintain consta conditions. The elimination of the stepped characteristic of Fig. 12.13 itself enables bett use to be made of adhesion. The transductor control can be arranged so that, should o axle slip, voltage is held corresponding to the speed of the non-slipping wheels and, course, a slip detection circuit can be arranged to reduce tractive effort in the event serious slipping occurring.

The locomotive illustrated in the Frontispiece is controlled by this means.

12.10. Thyristor control of a.c. locomotives

The advent of semiconductor switching devices, notably the thyristor, has enabled t control of traction motors, both in motoring and braking, to be revolutionised.

Figure 12.16(a) and (b) show the basic construction and control features of a thyrist



(a) Construction of silicon-controlled rectifiers.

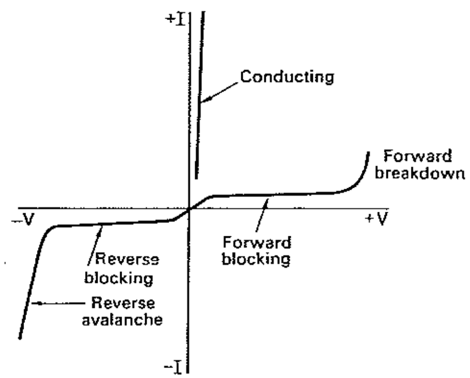


FIG. 12.16.(b) Construction and characteristics of silicon-controlled rectifiers (thyristors).

Its characteristics are that, under the voltages normally applied across the main terminals, it is in the "blocking" or non-conducting state. The application of voltage to the gate, however, causes the device to conduct in the forward direction and it will continue to do so as long as there exists an e.m.f. between the cathode and anode notwithstanding the removal of voltage from the gate.

In the case of a.c. locomotives, the tap-changing equipment of the transformer can be eliminated by using thyristors and the voltage applied to the motors may be regulated by modifying the firing angle of the thyristor. Thus the voltage received from the catenerary rises from zero at the beginning of the cycle to a maximum half-way through the first half-cycle (that is after 0·005 s when on 50 Hz supply) and falls to zero each half-cycle. If therefore the instant of initiation of conduction is delayed during the 0·01 s half-cycle, power is reduced. The power circuit of a locomotive comprises separately-fed traction motor armatures and separately-fed individually controlled field windings. Each armature circuit is fed from two rectifier bridges in series, each bridge being composed of two arms with several thyristors in parallel and two arms containing diodes. The principle of operation is that the voltage applied to each armature depends on the firing angle of the thyristors. The resulting current is measured and its value used to control the field current which is supplied through similar bridges. (The main field windings are arranged to carry about one-fifth the current at five times the voltage of those in a conventional series traction motor.) Up to half voltage one bridge per motor is used, the firing of the thyristors being gradually advanced until free firing (full-wave) conditions are attained. The fully retarded thyristors in the second bridge are then gradually advanced to add an increasing voltage from a second transformer secondary winding until both bridges are fully advanced and full motor voltage is attained. Further increase in speed is obtained by retarding the firing of the thyristors controlling the field (field-weakening).

Thyristor control is therefore notchless during both driving and braking, which greatly enhances the performance of a locomotive from the point of view of adhesion. The "shunt" characteristic arising from separate excitation provides a falling torque with increasing speed which quickly restores adhesive working. Where slip is persistent the field strength of the offending motor or motors is held constant and the armature current is reduced until wheelslip is arrested.

Overhead and third rail power systems always pose the threat of interference with adjacent telecommunications circuits and measures of suppression and immunisation are necessary.[16] For example, it has been suggested that a capacitor bank of about 1 MVAR on the train should produce a significant decrease in line harmonic content.[17] However, many of the initial fears and doubts about the effects of thyristor-controlled traction prior to its introduction into the 25-kV electrification networks of British Railway were subsequently found to be unduly pessimistic.[18]

Comparative tests carried out by O.R.E. locomotives fitted with conventional tap changes or thyristor control showed an 8% advantage in adhesion performance attributable to thyristor control.

12.11. Chopper control

A feature of the thyristor is that there is no way of switching it off once current has started to flow until and unless current ceases to flow due to some external factor. When

applied to a.c. systems of course current falls to zero at each half-cycle and is not restor unless and until the gate is re-energised. The thyristor is therefore well adapted f application to the control of a.c. systems but is not inherently capable of interrupting controlling a d.c. circuit. Nevertheless the inefficiency associated with rheostatic conti of rolling stock has led many administrations to adopt electronic control to the d system using choppers (hacheurs). There are several alternative circuits for achieving "switching off" but most chopper equipments which have been used in practice a based on two-thyristor resonant pulse circuits with a chopping frequency of arou 80 Hz although some systems range between 10 and 75 Hz.

In rapid-transit service energy savings of about 37% can be attributed to the adopti of chopper control. Of this, 23% is derived from regenerative braking, 9% from impro ment of control efficiency and 5% from decrease in weight.

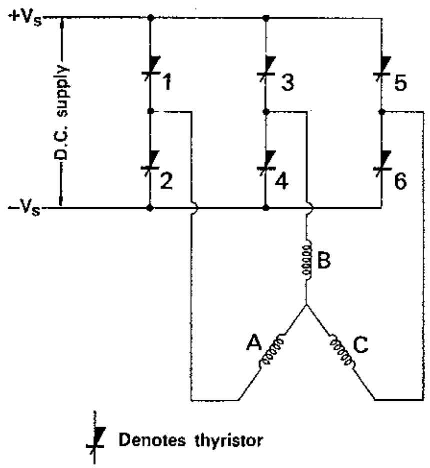
A series of trials on the adhesion performance of two three-car 3-kV multiple u trains, one fitted with conventional rheostatic control and the other with choppers, w carried out by the O.R.E. in order to evaluate adhesion performance. The basis comparison was the minimum distance required to accelerate from rest to maxim speed making full use of the available adhesion. Dependent on ambient conditions, t chopper control produced a reduction in starting distance from 5 to 23%.

12.12. Application of induction motors

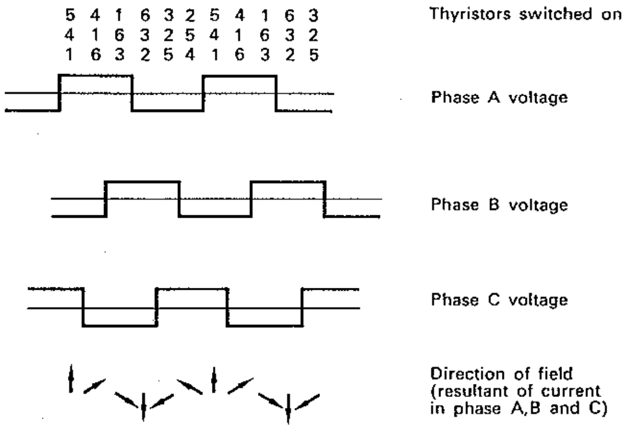
The series motor suffers from the fact that it comprises a commutator having brushes wear out and which introduces the risk of flashover. The induction motor, particularly the squirrel cage form, is of simple construction and involves no rubbing contacts. requires a polyphase a.c. supply, however, and losses are high if speed of rotation diff markedly from the synchronous speed characteristic of the supply. This difference known as "slip".

The advent of controlled semiconductors offers the possibility of varying the frequen of supply so that torque control during motoring and braking can be achieved by conti of slip frequency.[19] Thus the supply frequency would be adjusted sufficiently closely the speed of rotation of the motor by switching "on" or "off" of controlled silic rectifiers. Attempts at obtaining similar control in the past by using rotating machines transform the frequency of current supplied to the motors were technically successful b open to the objection that the total capacity of the machines installed was from two three times the horsepower actually developed. The losses associated with each trar formation of power were also considerable when taken in the aggregate. The silic control rectifiers (thyristors) are compact and their price is continually being reduce They are essentially switches rather than power converters so that their efficiency is hig A typical value of forward volt drop is 0·6 volt.

Whatever the form of supply it can easily be converted to d.c. and used to supply t circuit shown in Fig. 12.16. If the thyristors can be switched on for intervals as shov in Fig. 12.17 a square wave voltage will be applied to the stator phase windings. Becau these will be inductive, the current wave will be more nearly sinusoidal. However, tl problem of turning off the controlled rectifiers remains and commutating capacito may be used for this purpose. Various arrangements are possible and Fig. 12.18 shows method whereby commutating pulses are supplied to the controlled rectifier from tl



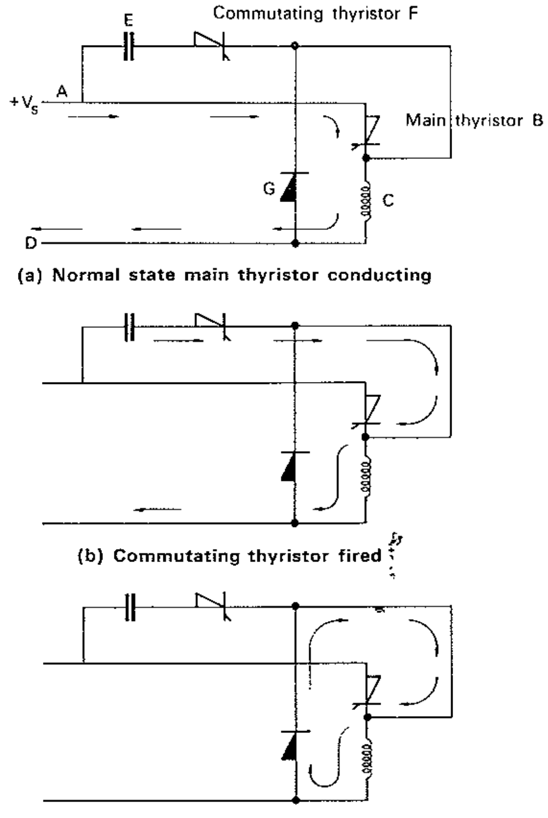
(a) Arrangement of thyristors to provide three-phase power.



(b) Switching sequence of thyristors.

FIG. 12.17. Generation of three-phase power from D.C. supply.

a.c. side. During the power part of the cycle, current passes from A through the control thyristor B, inductance C to neutral terminal D. A negative signal must, of course, be supplied to the gate of B for this to happen. Capacitor E is charged so that the potential is $V_s - V_r$ where V_r is the reverse voltage across B. When commutating thyristor F is fired, this potential is applied to the cathode of controlled rectifier B, thus subjecting this component to a net reverse voltage of V_r. Due to the inductance C, however, the current therein will continue but will be drawn instantaneously from the condenser E. B itself



(a) Normal state main thyristor conducting

(b) Commutating thyristor fired

(c) Commutating current decaying through diode

FIG. 12.18. "Switching-off" circuit.

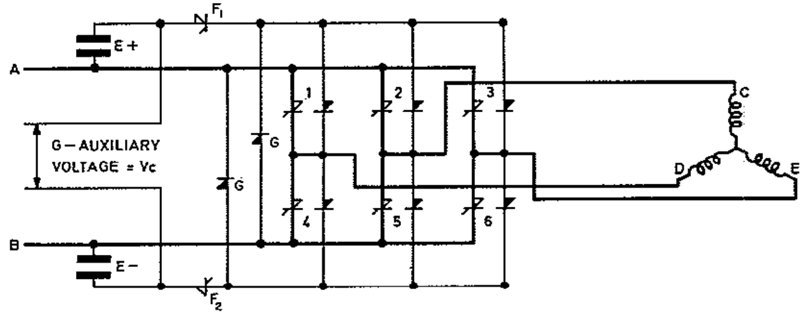


FIG. 12.19. Circuit used on experimental locomotive.

having negligible inductance, current will cease to flow as the negative voltage is applied and if this is maintained for sufficiently long, the P.N. junction will revert to the non-conducting state. Condenser E will continue to discharge through C until it becomes negative. Control rectifier F will then turn off. The current in C will then "free wheel" through simple rectifier G until it decays to nothing. In practice, of course, inductance C would be embodied in the load.

The frequency of supply will therefore be determined by the timing of the signals applied to the control rectifiers. For a three-phase supply, six circuits such as shown in Fig. 12.18 are combined together as shown in Fig. 12.19. The torque characteristics of an induction motor are shown diagrammatically in Fig. 12.20. It will be noted that the "no-load" condition corresponds to a rotational speed equal to the speed of rotation of the magnetic field. Thus a three-phase machine having six poles (two for each phase)

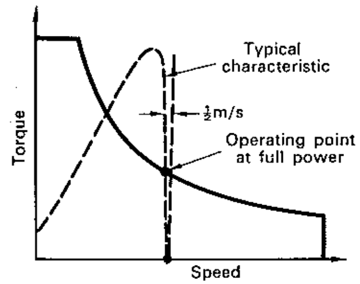


Fig. 12.20. Torque–speed characteristics of induction motor.

would rotate at 3000 rev/min if fed with current at a frequency of 50 Hz. As load is applied the motor slows down and the difference between its speed of rotation and that of the field is known as the slip frequency. The slope of the curve as it intercepts the zero load line is usually quite steep so that if load is suddenly thrown off, the motor speeds up only slightly. This is a particularly useful feature for traction because, under bad adhesion conditions, slip (in the mechanical sense) is limited quantitatively to the pre-existing value of the electrical slip.

The torque to be exerted by a motor at any time can be controlled by measuring the actual speed of rotation of the motor spindle and then supplying control pulses to the inverter at a slightly greater frequency. The difference in frequency would be equated to the slip frequency corresponding to the desired torque and a nominal slip (electrical) could be set equal to mechanical creep corresponding to maximum adhesion (i.e. by reference to some independent means for determining vehicle speed). Low adhesion conditions would then be automatically catered for by reduction in torque corresponding to actual slip.

The above is an example where considerable power may be controlled by comparatively weak signals and is thus a means for improving the characteristics of a propulsion system. Figure 12.21 shows an experimental diesel-electric locomotive using squirrel-cage motors.

Figure 12.22 shows a shunting locomotive fitted with an inverter and induction motors in use on the Swiss Federal Railways.

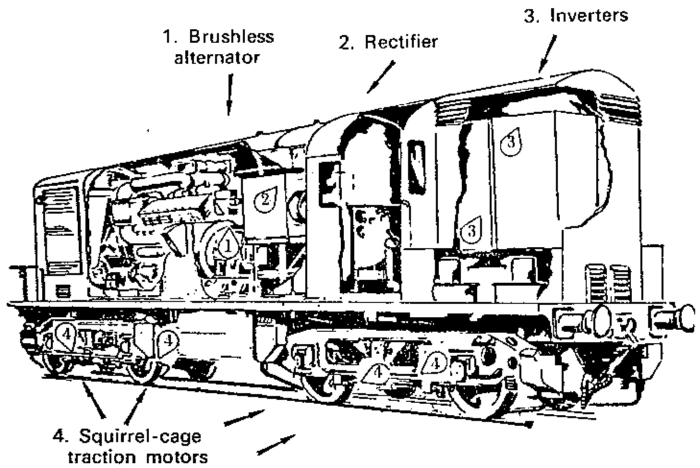


Fig. 12.21. The experimental Hawk locomotive.

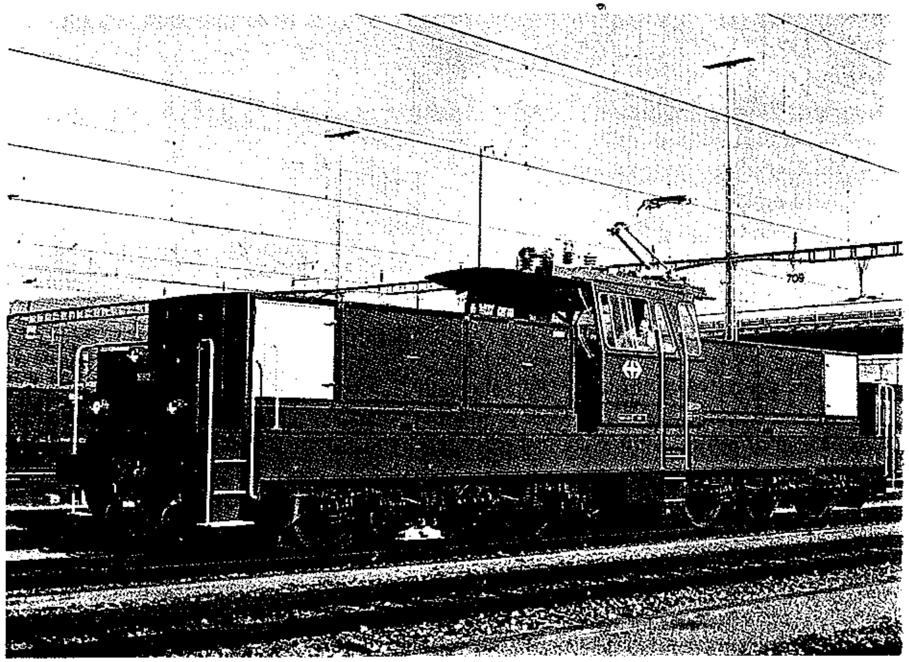


Fig. 12.22. Shunting locomotive fitted with inverter and induction motors. Swiss Federal Railways.

Transitors—pulse width modulation

The main problem to be overcome in the use of thyristors fed from a d.c. source is the provision of some form of reverse e.m.f. to achieve the switching-off event. Transistors possess the advantage that the current flowing is always controlled by the gate current but, until recently, have not been available with sufficient power for traction applications. However the I.C.T.S. system demonstrated at Kingston, Ontario (see 16.6), has operated with inverters of 650 kVA fitted with power transistors. The circuit is shown in Fig. 12.23(a). To achieve the required inverter rating, the transistors are arranged in series and

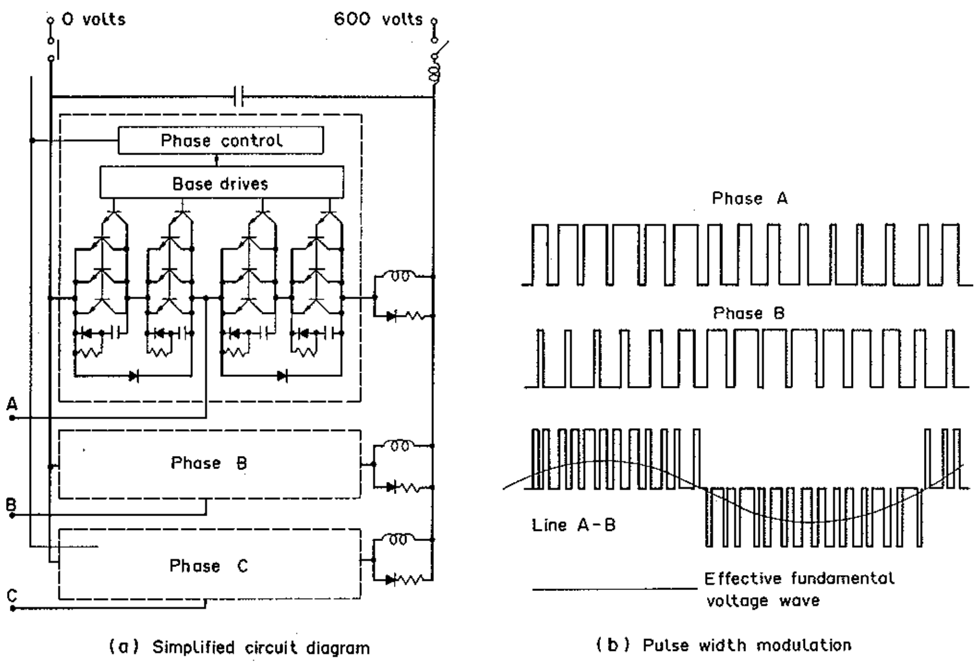


FIG. 12.23. Transistorised invertor.

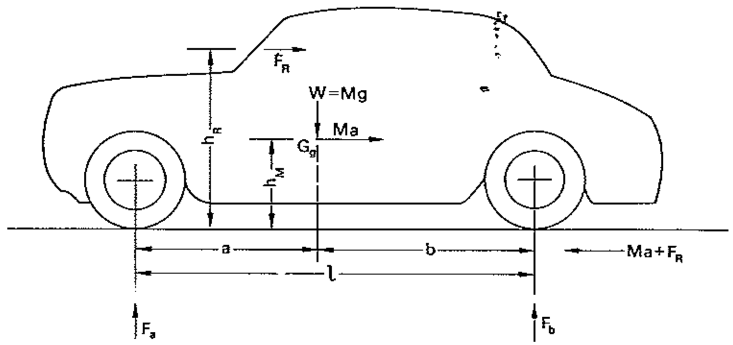
parallel connections. Basically, each of three separate circuits feeds one of the phases of the three-phase motor and embodies two switching elements which in turn connect the phase output to the positive supply rail and the negative supply rail.

Because of the ease with which transistors can be switched off, the voltage wave applied to the motor may be shaped using pulse-width modulation. This is done by selecting a

frequency for switching-on which is much higher than the motor's synchronous
quency for switching-on which is much higher than the motor's synchronous freque
Typically when the effective output is at a frequency of 10 Hz, that of switching-on transistors if 400 Hz. Figure 12.23(b) shows how the waveform is synthesised by var the duration of each pulse.

12.13. Wheel-slip control—torque reaction

The value of the coefficient of friction between rail and wheel varies widely from t to time and from place to place so that a driver or automatic-acceleration-control sys cannot anticipate the adhesion limit in advance. Usually, however, one motor will before another if only because of the effects of torque reaction.[20] Drawbar pull act along a line at some height above the rail surface and even in the case of a si vehicle, the inertia forces will act through the centre of gravity (Fig. 12.24). Thus the



$$F_a = \frac{(Wb - Mah_M - F_Rh_R)}{l}$$

$$F_b = \frac{(Wa + Mah_M + F_Rh_R)}{l}$$

FIG. 12.24. (a) Calculation of torque reaction (single vehicle).

a moment applied to the vehicle body which is counterbalanced by a redistributio wheel loads so that they will not all tend to slip simultaneously.

Consider the forces acting on the locomotive body, as shown in Fig. 12.24(b). tractive effort F_c acting at a height h_c above the rail will be balanced by the sum of tangential forces acting at rail level to give rise to a couple $F_c h_c$. This will result in leading bogie being unloaded by an amount given by $F_c h_c / l_1$, this load being transfe to the trailing bogie. In addition to this effect torque reaction may occur on the indivi bogies due to the reaction of the horizontal load on the bogie pivot, as denote equation (12.27):

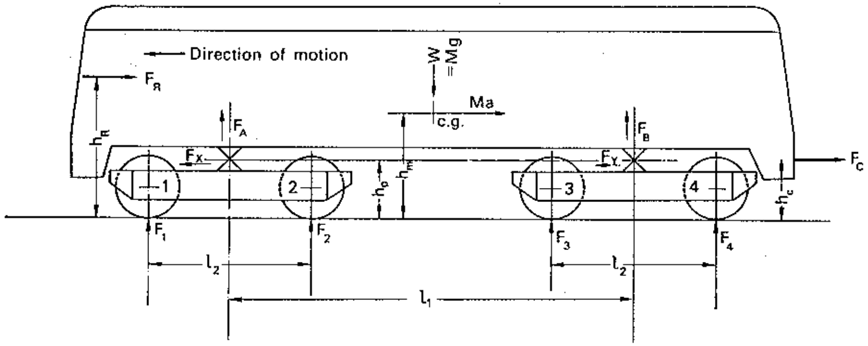


FIG. 12.24. (b) (Locomotive). Torque reaction.

$$\left.\begin{aligned}\Delta F_1 &= -\frac{F_c}{2l_1}\left[h_c+h_p\left\{\frac{l_1}{l_2}-1\right\}\right]\\ \Delta F_2 &= -\frac{F_c}{2l_1}\left[h_c-h_p\left\{\frac{l_1}{l_2}-1\right\}\right]\\ \Delta F_3 &= +\frac{F_c}{2l_1}\left[h_c-h_p\left\{\frac{l_1}{l_2}-1\right\}\right]\\ \Delta F_4 &= +\frac{F_c}{2l_1}\left[h_c+h_p\left\{\frac{l_1}{l_2}-1\right\}\right]\end{aligned}\right\} \qquad (12.27)$$

In a conventional design torque reaction can lead to a loss of adhesion of about 7%. Whilst the redistribution of load between bogies cannot be avoided, the axle loads within bogies can be equalised by transmitting traction and braking forces to the body through low-level traction bars, as sketched in Fig. 12.25. Alternatively, a compensating

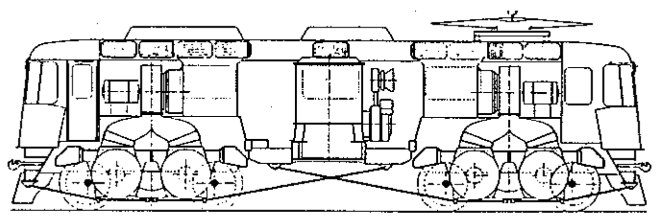


FIG. 12.25. Arrangement of low level traction bars.

moment can be applied to the bogie by a pneumatic cylinder which is pressurised at any instant in accordance with the value of F_c.[21] Some experimental results from a locomotive which was so equipped are presented in Fig. 12.27.[15]

When an axle commences to slip the motor will accelerate and the current will fall (When a driver notices this he will reduce the voltage applied to all the motors ui equilibrium is restored. The relationship between coefficient of friction and relat sliding usually takes the form shown in A of Fig. 12.26 in which an initially linear porti

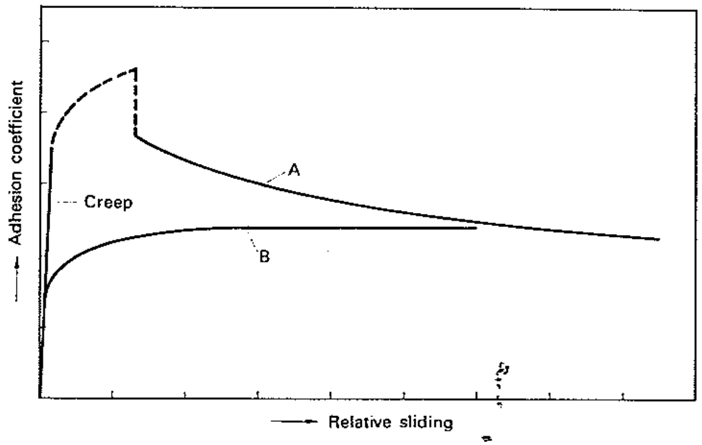


FIG. 12.26. Variation of adhesion with relative sliding.

curves over gradually until a maximum is reached after which there is a catastropl falling off. This is followed by a region where increasing speed of sliding is accompani by a falling friction coefficient. This behaviour is indicative of an unstable conditi which is inherently difficult to control. The breakaway point generally occurs at about 2 relative sliding. Under bad conditions of adhesion, the peak may not occur, the cur taking the form shown in B of Fig. 12.26. Whilst the tractive effort which can be exert in these circumstances must be severely limited, the situation is a stable one and operati at reduced tractive effort is feasible. In practice, adhesion coefficients vary widely ai results must be presented on a statistical basis as shown in Fig. 12.27.

It is obviously necessary to exercise such control as to prevent operation in the unstab region. This represents an aspect of the engine driver's duties calling for the highe amount of skill.

In the case of multiple-unit equipment which has a good proportion of its axl motored, the demand made on adhesion value to secure the desired amount of acceler tion is not great and it is usually sufficient to set the current control relay at such a val as to avoid slipping. Thus the automatic acceleration system acts also as an automat slip-protection device.

With locomotives, however, it is necessary to utilise the potential adhesion fully an were the above practice followed, the desired loads could not be hauled. Manual contr of acceleration is therefore usual in these circumstances. Even so, the demand exists f some aid to inform the driver of the imminence of slipping and of course this is essenti on automatic railways.

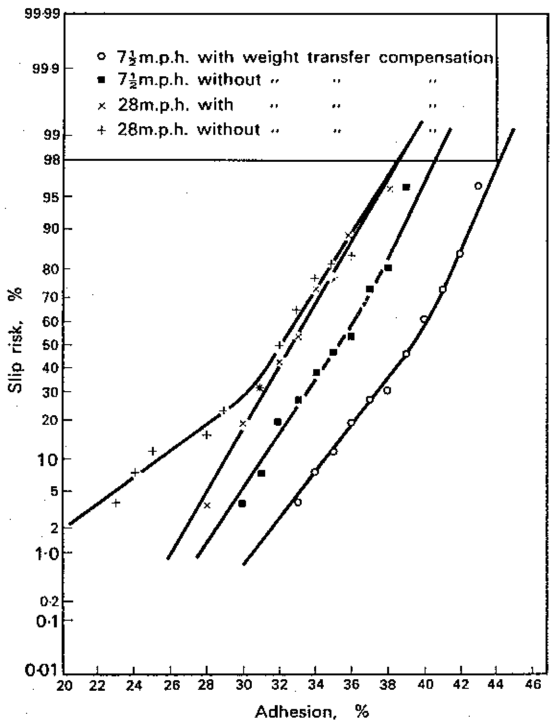


FIG. 12.27. Statistical presentation of adhesion values using slip risk curves. Variation of adhesion between wheel and rail.

Measured adhesion characteristics—O.R.E. tests

A series of observations of the incidence of slipping of steam locomotives on a gradient of 1 in 60 showed a strong correlation with relative humidity measured at rail level. Discriminant analysis of a number of variables indicated that water when present in liquid form was not conducive to slipping. A greater proportion of incidents occurred during rainy weather. Thus out of 4000 recordings it was established that 8% of all trains slipped in normal weather conditions, 5% of them severely, whereas 26% of all trains slipped under wet conditions, 16% severely. More detailed analysis showed that heavy rain (greater than 0·05 in.) had no significant effect on slipping and that the highest proportion of incidents occurred during conditions of slight rainfall or when there was a mere trace of rain (less than 0·005 in.).[22]

The value of the coefficient of adhesion is generally derived from the drawbar pull of a locomotive as recorded by a dynamometer car. Relative sliding is usually recorded simultaneously because the value of slip at maximum tractive effort has important implications for the design of control systems.

There are three basic procedures for carrying out adhesion tests, as follows: (a) constant speed; (b) accelerating; and (c) decelerating. From the experimental point of view

testing at constant speed is preferable because the instantaneous value recorded at drawbar represents the true tractive effort and is unaffected by the inertia forces wh would be introduced by any change in velocity of the locomotive. To explore the range of adhesion coefficient it is necessary to increase the value of the tractive eff until slipping occurs. This necessitates some form of loading device which can maint a constant speed whilst drawbar pull is increased. The mobile test units of British F embodied a feedback system. When these were used to test a locomotive fitted w notchless control high values of adhesion (43·3%) were maintained in a consist manner.[15]

When, as was usual until very recently, the control of a locomotive is achieved notching, the constant-speed test is handicapped by the stepwise nature of the locomot control. Thus, starting from a point where the tractive effort derived from the motor in equilibrium with the applied drawbar load, it is necessary to advance the locomot controller which results in a discrete increase in voltage applied to the motor an sudden change from one torque-speed curve to another: see Fig. 12.28. This will unlikely to match the coefficient of adhesion exactly and if it falls short of the point sliding the procedure must be repeated, whereas if it exceeds the value of torque requi to initiate sliding, unstable operation will result and the actual limiting value will not

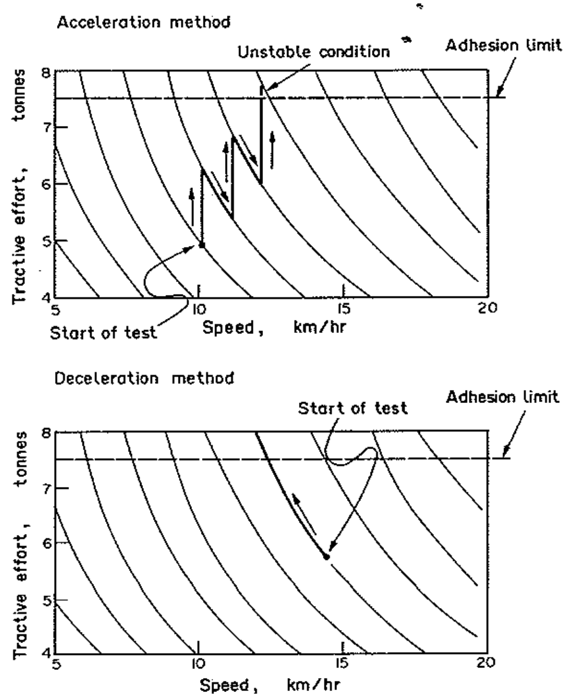


FIG. 12.28. Methods of adhesion testing.

recorded. Thus such a test can only determine the value of tractive-effort attainable from the notch immediately below the slipping point rather than the slipping point itself. Accelerating tests where train speed is not controlled and which most closely reproduce operating conditions behave similarly. On the other hand, if the controls are set just below the point at which sliding is expected to take place and the load is then gradually increased until this occurs, the true value of the coefficient of adhesion will be determined.[23]

Tests in the Jura, 1962

In a series of tests carried out between Pontarlier and Frasne[24] the deceleration method gave results which were usually from 5 to 8% higher than the acceleration method. These tests were designed to evaluate the effect of a number of design features on the attainment at maximum adhesion coefficient. The factors evaluated in one set of tests were: (1) speed; (2) type of current, a.c. or d.c.; (3) type of motor, i.e. direct or rectifier fed; (4) elasticity of transmission between motor and axle. One obstacle to the use of service locomotives for adhesion research is that they are usually conservatively designed so as to utilise comparatively low values of adhesion (about 24%) and are unable therefore to explore the full range of adhesion which is available for a good deal of the time. In the tests at Pontarlier two six-coupled shunting engines had their connecting rods removed so that the full power of the driving motor could be applied to a single axle. These were placed at the head of a train (their position was interchanged at intervals during the trials) which consisted of a dynamometer car, an electronics car and a freight locomotive which was equipped with a continuously variable control which could operate in the braking mode.

The results indicated that there was no difference in adhesion for the two speeds employed in the tests, namely 10 and 20 km/hr. The railway line on which the tests were carried out ran from France to Switzerland. For some tests it was fed from France at 50 Hz, 25 kV and for others from Switzerland at $16\frac{2}{3}$ Hz, 15 kV. The type of current fed to the motors was shown to have no effect on adhesion performance. The locomotive fitted with rectifier and d.c. motors gave better performance (14% higher adhesion) than the one fed with a.c. directly and this was attributed to the fact that, on any given notch, the variation of current with speed was less steep. The elasticity of the rotary drive mechanism did not affect the adhesion.

The value of relative sliding was measured by fitting "magslip" units to the driving axle and an idle axle and recording the difference in their output. Relative sliding usually attained a value of about 2·5% before adhesion began to fall off.

Other tests were carried out on main line locomotives of the BB type which normally had two axles in each bogie coupled to one another through gears. Comparative trials with or without coupling did not produce consistent evidence of the advantage of this practice.

During a period of 2 weeks testing values of coefficient of adhesion varying from 18 to 59% were recorded. A strong inverse correlation with relative humidity was observed.

The special test machine

Although a number of important conclusions could be drawn from tests on existi locomotives it was necessary to explore the frictional behaviour at high speeds involvi power in excess of any then existing locomotive. It was also necessary to isolate t phenomena occurring at the wheel–rail interface from the influence of the design featu of the locomotive as a whole. Accordingly a special test machine was adapted from t gas turbine locomotive illustrated in Fig. 12.29.[30] The original power unit was remov

FIG. 12.29. Gas turbine locomotive adapted for adhesion testing.

and, during tests, another locomotive was used to collect and moderate the power whi was supplied to motor or motors which were geared to a single axle situated at the cent of one of the bogies. In one set of tests the motor used was of a type usually used by t S.N.C.F. to drive the two axles of a "monomotor" bogie and in another, two 1200 h. motors of a type normally used to drive individual axles of express locomotives we geared to the single axle. The load applied to the test axle could be varied from 13 t 22 t using an oleo-pneumatic apparatus which had virtually infinite compliance so as minimise the effect of track imperfections on instantaneous axle load. Interchangeal test axles fitted with wheels of 0·92 m and 1·4 m diameter respectively could be fitted evaluate the effect of wheel diameter. Figure 12.30 is a cross-section of the test bo

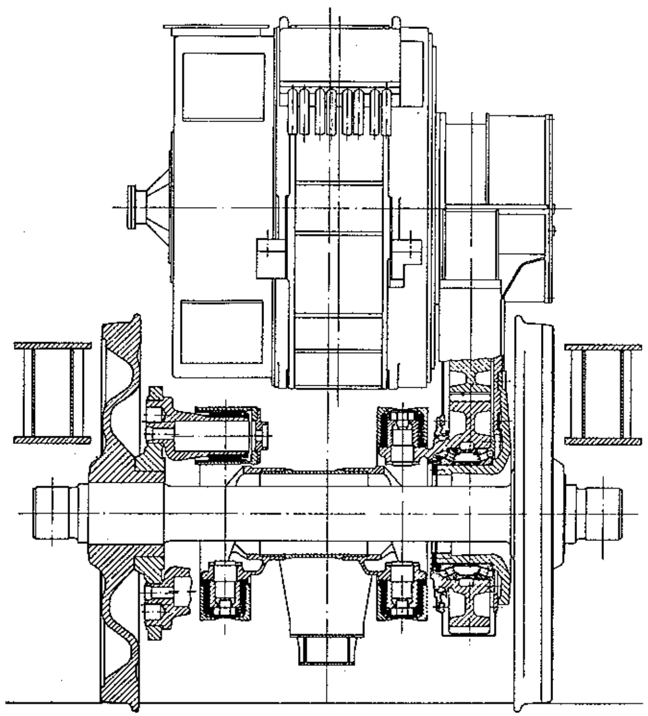


FIG. 12.30. Cross section of drive system of locomotive adapted for adhesion testing.

when adapted to a single motor driving 1·4 m driving wheels through a "Jacquemin" cardon shaft drive. Strain gauges fixed to the wheels enabled normal and tangential forces to be recorded continuously and confirmatory measurements of tangential force could be made by axle-boxes fitted with strain gauges.

This test machine has been used in a number of trials in Europe, and the conclusions of the trials at Pontarlier have been confirmed over a wide range of speed and wheel load.[31, 32, 34, 35] It has been confirmed that speed has little effect over the range 20 to 120 km/hr and that relative humidity reduces friction coefficient particularly when this exceeds 80%. The maximum value of friction occurs between 1% and 8% relative sliding and there is usually a wide "plateau" on the curve before a catastrophic reduction in friction occurs. Figure 12.31 presents a typical test result.

Other important results were that Amonton's Law is not obeyed over the practical range of load between wheel and rail (120 kN to 180 kN), the friction coefficient being lower for the higher load. The slip value at maximum friction was also higher for the lower load. Wheel diameter had no noticeable effect on friction coefficient. A relative speed of from 0·5 to 2 km/hr appears to be associated with maximum adhesion and may be related to the heating of the rail by friction.

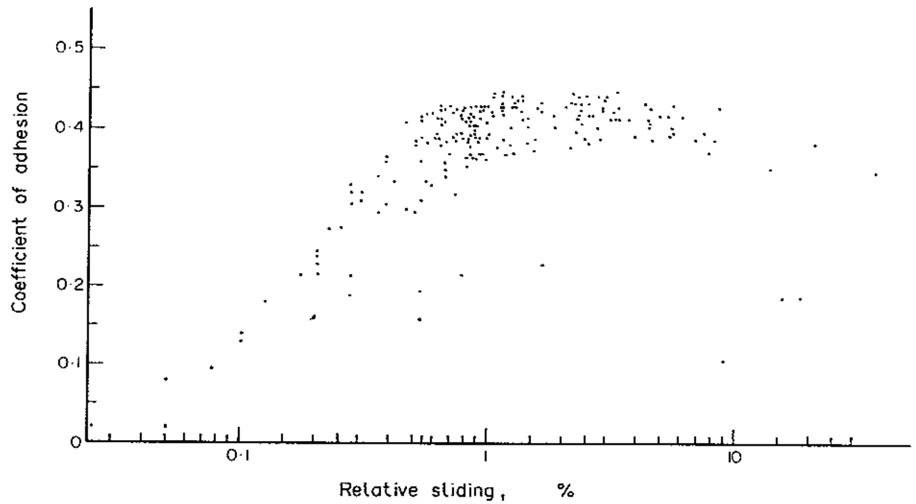


FIG. 12.31. Experimental measurements of traction and relative sliding.

Although the value of the stiffness of the transmission train between motor and whe rim did not affect the magnitude of the friction force it had a marked effect on the cidence of relaxation oscillations.[33]

It was noted that the tangential forces exerted by the two wheels on a single test a were seldom identical and that lower friction was attainable on sharper curves accordance with equation (12.28):[32]

$$\mu = \mu_0\left[1 - \frac{12{,}100}{R^2}\right] \qquad (12.$$

where R is radius of track in metres
and μ_0 = coefficient of adhesion measured on straight track.

Remedial measures

These are discussed in reference 23.

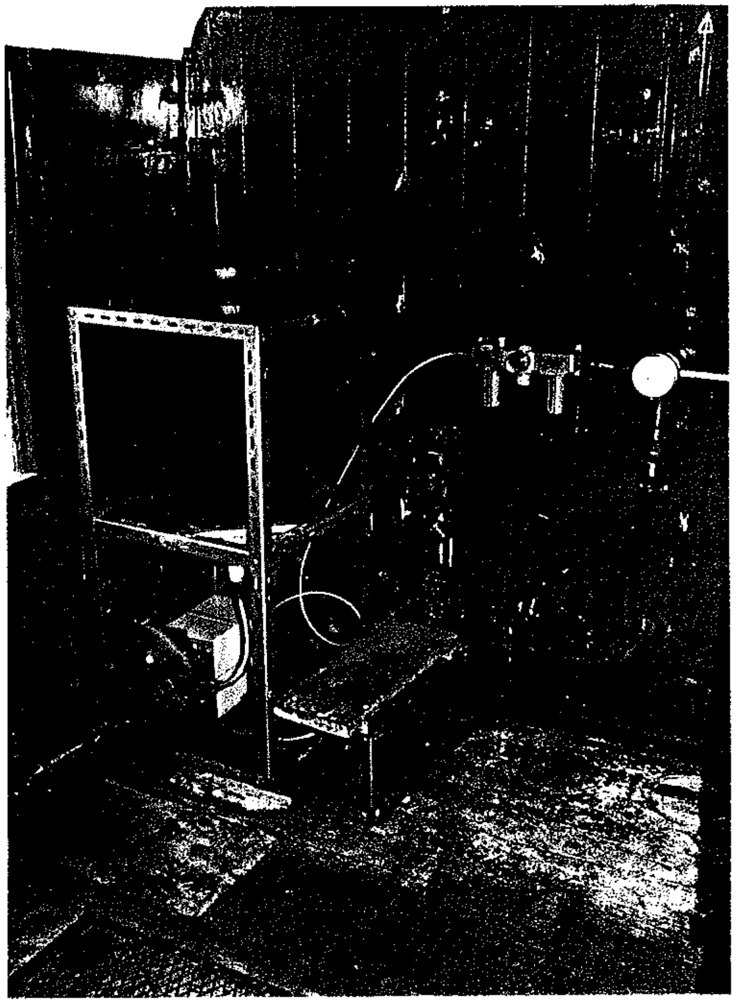
The powerful effect of alkali[22] is the most promising action so far recorded f increasing adhesion. The beneficial effects of soduim meta-silicate may be attributal to the reaction with natural oils to form soluble products.

The possibility may exist of transforming the surface composites by altering the natu of the emulsion by applying an appropriate emulsifier, perhaps a non-ionic detergent.

Experiments have been carried out wherein the rails and wheels were treated by electric arc. They indicated that the probability of sliding at very low coefficients adhesion was very much reduced.[25, 27, 28]

The open arc used in the above tests was noisy and could cause interference with oth circuits. Accordingly British Rail have experimented with the use of plasma-arcs. Each plasma torch had a continuous rating at the maximum current of 350 A of 30 kW with 80% argon and 10% hydrogen. Over a range of experiments made up to 48 km/hr (30 m.p.h.) the mean adhesion level was increased from 0·26 to 0·3 and, as in the case with the straight arc treatment, the skewed distribution characteristic of a number of isolated patches of severe contamination was eliminated.

A method of overcoming low adhesion conditions which can be regarded as standard practice is the application of sand to the track. Whilst generally effective when the sand



(a) Apparatus within vehicle for storage and application of Sandite.

FIG. 12.32. Liquid sanding (courtesy British Rail).

is correctly applied to the nip between the wheel and rail, this is often not the case a high speeds or in high winds. Excessive use of sand may clog the ballast or interfere wi the operation of track circuits. The factors governing the performance of the sandi method may be summarised as follows: (1) the quality and grain size of the san (2) the treatment of the sand and the handling it undergoes between the place of trea ment and the place of use; and (3) the design of the sanders and the optimum rate supply of sand from them.[29]

An improved method of sanding has been developed using a carefully chosen inorgan gelling agent (Laponite) to form a gel with water. At 2½% concentration (by weight) t gel is formed in about 5 minutes, after which it is mixed with an equal weight of sand form a thixotropic suspension which flows readily and can be distributed on to the hea of the rails through small-bore pipes.

A significant improvement in adhesion results from the application of the suspensi (known as "Sandite") which persists for periods up to 6 hours on lines subject to mai line traffic densities. Sandite is now used in several areas of British Rail where delays a rail and wheel-tread damage has occurred due to wheelspin and wheelslide, particula during the presence of falling leaves. For example, in the Leeds Division, commuter lin are now treated every evening during the autumn and winter months. Figure 12.32(shows the equipment installed within the rail vehicle to apply Sandite. The mixture drawn from the reservoir and applied to the rails by a compressed-air-operated peristal pump. Figure 12.32(b) shows the appearance of Sandite on the head of a rail.

(b) Sandite on head of rail.

When, during trials under operating conditions, attempts to apply tractive effort equivalent to a coefficient of adhesion of 0·2 resulted in an 87% slip-risk, the application of Sandite reduced the slip-risk to 5%.

Slip-detection circuits

Automatic devices are therefore employed which act on sliding taking place and these can be arranged either to (1) apply sand to the wheel, (2) reduce tractive effort, (3) increase loading on the offending axle or (4) the application of anti-slip brake.

The most important problem is therefore the detection of sliding by an instrument which is sufficiently sensitive to be effective but which is not affected by such factors as difference in wheel diameter or the normal acceleration of the train. It will be recalled

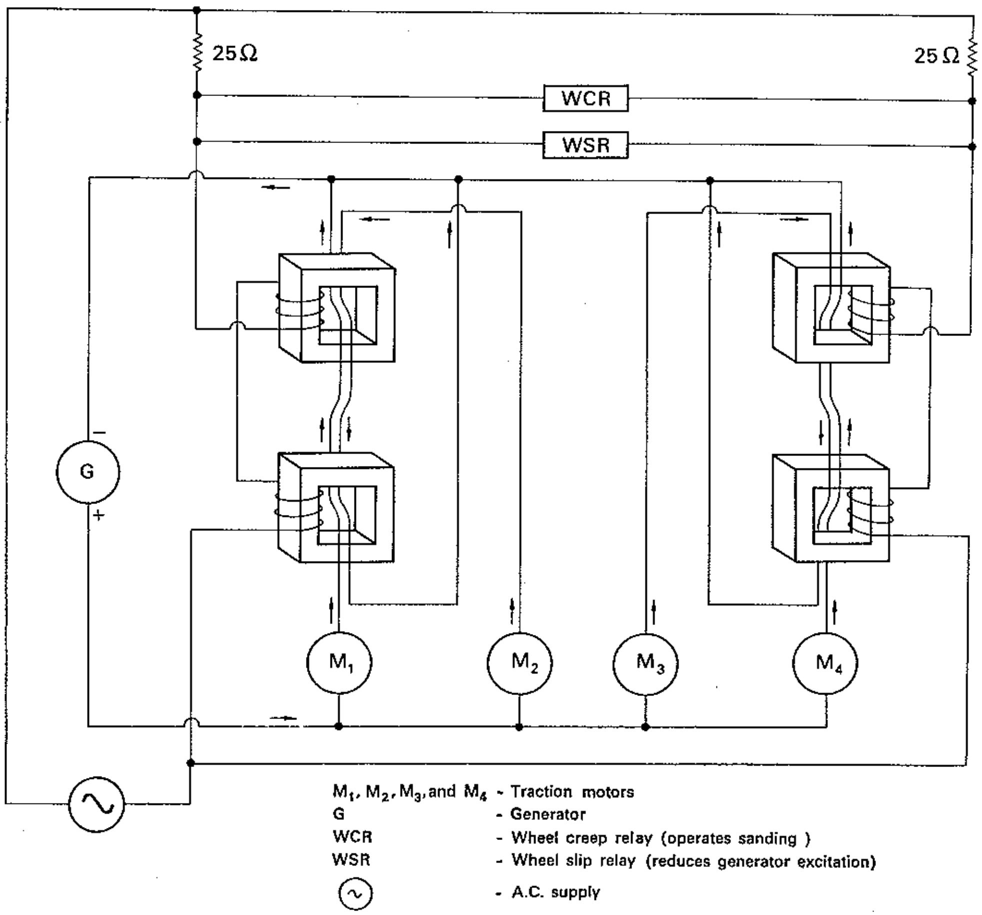


FIG. 12.33. Simplified circuit for transductor slip control.

that with a series motor both current and voltage across the armature change with sp
This fact may be used to indicate the increase in speed which accompanies sliding. W
two motors are in series, a comparison of voltages across their individual terminals
also provide an indication. This provides an excellent opportunity for the use of tr
ductors. The motor leads carrying d.c. are arranged to pass through transductor cc
those from two motors being connected in opposition as indicated in Fig. 12.33. W
current in each motor circuit is equal, the magnetic fields cancel out and the a.c. win
is unaffected. When, however, a difference occurs the core approaches saturation
corresponding increase in the a.c. current in the detection circuit. At starting the two
of motors are connected in series and wheel slip relays compare the voltage across
motor of a pair and transductors compare the current taken by each group. S
adjustment of the sensitivity of the apparatus is necessary to avoid false operation at
speeds whilst maintaining rapid action at high speeds. This is achieved by a third
winding on the transductor which carries a current proportional to generator voltag

Another method of control is to derive an error signal from axle-driven generat
The speed of each axle is measured and compared with the average speed of all a
Here again the sensitivity of the device is varied with speed, a 6 m.p.h. (2·7 m/s) differe
causing operation at low speeds increasing to 12 m.p.h. (5·4 m/s) at 100 m.p.h. (45 n

For a modern control system a digital input may be preferable. Figure 12.34 sho
method employed by the S.N.C.F.[38] A pick-up, consisting of a permanent ma

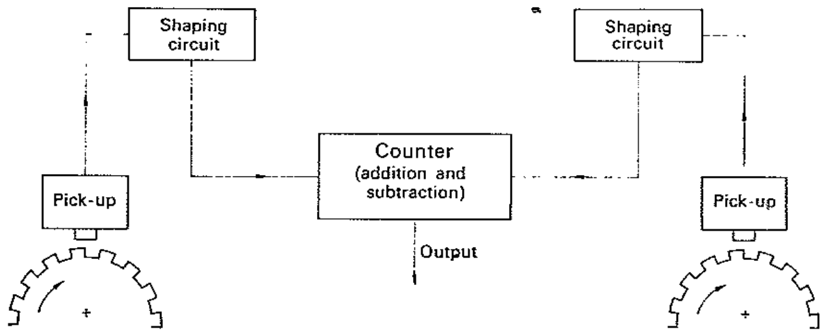


FIG. 12.34. Digital arrangement for slip control.

embraced by a coil, is mounted on the gear case of the main motors so that the pass
of the teeth of the gear drive causes an a.c. current to be generated having a freque
proportional to the speed of that particular axle. The output from each axle is sha
up to form a pulse of rectangular wave form. These are then fed to a counter unit
of opposed senses. Thus the shaped pulses from one pick-up are added and those f
the other subtracted so that when the speeds of the two axles are identical, zero ou
results. When slipping takes place, an output can be obtained which consists of a num
proportional to the integral extent of sliding. This can be converted to a rate measurem
by resetting at frequent regular intervals (about 100 metres) and taking the maxim
reading. Variation in tyre diameter is compensated for by cutting off one pulse i
pulses from the pick-up serving the smaller wheel set.

A refined transducer which senses gear-tooth position has been described.[39]

In view of the fact, typified in Fig. 12.31, that the maximum coefficient of adhesion is often accompanied by considerable relative sliding it is desirable to set an automatic slip control to a particular value of relative sliding. To do this it is necessary to know the true speed of the vehicle along the track as well as the peripheral speed of the wheels. Radar-based devices are now used to measure the time speed.

Anti-slip brake

It has been found that a mild application of the brakes may bring wheel spin under control. This may be attributed to a cleaning action by the brake shoes or alternatively to the restoration of stable conditions by reducing relative sliding to below the critical value. Two stages are generally provided and controlled by the same type of electronic control as is applied to the anti-skid device described on page 196 Chapter 13.

12.14. Control of diesel power

The torque-speed characteristic of a diesel engine is much flatter than that of a petrol engine and therefore the speed of the engine can be varied over a continuous range in accordance with the demand for power. This is accomplished by connecting the engine governor to the generator field. Control of power output is by means of engine speed. On a diesel-electric locomotive the main control is usually pneumatic so that position of the driver's control handle determines the air pressure which is led to the engine governor so as to vary its setting. Figure 12.34 shows a schematic diagram of an engine control system.

Suppose that load increases on the locomotive due to an increasing gradient; engine speed will fall and the governor will tend to increase the fuel input and therefore the power output. This, however, will also close an electrical circuit which will cause a rheostat in the generator excitation circuit to be adjusted by a miniature motor. Decrease in excitation will result in a decrease in generator output volts. Thus motor speed will decrease and the train will slow down. The corresponding reduction in load will cause engine speed to rise again, the electric circuit will now be opened and the generator excitation rheostat will remain in its new position until there is some further change in train-resistance, or until the driver, noting the falling off in train speed, deliberately advances his controller to compensate.

Whilst the principle of relating engine speed to power output is soundly based, the method of driving at constant power during the accelerating period is manifestly impossible due to adhesion and limitations on motor current at lower speeds. A system wherein the driver determined tractive effort, for example by setting a relay actuated by motor current, would be preferable. The engine speed would then be adjusted automatically to provide power most economically, governor setting being adjusted to output by some inbuilt control function. The set of curves shown in Fig. 12.35, whilst typical of engines as a class, do not necessarily represent precisely the performance of any particular engine at any particular time, therefore an "adaptive" control system which adjusted engine speed to power demand for minimum fuel consumption irrespective of vehicle speed is desirable.

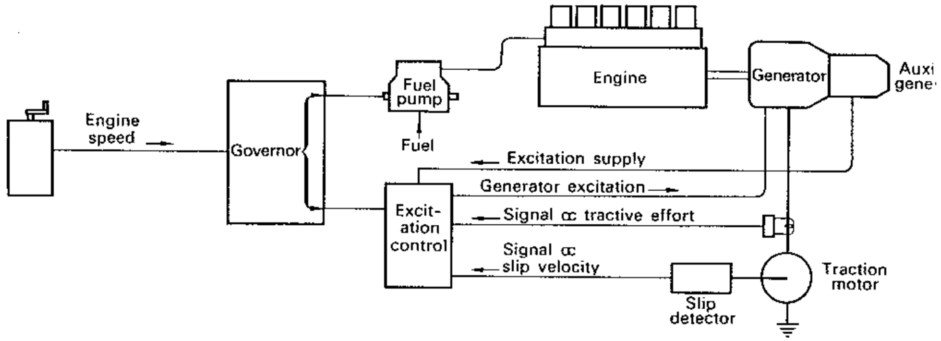


FIG. 12.35. Control diagram for diesel-electric drive.

12.15. Adaptive control

This serves to introduce the idea of an "adaptive" control and the subject c "autonomics". Supposing we have two variables each independent of one another an not under control. We can arrange independent feedback systems so as to stabilise som mechanism which is subject to their action. However, the resulting setting may not b the optimum. If we plot the two variables in the *X* and *Z* directions and construct

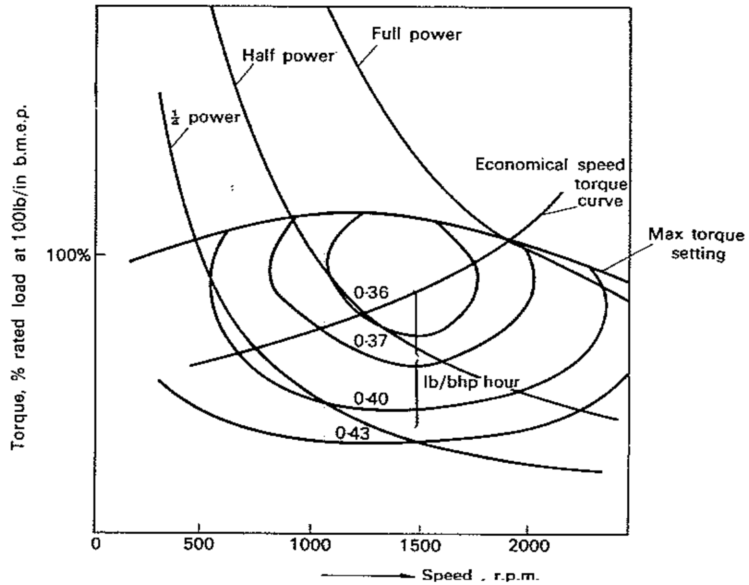


FIG. 12.36. Economical speed–torque settings for a diesel engine.

three-dimensional model in which the output quantity was represented in the Y direction, this would have the appearance of a hill (Fig. 12.36). The values of X and Z, corresponding to the peak of the hill, become the desired solution. Representing the system on a sheet of paper one uses contour lines as on a map. Most adaptive control systems embody a mathematical model wherein the predetermined relationship between variables is incorporated. One could, however, arrange for the machine to be subject to minor changes, the result of which would determine the next change. Thus a change leading to a lower value would be followed by one in the reverse direction. Any direction tending towards an increase would be followed until the increase stopped when another direction would be tried. This is equivalent to a man walking in a fog on the side of a hill. Even if he had no compass, provided he always walked in a direction leading upwards, he would eventually reach the summit. Thus this technique is known as "hill climbing".

An alternative would be to arrange for the equipment continuously to follow a small perturbation in both the X and Z directions so phased as to appear as a circle on the contour map. The value of the ordinate would be measured and integrated round the periphery. The mean height would be determined and the moment of the difference between the actual and the mean height calculated about every diameter. The centre of the circle would then be caused to move at right angles to the diameter giving maximum moment in the direction of upward tilt. For other adaptive sequences see reference 40.

Thus an adaptive control system can lead to a system wherein the input to a control system is itself part of a closed loop. Thus the characteristics of the system may themselves be changed in the light of changing circumstances. Such systems are known as "autonomic", that is, self governed.

12.16. Shock factors in acceleration

It is usual to impose limitations on the acceleration of rail-borne vehicles in the interests of passengers' comfort although this factor is seldom referred to in connection with road vehicles. It is the writer's view that rate of change or some higher derivative of acceleration is more critical because passengers can adjust themselves to changes in the magnitude and direction of the effective gravitational field, provided that they are afforded sufficient time to do so and that the element of surprise is minimised. One approach is to assume that there are certain "threshold" values of acceleration and its derivatives which do not disturb comfort. Figure 12.38 shows distance travelled and its derivatives against time for threshold values of rate of change of rate of change of acceleration, rate of change of acceleration and acceleration recommended by the S.N.C.F. Another possible solution of the comfort problem would be to find a function which increased smoothly from zero at $t = 0$ together with all its derivatives. This condition is met by a power law as developed in Fig. 12.39. It would, of course, be necessary to derive a similar law to govern the transition to steady state running at the end of the acceleration period.

A sinusoidal variation meets the stipulation that it contains no sudden changes which might take the passenger by surprise as well as being convenient mathematically. Figure 12.40 shows some curves constructed on this basis.

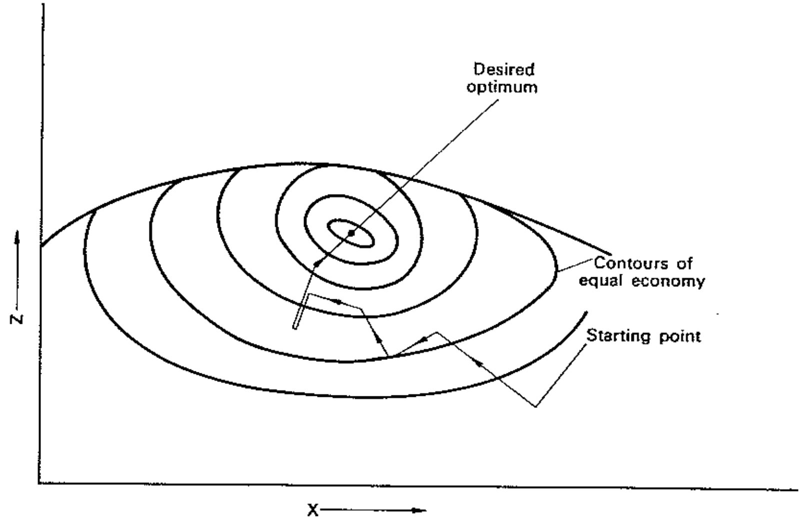


Fig. 12.37. Hill climbing technique.

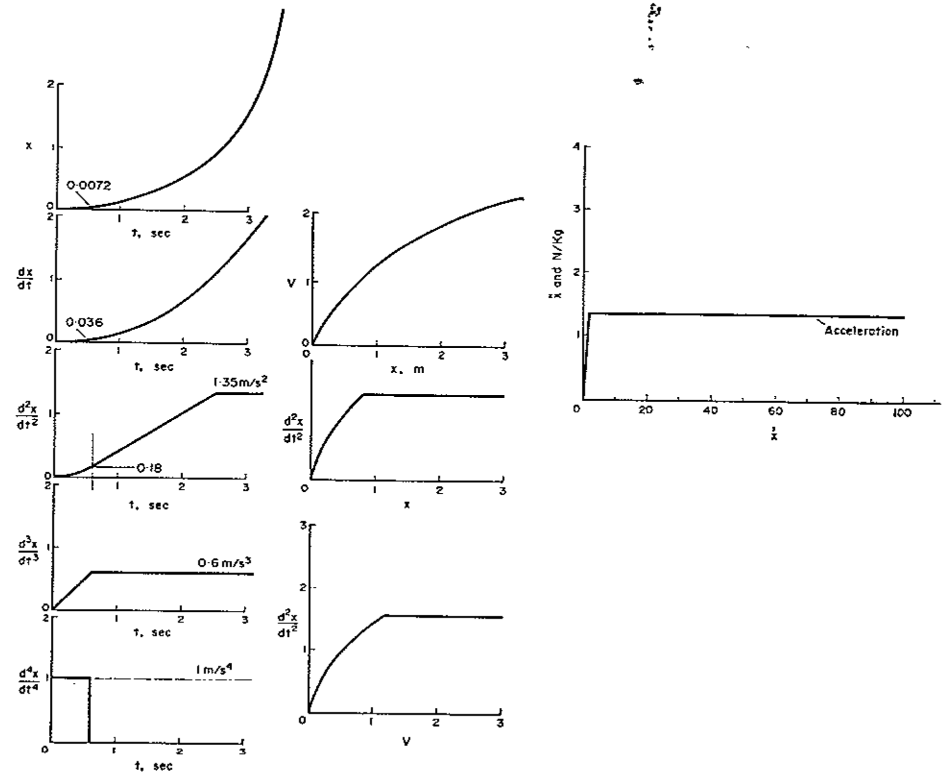


Fig. 12.38. Threshold values for acceleration.

Assuming that some threshold exists below which no adverse effects are experienced it might be supposed that Robertson's law of growth would provide a gentle transition. Figure 12.41 is based on this law which gives a result which does not differ markedly from the sinusoidal case.

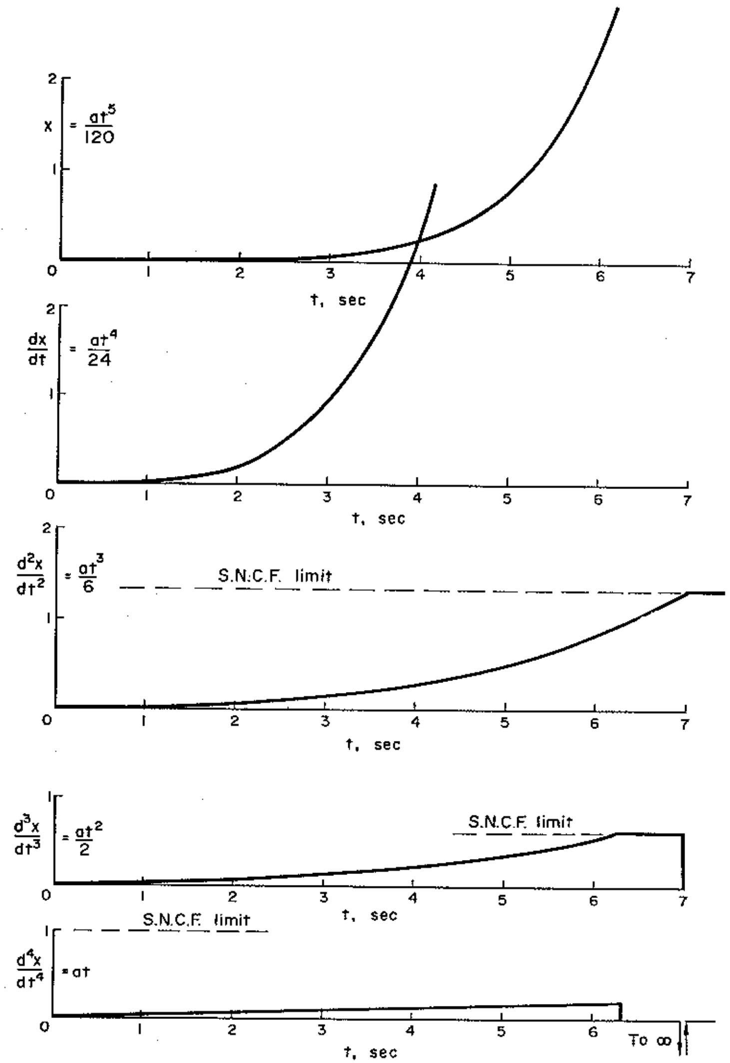


FIG. 12.39. Power law acceleration.

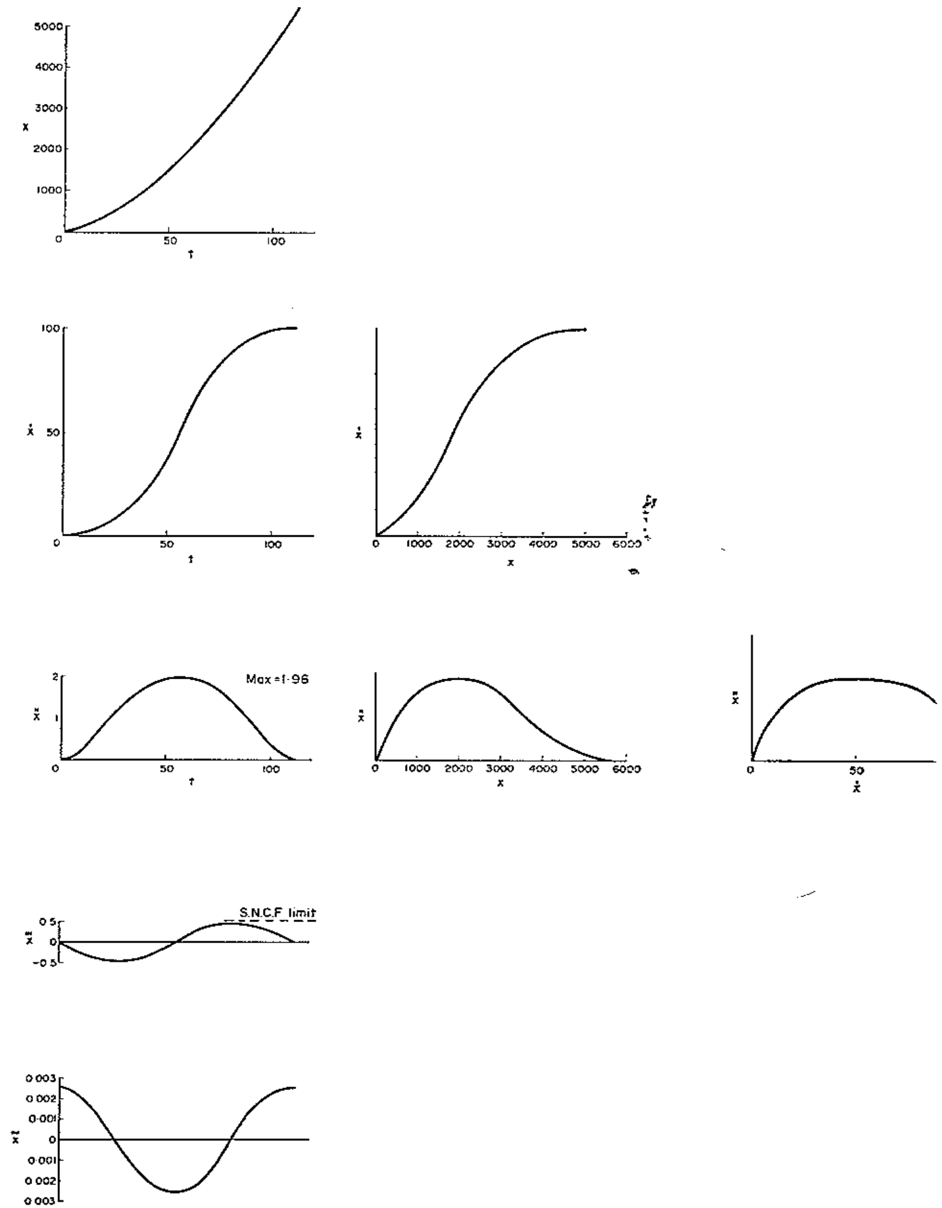


FIG. 12.40. Sinusoidal acceleration.

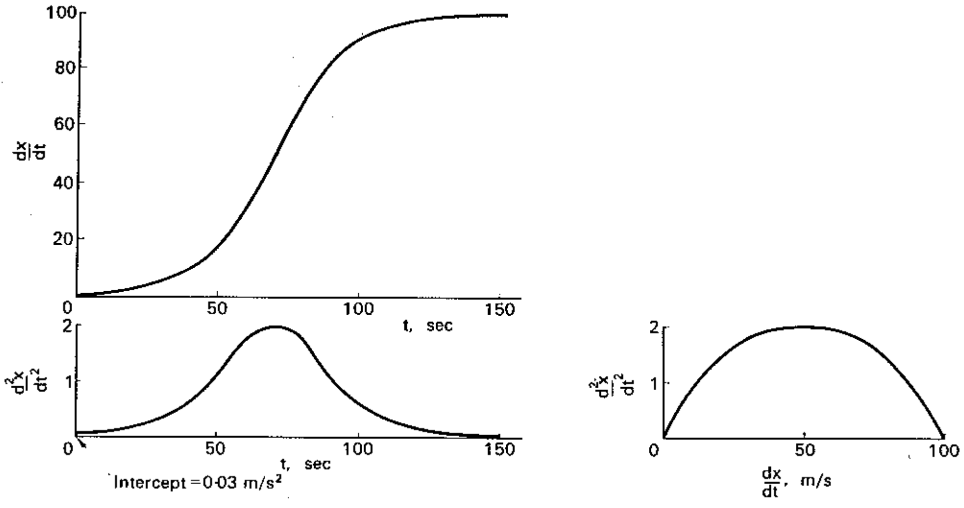


FIG. 12.41. Robertson's law of growth.

References

1. FOGG, A., Measurement of aerodynamic drag and rolling resistance of vehicles, *FISITA, Tenth International Automobile Technical Congress*, 1964, p. 3. Society of Automobile Engineers of Japan, Inc.
2. STEEDS, W., *The Mechanics of Road Vehicles*, Iliffe, London, 1960.
3. DOVER, A. T., *Electric Traction*, 4th ed., Pitman, London, 1963.
4. ANDREWS, H. I., The measurement of train resistance. *Proc. Instn. Loco. Engrs.*, vol. 44, p. 91 (1954).
5. LIPETZ, A. I., Air resistance of trains. *Railway Mechanical Engineer*, vol. 112, p. 129 (1938).
6. HAMMET, A. C., *The Aerodynamics of High-speed Ground Transportation*, Western Periodicals Company, North Hollywood, California, U.S.A.
7. DALBY, W. E., *Steam Power*, Arnold, 1900.
8. MAXWELL, J. C., On governors, *Proc. Roy. Soc.*, vol. 16, p. 270 (1868).
9. WELBOURNE, D. B., *The Essentials of Control Theory for Mechanical Engineers*, Edward Arnold, 1963.
10. ELLIS, F. E., History and prototype development. *Proc. Instn. Mech. Engrs.*, vol. 181, pt. 2A, p. 173 (1966).
11. ATKINS, A. J., Production design of an automatic transmission. *Proc. Instn. Mech. Engrs.*, vol. 181, pt. 2A, p. 186 (1966).
12. DAVIES, P. B., A study of continuously variable transmission. Ph.D. Thesis, University of Wales, 1962 (copy available in library of Institution of Mechanical Engineers, London).
13. MACMILLAN, R. H., The control of stepless variable speed transmission in automobiles and their possible application to regenerative systems. *FISITA, Third International Automobile Technical Congress*, 1964.
14. BARWELL, F. T., Safety and automation on electric and diesel-electric motive power units. *Bull. Int. Railway Congress Association*, vol. 38, p. 924 (1961).
15. LUCAS, A. W. and ROBERTSON, A. G., Control of tractive effort on electric locomotives. *Proc. Instn. Mech. Engrs. Convention on Adhesion*, 1964, p. 60.
16. DAVIS, R. M., Power diode and thyristor circuits. *I.E.E. Monograph Series* 7.
17. LITTLER, G. E., The harmonic performance of a high voltage electrified railway with single and multiple train load. *I.E.E. Conference Paper No.* 203, p. 105 (1981).
18. HERD, M. G., Effect of thyristor controlled traction on a 25 KV 50Hz electrification system. *Ibid.*, p. 87 (1981).

19. DAVIS, R. M. and BARWELL, F. T., The commutatorless diesel-electric locomotive. *Power Applications of Controlled Devices*, I.E.E. Conference Publications, Part 1, p. 133, 1965.
20. BORGEAUD, G., Weight transfer in a two-bogie locomotive and its compensation. *Proc. Instn. Mech. Engrs. Convention on Adhesion*, 1964, paper No. 5, p. 75.
21. LOOSLI, H. E., Adhesion between wheel and rail—results and experiments in an Alpine country. *Rail Transportation Division of the American Society of Mechanical Engineers. Report No. 82-RT-4.*
22. BARWELL, F. T. and WOOLACOTT, R. G., The N.E.L. contribution to adhesion studies. *Proc. Instn. Mech. Engrs.*, vol. 178, pt. 3E, p. 145 (1963).
23. BARWELL, F. T., Wheel to rail adhesion. *Proc. International Symposium on Contact Mechanics and Wear of Rail/Wheel Systems*, Vancouver, 1982.
24. O. R. E. Utrecht, QUESTION B.44. *Adhesion of locomotives from the point of view of their construction and operation.* Report No. 2. Adhesion Tests between Pontartier and Frasne, 1962 (1964).
25. Report No. 3. The development of spark cleaning for improving locomotive adhesion (1964).
26. Report No. 4. Statistical methods in the study of locomotive adhesion (1965).
27. Report No. 5. General report on the development of the practical application of sparking for improving locomotive adhesion (1966).
28. Report No. 6. Tests of rail sparking to improve adhesion, carried out on British Railways (1967).
29. Report No. 7. Use of sand to improve adhesion (1969).
30. Report No. 8. Equipping of a test machine with variable constructional characteristics (1969).
31. Report No. 9. Adhesion tests between Wadgasson and Hargarten (1971).
32. Report No. 10. Adhesion tests in 1971 (1973).
33. Report No. 11. Adhesion tests in autumn (1972).
34. Report No. 12. Comparative tests carried out in November 1975 between ÖBB thyristor locomotive class 1044 and SBB direct engine locomotives Re 4/4111 and ÖBB class 1042 with respect to their efficiency at the limit of adhesion (1976).
35. Report No. 13. Adhesion tests of spring 1975 with the test machine 18,000 converted to 16⅔ Hz traction (1977).
36. Report No. D769 [B44]E. Comparative tests concerning the adhesion of 3 kV & A.C. multiple-unit trains (thyristor controlled and conventional types) (1977).
37. Report No. 14. Synthesis Report. The current state of knowledge about wheel–rail adhesion (1978)
38. NOUVION, F., Electrical control devices for the control of adhesion. *Proc. Instn. Mech. Engrs. Convention on Adhesion*, 1964.
39. BOWLER, P. and SALES, G. A., A wheel speed transducer and slip-speed monitoring system. *I.E.E. Conference Publication No.* 203, p. 232 (1981).
40. MACMILLAN, R. H. and REES, N. W., *Autonomic Control Systems—Non-Linear Control Systems Analysis*. Pergamon (1962).

CHAPTER 13

Control of Braking

13.1. Forms of braking

Braking is the most important control function in transport because it determines the maximum capacity of a traffic lane.

The rate of braking is therefore an inherent characteristic of the control system as a whole and is determined by the requirements of that system, subject to the following limitations on its maximum value:

1. Deceleration should not cause discomfort or injury to passengers.
2. Physical factors, such as coefficient of adhesion between wheel and track, place a limit on maximum value of deceleration which may be obtained. These may vary from time to time.

The desired braking rate having been communicated to the vehicle by automatic means or determined on the vehicle by the driver or automatically as a result of information received, means have to be provided to regulate the braking apparatus to comply with the desired value. Figure 13.1 shows how the rate of energy dissipated is dependent on speed.

The physical means of braking generally available are as follows:

- Friction braking
  - on tread
  - off tread
    - expanding
    - disc
- Magnetic
  - track
  - electrodynamic
- Electric machine braking
  - resistance
  - regenerative

Auxiliary methods of dissipating kinetic energy, such as counter-pressure braking using engine cylinders or drag in hydraulic converters, can be used to reduce the need for friction braking.

With the exception of the track and electrodynamic brakes, all the above means are limited by the product of the coefficient of adhesion between wheel and track and the

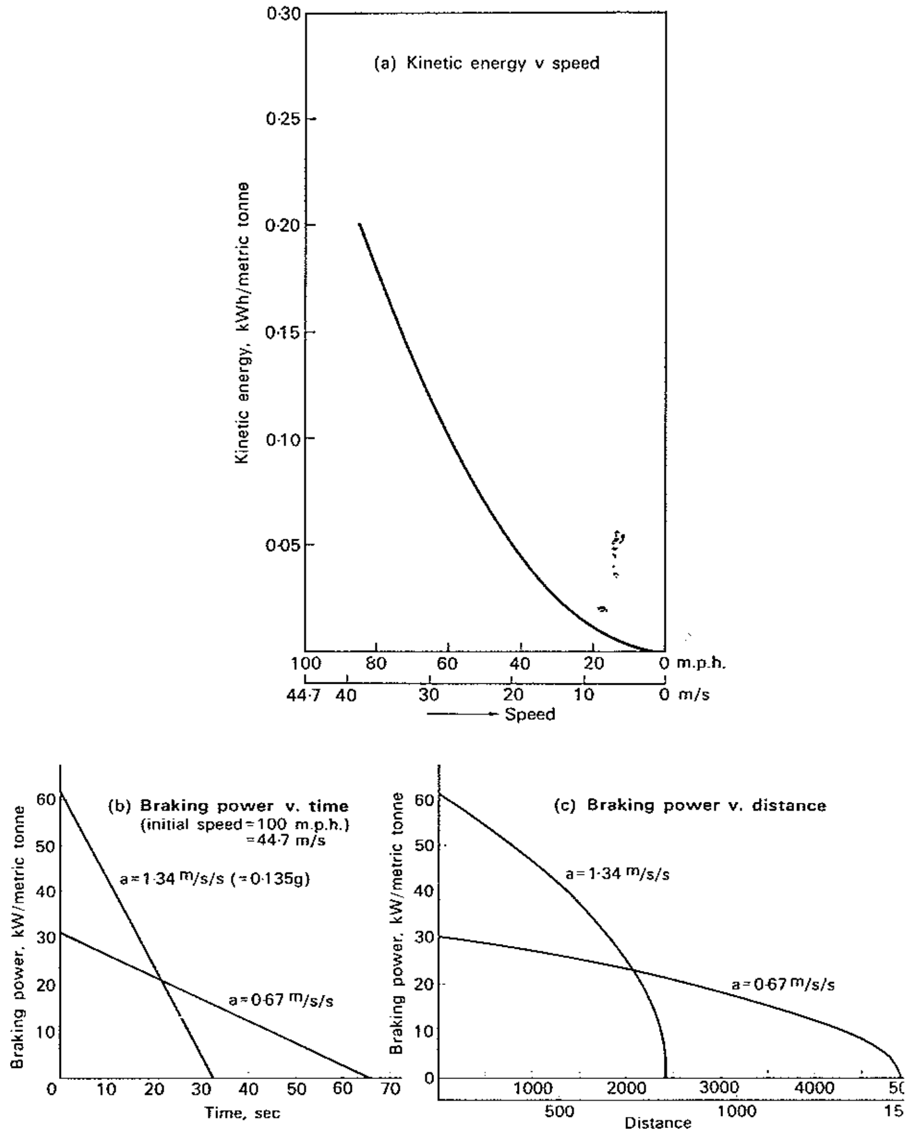


FIG. 13.1. Variations of braking energy with speed.

wheel load. If all the weight of the vehicle is taken on the braked wheels, the maxim braking force available is μMg and therefore the maximum rate of deceleration is μ itself is not a constant value but varies with creep and relative sliding as shown Fig. 12.31. Should the maximum be exceeded, the wheels will slow up. This must avoided for three reasons. First, the value of μ and therefore of deceleration is reduc and braking distance increased. Secondly, lateral control of the vehicle is lost, with the risk of sideways skidding (differences between front and rear axles are of importance here, see Chapter 14). Finally, when the wheels have come to a stop with the vehicle still moving, wear and frictional heating are concentrated on a narrow area. Thus flats are worn and the risk of tyre bursting increased.

With the friction brake, the task of ensuring that wheel-track adhesion is not exceeded is complicated by the fact that the frictional property of the brake-shoe material is a variable quantity. Sources of variation for off-tread brakes may be relative velocity and temperature (brake fade) whereas "on-tread" brakes may be affected by weather conditions.

The effect of a constant acceleration (a) on an object contained within a vehicle is to displace the vector representing gravitational force into an inclined position so that a passenger standing up in a vehicle experiences the sensation of the floor being inclined. The equivalent gravitational force is, of course, increased to $\sqrt{a^2 + g^2}$ which is not usually noticeable.

The inclination of the gravitational vector is the basis for the operation of the retardation controller which has been used for many years on electric trains of the L.T.E. The basic sensing element is a glass U-tube containing mercury (Fig. 13.2). This tube is

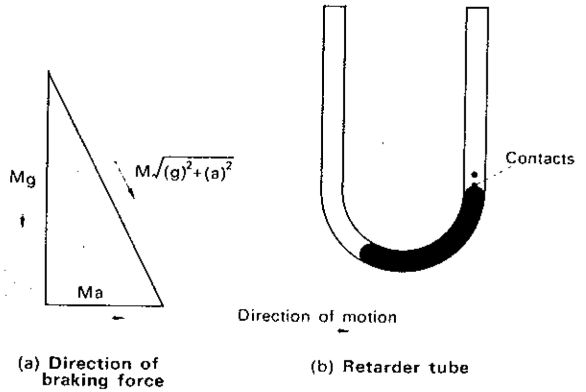


FIG. 13.2. Control of retardation.

mounted rigidly to the car body with the U-form lying in the vertical plane containing the direction of motion. The mercury will act as a pendulum and will rise in the limb nearest the front of the train by an amount proportional to the deceleration. Contacts are arranged so that the circuits will be closed as soon as a predetermined degree of deceleration is obtained.

The method has the advantage of compensating for the tractive forces associated with gradients. If a vehicle is descending an incline, the brakes will be required to absorb potential as well as kinetic energy. The effect of the inclination of the mercury tube is to cause the contact to be made at a correspondingly lower value of deceleration so that total braking force remains at the prescribed value.

13.2. Magnetic track and electrodynamic braking

Magnetic track brakes are commonly fitted to light electric vehicles. When not in they are usually suspended above the rails at a height of 0·1 m (4 in.). When required action they are lowered by pneumatic cylinders and energised by batteries carried on vehicle. There are broadly two types of construction, one in which single-pole pieces mounted on each side of a single coil as shown in Fig. 13.3(a), or the articulated typ which the magnet core is subdivided into a number of intermediate elements each which can adhere to the rail surface independently of its neighbour as illustrated Fig. 13.3(b). This helps to compensate for irregularities in the rail profile.

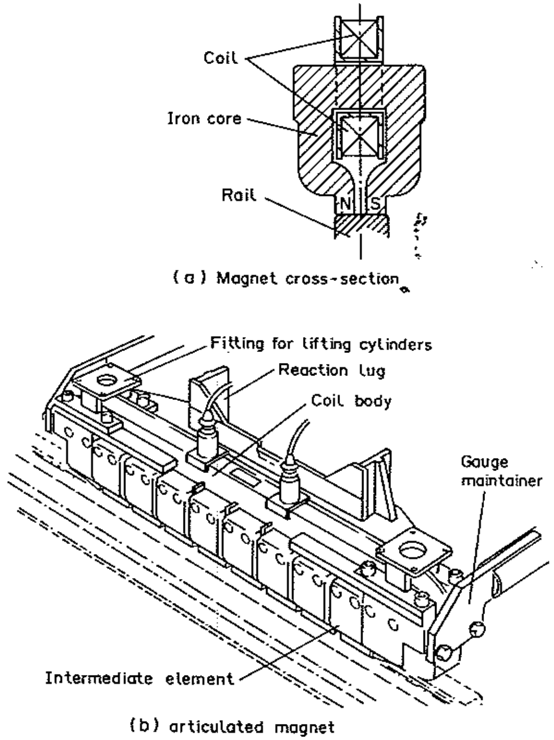


FIG. 13.3. Magnetic track brake.

Magnets are usually about 1 metre long and weigh about 200 kg. They may be wou for various voltages and consume about 1 kilowatt. Brakes are usually released wh speed has fallen below 50 km/h (31 m.p.h.) because friction increases sharply at low spe At 50 km/h a typical value of coefficient of friction is 0·20 giving a braking force of 10 k (1 ton-force) which corresponds to a normal attraction between wheel and rail of 50 k (5 ton-force).

The electrodynamic (eddy current) brake differs from the track brake in so far as there is no contact between rail and brake shoe but an air-gap of about 0·07 m (2¾ in.) is incorporated. The magnetic segments are wound with alternate north and south polarity so that flux passes across the air-gap along the rail and then across the next pole. When there is no relative motion between magnet and rail there is simply a normal force between them. When there is relative motion, however, voltages are generated leading to the formation of eddy currents. These distort the magnetic field, introducing a transverse component to the force acting on the magnetic poles. This is so directed as to oppose motion and thus to provide a braking action. The kinetic energy of the vehicle is therefore converted into thermal energy within the rail during the braking process.

Figure 13.4 shows the results of tests on a laboratory test stand. It will be noted that the braking force attains a maximum value of about 11 kN per metre (7400 lb force/foot) at a speed of about 75 km/h (47 m.p.h.).[1]

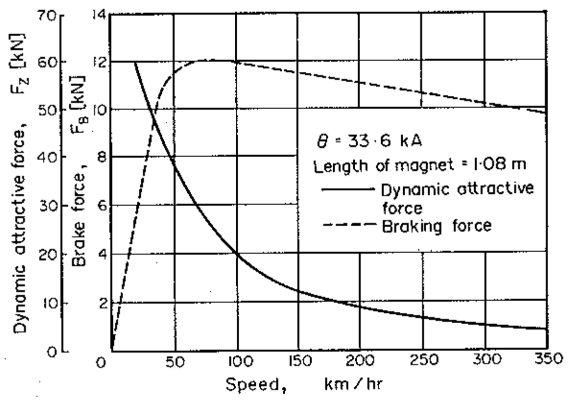


FIG. 13.4. Attractive force and braking force curves for eddy current brake.[1]

Greater flexibility, particularly at low speeds, can be achieved if the magnet is provided with a three-phase winding so that a travelling field is set up in the direction of motion. The speed of this field relative to the winding will be 2 *Pf* where *P* is pole pitch and *f* is the frequency of supply. This would have the characteristics of a linear motor in the plugging region (see Chapter 17, Fig. 17.35). Although subject to much experimentation, notably in France, the electromagnetic braking systems have not yet been applied in practice, chiefly because of the risk of overheating the rail.

13.3. Physiological aspects

The physiological reaction of passengers to accelerating forces has not yet been sufficiently studied but it seems most likely that it is change of acceleration rather than acceleration itself which is most important. A passenger can adjust his body posture to a

uniform acceleration which corresponds to a change in magnitude and direction o gravitational field within which he finds himself without much difficulty and, if stays constant, he will not experience discomfort. If the direction changes frequentl constant adjustments will lead to fatigue, and perhaps more important, if it change rapidly, adjustment may be insufficiently rapid to avoid injury.

The American word ROCOC is sometimes used. Its derivation is "Rate of Chan either acCeleration or deCeleration". For railway work, recommended limits[2] normal braking are 1·37 m/s² with a maximum rate of change of 2·35 m/s³. For emerg braking 2·35 m/s² deceleration and 9·0 m/s³ may be tolerated.

On the highway much higher values than these are commonplace and there w appear to be no objection to utilising the available adhesion to the full, provided shock could be avoided on application and release of the brakes.

Wilson[3] has reported on drivers' and passengers' reactions to braking from 70 m (30·6 m/s) as shown in Table 13.1.

TABLE 13.1.

REACTIONS OF PERSONS TO BRAKING DECELERATION

Average deceleration (m/s)	Reaction of driver	Reaction of passenger
2·65	Comfortable stop	Comfortable stop
3·43	Undesirable	Undesirable but not alarming
4·22	Very undesirable. Regarded as emergency stop	Severe and uncomfortable (may inflict injury if passengers unprepared)

The Road Research Laboratory,[4] in commenting on these results, point out these were average decelerations; maximum decelerations were greater perhaps by 40%. However, the remark in the brackets in the last line of column 3 indicates that of change of deceleration may be much more important from the passenger's poin view than deceleration itself. The consideration underlying the construction of Figs. 1 to 12.41 are even more important in braking than in acceleration. The limit for pu service vehicles braking from moderate speed is given as 4·9 m/s² although half value may cause injury to standing passengers.

13.4. Control of slip (slow up)

A form of control originally used in aircraft landing gear but applicable to r vehicles is known as the "Maxaret" system. Here the rotational deceleration of the wl is measured and the intensity of brake application is moderated so as to limit quantity. This method is an indirect one in so far as an assumption has to be m regarding the maximum adhesion likely to be available. The value of deceleratio then set so that this will not be exceeded. If conditions are such that "slow up" oc even at this limited value, this will become apparent by a higher rate of deceleration of the wheel and the control will therefore reduce the intensity of brake application.

Similar principles have been applied to railway braking but the equipment is set so that it does not operate until the axle is slowing up at a distinctly higher rate of deceleration than that corresponding to the deceleration of the train itself. When it operates, brakes are released completely and then reapplied. Thus the device may have the effect of increasing braking distance; also, if the braking torque is very nearly balanced by adhesion, the wheel will "slow up" at a rate only slightly in excess of the general rate of the train so that the device will not be actuated. In certain circumstances, on a long gradient for example, the axle may come to a standstill causing flats to be worn on the tyre.

Direct monitoring of available adhesion is therefore preferable and would permit the maximum safe braking rate to be applied at any time. Where idle wheels are available, speeds may be compared and similar techniques employed as those described in Chapter 12 for monitoring adhesion during acceleration.

Another source of variation in braking torque is the change of coefficient of friction between brake shoe and tyre, disc or drum. This can be monitored by reaction of the brake itself and brake application pressure controlled to give the required braking force. This will only be effective of course in the presence of an adequate coefficient between wheel and track.

Modern equipment, either multiple-unit or locomotive hauled, is fitted with anti-skid control. Referring to Fig. 13.5, each axle is fitted with a device which measures the exact

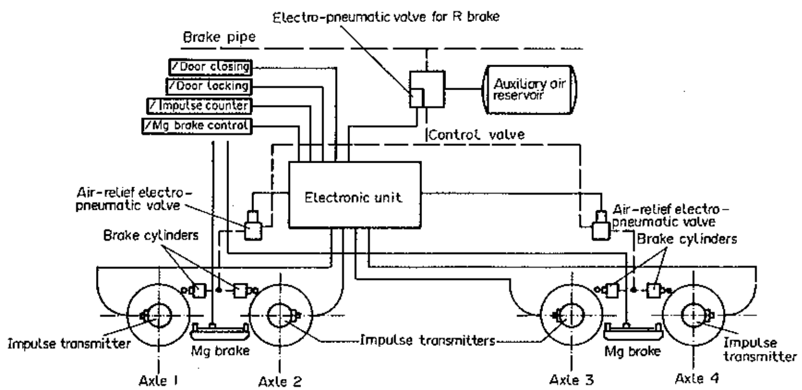


FIG. 13.5. Automatic railway brake control system.

speed. This information is processed in a microcomputer so as to quantify the extent of any skidding or "slow up" (i.e. wheels slowing down at a faster rate than that corresponding to the train speed). Braking action is modified on this basis, the retarding torque being reduced in response to momentary adhesion conditions so as to avoid skidding. Various criteria may be used to evaluate slip. Thus one method may be a direct comparison between the rate of rotation of two axles (adjusted for variation in wheel diameter)

and another is the difference in speed of any single axle and the mean of all the c axles. In the remote event of all axles skidding simultaneously by the same amou limit may be set at a value of deceleration slightly above the maximum value expe during normal working. A limited amount of skidding is allowed to occur so tha approximation may be made to the attainment of optimum adhesion values.

Whilst the retardation which can be applied to a vehicle should be independe loading being dependent solely on coefficient of adhesion, an individual vehicle provided with any form of monitoring "slow up" will be unable to withstand as n braking torque when empty as when laden. Some modern vehicles may carry three t their own weight so that a four to one variation in braking torque may have to be allc for. It is assumed that for practical purposes wagons are either empty or nearly loaded so that it is considered satisfactory in this country to arrange for two stag braking only the change-over taking place at 65% capacity. Figure 13.6 shows a sy

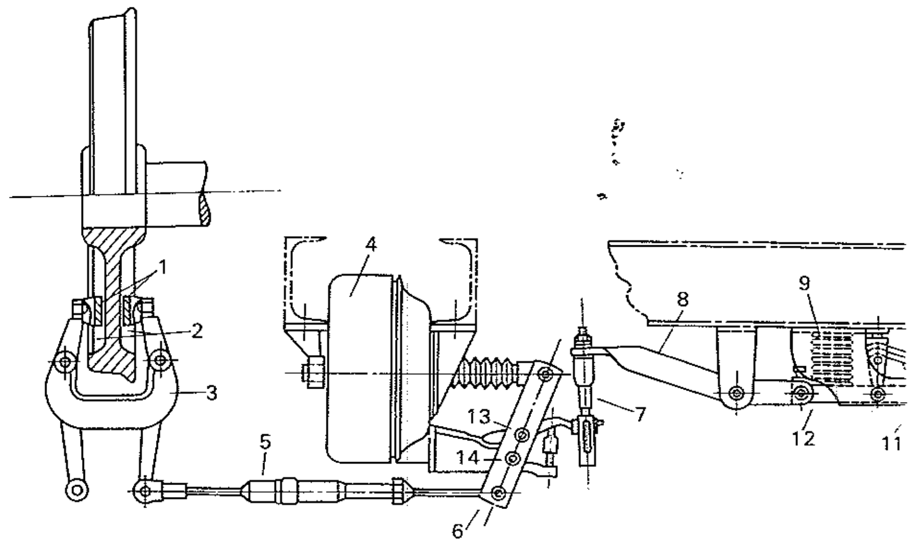


FIG. 13.6. Automatic two-stage disc brake showing mechanism as for vehicle in tare condition with brakes off.

of disc brakes recently developed for application to four-wheeled mineral wagon British Railways.[5] Here the control is derived from the reaction at the point of att ment of one of the laminated springs (10). The reaction of the load at this poi transferred to pivotted beam (12) which is supported at the other end by helical spr (9). When load is below 66% capacity the strut containing fulcrum (13) remains eng in the position shown so that the brake cylinder operates through a one to one lever As the wagon is loaded, connecting lever (8) forces connector (7) downwards so t on the load exceeding the prescribed amount, fulcrum (13) becomes disengaged fulcrum (14) becomes effective, imposing a 2½ to 1 leverage on the system with a co quent 2½ times increase in braking torque.

Another form of automatic adjustment commonly found in braking systems functions in order to combat the effects of wear. Most braking systems allow a little free space between the friction pad and the rotating element. This should be kept as small as possible in order to minimise waste of time and energy when the brakes are applied. The p inciple of operation consists of determining a zero with brakes on and then limiting the amount of travel relative thereto of the block and its immediate attachments when brakes are released, whilst allowing the brake piston or other operating feature to travel fully back to its inactive position. Thus any wear that has taken place is compensated for by the difference between these two movements.

13.5. "On-tread" braking

This is now confined to railway practice and interest in Britain relates to two materials, cast iron and resin-bonded non-metallic blocks, although wooden blocks are sometimes used in overseas metropolitan systems.

Cast iron is still the most commonly used material because of its cheapness and reliability. It has, however, the disadvantage that the brake dust can be objectionable owing to its general dirtiness and its effect on electrical equipment. Cast iron has the advantage that its coefficient of friction does not alter materially as between wet and dry conditions but the disadvantage that the coefficient of friction is markedly dependent upon speed, increasing as speed is reduced. This may lead to a very uncomfortable stop so that drivers tend to release the brake immediately before stopping, thereby increasing braking distance. If brake pressure is set so as to avoid "slow up" at moderate speeds, retardation will be inadequate at high speeds. Some form of automatic adjustment is therefore necessary. Two-stage brake systems are sometimes used wherein higher pressure is applied over the higher speed range.

The frictional properties of non-metallic blocks are very much dependent on their composition and manufacture. They possess the advantage that their frictional properties do not vary with speed. The disadvantage that the coefficient falls off with rainy weather is not as great as appears at first sight because adhesion between wheel and rail varies in a similar manner so as to reduce the risk of "slow up". Improved forms of automatic control which are able to compensate for variations in brake block friction are required to suit traffic conditions of the future and in these circumstances there would appear to be little advantage in retaining the cast-iron block.[8]

13.6. "Off-tread" braking

Expanding shoe brakes are used extensively in automobiles and can be arranged to have an amplifying effect if the shoe is in "leading" position. They are often non-linear in the relationship between applied force and braking torque and are not particularly suitable to automatic control because of the effect of geometrical changes arising from wear. Disc brakes have attractions for both road and rail because their characteristics are unaffected by wear and because they afford better opportunities for heat dissipation.

The falling off in coefficient of friction of the cast-iron block renders it unsuitable for modern high-speed multiple-unit trains and non-metallic blocks tended to polish the

tyres, depriving them of a measure of adhesion. Disc brakes for electric motor coa(were therefore devised as shown in Fig. 13.7.[6] Whilst when stopping from mode speeds (below 20 m/s) performance was inferior to cast iron, an improvement in stop| distance of 40% was achieved when braking from a speed of 44·7 m/s. The risk of w "pick up" remained and the operation of the wheel-slip protection control was show be of critical importance.

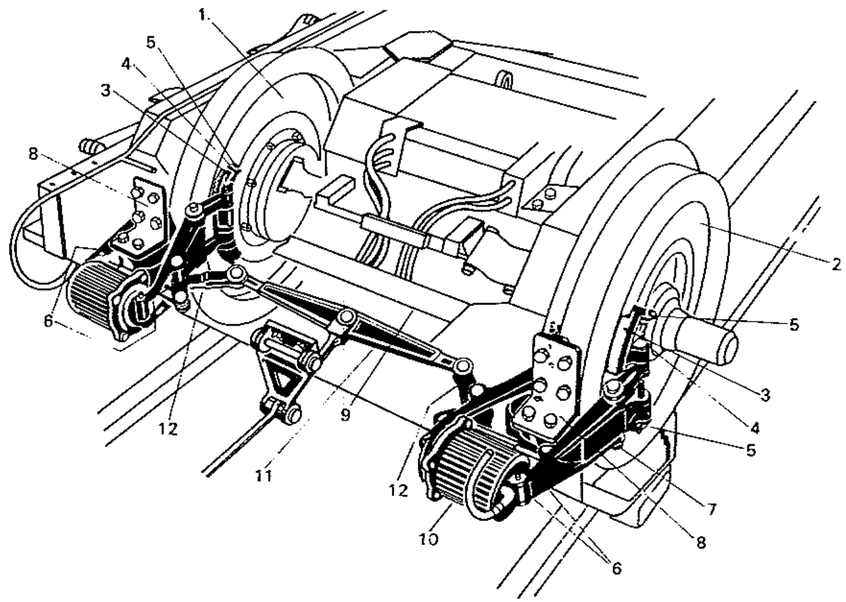


FIG. 13.7. Disc brake system for motor bogie.
1 Inner disc. 2 Outer disc. 3 Pads. 4 Steel shoes. 5 Pad securing clip. 6 Caliper levers. 7 Yoke. 8 Bracket. 9 Headstock of bogie frame. 10 Air cylinders. 11 Equalising beam. 12 Bell crank.

13.7. Servo actuation

The pedal force required to achieve maximum deceleration in automobiles ranges f 200 to 400 newtons. However, sensitivity can be improved by servo systems which am| pedal pressure whilst retaining the possibility of direct actuation in the event of fai of the amplifying device. The source of energy usually employed is the depressioı pressure in the inlet manifold.

Figure 13.8 shows in diagrammatic form a system used on passenger automob Pedal pressure is applied to the left-hand side of piston A. In the absence of any amplif tion effect this would be transmitted to the right-hand side of this piston and direct tc brake actuating cylinders Additional force, however, can be made available by differ(in pressure across piston B which is connected at the right-hand side to the inlet manif In order that the desired amplification shall take place, some control is necessary of

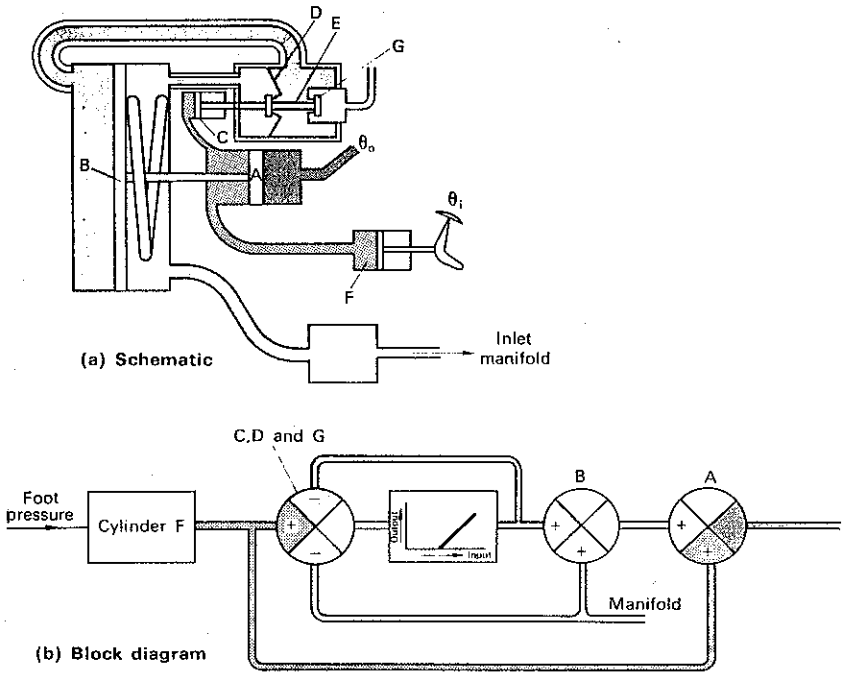


FIG. 13.8. Automatic brake servo system.

pressure on the left-hand face of piston B. Fluid from the main cylinder which corresponds with the pedal force actuated by the driver is admitted to an auxiliary cylinder where it acts on piston C. This is balanced against flexible diaphragm D which is connected on each side to the chambers on the left and right of piston B. Thus the pressure difference across piston B is compared with the pedal-actuating force. When this force is increased relative to this difference in pressure, the spindle E will move to the right, opening valve G which will admit air from the atmosphere to the left-hand side of piston B augmenting the brake actuating force. When equilibrium has been restored the pressure difference across D will balance the oil pressure acting on piston C, spindle F will move to the left, closing valve G.

13.8. The compressed-air brake

The continuous automatic compressed-air brake introduced by George Westinghouse in 1870 ranks second only to the governor of James Watt as a milestone in the development of control engineering. The essential feature of the system was the "Triple" valve which provided a servo action releasing compressed air from a reservoir to the brake cylinders as the result of a falling in pressure of air in the train pipe. This pipe acted as a control line during application of the brakes and at other times as an energy source. Each

vehicle was provided with at least one compressed air reservoir so that energy \
available for applying the brakes in the event of the main source being cut off.

An air compressor is provided on the locomotive which was set to provide air t
main reservoir thereon at a pressure of about 90 lb/in² ($6{\cdot}2 \times 10^2$ Pa). To release
brakes pressure was applied to the train pipe which charged all the subsidiary reserv
distributed between the different vehicles. The triple valve (Fig. 13.9) consisted of a sl

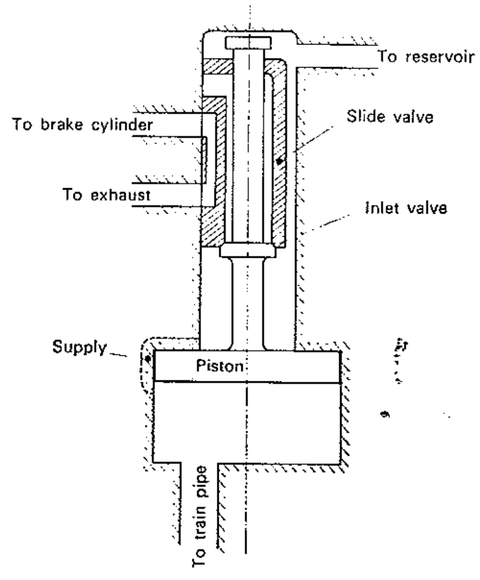


FIG. 13.9. Triple valve.

valve which was driven upwards or downwards by a piston rod. Slight relative moti
was allowed, however, so that a downward movement of the piston opened a feed va
connecting the main chamber to the interior of the slide valve. In the running positi
air pressure in the train pipe forced the piston up to its highest point leaving the fe
valve closed and the brake cylinder connected to the exhaust. The brake itself was h
off at this stage by a spring. Air flowed past the piston via the supply groove entering t
reservoir by way of the main chamber. If there was a reduction in brake pipe pressure t
piston would move downwards, first opening the feed valve and then moving the sli
valve itself. This would first of all cut the brake cylinder away from the exhaust and th
the interior of the slide valve. As a result of charging the cylinder, the reservoir press
would fall until it is balanced by that in the train pipe. The piston would then begin
move upwards again to close the inlet valve. A further reduction in train pipe press
would open the valve again increasing cylinder pressure still further. There were tri
valves on each vehicle on the train so that brakes were applied almost simultaneously
a controlled amount. It is to be noted that intensity of brake application was proportio
to reduction in train pipe pressure. To release the brake the train pipe had to be recharg
with pressure so that the piston moved to its highest point immediately thus releasing

pressure from the brake cylinder. The straight Westinghouse system thus suffered from the disadvantage that a driver once having made an application could not reduce this in intensity; he could only release and reapply.

In order to overcome this difficulty a Swiss Company, Bührle, introduced a valve known as the Oerlikon Control Valve or "Distributor", the principal features of which are sketched in Fig. 13.10. This design embodies tension-free diaphragms and hard rubber flat-seated valves. It is therefore much more easily maintained that the triple valve.

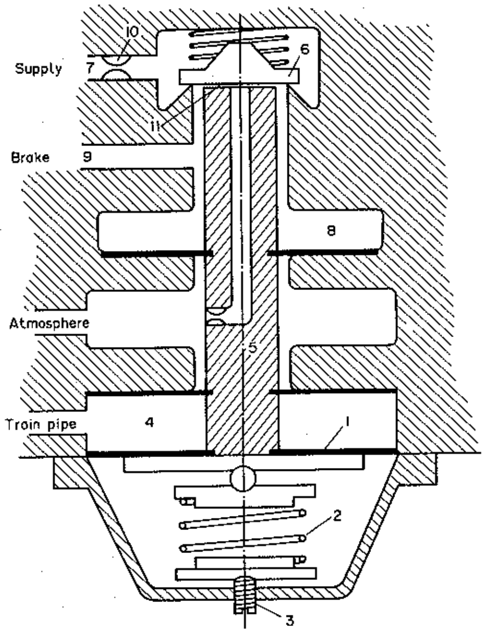


FIG. 13.10. "Distributer" valve, principle of operation.

The method of operation is as follows: When the brake is off, full pressure from the train pipe in chamber 4 acts on diaphragm which is balanced against spring 2, which can be set at the desired value by adjusting screw 3. When pressure in the train pipe and in chamber 4 falls off, the balance is destroyed and spring 2 forces spindle 5 upwards. This lifts valve 6 which admits air from the supply pipe 7 to the brake actuation system through channel 9. Valve 11 is also closed. Pressure now builds up in chamber 8 which gradually overcomes the balance of spring force causing spindle 5 to descend so closing valve 6. Stable operation is secured by the action of choke 10 and an expansion reservoir (not shown). The pressure in pipe 9 is not usually applied direct to the brake cylinder but to a pressure relay which acts as an amplifier. This pressure may also be adjusted automatically to compensate for variation in vehicle loading. The diaphragms and valves are so proportioned that, when pressure in chamber 4 has dropped below the set brake-pipe

pressure by 0·15 MPa (21 lbf/in.^{2}), the maximum braking pressure is applied to t brake-operating cylinder.

Unlike the triple valve, the control valve responds to a limited increase in train pi pressure (after a partial brake application). The increase in pressure in chamber 4 w compress spring 2 and cause spindle 5 to descend which will close valve 6. Further dow ward movement of spindle 5 will open valve 11 so as to connect the brake supply pressu to atmosphere. Thus brake cylinder pressure will fall until a new balance is achieved.

One of the most important features of the system is that, should a train break into tw the train pipe will be ruptured, pressure will fall to zero and the brake will be appli on both parts of the divided train. Any emergency action leading to a rapid reduction train pipe pressure will cause the piston to travel its full stroke which will cause the sli valve to pass completely away from the brake cylinder feed port so that this is put in direct contact with the reservoir giving a full application.

There have been several modifications to the air-brake system including the two-pi arrangement but the main objection to its further development lies in the fact that control action takes place by means of pressure variation in a long pipe. This is bound cause delay between the initiation of braking at the beginning of a train and the corr ponding action at the end. This would be particularly serious were complete automa operation to be applied.[9]

13.9. The electropneumatic (E.P.) brake

Control action is fastest when done electrically and in one of the most extensively us systems, the electropneumatic brake, the original air brake system is retained. The tr pipe is now concerned only with the supply of energy and is kept fully charged in norn operation. The triple valve therefore remains in the release position. Should train p pressure fall due to parting of the train or the driver moving his control to the positi calling for air operation, the triple or distributer valve will resume its normal functi The electropneumatic functioning is obtained by means of a brake valve assem mounted in parallel with the triple valve. An application valve connects the reserv directly with the brake cylinder. This is electromagnetically operated so as to provide immediate response throughout the train. In one system the degree of brake applicat was proportional to the time that the driver held his brake controller in the applicat position. Thus the actual voltage applied to the application valve was not critical. second valve, known as the "holding valve", connected the brake cylinder with exhau This valve was normally open but was closed when brakes were applied, i.e. whene the driver's brake handle is in either the Application or Holding positions. To moder intensity of application, the driver has merely to place his handle in the "Release Running" position for a few seconds before restoring it to the Holding position. T electric train line consisted of five wires whose function is as follows:

1. feed to application valve.
2. feed to holding valve.
3. common return.
4. interlock.
5. positive.

The interlock switch provides a safety feature connecting the brake cylinder to atmosphere in the event of any of the electrical circuits becoming defective. From the control point of view the important characteristic of this form of the E.P. brake was that brake pressure was proportional to the time that the application and holding valves were energised.

In another form of E.P. brake the intensity of the brake application is graduated into seven stages by a combination of three solenoid valves, as presented in Table 13.1.

TABLE 13.1. CODE FOR SEVEN-STAGE E.P. BRAKE

Braking stage	Equivalent piston size		
	Small	Medium	Large
1	+		
2		+	
3	+	+	
4			+
5	+		+
6		+	+
7	+	+	+

+ crosses indicate chamber open to pressure.

Each solenoid valve admits pressure to one of three chambers each fitted with a diaphragm which acts on a common spindle. The effective areas of the chambers differ from each other. The forces acting on the spindle are thus added together and the total force is translated into brake pressure by a proportional brake valve.

The various pneumatic and electrical control functions may be combined into a single control assembly. This will usually contain the three solenoid-controlled valves, the control valve or distributor (which automatically introduces straight air-brake operation in the event of failure of an electrical circuit) as well as the proportional valve, pressure controller and brake operating relay. These may be combined with the electro-dynamic brake so that all braking is responsive to a single control action by the driver. An analogue signal proportional to the degree of braking force produced by the motors acting as generators is fed to the air-brake control assembly which determines the amount of supplementary braking required (if any) and admits the necessary amount of air to the cylinder.

13.10. The vacuum brake

Another form of brake still widely used in the United Kingdom and India is the vacuum brake. This differs from the compressed air brake in that train pipe pressure and cylinder pressure vary in the same sense and not in the reverse sense. Brake cylinders are usually mounted vertically, the brake being kept "off" by the weight of the piston when there is no pressure difference across this. A vacuum is maintained both above and below

the piston in the running position but the connection from the train pipe to the uppe part of the cylinder is provided with a check valve which prevents the flow of air fror the pipe to the cylinder. On air being let into the train pipe, either by the driver or by th parting of the train, this acts on the lower side of the piston causing this to rise and t apply the brakes with an intensity proportional to the absolute pressure in the train pipe

In this simple form, all the air required to enter all the brake cylinders throughout th train had to pass through the driver's valve with the result that there was considerabl time lag between the brakes coming on at the front and back of the train. Baldwin *et al.*[reports severe snatches due to this cause. Dumas of the Great Western Railway therefor introduced in the year 1910, a direct admission (D.A.) valve which acts as a servo device It comprises a diaphragm to which the train pipe is connected below and the brak cylinder above through a choke. On an increase in pressure in the train pipe, the dia phragm rises and admits air from the atmosphere to the brake cylinder until the two ar balanced. Thus the valve acts to speed up the application down the train. The effect of th restriction between the top of the valve and the diaphragm is that if the fall in train pip pressure is very rapid, due to an emergency application for example, the pressure abov the diaphragm would not be able to follow it so as to close the valve. Thus cylinde pressure could be allowed to rise above train pipe pressure instantaneously in order t provide maximum braking effort. Unlike the triple valve, the D.A. valve plays no pai in the release of the brakes, this being effected through a non-return valve connected i parallel therewith. Thus, after a partial application, a driver can obtain a partial releas by increasing vacuum without having to release completely as in the Westinghouse brak

The vacuum brake can, of course, be provided with electric actuation as easily as th Westinghouse but its main disadvantage is the time taken to release the brakes after the have been fully applied. This may be as much as 3 minutes for a long train. This is, c course, determined by the volume to be evacuated and it is possible that this will be muc reduced by the application of disc brakes.

For railcars, where a second pipe can be used to couple all vehicles to the exhauster, reservoir is provided on each vehicle which assists the release of the brakes.

Although more than one major railway system at present operates main-line traffi using the vacuum brake, the modern pneumatic or electropneumatic brake has so man advantages that eventual conversion is to be expected. Thus in 1965 British Rail com menced the conversion of the railway from vacuum brakes to air brakes. It has for som time been the practice for locomotives to be fitted with air brakes and with separat equipment for operating the train brakes on the vacuum system. A dual braking syster was now required, however, and the system shown in Fig. 13.11 has been developed b Messrs. Davies and Metcalfe. The basic control is achieved using the automatic air-brak valve which operates the locomotive brakes and the train brakes when these are ai operated. Where vacuum-braked vehicles are hauled they are controlled through th vacuum train-pipe wherein the intensity of the vacuum is related to the pressure in th main air train-pipe by the air/vacuum relay valve. In the event of the parting of a trai fitted with the vacuum brake, the loss of vacuum would be translated into an equivaler pressure change through the vacuum/air relay valve. The direct-air brake supplies ai directly to the locomotive brakes.

A comprehensive account of current British practice of railway braking is presente by Broadbent.[10]

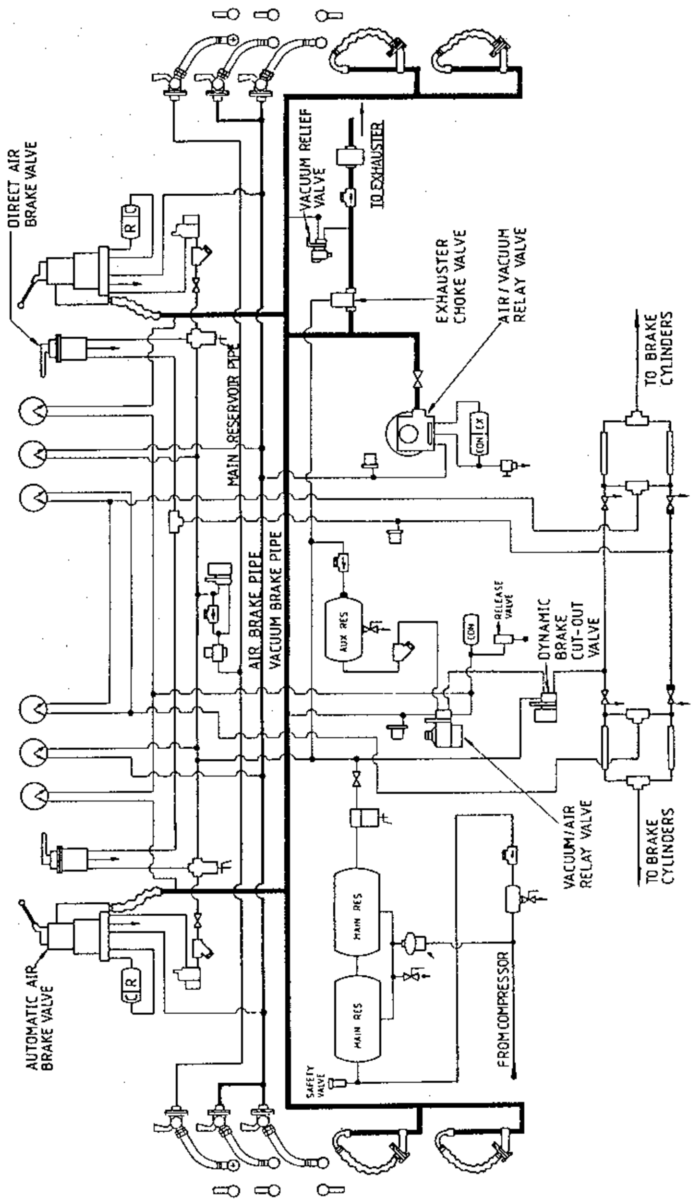


FIG. 13.11. Metcalfe-Oerlikon dual air/vacuum automatic brake equipment.

References

1. BERNDT, P., KROGER, U. and SAUMEBBER, E., The principles of operation of track brakes (Magr and Eddy Current)—Interaction with the track. *Institution of Mechanical Engineers Confer Publication* No. 11, 1979, p. 229.
2. Electric Railway Presidents' Conference Committee, Bulletin No. 3, 1932.
3. WILSON, E. E., Deceleration distance for high-speed vehicles. *Proc. Highw. Res. Bd., Wash.*
4. *Research on Road Safety*, H.M.S.O.
5. MADDISON, T. B., The development of vacuum-operated disc brakes for freight trains. *Proc. In Mech. Engrs. Convention on Railway Braking*, 1962, p. 37.
6. SYKES, W. I. A., Disc brakes for high-speed railway rolling stock. *Proc. Instn. Mech. Engrs. Conven on Railway Braking*, 1962, p. 28.
7. BALDWIN, T., PEACOCK, D. W. and SCALES, B. T., Problems arising with continuously braked frei trains. *Proc. Instn. Mech. Engrs. Convention on Railway Braking*, 1962, p. 12.
8. WISE, S. and LEWIS, G. R., Composition brake blocks and tyres, Institution of Mechanical Engine Railway Division Preprint, 1970.
9. BUHLER, The dynamic behaviour of the air-brake and its simulation by analogue computers, *Bull. Railway Cong. Assoc.*, vol. IV, p. 59 (1967).
10. BROADBENT, H. R., *An Introduction to Railway Braking*, Chapman and Hall, London, 1969.

CHAPTER 14

Steering — Directional Stability

14.1. Steering

Next to braking, the most important aspect of control in transport is "steering", whether in the narrow sense of a ship's response to the movement of the helm or in the broader sense, a navigator's plotting of a course. In the narrow sense, which will be dealt with here, there are two aspects to steering, one the maintenance of a vehicle on a set course and two, the action of changing that course. Taking a railway axle running on a straight track as an example (Fig. 14.1), the wheels are provided with coned tyres and

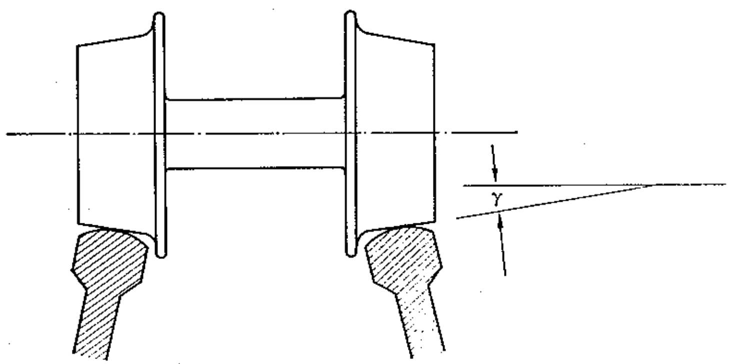


FIG. 14.1. Coning of railway wheels.

with flanges, the latter being so arranged that a substantial clearance exists between them and the gauge face of the rail. If the axle is placed centrally on a straight track with its axis at right angles to the rails it will continue to roll in a straight line. If, however, it is assumed to be displaced laterally to the left, the diameter of the contact circle will be greater on the left and lower on the right. It will therefore tend to move faster on the left than on the right and the axle will consequently become inclined so that movement of the axle will occur towards the right. Thus there exists a self-correcting action. When the centre point of the axle becomes coincident with the centre line of the track, the axis will be inclined to the perpendicular to the rails and thus it will overshoot and repeat the same sequence of movements but with displacement to the right. Thus an axle will follow a sinusoidal path along the track.

It is usual to provide more than two axles on a vehicle and to mount them in pairs subsidiary trucks known as bogies. Often these comprise a further elastic element kno as a "bolster" between the bogie frame proper and the body frame. This may be connect to the bogie frame by vertical swinging links so that the vehicle body is suspended though on a pendulum having a period of about 1·5 seconds, thus providing a furtl degree of freedom and permitting further possible interaction between "steering" a "suspension". The number of possible modes of energy exchange may become very lar and difficult to treat mathematically. However, it can be reported from experience th bogies sometimes exhibit instability, hunting violently from side to side.

Vertical control or suspension is not normally regarded as a branch of control enginee ing but it is of great importance with regard to steering because of the interchange energy between the vertical and horizontal modes of oscillation.

Directional stability implies the ability of a vehicle to continue on a defined path wi only a slight deviation irrespective of the magnitude of disturbing influences. Tl deflection and rolling characteristics of the tyres will therefore determine the behavio of the vehicle.

Suppose a pneumatic tyre is rolling in the direction XX^1 as shown in Fig. 14.2 a subjected to a force F acting in the YY^1 direction, this will cause slip in the Y directi so that the actual path of the tyre will be represented by the dotted line which is inclin to the XX^1 axis at an angle θ. This is known as the slip angle.

The relationship between θ and F depends on the construction of the tyre, degree inflation, applied load, etc.,[1, 2] but for values of θ below 10° a linear approximati may be made so that

$$F = K\theta. \qquad (14.$$

Because of the hysteresis of the material, tractive forces, etc., the reaction to F, indicat as F_1 in Fig. 14.2, will act through a point displaced behind the hub by an amount known as the "pneumatic trail". A torque is therefore set up with magnitude Fq whi tends to cause the wheel to swivel into the plane of motion. When θ is deliberately impos

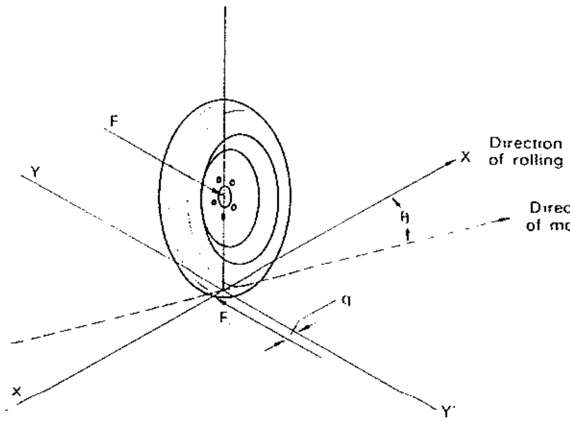


FIG. 14.2. Definition of slip angle and pneumatic trail.

so as to steer the vehicle into a curve, the sum of Fq over all "steerable" wheels represents the torque which must be applied to steering gear.

A four-wheeled vehicle steered by the front wheels only will change its heading as the result of the lateral force F acting on the mass of the vehicle to overcome its lateral and rotational inertias. The rate of acceleration and therefore the magnitude of F will depend on the linear velocity of the vehicle. If the wheels are directed at an angle Φ to the central axis of the vehicle their actual motion will take place in the direction $\Phi - \theta$ depending on the magnitude of F, therefore vehicles tend to understeer at speed as indicated in Fig. 14.3.

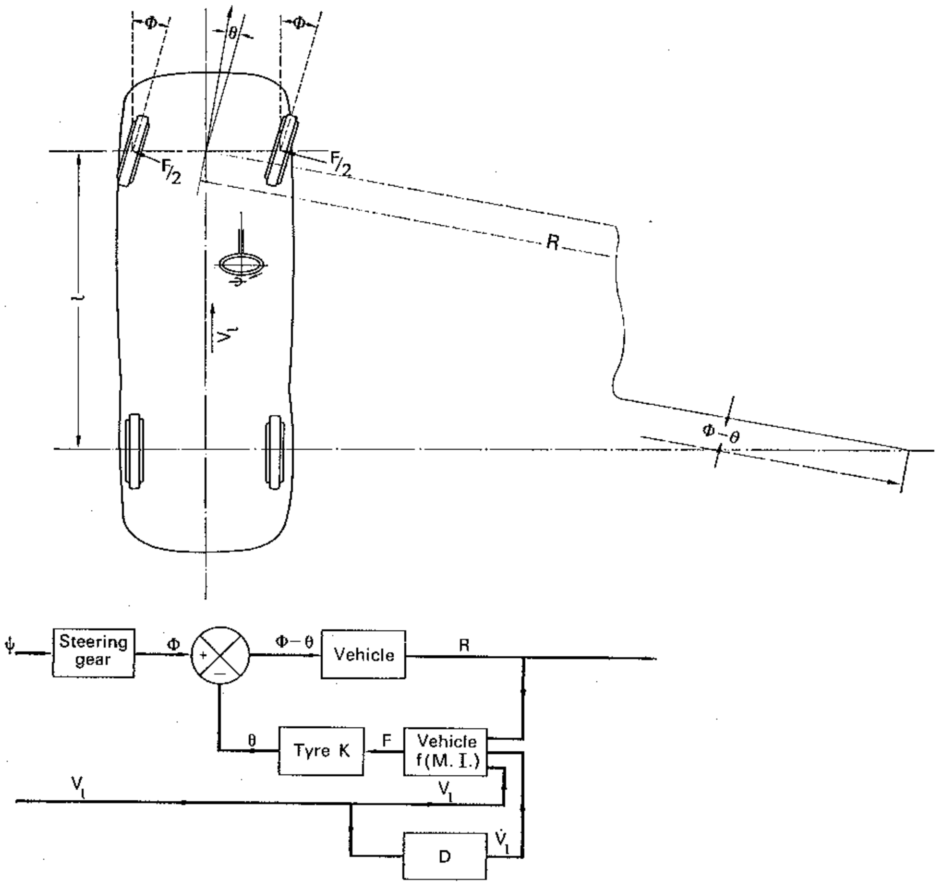


FIG. 14.3. Diagram illustrating understeer.

The control is of the "integrating" type in so far as the rate of change of the heading of the vehicle is proportional to the input, i.e. angle of the steering wheel, multiplied by distance travelled.

The magnitude Fq may become so great at speed that it becomes difficult for a driver to exert sufficient control over the magnitude of angle. Power assistance is therefore

available on some vehicles which amplifies a driver's responses without introducing an significant change in the transfer operator connecting steering-wheel position and vehicl behaviour and above all avoiding any destabilising feature. This is achieved by introduc ing a position servo which has the characteristic that rate of change is proportional t error. The principle is illustrated in Fig. 14.4.

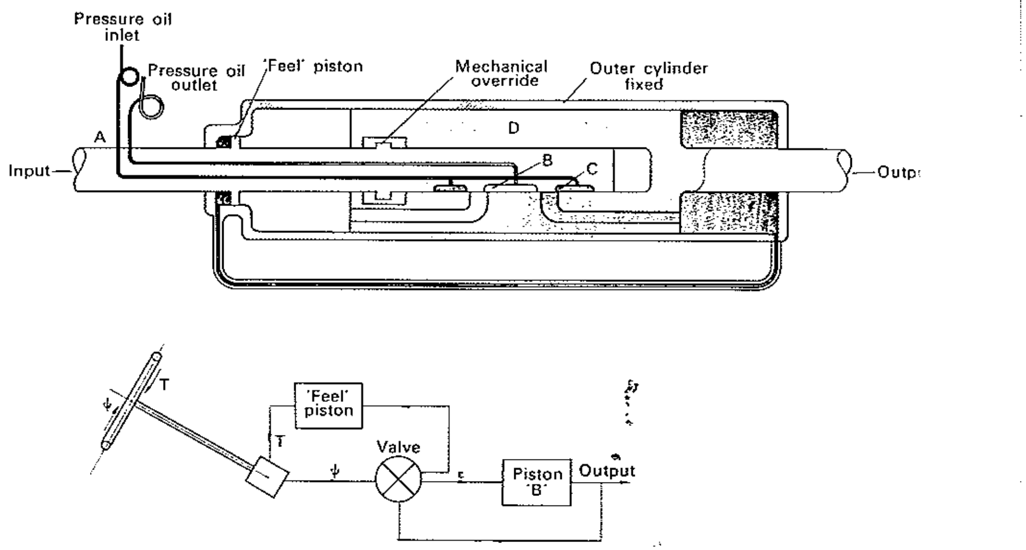


FIG. 14.4. Principle of power-assisted steering.

Movement of spindle A causes ports B and C to open allowing access and egress pressurised oil to the main cylinder. As piston D responds to this pressure the moveme tends to close the ports. Assuming that the resistance to motion is small compared wi the dimensions of the power device, the rate of movement of the piston is proportional the opening of the valves. Thus

$$\frac{dx}{dt} = K\epsilon. \qquad (14.$$

The arrangement of the system is such that, in the event of failure of the hydraul apparatus, the integrity of the manual system is maintained.

14.2. Directional stability

Consider a vehicle such as is shown in Fig. 14.5 which, due to some transient influenc a gust of wind for example, has been caused to take up a position with its centre li inclined to the direction of motion by an angle θ. Setting aside any action by the driv in adjusting the steering gear, a stable situation would be one in which, as time went o θ became reduced to zero. Instability would connote an increase in θ leading ultimate to disaster.

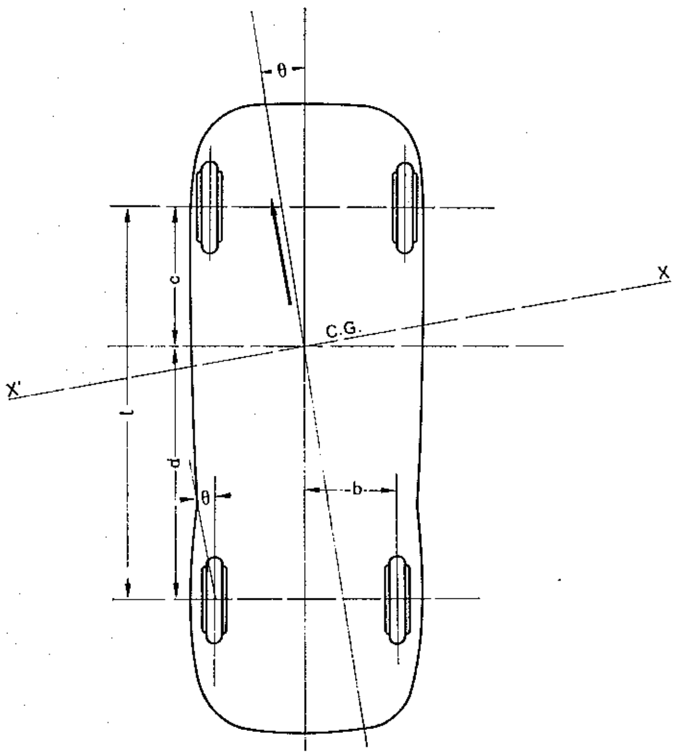


FIG. 14.5. Notation for directional stability.

Let the origin of the axes of references be located at the centre of gravity of the vehicle O. Let the gauge be equal back and front and of value $2b$ and let the wheel base be $c + d$ where c is the distance of the leading axle from the centre of gravity. Let the tyre characteristic of the leading axle be K_L and that of the tyres on the trailing axle K_T.

The moment about the centre of gravity will be

$$\sum Fy = \sum K\theta y = \theta\{K_L(y_1 + y_2) + K_T(y_3 + y_4)\} \qquad (14.3)$$

where

$$y_1 = c \cos\theta + b \sin\theta, \qquad (14.4)$$

$$y_2 = c \cos\theta - b \sin\theta, \qquad (14.5)$$

$$y_3 = -d \cos\theta + b \sin\theta, \qquad (14.6)$$

$$y_4 = -d \cos\theta - b \sin\theta. \qquad (14.7)$$

Thus clockwise moments which would tend to increase θ and would therefore lead to instability are as follows:

$$2\theta\{K_L c \cos\theta - K_T d \cos\theta\} \qquad (14.8)$$

or

$$2\theta \cos\theta \{K_L c - K_T d\}. \qquad (14.9)$$

Bradley and Wood[3] showed that during braking it was essential that the rear wl should not become locked if loss of directional control was to be avoided.

The magnitude of the lateral force will be $2(K_L + K_T)\theta$. If the product of K_L and less than $K_T \times d$ the system will be stable at all speeds. Thus for identical tyres front back, the centre of gravity should be forward of the geometrical centre of the wheel l The condition and degree of inflation of the rear tyres is therefore more important is the case of those at the front.

This account is oversimplified because it neglects the rotational inertia of the ve about the vertical axis. If this is taken into account, stable operation is possible bel critical speed given by

$$V = \sqrt{\frac{2K_L K_T}{M(K_L c + K_T b)}} \qquad (1\cdot$$

where M is the mass of the vehicle.

For further information on directional stability the reader is referred to account Rocard[4] and to Segal.[5]

The quality of adhesion between wheel and road is therefore even more importan directional control than it is for driving and braking. The value of the coefficier friction is not sensitive to the direction of the force applied and the laws governin magnitude as a function of velocity of relative sliding, previously discussed in relatic propulsion, also apply to steering. It will be recalled that the most important charact tic was that friction increased with relative sliding until a maximum was reached, ; which there was a marked falling off to very low values. In the case of pneumatic t adhesion depends on the formulation of the rubber used to form the tread, the t pattern, the material and texture of the road surface and the nature of any contamina material on that surface.

Under favourable conditions on clean dry roads, coefficients of 0·8 can be achieved. under adverse conditions, when considerable water is present, the value may fal practically zero. This is due to the hydrodynamic pressure of water trapped between and road. Figure 14.6 from ref. 6 shows three zones of action. The mechanism of pres generation has generally been regarded as being based on Reynolds-type hydrodyna films but an alternative view has been published by Wallace and Trollope.[7]

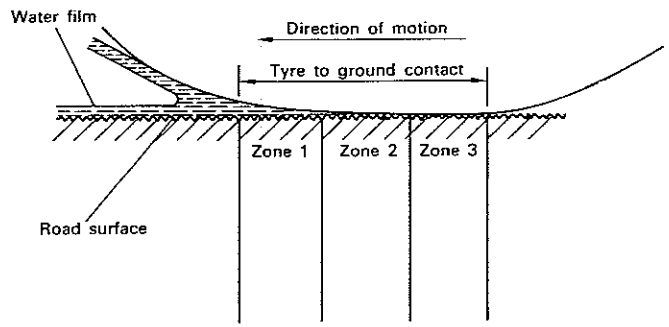


FIG. 14.6. Schematic representation of "three-zone" concept.

The extreme case of loss of adhesion is known as "aqua-planing" and the development of tread patterns to facilitate the escape of water enables some improvement to be achieved. There appears to be no complete solution and the only way of achieving safe operation under flooded road conditions remains to reduce speed drastically.

14.3. Hertzian contact

Whilst the action of a steel tyre on a steel wheel is not qualitatively dissimilar to that shown in Fig. 14.1 it is quantitatively very different. Contact conditions are governed by the laws of Hertz which predict an ellipse of contact having semi-axes a and b whose lengths are determined by

$$\left[\left(\frac{1-\nu_1^2}{E_1}+\frac{1-\nu_2^2}{E_2}\right)\frac{3}{2}\frac{W}{1/r_{11}+1/r_{12}+1/r_{21}+1/r_{22}}\right]^{\frac{1}{3}} \quad (14.11)$$

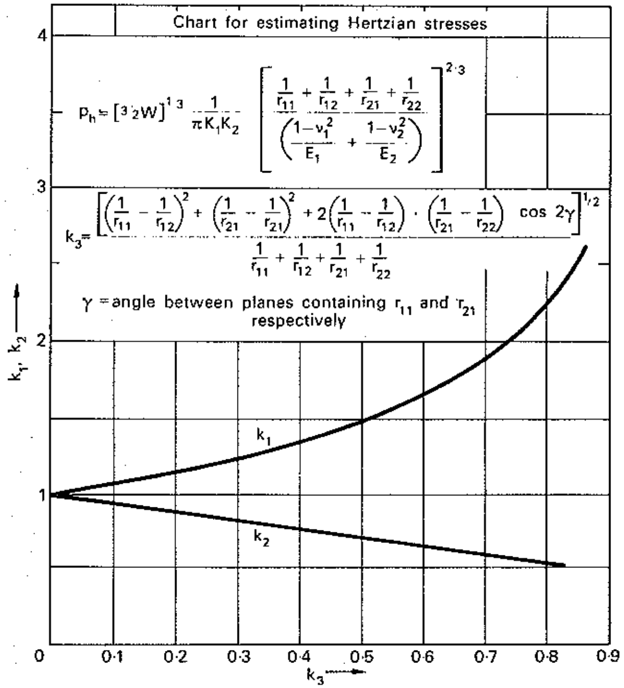


FIG. 14.7. Constants for calculating Hertzian stress.

multiplied by a constant K_1 and K_2 respectively which can be derived from Fig. 14.7.
ν_1 and ν_2 are the Poissons ratio of the wheel and tyre materials respectively,
E_1 and E_2 are the Youngs moduli,
W is the applied load,

and r_{11} is the radius of the upper body in radial plane,
r_{12} is the radius of the upper body in the axial plane (can be taken as infinity),
r_{21} is the radius of the lower body in the radial plane (in the case of the rail thi infinity so that $1/r_{21}$ is zero),
r_{22} is the radius of the lower body in the axial plane (i.e. radius of camber of rail he

For crossed cylinders of identical material equation (14.11) simplifies to

$$K\left[\frac{3W}{E}(1-\nu^2)\frac{1}{1/r_{11}+1/r_{21}}\right]^{\frac{1}{3}}. \quad (14.$$

Stress will be distributed over an ellipse of contact so as to be represented by ordinates of a semi-ellipsoid constructed on that surface. The maximum pressure compressive stress will be 1·5 times the average pressure or $1{\cdot}5W/\pi ab$.

When a tangential force is applied, elastic deformation of the two surfaces will oc in such a way that a certain amount of slip will be unavoidable. This will naturally oc near the periphery of the ellipse of contact where the normal pressure is lowest. As tangential force is increased this area is progressively reduced in size until when tangential force reaches its limiting value, sliding occurs.

When the system is in motion, the slip area occurs at the edge of the contact zone whi in association with the elastic deformation of the adhesion area, leads to a difference the peripheral speed of the wheel and the rate of progression of the vehicle along track. In the case of traction the wheel will travel in one complete revolution a dista along the rail which is slightly less than the circumference of the wheel. In the case braking it will travel further. The difference between the actual distance travelled and actual motion of the circumference of the wheel is known as the "creep ratio". Cree numerically equal to the value of relative sliding as follows:

$$\frac{\text{Speed of sliding}}{\text{Peripheral speed} - \text{sliding speed}} \quad \text{for traction}$$

and

$$\frac{\text{Speed of sliding}}{\text{Peripheral speed} + \text{sliding speed}} \quad \text{for braking.}$$

Carter[8, 9] gave an expression for creep which is equivalent to

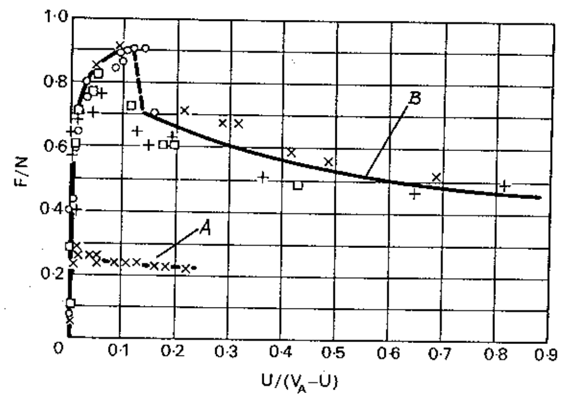
$$\xi = \mu\sqrt{\frac{4P}{G}(1-\nu)\left(\frac{1}{r_{11}}+\frac{1}{r_{22}}\right)} \times (1-\sqrt{1-F/F_{\max}}) \quad (14.$$

where ν = Poissons ratio,
G = modulus of rigidity,
P = normal load per unit breadth of cylinders,
F = tangential force and $F_{\max}$ its maximum value,
μ = coefficient of friction corresponding to bulk sliding between the surfaces

An approximation due to Carter is as follows

f = tractive force per unit creepage (14.
= 3500 $\sqrt{\text{Radius of Wheel (inches)} \times \text{Load on Axle (pounds force)}}$.

Figure 14.8 shows some actual values of creep determined at the Rugby testing station of British Railways and Fig. 14.9 actual imprints of railway vehicles determined by Andrews.[10]



A, from test plant.
B, from laboratory rig.

FIG. 14.8. Creep values.

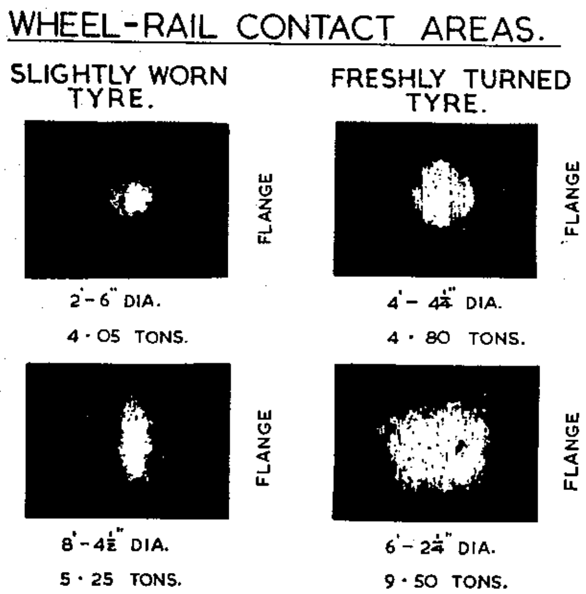


FIG. 14.9. Imprints of railway wheels (after Andrews) 0·54 × full size.

14.4. Running of coned wheels

The importance of creep to the railway control engineer is two-fold. Firstly, in acceler tion and braking it affords an opportunity of regulating applied torque within the value c the coefficient of friction so avoiding gross sliding. Secondly, it is important in connectio with directional stability in so far that it permits stable running when the axle is not a right angles to the track. It must be realised that the direction of creep is determined b the direction of applied force rather than by the direction of motion.

Some railway vehicles have cylindrical tyres but the vast majority have coned wheel as described previously, so as to support an inherent steering action. Let r = radius c wheel, $2b$ = distance between centres of contact ellipses and γ = angle of conin expressed as a vulgar fraction, i.e. 1/20. Let y = the lateral displacement of the centre c the axle from the centre of the track. Then the effective radii of the right- and left-han wheels are respectively $r - \gamma y$ and $r + \gamma y$. Then for an average rate of progression an neglecting creep, the axle of the right-hand wheel will roll $(r - \gamma y)V/r$ and the left han will roll $(r + \gamma y)V/r$. The left hand will therefore gain on the right at a rate $(2 - \gamma y)V$ so that the axle slews with an angular velocity

$$\frac{d\theta}{dt} = \gamma y V/rb$$

but $$V = \frac{dx}{dt}, \quad \therefore \frac{d\theta}{dx} = \frac{2\gamma y}{rb} \qquad (14.1$$

but $dy/dx = \tan\theta$ which for small angles equals θ,

therefore $$\frac{d^2y}{dx^2} = \frac{d\theta}{dx} = 2\gamma y/rb, \qquad (14.1$$

$$\ddot{y} - 2\gamma y/rb = 0, \qquad (14.1$$

$$y = A \sin\sqrt{\frac{\gamma x}{rb}} \qquad (14.1$$

where A is the maximum amplitude determined by the initial conditions. The wavelengt of the oscillatory path will be $2\pi\sqrt{rb/\gamma}$. Thus for a 42-inch wheel (1·165 metres) on 4-fe 8½-inch track (say 59 inches or 1·5 metres contact spacing and 1/20 coning) we hav $2\pi\sqrt{rb/\gamma}$ as 19·2 metres. Thus at 100 m.p.h. (44·7 m/s) the frequency of disturbanc would be 2·3 Hz.

However, whilst adequate for low speeds, this treatment neglects the effect of the ma of the axle. The acceleration both rotationally and laterally of this mass will give rise t forces at the wheel periphery which will cause creep at the interface. This will have slight tendency to modify the trajectory of the bogie but more important, will introduc an element of speed dependence to the system.

Should the axle take up a radial position when traversing a curve radius R, good steerir will occur when the axle displacement of the centre line is such that the effective radii c the coned contact paths is equal to the ratio of the radii of the inner and outer track $br = R\gamma y$. Thus if lateral displacement is limited to ¼ inch (0·00635 metre) the minimu radius for smooth running will be 1210 metres for the 42-inch wheel-set considered abov

14.5. Inscription within sharp curves—steering by flanges

Given adequate radius of curvature and degree of superelevation, high-speed running will not involve continuous flange contact and serious wear need not occur. However, when vehicles are required to traverse sharp curves, flange force becomes of great importance.

Consider the four-wheeled vehicle shown in Fig. 14.10 and neglecting creep and wheel tread conicity, the flange of the outside leading wheel will impinge on the side of the rail

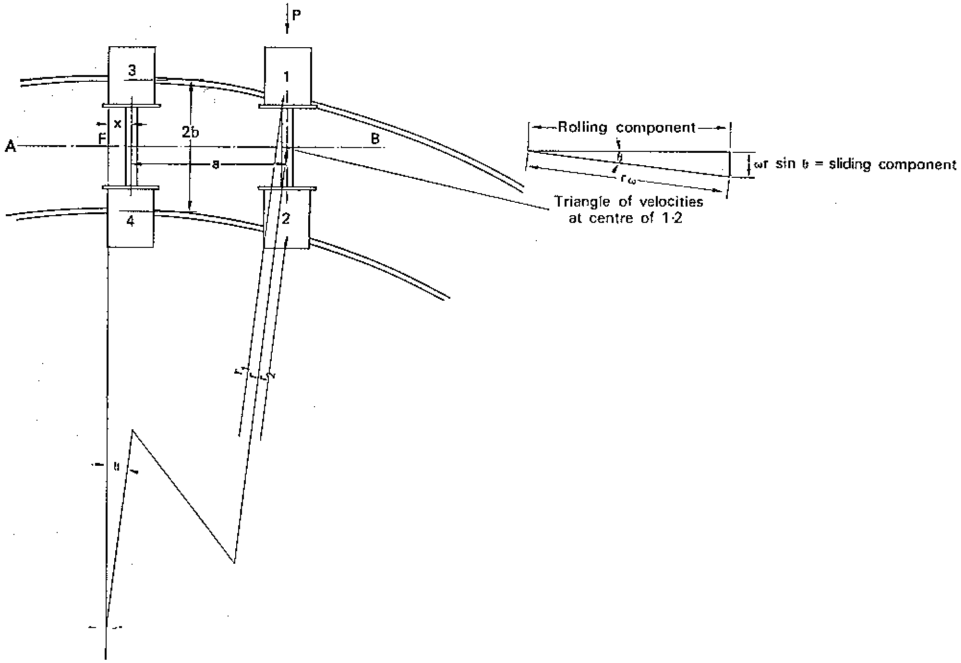


FIG. 14.10. Notation for calculation of flange forces.

and the trailing axle will tend to take up a radial position. Under steady state conditions the vehicle can be regarded as rotating about the centre of the curved track or instantaneous centre with a constant angular velocity ω. The linear velocity of any part of the vehicle will be ωr and will be directed at right angles to the line joining the part with the centre. It is desired to know the attitude taken up by the centre line of the vehicle. Suppose this to be AB and drop a perpendicular from the instantaneous centre. The distance from the centre of the rear axle will be denoted by x.

The relative motion of each wheel centre will be made up of two components, one due to rolling action and directed along a line parallel to AB and one, of as yet undefined direction, due to sliding of the wheel surface over the rail. It is assumed that each wheel is loaded equally by force W and that the coefficient of friction is constant at value μ.

Taking axle 1, 2, the velocity at the centre will be made up of a rolling action and a sliding component operating at right angles to AB. However, the actual velocity of 1 will

be greater than 2 because of the increased value of r. Assuming, therefore, that b wheel treads slip by an equal amount, the component due to rolling will be $\omega r \cos\theta$ a that due to sliding $\omega r \sin\theta$.

At wheel 1 the triangle of velocities will consist of a hypotenuse tangential to the r a rolling component equal to $\omega r \cos\theta$ and a sliding component which can be shown be normal to 1 F. Similarly the sliding velocities at the wheel treads can be shown to directed at right angles to a line joining them to a point F. In the absence of braking a traction forces, point F can be shown to lie on the centre line of the vehicle.[11]

The direction of action of friction forces is directly opposed to that of their relat motion so that the sliding component of the velocity vector diagram indicates the dir tion of the friction force. Their magnitude is known as equal to μW where W = load each wheel. It is therefore possible to evaluate the steering or flange force by tak moments about the intersection of their normals and by summing their compone acting parallel to the flange force.

It is customary to consider the flange force to act independently of the tread forces a this is represented by P. This force is responsible for overcoming the friction forces at the treads and can be evaluated as follows:

Taking moments about F

$$P(x+a) = \mu W_1\sqrt{(x+a)^2+b^2} + \mu W_2\sqrt{(x+a)^2+b^2} + \mu W_3\sqrt{x^2+b^2} + \mu W_4\sqrt{x^2+b^2} \qquad (14.$$

$$= 2\mu W\left(\sqrt{(x+a)^2+b^2} + \sqrt{x^2+b^2}\right) \text{ when } W_1 = W_2 = W_3 = W_4 =$$

Equating transverse forces

$$P = \mu W_1\frac{x+a}{\sqrt{(x+a)^2+b^2}} + \mu W_2\frac{x+a}{\sqrt{(x+a)^2+b^2}} + \mu W_3\frac{x}{\sqrt{x^2+b^2}} + \mu W_4\frac{x}{\sqrt{x^2+b^2}}$$

$$= 2\mu W\left(\frac{x+a}{\sqrt{(x+a)^2+b^2}} + \frac{x}{\sqrt{x^2+b^2}}\right). \qquad (14.$$

In the case of a four-wheeled high-speed freight wagon, a equals 4 metres whereas b c be taken as about 0·75 metre and an approximate solution due to Jenkins[10] is

$$x = \frac{b^2}{a}$$

and in this case F will lie 0·14 metre behind the rear axle. Therefore, from equat (14.20) the flange force will be

$$\mu W\left\{\frac{2\sqrt{(x^2+a^2)}+b^2+\sqrt{x^2+b^2}}{x+a}\right\}$$

or

$$P = 2{\cdot}86\mu W. \qquad (14.$$

Thus, with a coefficient of friction of 0·25 a lateral force of some 7 tons can be exer by a 20-ton four-wheeled vehicle. It must be noted that this arises from friction for only, neither radius of curvature nor speed being taken into account. Such a treatment can be elaborated but there can be little justification for this in view of the drastic assumptions made, notably that coefficient of friction is independent of velocity of sliding.

Vehicle designers treat more complex cases by various methods[12, 13] and notably by the Heumann diagram.[14–15] These are static treatments of an essentially dynamic situation and their main use lies in estimating the forces applied to the vehicle structure during movement at slow speeds over curves which are in any case too sharp for high-speed working.

14.6. Vehicle ride quality

Axles are seldom attached to vehicles with so little restraint as to permit oscillation acccording to equation (14.18) although the situation might be approximated to in the case of a badly worn vehicle. Modern locomotives and passenger and goods vehicles now tend towards the BB arrangement, i.e. with two-axle trucks each pivoted at the centre. Because the pivot-point is at the centre, the trucks are in a state of neutral equilibrium, i.e. neither stable nor unstable. If the pivot could be placed ahead of the mid-point between the axles, stability would result, but because most railway vehicles are required to operate in two directions, this would have the opposite effect when the vehicle was running in the opposite direction.

The case where two axles are constrained by the vehicle to remain mutually parallel is therefore of considerable importance.

Reverting to Fig. 14.5, c is now equal to d and $K_L = K_T$ = creepage coefficient ξ/F. Again, consider the pivot as the origin and let the bogie be inclined at an angle θ to the centre line of the track. Then because $c = d$ and $K_L = K_T$ there is no moment tending either to increase or decrease θ. The whole vehicle is, however, moving laterally, i.e. crabwise, at velocity

$$\theta\frac{dx}{dt} = \frac{dy}{dt}.$$

Consider one axle inclined at an angle θ to the centre line of the track but constrained to move mainly in the longitudinal direction but with some provision for lateral movement. Firstly, if its axis is at right angles to a straight track but its centre displaced an amount z towards the right then, if free to move on the axle, the right-hand wheel would rotate at $V/2\pi(r + \gamma z)$ and the left-hand wheel $V/2\pi(r - \gamma z)$, but with a solid axle, both wheels are required to roll at the same speed. If the difference is large, relative sliding will occur often, made obvious by a high pitched screeching noise on a dry day but for small values of z this difference is accounted for by creep as though one wheel were driving and the other braking. Thus forward-creep ratio of one wheel equals $\gamma z/r$ which is numerically equal to the backward-creep ratio of the other wheel. The force exerted at each wheel is $\gamma z/r\, f(\xi)$ giving rise to

$$2\frac{\gamma z}{r} f(\xi) b.$$

Similarly lateral motion at the rate dz/dt will give rise to creep ratio $(dz/dt)V$ as averaged between the two axles.

Although the expression for creep is non-linear in form, a linear approximation justified for small movements and hence we may write $\xi = KF$. Thus for an axle norı to the direction of motion but with lateral displacement z, turning moment equals

$$b\frac{\xi}{K} = 2\frac{b}{r}\gamma\frac{z}{K}. \qquad (14.$$

Let the distance from each wheel contact point to the centre, namely $\sqrt{b^2 + c^2}$, be writ l. Then when the angular displacement varies, turning moment equals

$$\frac{d\theta}{dt}\frac{1}{V}\frac{1}{K}l^2. \qquad (14.$$

Therefore equalising moments we have

$$4\frac{d\theta}{dt}\frac{l^2}{KV} - 4\frac{b}{r}\gamma\frac{z}{K} = 0, \qquad (14.$$

but equating velocity of lateral movement, $\theta V = dz/dt$,

$$\frac{d^2z}{dt^2} + \left(\frac{l}{V}\right)^2 - \frac{b\gamma z}{r} = 0, \qquad (14.$$

but $V = dx/dt$,

$$\therefore \frac{d^2z}{dx^2} = l^2 - \frac{b}{r}\gamma z = 0, \qquad (14.$$

$$z = z_{max} \sin\sqrt{\frac{b\gamma}{rl^2}}\,x \qquad (14.$$

recalling that $l^2 = b^2 + c^2$

$$z = z_{max} \sin\sqrt{\frac{b\gamma}{r(b^2 + c^2)}}\,x \qquad (14.$$

and the wavelength of the sinusoidal bogie path

$$2\pi\sqrt{\frac{(b^2 + c^2)r}{\gamma b}}. \qquad (14.$$

Thus for a standard gauge vehicle with 3 feet 6 inches (1·07 metres) diameter wheels a 8 feet 6 inches (2·6 metres) bogie wheel base the wavelength would be 116 feet (35 metı and the frequency

$$\frac{\text{miles/hour}}{79}\ \text{Hz or}\ \frac{\text{metres/second}}{35}\ \text{Hz}$$

Thus, depending on the initial conditions, the bogie pivot will oscillate from side to s at a frequency corresponding to the speed of the train. At some speed this could coinc with some characteristic natural frequency of the suspension system so that oscillat may be forced.

The treatment applies strictly only to the circumstance when the vehicle moves along the track at negligible speed. As soon as speed becomes significant, inertia forces will introduce terms of the form $d^2\theta/dt^2$ into equation (14.22) which will alter the wavelength. Other terms should also be included to represent pivot friction. Inequalities of loading between wheels will also affect equilibrium. The material used to construct the bogie frame will be deformable so that the axles may not always be perfectly parallel. The most important qualification, however, is that relating to tyre profile which, in the worn condition, ceases to be truly conical and can have the effect of increasing the angle of coning during part of the cycle.

The relationship between force and displacement may be non-linear and certainly when bulk sliding occurs under clean, dry conditions, the coefficient of friction may fall with velocity. The fact that silent, smooth running is more likely to occur in damp weather is adduced as supporting evidence of the importance of the force–displacement relationship occurring in the contact zone between wheel and rail.

Bishop(17) has carried out model experiments in which he demonstrates that, whilst a bogie may be quite stable at low speeds, it may become unstable at a definite critical speed. The application of damping of rotation about the bogie pivot had no effect on the critical speed but the introduction of springs whose action tended to restore the bogie to its central position led to considerable improvement.

The quantitative treatment of the control of vehicle oscillation is still the subject of much research and notable advances have been made by Wickens and his co-workers of British Railways.(18–22) Using a linearised expression for creep forces they set up equations of motion for a wheelset taking into account longitudinal and lateral creep forces and spin creepage. Adding inertia terms and coefficients governing the suspension displacements and velocities to quantify the forces acting between wheelsets and body, the equations governing the lateral motion of a complete two-axle vehicle were set down. These comprise seven homogeneous differential equations. Examination of the matrices of these equations shows from their asymmetry that the lateral dynamic behaviour of railway vehicles can be non-conservative. Therefore the assumption made earlier that bogie hunting was explained simply by the coincidence of the frequency of lateral excursions of the axle with the natural frequency of some mode of vibration of the vehicle structure was an oversimplification. The linearised equations indicate the critical speed at which instability will set in. This is not, however, related to a natural frequency as in a forced vibration when it might be expected that a further increase in speed would remove the forcing frequency from the natural frequency so as to bring about a resumption of quiet running. It is simply a boundary below which running is stable and above which it is unstable. The actual motion is a non-linear oscillation extracting energy from the source propelling the vehicle to feed lateral oscillations of amplitude, which increases until a limit cycle is set up which will persist as long as operation is attempted in the unstable region.

The value of the linearised treatment is therefore to set the boundaries within which stable operation is possible and to identify constructional parameters which may be modified at the design stage to increase the range of stable running. Some factors such as effective conicity which depends on rail profile as much as on tyre shape, creep factors and load variations are beyond the control of the designer and suspension parameters such as lateral and yaw stiffnesses and lateral damping must be chosen to ensure stability.

14.7. Motion on curves at speed

Motion on curves is of course accompanied by centrifugal force which is usı compensated for by some form of banking. In the case of the railway the outer r; raised above the inner by an amount known as superelevation. For perfect balan(gravitational and centrifugal forces, the slope amounts to tan $^{-1}V^2/gr = \phi$ radians (metres difference in height for standard gauge,

$$0{\cdot}146V^2/R \qquad (1\cdot$$

where V is in m/s and R is in metres. The maximum value of superelevation us(Britain is 0·15 m but it is customary to permit trains to operate at speeds whicl equivalent to a superelevation of 0·26 m (where rails are continuously welded). In (words the permitted "cant deficiency" is 0·11 m.

In considering higher speeds which may be required in the future, centrifugal force impose severe limitation on the use of existing rights of way. If we assume sm transitions and complete compensation, the passenger will not be aware of any la force but simply that the force of gravity will appear to be increased by a factor e to sec ϕ. If we assume that he can tolerate an increase in apparent weight of 20% th

$$V_{max} = 2{\cdot}55\,\sqrt{R}\ \text{m/s.} \qquad (1$$

This is approximately twice the speed which could be obtained on existing track normal cant deficiency. The angle of superelevation would be 33·5° thus necessit unconventional construction.

14.8. Radial-type bogies (trucks)

As demonstrated in Section 14.5 a bogie, when acted upon by the frictional f(between wheel and rail, will set itself in such a position relative to the track that siderable force will be required to act between the outer flange of the leading wheel: the track. This frequently leads to excessive wear and can cause derailment.(23–26)

When designing a bogie there is a conflict between the factors conducive to maxi: stability on straight track and those leading to minimum force on curves. For high-s passenger trains consideration of passenger comfort limits the degree of track curv: to that which can be negotiated by a conventional bogie, but when freight is han particularly on curved mountain railways, flange and rail wear can become exce: A number of designs have therefore been proposed which allow the axles of a bog pivot relative to the frame of that bogie so as to take up a radial orientation relati the curve, thereby reducing the angle of attack of the leading axle.

The two axles are usually connected by a relatively stiff linkwork which equalise angular rotation of the axles relative to the bogie frame. In the Scales (Devine Man turing Company) bogie the linkwork is connected to the main frame of the vehic that the position of the axles is directly related to the radius of curvature, as characte by the angle made by the bogie centre line and the vehicle body. A typical radial bo; illustrated in Fig. 14.11.

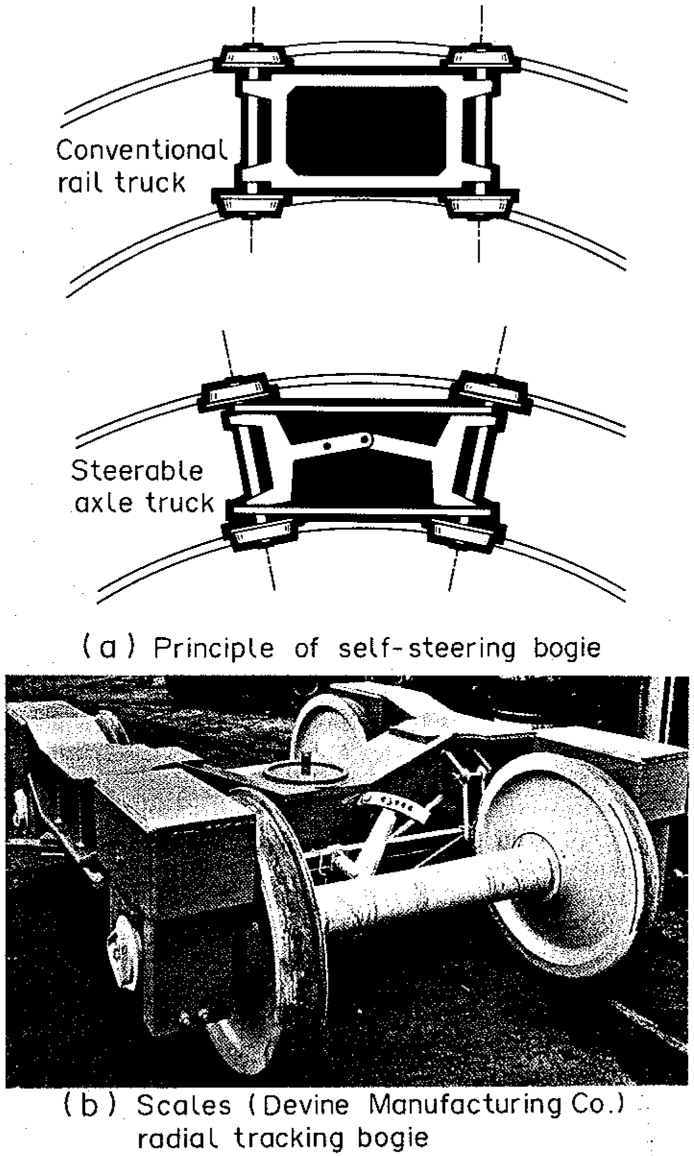


FIG. 14.11. Radial bogies.

Whilst radial trucks have been most frequently used for freight service, an example of use for passenger work is cited in Chapter 16 in connection with the linear motor-propelled trucks on trial at the test installation near Kingston, Ontario.

14.9. Effect of oscillation on passengers

The first object of the control system of a passenger vehicle must be to protect passengers from undesirable and harmful motions. Mention has been made in Chapte of limiting values for ROCOC but less intense forces may be harmful if applied peri cally. Thus experiments at Wright Field wherein human volunteers were shaken on vil tion tables showed that severe pain could be experienced at certain frequencies separ by other frequencies which gave rise to no comparable discomfort. This is attributabl the forcing of vibrations in certain organs. Thus the intestines have a natural frequenc about 5 Hz, the spine and ribs about 11, the eyeballs about 75 and the jaw about 100

Figure 14.12 summarises some data published by Guignard and Irving.[25] When degree of distress of the subject was measured by comparing the rate of respira during vibration with that at the rest period, significant changes were noted only at 4·0, 4·8 and 9·5 to 13·5 Hz, i.e. at frequencies which may correspond with resonanc certain physical systems.

This data relates to experiments carried out under extreme conditions and prov little guidance to a designer who is concerned with the construction of vehicles which be accepted as "comfortable" by the general public. The data which was available in literature presented a somewhat confusing picture partly because of the difficulty relating the nomenclature and methods used by the different investigators and pa because little guidance is provided regarding variations in vibration tolerance as betw different individuals.[25, 26]

For a number of years the quality of ride of a rail vehicle was assessed against a "i index" which purported to provide a scale which is related to the perception of com on the part of an individual passenger. The evidence upon which the index is base

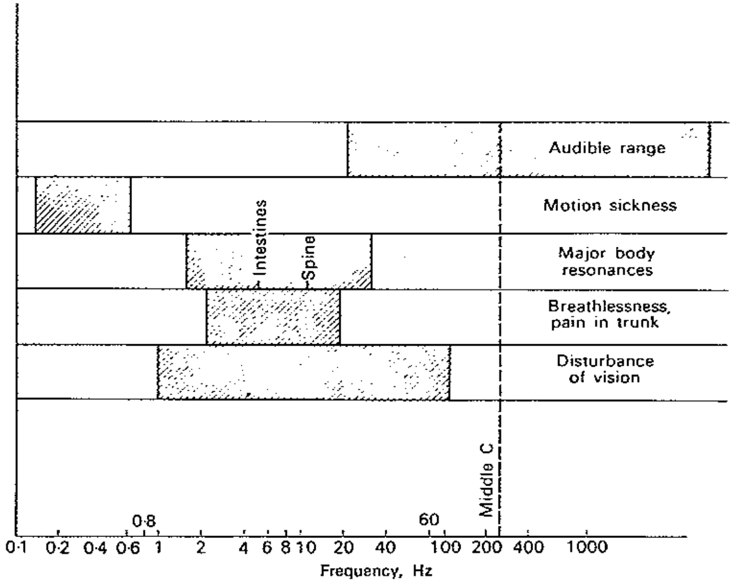


FIG. 14.12. Frequency response of human body.

questionable not only because of the paucity of experimental data but because of a presumed relationship between time of exposure to a vibratorary situation and its perceived intensity, which has no foundation in fact. Similar considerations apply to the draft proposals of the International Standards Organisation.[28, 32]

The ride-index technique appears to be faulty on two main counts. First, the numerical scale is one that tends assymptotically with increasing level to a value of about 6 whereas data are now available which indicate an almost linear variation of subjective response with physical stimulus. Secondly, the proposed method of summing two or more components by adding the tenth powers of the individual values and taking the tenth root of the results can hardly be justified from the relatively crude data from which it seems to have been derived.

In the light of the foregoing a programme of research was undertaken at the University College of Swansea which was based on the employment of psychological techniques to access the reaction of subjects to vibration environments both in the laboratory and in actual transport situations.

It is important to study the factors which will have to be embodied in new systems for them to be regarded as being acceptable to a wide clientele. In a series of studies carried out by Clarke and Oborne at the University College of Swansea, passengers in trains, cross-channel hovercraft and helicopters were interviewed and were required to assess the relative importance of various aspects governing comfort, notably temperature, ventilation, noise and seat comfort. Seat comfort, ventilation and temperature were rated as being more important than noise and vibration.

Probably the most difficult thing to determine, however, is the quality of ride which would be tolerable. Over the years many tests have been carried out to determine human reaction to vibration over all or part of the relevant frequency range. However, the published results reveal what can only be described as remarkable inconsistency. Accordingly a series of investigations was set up in the author's department in co-operation with British Rail and supported by the Science Research Council in order to ascertain what would be a reasonable basis for design of a novel system. Investigations were carried out on operational vehicles as well as within the laboratory.[29] As far as the tests in the field were concerned these were based on a questionnaire, but at the same time compact vibration recording equipment was devised so as to provide an objective measure of vibration to which the passenger was being subjected whilst replying to the questionnaire. Six piezo transducers were used to provide signals for vertical, lateral and fore and aft directions. These were mounted on the floor of the vehicle in several positions to give a reasonable indication of the overall vehicle movements. A recording was obtained by multiplexing six channels of acceleration information on to a battery tape recorder via a specially built encoding package.

Considerable semantic difficulty was experienced in providing passengers with expressions whereby they could describe the vibration situation.[30] There is no word in the English language relating to vibration which is analogous to "loudness" for noise or "brightness" for lighting. After a period of laboratory studies it was finally decided to use the line-rating method, as shown in Fig. 14.13.

A diagram was printed in the questionnaire which consisted of a line 10 cm long. At each end descriptions were given relative to the phenomenon being studied and the subject was required to make a mark on the line to indicate the severity of what he was

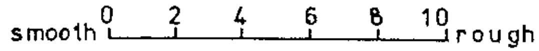


FIG. 14.13. Method of line rating.

experiencing at the time. Attempts were made to quantify the end points by such de tions as "smooth" being equated as a complete rest, and "rough" correspondi travelling in an old automobile on an unmade road.

In order to validate the techniques used for obtaining ratings in the field survey laboratory experiments were carried out within the university. A series of tests was in which human subjects were exposed to vertical vibrations over the frequency ra 1–70 Hz and over a range of vibration amplitudes corresponding roughly to the corded in public service vehicles. One important finding was that in contrast t situation with noise, the response to variation in amplitude of vibration was prac linear.[29] Thus in the expression

$$\psi = K\phi^n \qquad ($$

n approximates to unity over the frequency range 0 to 10 Hz

where ϕ = objective magnitude,
ψ = subjective magnitude,
k = a constant.

As was expected, tolerance to various amplitudes of vibration varied with freq and it was possible to plot curves of equal comfort whereby the amplitude necessa providing a given sensation was plotted against its frequency.

In a set of experiments designed to establish the degree of variation betwee sensitiveness to vibration of different individuals and to explore the possibility tha might vary with the passage of time, twenty individuals were subjected to sinu vibration stimuli in the vertical axis while standing on a vibrating platform.("standard" vibration at a given frequency was first applied. The vibrator was then another frequency and the subject was required to adjust an amplitude controller he judged that the sensation he experienced was equal to that of the standard. By re ing this procedure for a number of frequencies and plotting the magnitude of the adj amplitudes, equal sensation contours could be produced. Each test was repeated a time gap which varied from 1 to 66 days in order to assess the possibility that the res of an individual to vibration stimuli might vary with the passage of time. Result reproduced in Fig. 14.14 from which it will be gathered that there is a conside variation between the response to vibration of different individuals but that the from each individual does not vary with time.

Table 14.1, which is based on a series of observations made on both train and h craft, shows that there was no correlation between a passenger's evaluation of co and the duration of his or her journey.

It is important that the variability of passengers' responses to vibration shoul taken into account when designing vehicles. (To design vehicles on the basis of mea

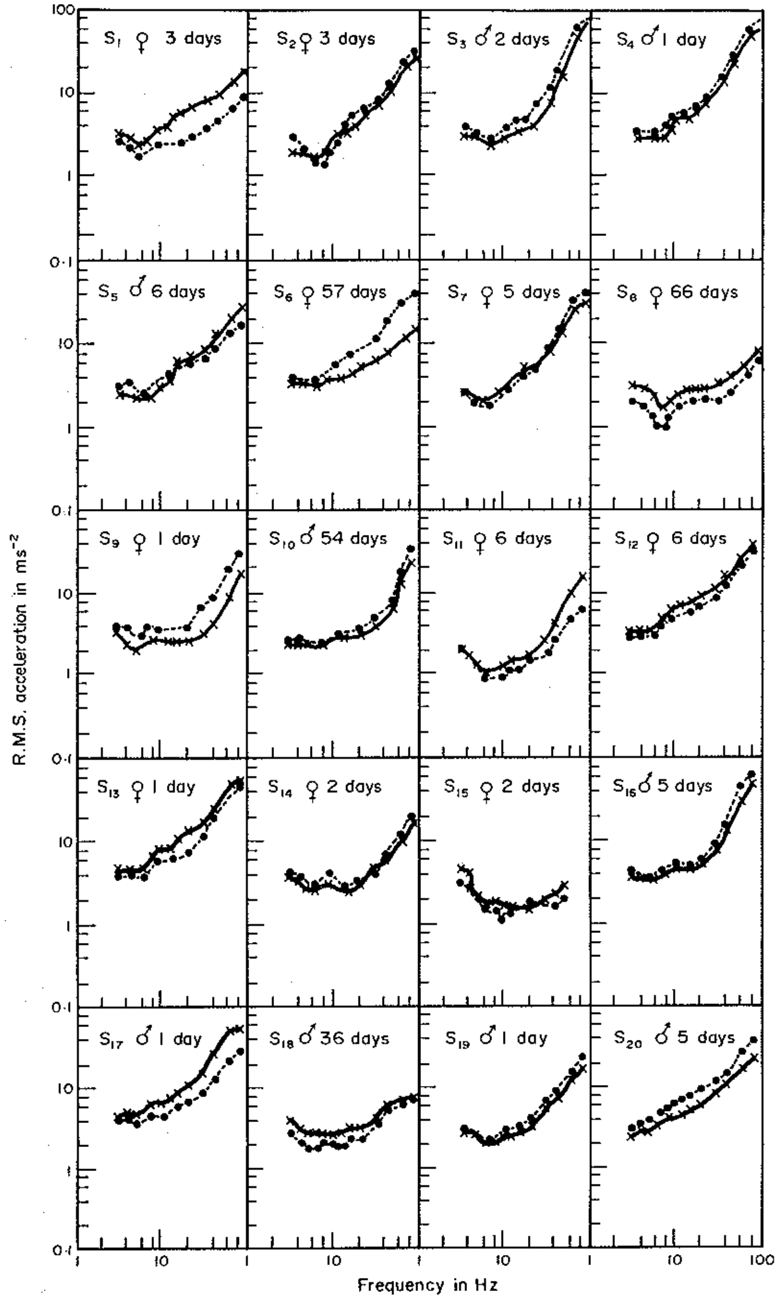


FIG. 14.14. Individual equal sensation contours obtained during two sessions.

TABLE 14.1. CORRELATION COEFFICIENT BETWEEN PASSENGERS' RATINGS OF COMFORT AND DURATION OF JOURNEY

	Train	Hovercraft
Vibrating intensity	−0·02	—0·07
Overall comfort	−0·01	−0·06

to condemn half the passengers to an uncomfortable ride.) In addition to th values. Fig. 14.15 shows a line embracing 95% of the population.[38]

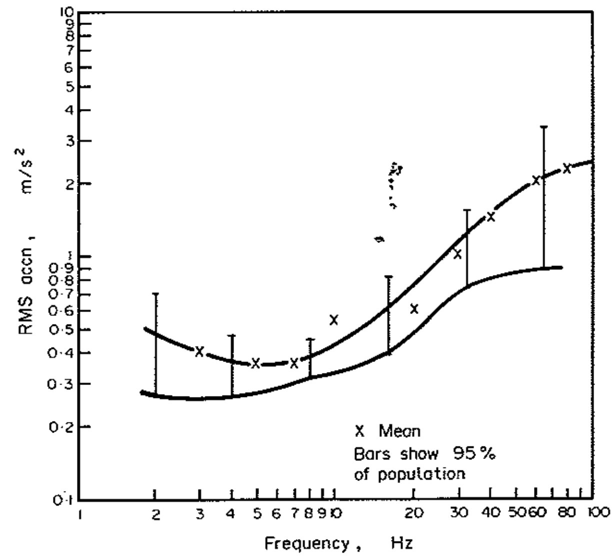


FIG. 14.15. Desirable standard for passenger ride comfort.

Rate of change of acceleration

Another important consideration from the comfort point of view is the rate of of acceleration. Loach and Maycock[30] recommend a figure of 0·372 m/s³. Anotl of looking at this problem is to consider the rate of rotation of the equivalent ; tional vector. Loach and Maycock's figure would appear to be equivalent to a rotation of 0·38 radian per second.

Regarding rate of change of acceleration in the direction of motion, a tradi accepted limit derived from tramcar practice is 1·37 m/s³. This is known as RO(

As far as linear acceleration is concerned the figures for Morgantown are 0·61 m/s² service braking and 2·94 m/s² emergency braking rate, which is probably the highest value appropriate to a guided system. Of course much higher rates of braking are experienced in road vehicles.

For any given speed the configuration of the route must be determined by the reaction of passengers to accelerating forces. Civil engineers usually consider that the centrifugal forces of passengers can be balanced by canting the track and that passengers can tolerate a limited amount of "cant deficiency". Thus in British practice a cant of 6 in. was provided (which with a standard gauge is equivalent to 6° tilt) and a further 3½ in. is permissible, making a total tilt of about 9°. The provision of a track complying with these limits but which would be suitable for a speed of 100 m/s (224 m.p.h.) would necessitate a minimum curvature of 6·020 metres which would be expensive in difficult terrain.

An example of a modern railway designed for high speed is that constructed by S.N.C.F. between Paris and Lyon. Here speeds of 270 km per hour (168 m.p.h.) are envisaged. The minimum radius of the curve is 4000 metres and the maximum superelevation is 180 mm. A cant deficiency of 85 mm corresponds to a speed of 300 km/h (186 m.p.h.).

By allowing the vehicle to tilt, passenger comfort can be secured on sharper curves. Thus British Rail's APT can tilt through 9° which is in equilibrium on a curvature of 4090 m at 100 m/s (226 m.p.h.). Table 14.2 shows the speeds attainable by different forms of guided transport on curves of different radius.

TABLE 14.2. SPEED ON DIFFERENT RADII AS LIMITED BY PASSENGER COMFORT

Radius (m)	*Conventional Train* 6 in. cant plus 3½ in. cant deficiency on 4 ft 8½ in. gauge			*Advanced Train* 6 in. cant plus 9° swing			*H.S.G.T.* 12·5° cant plus 12·5° swing		
	m.p.h.	km/h	m/s	m.p.h.	km/h	m/s	m.p.h.	km/h	m/s
200	40·9	65·8	18·2	49·5	79·7	22	67·6	108·8	30·2
400	57·9	93·1	258	70	112·7	31·3	95·6	154	42
1000	91·4	147·2	40·9	110	178	49·5	151	243	67·6
2000	129	208	57·8	156	252	70·0	214	344	95·6
4000	183	294	81·8	221	356	99	302·5	487	135

In the U.S.A. the new Federal Track Standards call for six classes of quality of the permanent way. The best, class six, which is considered suitable for a speed of 110 m.p.h. (49 m/s), calls for an alignment of 0·5 in. (0·013) on a 62-ft (19 m) line. The most favourable interpretation of this would be a sinusoidal wave form 0·018 m in amplitude with a pitch of 38 m. At the present-day operating speed, frequency would be 1·3 Hz, r.m.s. acceleration 0·848 m/s² and the rate of increase would be as the square of the speed.

Assuming that a desirable value would be 0·2 m/s² the vehicle suspension would h: attenuate track input by a factor of 4·1. At three times this speed a factor of 36 wo required. Thus air and magnetic suspensions which might reduce the input to the su sion system by allowing the thickness of the air gap to vary so as to compensate for irregularities would have an advantage.

As far as unconventional systems are concerned, Fellows[34] reports that 23-n reinforced concrete beams cast in steel moulds were straight within ±0·003 m ver and 0·005 m horizontally. Again therefore the advantage appears to be with the u ventional system.

References

1. BRADLEY, J. and ALLEN, R. F., Factors affecting the behaviour of rubber-tyred wheels on ro: faces, *Proc. Inst. Auto. Engrs.*, vol. 25, p. 63 (1930–1).
2. GOUGH, V. E., Cornering characteristics of tyres, *Automobile Engineering*, vol. 44, p. 137 (
3. BRADLEY, J. and WOOD, S. A., Some experiments on the factors affecting the motion of : wheeled vehicle when some of its wheels are locked, *Proc. Inst. Auto. Engrs.*, vol. 25, p. 46 (1
4. ROCARD, Y., *L'Instabilité en Mécanique*, Masson et Cie, Paris, 1954. See also *Dynamic Inst* Crosby Lockwood, London, 1957.
5. SEGEL, L., Theoretical prediction and experimental confirmation of the response of the auto to steering control, *Inst. Mech. Engrs., Auto. Div. Proc.*, No. 7, p. 310 (1956–7).
6. ALBERT, B. J. and WALKER, J. C., Tyre to wet road friction at high speeds, *Inst. Mech. Engrs. Div. Proc.*, vol. 180, pt. 2A, p. 105 (1966).
7. WALLACE, K. B. and TROLLOPE, D. H., Water pressure beneath a skidding tyre, *Wear*, vol. 13, (1969).
8. CARTER, F. W., On the action of a locomotive driving wheel, *Proc. Roy. Soc.* A, vol. 112, (1926).
9. CARTER, F. W., On the stability of running locomotives, *Ibid.*, vol. 121, p. 586.
10. ANDREWS, H. I., Contact between a locomotive driving wheel and rail, *Wear*, vol. 2, p. 468
11. JENKINS, R. B. M., Railway vehicle flange forces, *Mathematical Gazette*, vol. 46, p. 356
12. LOMONOSSOFF, G. V., *Introduction to Railway Mechanics*, Oxford University Press, 1933.
13. PORTER, S. R. M., The mechanics of a locomotive on curved track, *Proc. Instn. Mech.* vol. 126, p. 457 (1934).
14. HEUMANN, *Organ für die Forschung d. Eisenbahnwesen*, vol. 84, p. 465 (1930).
15. MEINCKE, F., *Kurzes Lehrbuch des Damplokomotivebaus*, pp. 169–83, Springer, Berlin, 1931.
16. KOFFMAN, J. L., Bogie design for modern wagons, *Modern Railways*, vol. 25, p. 304 (1969).
17. BISHOP, R. E. D., Basic causes of rough riding on railways, *Engineering*, vol. 194, p. 29 (196
18. WICKENS, A. H., The dynamics of railway vehicles on straight track, fundamental conditi lateral stability, *Proc. Instn. Mech. Engrs.*, vol. 180, pt. 3F, p. 29 (1965).
19. BOOCOCK, D., Steady-state motion of railway vehicles on curved track, *J. Mech. Engng. Sci.*, p. 556 (1969).
20. WICKENS, A. H., GILCHRIST, A. O. and HOBBS, A. F. W., Suspension design for high perfo two-axle freight vehicles, *Proc. Instn. Mech. Engrs.*, vol. 184, pt. 3D, p. 22 (1969).
21. HOBBS, A. E. W. and PEARCE, T. G., The lateral dynamics of the linear induction motor test *A.S.M.E. Trans. G. Jnl. of Dynamics Systems, Measurement and Control*, vol. 96, p. 147 (197
22. WICKENS, A. H. and GILCHRIST, A. O., *Railway Vehicle Dynamics—the Emergence of a Pr Theory*, Council of Engineering Institutions, MacRobert Award Lecture (1977).
23. ANDERSON, E., Running properties—recent advances in bogie design technique, *Proc. X American Railway Congress*, 1981.
24. HENDRICK, J. K., WORMLEY, D. N., KIM, R. R., KAR, A. K. and BAUM, W., Performance li rail passenger vehicles. Conventional radial and innovative trucks, *U.S. Department of Transpo Report DOT/RSPA/DPA-50/81/28*, March 1982.
25. GUIGNARD, J. C. and IRVING, A., Effects of low frequency vibration on man, *Engineering*, v p. 364 (1960).
26. BARWELL, F. T., Problems of support, guidance, and propulsion involved in high speed tra systems, *Monthly Bulletin of the International Railway Congress Association*, vol. 44, p. 821 (1

27. Hanes, R. M., Human sensitivity to whole-body vibration in urban transportation systems - A literature review, *Transportation Programs Report, APL/JHU*, May 1970.
28. Barwell, F. T. and Clarke, M. J., Passengers' reaction to vibration, *Proc.I.Mech.E. Conference on Passenger Environment*, p. 48, 1972.
29. Clarke, M. J. and Oborne, D. J., Techniques for obtaining subjective response to vertical vibration, *Proc. 1975 Ride Quality Symposium, NASA. TMX-3295. DOT-TSC-057-75-40* (1975).
30. Clarke, M. J. and Oborne, D. J., Reaction of passengers to public service vehicle ride, *Ibid.*
31. Oborne, D. J. and Humphreys, D. A., Individual variability in human response to whole-body vibration, *Ergonomics*, vol. 19, p. 719 (1976).
32. Oborne, D. J., A critical assessment of studies relating whole-body vibration to passenger comfort, *Ibid.*, p. 751.
33. Barwell, F. T., Some contributions of university and related research to the development of railway technology, *Proc. XV Pan-American Railway Congress*, 1981.
34. Fellows, T. G., High speed surface transport, *Instn. Mech. Engrs. Railway Engineering J.*, vol. 3, No. 2, p. 4 (November 1973).

CHAPTER 15

Automatic Railways

15.1. General principles and early history

It will be apparent from previous chapters that all the functions which have to t formed in controlling the movement of a collection of vehicles can be fulfilled matically. Advantages which might be expected to follow from automation to a gre: lesser degree are (1) improved safety due to elimination of human error, (2) inc capacity due to omission of time margins introduced into existing systems to all human reaction time, (3) more rapid accommodation to changes in demand beca speed and accuracy of computer-based decisions and (4) improved productivity b operators would be relieved of routine duties often requiring special skills and train as to be free to operate with intelligence on information received.

Trapeznikov[1] emphasised that the main function of a human being in any proce that of introducing control information into his environment. This implies that the be little justification for setting people to undertake tasks which involve processes are so simple and repetitive that electronic devices are adequate for the task. He : "In a correctly functioning system the high frequency 'noise' generated at a lower l suppressed there. For instance, if a central computer is loaded with detailed proble lower levels or a manager of a large concern deals with problems which could be sol lower levels, then the system is not successful."

It is to be expected that the railway should be the first transport system to be auto because of the following considerations:

1. Track and vehicles are under the same ownership and management.
2. The track is used exclusively for one type of operation.
3. Railways (at least in future) are concerned with large flows of traffic betw limited number of terminals.
4. The solution of the technical problem of guidance both for single vehicles and s of vehicles is inherent in the system.
5. The large measure of automation already applied to the various sub-systen signalling, power control, braking.
6. The increasing speed of inter-city services and intensity of urban transport ar ing excessive demands on the speed of observation and action of the drivers

An automatic system of a sort could be achieved by putting together all the auto sub-systems of present-day railway practice but this would not necessarily provide tl solution. Perhaps it is necessary to draw a distinction between the concept of "aut trains" and an automatic railway, i.e. between an atomic and a systems app

The former approach is one most naturally adopted when existing systems are converted from manual to automatic operation whereas the latter is more appropriate when an entirely new system is being designed.

The automatic railway can be regarded as a "kinetic control system", i.e. a system, the purpose of which is to control displacement, velocity, acceleration or any higher time derivative of the motion of the controlled member. In the case of the railway, the easiest way of representing the desired pattern of behaviour is the graphical timetable. It is usual to produce such diagrams using straight lines to represent the distance/time relationship of individual trains. This would imply infinite acceleration and deceleration connected by periods of running at constant speed. Strictly, the diagrams should be modified to allow for finite values of acceleration, deceleration and economic speed. The simple representation, however, can be regarded as an instruction given to the railway system, the function of various agencies in that system being to operate the service in as close conformity with this diagram as is possible.

In the case of a train the vehicles will move forward by the application of torque to the wheels of the locomotive which will be resolved into tangential force at the wheel rim balanced by horizontal force at the bearing centres. Such a force cannot exceed the product of the normal force between wheel and rail and the coefficient of friction at the interface. Application of the torque in excess of this will simply result in the acceleration of rotation of the axle without a corresponding acceleration of the train. The rate of acceleration will be determined by the mass of the train and the applied force less any other resistance to motion either of a frictional character or due to gradient. The factors that enter into train resistance are complex; first, there is the energy dissipated at the tyre, secondly, friction in the bearings and, thirdly, intermittently acting frictional force between flange and rail and the absence of a differential will produce enhanced resistance on curves. All these things will vary from time to time on the railway. The mass of the train will obviously be affected by the traffic offering. Perhaps the number of vehicles will be changed or the mass simply varied due to variation in pay load. The coefficient of adhesion is known to vary widely from time to time depending on weather conditions and degree of contamination of the rail surface. Temperature will cause variation in rolling resistance due to changes in viscosity of lubricant and variations in weather will also affect the resistance due to curvature where that curvature is such that a certain degree of sliding of wheel on rail is necessary in order to enable the curve to be negotiated. Furthermore, air resistance will be determined by the magnitude and direction of wind. These variations may set a twofold task to the Control Engineer. In the first place the control of the motive power itself may be relied upon to allow for changes in resistance so that journey time is unchanged. However, this may not always be possible, it certainly will not always be economic and delayed running may well occur. It is then the function of the automatic system to take these delays into account and regulate the passage of other trains so as to avoid conflict and yet to retain as close an adherence to the working timetable as possible.

15.2. Analysis of human contribution under the present system

To appreciate the reasons underlying the operation of most railways it is necessary to go back into history and political and commercial decisions which preceded technical

solution. For example, taking the channel tunnel proposals, these are based on economic valuation of potential traffic followed by a designed traffic pattern basec current technology and geographical and political considerations, thus envisaging types of traffic, the most important being international freight, secondly, multiple-services between capitals, thirdly, car ferry service and, finally, international sleeping traffic. Another instance of basic data was provided by the *Report on the Reshapin British Railways* which embodied, for example, diagrams of "Flows of Freight Tr favourable to rail but not on rail".

Management having specified the traffic to be offered, the engineers will then detern the rolling stock, motive power and track in order to carry this. The next action is preparation of a timetable. This is one of the two basic control documents goveri railway operation, the other being the "Rule Book". The important characteristic of timetable is that it is reproduced widely and in many forms. Working timetables are g to drivers and guards, to signalmen, station-masters and, in different forms, even general public. Thus, if all act according to timetable and in conformity with the stanc procedures laid down in the rule book and if that timetable is a realistic document, the system should literally go like clockwork. Indeed the basic presumption is th system of standard time should exist. In early days, special steps had to be taken by railways themselves to standardise times throughout their systems. Time is ultima corrected by astronomic observations so that the basic control to which railway opera is related is in effect the rotation of the earth about its own axis.

The preparation of the timetable is therefore the basic programming operation bearing in mind the various speeds of trains, complex layouts, junctions and so forth. generally been carried out by men of great experience and modified with reluctanc irregular intervals. The essential operation of forming a timetable consists of the pre tion of the position of various trains at any given time and the avoidance of conflic situations. This is admirably carried out by means of a digital computer.[2] The autom railway has therefore to comply with this overall timetable and to have within itself m for correcting or at least allowing for errors such as departure from the timetable o sioned by such events as locomotive failure, adverse weather, passenger emergency,

The existence of this generally available timetable or programme enables m different human agents to make their contribution. Station-masters will arrange stations to be open, booking clerks to be available to receive the first passengers and signalmen to be on duty in the box. Shed-masters will arrange for drivers to be avail and locomotives ready for service. Yard-masters will see that rolling stock is ready use. Thus a great multiplicity of actions can take place in accordance with the Ma Plan although there is no one action which is the obvious initiating action for the oth The motive-power unit which is the focus of all controlling action and prime concen tion of energy can, however, be taken as a basic reference point. At a certain time shed-master will communicate with the signalman who will clear the line from the de to the carriage sidings, their actions being interwoven of course with the movemen other trains.

The timetable states that a train will start at a certain time, either with passenger freight. Some human agency is required to control this action. This is not usually lef the drivers in control of the trains who are themselves controlled, firstly by the gu whose instructions might be called "operating data" and secondly by the signalmen wl instructions are overriding and might be termed "safety data". The guards will not themselves initiate action unless it is indicated to them by station-masters or their representatives that the journey should commence. They then carry out a safety function ensuring that all passengers are safely within the train and all doors are closed. Then they instruct the drivers to proceed, subject to signals. The driver would not, however, proceed unless he had a corresponding indication through the signalling system in which case he would moderate the speed of the train and its arrival time in accordance with the basic working timetable or with subsequent instructions, varying the basic timetable either in accordance with traffic requirements or speed restrictions arising from work on the line. It must be borne in mind that it is quite possible for a complete operation of a train to take place under control of a driver without any intervention from the signalling system. The signalling system only operates to protect him against conflicting movements with other trains.

Consider the action of a signalman in the traditional block system. Knowing from his working timetable that a train is about to depart and having had this confirmed by the station staff, the signalman at the initiating box communicates with the signalman in the next box and carries out the procedure described in Chapter 3, Section 3. It is to be emphasised that positive action has to be taken to prove the availability of a block before any train actually enters. Proof of clearance is based on the signalman's observation of the tail indicator confirming that all vehicles which have entered a section have also left it. Another duty of a signalman is to make a permanent record of trains passing.

At junctions, station roads, sidings, etc., the duty signalmen have to set routes according to the working timetable or, within station limits, in accordance with requests made to them by the operating staff. Audible signals are used as one means of communicating requests from the locomotive driver to the signalman.

In the mechanical box the signalmen have themselves to know the interlocking sequence and to pull levers in the right order. In some (more modern) electric boxes, route setting is carried out by relays and the signalmen have merely to indicate the required route by operating two control knobs, at "entry" and "exit" of the required route respectively as displayed on an illuminated diagram in his box.

In contrast to the manual block, the so-called automatic system described in Chapter 3, Section 7 relies upon track circuits to monitor the occupancy of a section continuously. The signals are normally at clear and go to danger when the section ahead is occupied. Thus signalmen are no longer required to progress trains from block to block and the number of signalboxes can be reduced considerably, finally being confined to those concerned with "regulating", i.e. determining priorities in the event of the day's working not conforming precisely with the working timetable.

The control actions necessary to run a railway can be divided into three levels. First comes the train-operating requirement carried out by drivers in co-operation with the guards in accordance with the working timetable. Then comes the action of the signalmen which is essentially concerned with safety and, lastly, the overall railway operating control which is the responsibility of the regulating signalmen and control officers. The latter are very largely concerned with ensuring the presence of men and their reliefs at the various parts of the system.

The decision must be taken whether to devise a single automatic control system to carry out all three functions or whether to perpetuate the traditional system of having

three independent systems with, of course, provision being made for the second signalling system to override the others.

15.3. Possible systems for intensively used passenger lines

Important routes will almost certainly be track circuited throughout and will be pr vided with colour-light signals; many functions will already be automatic—for examp the action of the signalman in route setting, the working timetable being put into sor mechanical form such as punched paper tape. The locomotives, or more probably in tl case the multiple-unit cars, could be arranged to operate continuously under the contı of the signalling system, that is, they would be arranged to go unless a stop signal int vened. The means of communication between the fixed and the moving parts of the syste is therefore of vital importance.

The most commonly used means of communication of information to a train is t coded track circuit which has the advantage that a train cuts out any possibility of tl signal going to a subsequent train. Higher frequencies may be injected into the rail f further communication functions but a limit to this arises because of the high attenuati of steel rails, therefore these can only be used for providing local information.

For main-line working where very long blocks are envisaged, it is attractive to consic separate conductors because these can be designed to operate effectively over mu longer distances. If a system is essentially land based, i.e. wherein commands are pass to a train depending on its measured position, positional accuracy of stopping will ensured, but if an element of programming of the train is performed in the vehicle itse then precise methods of measuring position or distance travelled are required. One way meeting this problem is to calibrate the track using, for example, a conductor which displaced from side to side in a zigzag formation at a constant pitch. Thus by using counting device on the locomotive, the distance will always be known with great precisic This would be more accurate than measuring the speed and integrating.

Another important consideration is the manner in which the economy of operation distinct from safety is achieved because there will be more than one way of driving train.[3] A simple rule would be that the train would accelerate at the maximum ra permissible, limited on the one hand by passenger comfort and on the other by t tractive effort available at the motors, taking into account the adhesion limit. It wou then continue at a constant speed until the point was reached when braking had commence for the train to stop at the required place. This would give the maximu average speed but would not necessarily be as economical as a slower run embracing certain amount of coasting.

Instructions given to a train could be "start", "proceed at speed not greater than" a "stop at a particular point". This final order would require a certain amount of equi ment in the motive-power unit itself. Speed would be measured and braking distar computed from this. When the train has approached the stopping point to within estimated distance, the braking system would come into play so as to bring the train tc stop at this point. It is not necessary that the instruction should be given at the actı point where braking would be initiated. It may be given earlier provided that the actı distance is known and accurate measure of progression of the train along the track provided.

The action required for safety purposes, bringing the train to a stop at a predetermined point, requires to comply with the full standards of "fail to safety" and must override all other controls. The secondary control and communication equipment which would govern actual operation, whilst necessarily reliable, need not possess this characteristic.

Sufficient examples of automatic trains already exist and sufficient automation of the signals is in the process of being developed to be sure of the success of automatic driving. A more difficult function to reproduce automatically is that of the controller whose functioning would require many channels of communication leading to a central computer. The system would be so complex that the principles of "management by exception" would be advantageous. At every stage there would be a comparison between programme events and actual events and only those differences between them which could not be cleared locally, i.e. by increasing the speed of a train, should be forwarded to the central computer. This computer would compare the existing situation with the timetable situation and redesign that timetable to deal with the situation. It is possible that an overall communication system would be needed which was quite separate from that used to communicate ordinary operational controls.

15.4. Existing installations

All the examples of automatic railway operation in existence at the time of publication of the first edition of this book could be regarded as consisting of means for automating the actions previously performed by the drivers rather than constituting complete automatic systems. They represented systems which are now termed Automatic Train Operation (ATO). Nevertheless, they present a considerable diversity of design and it may be useful to identify some characteristics upon which comparison may be based.

Any motive-power system must operate within certain constraints such as the adhesion limit, the limit imposed on rates of acceleration and deceleration by consideration of passenger comfort and the limit to speed imposed by the configuration of the route. Some systems are compared in Table 15.1.

Another system of classification can be based on the means used to convey instructions to a train and the method of closing the "feed back" loop. Table 15.2 presents a comparison of the same systems from this point of view.

The Post Office Railway,[4] which was opened for traffic in 1927, has operated automatically from the outset. It is $6\frac{1}{2}$ miles long and is used exclusively for the carriage of mails between two London railway terminals and seven postal establishments. Control is achieved by the switching of current "on" or "off" at the conductor rail. On the motive power units themselves, d.c. series wound motors are connected in series with solenoid-operated brakes so arranged that the removal of current applies the brakes. The trains are therefore either propelled or braked, there being no coasting period. No control equipment is provided for the motors, excessive currents at starting being prevented by resistances connected permanently in series. Voltage, normally 440 V, is reduced to 150 V in the station area where slower movement of trains is required.

The line is track circuited throughout and the presence of a train in one section causes the power supply to be removed from the conductor rails of the next section in the rear. In stations having a simple layout, operation is entirely automatic, with trains arriving automatically at a selected receiving berth and being dispatched by station staff operating

TABLE 15.1. RECOGNITION OF CONSTRAINTS IN VARIOUS AUTOMATIC RAILWAYS

Constraint	Railway					
	Post Office [4]	London Transport [7]	Carol Mine [5]	New York Subway [6]	D.B. (Schmitz & Kilb) [11, 12]	R.A.T.P. [8, 10]
Adhesion	None	Current limit relay	Anti-slip relay	Current limit relay		
Power	Automatic starter	Motor characteristic	Engine governor			Three levels of traction determined by excess of time over 0·5 sec.
Speed	Not controlled but "notch lock" used to regulate start and stop	Coded track	Frequency responsive speed governor acts on generator field brake application if error exceeds 2 m.p.h.	Axle-driven generator output compared with track code to demand either power or coasting events	Actual speed compared with variable frequency signal in inductive link	Time to traverse coil at normal speed = 0·5 sec.
Braking rate	Initiated by conductor rail becoming dead —one rate only	Mercury tube	Electromagnetic brakes for service stops	Speed reduced to 6 m.p.h. Maintained at this by alternate power and braking until code removed altogether	Braking target given in digital form. Rate regulated by function generator	Two levels of braking determined by shortfall of time below 0·5 sec

TABLE 15.2. TRANSMISSION OF INFORMATION IN AUTOMATIC RAILWAYS

Sense of information	Railway					
	Post Office	London Transport	Carol Mine Railway	New York Subway	D.B. (Schmitz & Kilb)	R.A.T.P.
To Train	Voltage in power circuits	By induction: (a) for safety from coded track circuits (b) for control of operation from local high frequency in rails, 15 kc for coasting 3 kc for 30 m.p.h. 2·5 kc for 25 m.p.h. 2 kc for 20 m.p.h. 1·5 kc for 15 m.p.h. 1·0 kc for 10 m.p.h.	By induction from coded track circuits, i.e. $37\frac{1}{2}$ c.p.m.—service brake, 75 c.p.m. = 7·5 m.p.h., 120 c.p.m. = 15 m.p.h., 180 c.p.m. = 30 m.p.h., 27 c.p.m.—reverse, also 960 Hz carrier system modulated at about 100 Hz	$91\frac{2}{3}$ Hz through running rails modulated as follows: 180 c.p.m.—brake to 16 m.p.h., 270 c.p.m.—accelerate to 30 m.p.h., 0—brakes heavily applied. In loop 75 c.p.m. operates doors	"Permitted Speed" and "Distance from Stop" fed by induction from coded looped conductors. Accurate position by counting loops in conductor at 100-m intervals	Conductor carrying 5-amp 150-cycle current is transposed at intervals corresponding to 0·5 sec of motion of train at prescribed speed. Faster or slower running operates brakes or power control respectively. Lost or Gained time indicators at station platforms
From or about train	Track circuit and current flow	Track circuit	Track circuit	Track-circuit ultrasonic presence—detection control of door opening	Track circuit or axle counting	Track circuit

a "train-ready" switch. At stations having a more complicated layout, movements a controlled by interlocking frames wherein the operation of points is co-ordinated wi power connections.

The approaches to the station include a section having a 1 in 20 gradient which de celerates trains which are brought to a standstill on "braking sections". The final part c the journey into the receiving berth itself is made under the action of a "camshaft control which is activated by the appropriate track circuits. After a short period, th camshaft reduces voltage from 440 V to 150 V. This produces a final speed of 8 m.p.h (3·6 m/s) which is slow enough for the cars to make an accurate stop at the unloadin berths which are indicated by dead sections.

Although from the control point of view the system is exceedingly crude, successfu operation over a period of forty years can be adduced as evidence of the feasibility c automatic railway operation.

The Carol Lake Branch of the Quebec, North Shore and Labrador Railroad[5] use coded track circuits for operation as well as for safety. Heavy trains containing iron or operate entirely without drivers. Four trains are continuously in use, one loading, on unloading and two in transit. The line is single but with a passing loop wherein empt trains are required to stop to allow the free passage of loaded trains. In this installatio particular attention has had to be paid to the precise positioning of cars at loading an unloading points and it is doubtful whether manual operation could have achieved th required accuracy.

The New York Subway[6] has experimented with automatic control and in 196 commenced to operate a shuttle service between Times Square and Grand Centra Station in the City of New York. The distance between these points is only half a mil (800 metres) and a single train made upwards of 144 round trips per day. The departu times were controlled by a predetermined programme punched into a perforated tape.

The track was divided into insulated sections, each provided with two track circuit One, a single rail circuit of conventional form, provided for train detection and broken ra protection and a second was fed with a.c. at a frequency of $93\frac{2}{3}$ Hz pulsed alternatively a either 75, 120, 180 or 270 cycles per *minute* for control purposes as indicated in Table 15.2

On receipt of the "180" code, a train operated normally at 13·4 m/s under the action c a governor but on reaching a point situated approximately 165 metres from the entranc to the platform it encountered a 180 code which caused speed to be reduced. It the entered a section with an "O" code when a full brake application was made, bringing th train to rest. The deceleration during receipt of the "180" code was controlled by track side equipment at four stages. Proximity detectors were provided which were spaced ou proportionally to the desired speed of the train. The output was fed to timers having fixe time settings so that the speed of the train was measured. "Train-stop" devices were pro vided at the track side which could engage an arm on the train to produce a full air-brak application. If the speed of the train as measured by the proximity detectors did nc exceed a predetermined value, these devices were retracted so as to avoid any additiona application of the brakes.

The position of a train, after it has been brought to a stop at a platform, was detecte by ultrasonic detectors positioned overhead and pulsed at 20 kHz. When it had bee proved that the train was correctly positioned a "75" code was applied to a local inductiv loop which caused the doors to be opened.

To restart the train in the opposite direction the following sequence took place. Firstly, a minimum station stop timer ensured adequate time for passengers to leave and enter the train, then the "automatic dispatcher" started the train immediately if it were behind schedule and on its scheduled time if it was on schedule. The "75" code was the doors close after which the "270" code was applied so that the train would accelerate to 13·4 m/s.

Whilst it will be noted that this is a completely automatic system, it lacks the flexibility and sophistication necessary for operation of anything more complex than a simple shuttle service. In particular, the method of control of deceleration would be unsuitable if stops had to be made at different places under the action of a signalling system.

The London Transport System[7] embodies the desirable flexibility and, after preliminary trials on the District Line commencing in 1963, was applied to the operation of the Hainault Loop between Hainault and Woodford in 1964 and to the Victoria Line on its opening in 1969. This was the first major transport undertaking to be planned for automatic operation from the outset.

The function of the control system may be divided into two parts for convenience: firstly, that of securing safe separation of trains and, secondly, that of bringing them accurately to a stop at the required platform.

The signal system which is "fail safe" is based on coded track circuits. Current of not less than 4 amperes at a frequency 125 c/s passes along one rail and through the leading axle to the other rail. Coils mounted at the front of the train in advance of the leading axle sense this current which is recognised by a tuned circuit. The current supplied to the track circuit is coded by tuned pendula at 180 or 270 pulses per *minute*, the former permitting a train to proceed at a speed not exceeding 11 m/s without the application of motors. The latter code brings into operation a speed-control system embodying a mechanical governor which regulates the speed of the train to 11 m/s, applying power where necessary. A code at 420 pulses per minute permits unrestricted running within the capacity of the traction equipment. Thus if a train is following another separated by four insulated-track circuit sections it will receive the 420 code and proceed normally. If it enters the third section behind the preceding train it will receive the "270" code so that its speed is reduced to 11 m/s. If it enters the second section the "180" code will apply. Power will be cut off but the brakes will not be applied unless speed rises above 11 m/s. If it enters the first section behind that occupied by the previous train, no code will be sensed and the brakes will be applied immediately. Incidentally this section is coded at 120 pulses/minute for train detection purposes but the train-borne equipment is not sensitive to this.

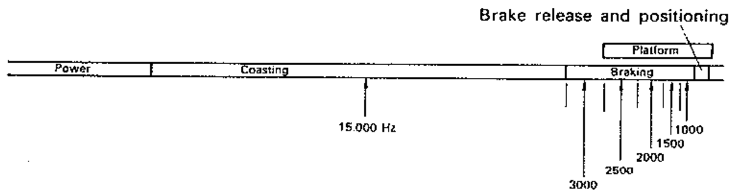


FIG. 15.1. Driving control system—London Transport.

Because the electropneumatic brake is not arranged to fail safe, a "trip valve" connected to the train pipe. When no code is received by the train, this valve is open causing the brake valves on each car to operate.

Only one attendant is allocated to each train and is responsible for opening and closi the doors. He is assisted in his observation of the rear of the train by closed-circ television, the display tube of which is mounted above the platform close to the att dant's compartment (Fig. 15.2). Once he has satisfied himself that it is safe to proceed presses duplicated starting buttons after which operation is entirely automatic until t train has come to a stop at the next station. Between stations the train proceeds at a s speed, stops if signals are at danger and restarts when signals clear without the interv tion of the attendant.

FIG. 15.2. Platform television, King's Cross.

Operational, as distinct from safety commands, are conveyed at "spots" on the runn rail about 3 metres long which convey signals at frequencies ranging from 1 to 20 k During a normal run, the first spot likely to be encountered would be at 15 kHz wh causes the motors to be switched off to initiate "coasting"; 20 kHz is used to apply brakes in the event of an occupied section ahead.

The most important use of audio-frequency spots is for controlling the stopping action at a station. As shown in Fig. 15.1, six spots are provided with frequencies ranging from 30 kHz down to unity. These frequencies are interpreted as instructions to limit speed to 4·5 m/s per kHz of signal received. These spots are spaced out along the track to correspond with the normal braking curve based on a deceleration rate of 0·94 m/s/s. The speed of the train is continuously measured by a tachometer generator producing an output frequency of 1 Hz for each 4·5 m/s of train speed. The signals from the command spots are sensed by coils carried on the train, amplified by a transistor amplifier and compared with those from the tachogenerator. Braking rate can be controlled at three levels, 0·66 m/s/s/, 0·94 m/s/s/ or 1·16 m/s/s/ using one of three mercury switch retarders of the type described in Fig. 13.2. If the control and measured frequencies are in correspondence, the median rate is maintained; if the train is going too fast, then the higher rate is applied and if too slow, braking rate is reduced to the lower value. When speed has been reduced to 1·8 m/s, control of braking is removed from the mercury switches and brake pressure is reduced to a moderate value to avert a jerk on stopping and yet is sufficient to hold the train on any gradient which may exist.

The tachometer generator is also used to feed an electronic governor which operates to control the speed of the train during receipt of the "270" code by motoring or using the E.P. brake. This control works within the limits of operation of the mechanical governor which acts as an overriding fail safe device operating directly on the straight air brakes.

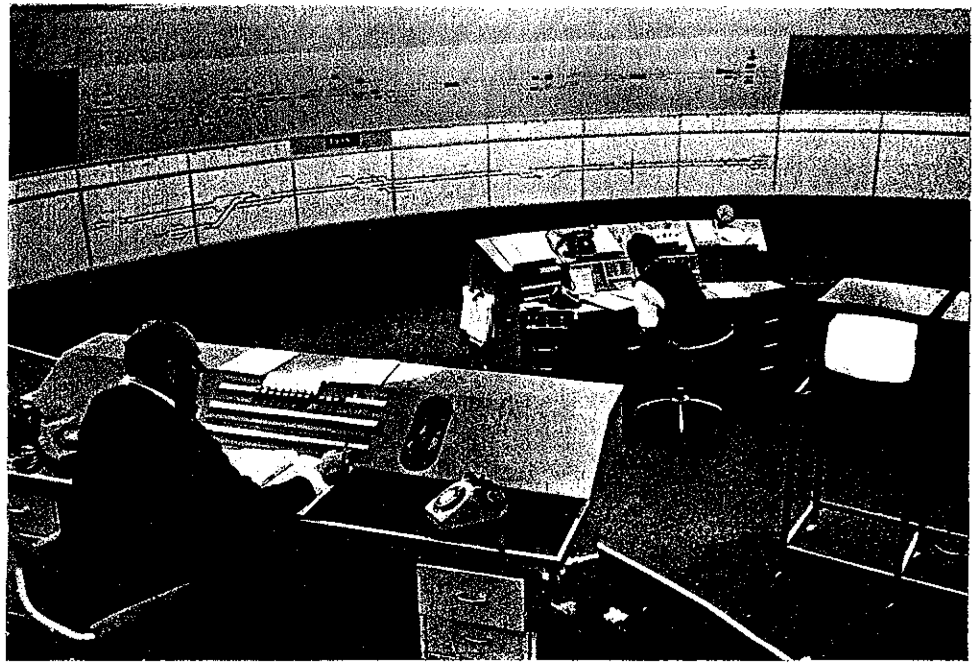


FIG. 15.3. Central control room, Victoria Line.

Operational control is by programme machines. "Identra" apparatus (see p. 98) is fitted enabling a train to be identified and the appropriate route selected. The central control room is shown in Fig. 15.3.

The Victoria Line has now operated for over ten years without a single "wrong side" failure of the safety equipment. The following extract from a lecture entitled "London Transport Experience with ATO on the Victoria Line" by Ware[18] illustrates the unexpected troubles which are inseparable from the commissioning of any large or complex engineering system.

> As the system had been tested on the Hainault–Woodford line no serious troubles were expected but inevitably the unforeseen occurred. It had been many years since a new tube line had been built and the September opening of the Walthamstow–Highbury section meant that warm, wet air was drawn into the tunnels producing massive condensation. The tracks were almost awash and current leakage causing distortion of the coded track-circuit signal was so serious that emergency brake applications were rife. With the wet track this produced many flatted wheels and consequent complaints from people living above the line were received.
>
> Various techniques were necessary to dry out the tunnels, including running one train backwards and forwards on each line under full possession. Eventually insulating pads were inserted under the chairs to improve insulation between rail and earth.
>
> The condition of the current rails was also rather poor and the consequent arcing at the collector shoes produced further interference, necessitating some modifications to the safety unit input circuitry. Adjustments to the code feeding arrangements were also necessary at various points to avoid gaps in the code pick-up.
>
> A further irritating problem concerned the "run-back detector"—a device to ensure that an emergency brake is applied if the train moves in the wrong direction. When a train reverses legitimately it has to be set for the new direction of travel and until forward movement occurs this is "held" by a magnetically latched reed switch. It was found that when a substantial current was drawn from a substation behind the train by another train ahead of the stationary one, the magnetic field surrounding the negative rail was sufficient to trip out the memory reed. This apparently unlikely situation turned out to occur at all the South end reversing points and a break in the negative rail under the car concerned had to be provided.
>
> The station braking control system is fairly simple in concept and uses some well-established techniques such as the mercury filled U-tube with electrical contacts for controlling the discrete braking rates (see Fig. 13.2). These mercury retarders are located on the front car and therefore adjust a braking rate to suit the gradient under that car. This gave a particular problem at Finsbury Park where existing station tunnels and platforms were used which have gradients varying over the platform length from 1 to 50 up to 1 in 50 down; the gradient at the front car can be very different from the average under the train.
>
> An *ad-hoc* solution was attempted, command spots being moved to try to eliminate the tendency to overbraking which could either produce short stopping or, if brakes were released, an over-run. These were not wholly successful and inaccurate stops still occur at Finsbury Park. In general, however, the station stopping system is quite adequate, the multiplicity of command spots giving a fairly fault-tolerant arrangement.

In spite of the additional equipment required for Automatic Train Operation, the Victoria Line has proved to be the most reliable of London's Underground Lines. Sample figures, based on a period of four weeks ending on 21st November 1981, showed the Victoria Line as suffering 4·9 incidents for 10,000 kilometres run whereas the long-established lines varied from 7·4 for the Central Line to 19·0 for the Bakerloo Line.[19]

Stockholm Sparväger[9] has approached the problem of operational control in a slightly different manner from London Transport. As in most systems reliance for safety is placed on the coded track circuit. "Spots" governing the onset of "coasting" and for initiating braking to a station stop consist of loops located between the rails and supplied by wayside impulse generators with current of 4 to 6 kc/s frequency. In contrast to the Victoria Line, where it will be recalled that braking rate was continuously monitored by wayside equipment, the Stockholm system employs a single spot situated 325 metres before the required stopping point. Vehicles are fitted with tachometer generators which indicate distance run as well as speed. When a train passes the brake initiating spot, the signal received starts a counter which thereafter subtracts pulses derived from the tachogenerator from a number representing 325 metres. The remainder, therefore, represents the distance remaining to be travelled. The square root of this number is determined by a simple binary-digital device and converted into a voltage which is compared with that representing the actual speed derived from the tachogenerator. If the actual speed is higher than desired, full braking is applied using the E.P. system. If speed is lower, half braking will be called for.

Thus the London and Stockholm systems are alike in that actual and desired speeds are compared and an error function derived during the braking period. In the Stockholm system this is done continuously, the desired speed being derived from position as measured by the tachogenerator. In the London system this is done at discrete spots where an indication of desired speed is transmitted from lineside apparatus.

The Glasgow Cable Subway was electrified in 1935 and re-equipped in 1982 for automatic operation. Safe operation is ensured using a conventional block system with two-aspect colour-light signals and trip cocks. The automatic train operation system is primarily concerned with the stopping of trains for operational as opposed to "anti-collision" purposes. It provides for restarting of trains after halts either at stations or at intermediate signals and also applies speed restrictions where these are required for permanent-way repairs. Track-side "balises" are provided which are self-powered from a signal transmitted by a passing train. Ten commands are possible. Three are sufficient to bring trains to a halt at the desired point. Three levels of speed restriction can be selected and the remaining commands are start-permission, automatic restart, and a command which instructs the train to travel at system maximum speed.

Hong Kong ATO and ATP systems

The basic principles of the Victoria Line control system have been applied in Hong Kong in an updated form, particularly with regard to the method of braking at stations. In place of the six control spots installed on the Victoria Line or the single-spot situation 325 metres from the stopping point on the Stockholm Sparväger, two spots are provided, one where the fastest train would be required to commence braking and another situated half-way to the stopping point. As in the Swedish system the distance travelled is determined by counting pulses produced by a tachogenerator. Instead of the "square-root" determination, however, the speed–distance profile is stored in a form of microcomputer known as a Programmable Read Only Memory (PROM). Variation in wheel diameter affects the accuracy of measurement of distance by the tachogenerator. The mid-point spot is used to calibrate the system at each brake application. This was achieved by recording the number of distance pulses detected by the train during its passage between the first two spots and so controlling the braking rate during the remainder of the stopping phase as to produce the same number of pulses (and thus the same distance). The trains therefore come to rest at the desired point, irrespective of variations in wheel diameter.

Another difference between the Hong Kong Metro and the Victoria Line is tha rails of the latter are carried on wooden sleepers whereas in the former system the are mounted directly onto reinforced concrete. An investigation of the possible effec the control system of the metal bars embedded in the track supporting structure has reported.[20]

A distinction is made between "Automatic Train Operation" (ATO) and "Auton Train Protection" (ATP). Trains can also be operated manually. The ATP is "fail s and can override the ATO or manual commands. It is set to come into operation slightly higher speed than the maximum permitted for ATO at any stage of operation.

Both systems are based upon the transmission of amplitude modulation signals f the track through bogie-mounted antennae. These signals are decoded and indica specific speed. Should this speed be exceeded by more than a specific margin, the equipment initiates a brake application which cannot be released until the train stop

When operating under ATO control, the train has four available command signal the propulsion gear, namely: off, series, parallel, and weak field. There are sixty- braking steps which are electrically controlled on an "energise-to-release" basis in o to ensure fail-safe characteristics. Brakes on each car are controlled by a pulse-w modulated signal which is transmitted through a pair of train wires. This sign generated by an encoder from an input from the ATO system and is read by a dec on each car which feeds a current demand to the power equipment or an analogue si to the friction brakes. The decoder also receives a "feedback" signal from the dyn brake and adjusts the friction brake to ensure that the total braking effort is suffic particularly as the dynamic brakes become less effective with falling speed. The dec also limits the amount of "jerk" during service braking.

Use of separate conductors. As pointed out in Chapter 9, the coded track circuit pre limitations to the variety of information which can be transmitted and in all the sys described above, some additional apparatus has had to be provided to convey al information necessary for complete automatic operation. Accordingly a numbe administrations have recognised the advantages of providing separate conductors i track to transmit information by means of an inductive coupling. The British Rai developments have already been referred to in Chapter 9 and other administrations have selected this method are the RAPT,[8, 10] the Deutsche Bundesbahn[8, 10] Tables 15.1 and 15.2), the Bay Area Rapid Transit District (BART) and the Hamb Hochbahn.

The wealth of information which can be collected and transmitted by this system c up many alternative possibilities of system design.

Bay Area Rapid Transit system (BART)

The BART system, which embraces an underwater tunnel 3·6 miles (5·8 km) long a the bay from Oakland to San Francisco, was the first U.S. system to be designed in r years and automatic operation was planned from the outset. The first section was op to revenue service in 1972 and it is interesting to compare the control system with th the Victoria Line opened in 1969. The primary control is by a digital code transm hrough the running rails and detected inductively. The track is shorted at one end of ach section and an inductive loop about 3 metres long at the other end induces a current of about 100 milliamperes peak-to-peak. (Note that this compares with the current of 4 amperes of the Victoria Line.) The voltage is not controlled directly and may be as low as 1·5 V. The inductive loops consist of 50 turns and have a power of 5 watts. The codes make up a binary system and there are four different sets of frequencies to prevent interference from block to block. They range between 5000 and 10,000 Hz. There is a six-bit code made out of 100 × × × × where × can be either "0" or "1". The last digit distinguishes movement from non-movement; zero or stop has an "0" at the end. All other speeds have a "1". The speed increments which are coded are 18, 27, 36, 50, 70 and 80 m.p.h. (29, 43, 58, 80, 112 and 129 km/hr).

The mode of train operation is that, on receipt of a change in the speed code, the train accelerates or decelerates at maximum rates until it reaches the desired speed (sometimes operation on falling gradients tends to be jerky). These rates are 2·7 and 3·0 m.p.h./s respectively (1·20 and 1·34 m/s²).

In addition to controlling the train the induced currents are used for train detection, a relay being installed at the end of the section remote from the induction loop which is set to respond to currents of the appropriate frequencies. This system has not always functioned perfectly (due to loss of train shunt) but has the advantage over its predecessors in not requiring insulated rail joints. Its operation has been reinforced by the introduction of an additional safeguard known as "Sequential Occupancy Release" (SOR). Sequence control is based on apparatus provided at the entrance to each block which notes the situation in the preceding or succeeding blocks as well as in the block concerned. Thus the apparatus may monitor the progress of a succession of trains through the vicinity of the block section. It would not give a clear indication for a train to enter a section, even if the track circuits were favourable, unless the previous train had been detected as having passed into the section beyond. SOR can be regarded as the electronic equivalent of the "block system" described in Chapter 3.

The foregoing relates to the "safety" system and, like the Victoria Line, a second system is used to bring trains to a stop at the required platform. A set of wires is mounted on the insulated cover of the conductor rail (which is energised at 1000 volts). These are transposed at intervals of 1 foot (0·3 m) and fed by a carrier wave at 7200 Hz. The coil is 740 feet long (226 m). Due to the transposition of the wires, the carrier is modulated to correspond with train speed. A "down-counter" is provided on the train, which is pre-set according to the length of the particular train. The speed of the train is compared with the output from the down-counter and controlled by release or reapplication of the brakes. Trains are of widely varying lengths (2 to 10 cars) but are all arranged to stop at the centre of the platform. The stopping system is extremely accurate.

A second control circuit mounted in the conductor rail covers approximately 1 mile from each station approach.

The train reports its composition and intended destination using a 32-bit code which operates the points for route setting. The computer can override this and operate the switches directly.

Also on the cover boards are two further circuits, one which controls the doors and the other which detects and repeats the door position.

The main control room is provided with a complete visual display of the state of the

system in considerable detail. Normal operation is by computer but the operators can intervene at any time. Drivers normally have no duty to perform but, in emergency, can drive trains manually up to a maximum speed of 20 m.p.h. There being no visual signals, they have to rely on sight to assure themselves that the track is clear. It is proposed in future to provide them with cab indicating instruments to enable them to drive at higher speeds during emergencies.

A "dead man's" feature requires drivers to operate the "close door" control during a station stop even when the doors are under automatic control. The doors will not close and the train will not start unless this has been done.

The main verbal communication is by f.m. radio using a leaky wave guide through the tunnels and line-of-sight transmitters in the open air. The route of the system is about one-third in tunnels, one-third on elevated structures, and one-third at ground level. The route length is 71 miles (114 km).

On 2nd October 1972 a train over-ran the end of the line at a station called Freemont. This mishap was traced to a fault in an electronic component wherein a silicon chip had come into contact with the side of its container in such a way as to maintain resistance of about 8 ohms. The faulty component was only required to function when a train crossed points and the fault escaped detection because of the infrequency of this event. However, in this particular case the speed was increased by the short-circuit instead of being reduced because of the points. Fortunately the driver was sufficiently alert to minimise the effect of the error.

São Paulo and Washington Metros

The BART system has now been in full operation since 1976 and a high degree of reliability has been achieved. The control system has been adopted as the basis for that used on the São Paulo and Washington Metro systems which have achieved a very high standard of reliability, as evidenced by a "mean time between failures of 1993 hours for the Automatic Train Operation as compared with a specified performance of 1364".

Regenerative braking has been successfully applied at São Paulo which has reduced energy consumption to 3·4 kW.h/car kilometre and greatly extended the life of brake linings.

15.5. Systems for high-speed working

The major developments on main-line railways at the present time are intended to enhance safety when operating at very high speeds (50 to 65 m/s) but the distinction between main-line and suburban practice may become less apparent as new systems such as B.A.R.T. are designed for maximum speeds of up to 35 m/s.

As an example of the potential capacity of separate conductors, time multiplexing may be used wherein different combinations of two frequencies can be transmitted. By this means it is possible to transmit one million bits of information during a message period of 20 milliseconds.[13] During this period a train travelling at 50 m/s will have covered a distance of only 1 metre. In order to ensure the integrity of the system the message can be repeated and passed on for further processing only when two identical signals have been received.

Because so much information can be transmitted whilst the train is moving through so short a distance, it is possible to consider the use of "beacon" devices. However, the operational situation may change when the train is between beacons and the opportunity for precise measurement of the position of the train, which can be derived from crossing conductors at intervals of, say, 100 metres, would be lost. European opinion appears therefore to favour the use of continuous conductors.

A combination of discrete and continuous information may be justified. Calculations of the number of components necessary for a continuous high-frequency information transmission system designed to meet O.R.E.[15] requirements in terms of information content for a double line of route 80,000 metres in length led to the conclusion that "Mean Time Between Failures" would be about 250 hours, therefore simplification of the circuits might be justified using beacon devices to transmit constant information such as changes in gradient or permanent speed restrictions.

Granted that the system must conform to some programme, the problem arises as to where that programme should be situated, whether on the train, on the track or partly on one and partly on the other. The view taken by London Transport[7] was that a "programme on train" concept requires "calibrating" instructions from the track and that, if a train were required to be diverted from its time-tabled route or to be reversed in direction, someone would have to have access thereto in order to make the consequential programme changes. With the "programme on track" system actually adopted, no difficulty arises from unscheduled working, the programme always being in readiness for any train which comes along.

Whilst eminently satisfactory for an operation wherein identical trains are required to conform to regular schedules, the system lacks flexibility for dealing with a variety of trains which may be required to operate at different speeds. Given the excellent communication facility provided by separate conductors, however, the dilemma does not arise because any apparatus provided for storing programmes on the train can be reset by remote control.

It is probable that most installations of the near future will rely on well-tried components for an overall "fail safe" system which will accordingly perpetuate the fixed blocks required by track circuits, axle counters or other means of train detection. This implies "fixed block" as opposed to moving block and control by progression from block to block rather than by the "train-order" system.

15.6. Moving block systems

When continuous conductors are used and provided with loops for distance measurement, it is possible for a train to maintain a record of its position at all times. This can be transmitted in digital form to a track-side or central computer where it can be regarded as a form of presence detection, or it can be transmitted generally along the continuous conductors for the information of following trains.[13, 16, 17] Thus a following train can have information about all trains from which the position of the nearest train could be derived. This can be compared with its own position and any braking programme

necessary to avoid collision can be applied. This can constitute a moving-block system without wayside equipment, which would maximise line capacity and would provide the greatest possible flexibility. A central computer is only required to exercise operational control, i.e. to determine economic speeds or modify schedule information. Trains can be addressed either by their individual codes or by the reference number of the section of the track where they happened to be at the instant of the communication. If reasonable provision is made for storing information on the trains, the intervals between messages from the central computer can be prolonged with consequent easing of the load on the communication circuits. Such a development depends upon the availability of electronic components and systems of proven "fail to safety" characteristics although the principles could be embodied in an "operational" system working within the constraints imposed by a more conventional "safety" system. Thus it is possible to control "service" braking on an optimised operational basis which on occasion can be overridden by "emergency" braking.

15.7. The relative cost of different systems of signalling and control

The cost of three basic types of signalling and control systems when applied to an existing North American inter-city railway route (New Orleans to Los Angeles over the Southern Pacific Railroad) has been estimated.[21] Account was taken of the existence of a large number of coded track circuits already installed along this route. Table 15.3 summarises the results.

TABLE 15.3. RELATIVE COSTS OF DIFFERENT SIGNALLING SYSTEMS
(U.S. dollars, 1980 level)

	Type of system		
	Intermittent	Continuous in-rail type	Inductive loop
Capital cost of fixed equipment	$7,600	$50,000	$90,000
Cost of locomotive equipment	$25,000	$50,000	$115,000

References

1. TRAPEZNIKOV, V. A., Control economy, technical progress. *Proc. 3rd Congress of the International Federation of Automatic Control, London*, 1966.
2. MILLARD, S., The production of working timetables, passenger station workings, locomotive and train crew schedules by computer. *Proc. Symposium on the Use of Cybernetics on the Railways*, p. 133, International Railway Congress Association, Brussels, 1963.
3. BARWELL, F. T., COALES, J. F. and BARTON, H. H. C., Application of automatic control theory to railways. *Proc. Instn. Mech. Engrs.*, vol. 179, pt. 3A, p. 8 (1964).
4. MEW, G. M., The Post Office Railway. *Proc. Instn. Mech. Engrs.*, vol. 179, pt. 3A, p. 24 (1964).
5. FREEMAN, S. W., Railway automation in North America, safety and control aspects. *Proc. Inst. Railway Signal Engineers*, p. 123 (1962).
6. PARKINSON, J. A., Automated subway-train operation in New York. *Proc. Symposium on the Use of Cybernetics on the Railways*, p. 154, International Railway Congress Association, Brussels, 1963.

)ELL, R. and MANSER, A. W., Automatic driving of passenger trains on London Transport. *Proc. nstn. Mech. Engrs.*, vol. 189, pt. 3A, p. 24 (1964).

ЛAJOU, J., AUDINOT, P., CHOLLEY, J. and MAGNIEN, C., L'expérience de la régie autonome des ransports parisiens en matière de commande automatique du mouvement des trains. *Proc. 1er ;ymposium International sur la Régulation du Trafic, Versailles*, preprint no. 5, p. 5 (1970).

<EKONIUS, D. and KALLBERG, N. O., The automatic pilot for underground trains in Stockholm. *'roc. Instn. Mech. Engrs.*, vol. 189, pt. 3A, p. 71 (1964).

ŁUHLMANN, H., Automatic driving of trains, *Proc. Instn. Mech. Engrs.*, vol. 189, pt. 3A, p. 106 (1964).

;CHMITZ, W., Möglichkeit der selbsttätigen Steuerung eines Zuges in Verbindung mit einer modernen ;ignaltechnik. *Neue Technik*, no. A2, p. 70 (1964).

<ILB, E., Grundästzliche Planung der selbstättigen Fahr und Bremssteuerung für Eisenbahnzuge. *Neue 'echnik*, no. A2, p. 19 (1964).

;CHMITZ, W., The possibility of driverless train operation in relation to modern signalling. *Proc. ;ymposium on the Use of Cybernetics on the Railways*, p. 188, International Railway Congress \ssociation, Brussels, 1963.

3ARWELL, F. T. and OGILVY, H. H., Communications and their effect on railway operation. *Proc. nst. Railway Signal Engineers*, p. 135 (1966).

\NON., *Transmission of Information between Rail and Motive Power Unit*, O.R.E. A46, Interim Report No. 3.

-AGERSHAUSEN, H., Use of electronic computers for driverless train operation and centralised control of the traffic on the line. *Proc. Symposium on the Use of Cybernetics on the Railways*, p. 295, International Railway Congress Association, Brussels, 1963.

FRICKE, H., Running under electrical control through the identification of the position of trains using elective systems of signalling. *Proc. Symposium on the Use of Cybernetics on the Railways*, p. 302, Railway Congress Association, Brussels, 1963.

VARE, D. K., London Transports' experience with ATO on the Victoria Line. *Railway Engineers 'orum*, Discussion on Automatic Train Operation, 11th March, 1982. Institution of Electrical Engineers.

SMITH, V. H., Fully automatic controlled trains. Institution of Railway Signal Engineers, Advance paper, 3rd March 1982.

'SO, S. K., LAM, F. K., CHAM, F. H. Y. and EDGELEY, R. K., Attenuation effects of under-track einforcing on automatic train protection signalling systems. *Proc. I.E.E.*, vol. 128, pt. B, p. 92 1981).

Evaluation of Signal/Control System Equipment and Technology, Task 5. Economic Studies, Federal Railroad Administration, Washington, D.C., U.S. Department of Commerce, National Technical nformation Service, PB81-190209.

CHAPTER 16

Some Modern Rapid Transit Lines

16.1. Need for guided transport

The tendency for increasing reliance to be placed upon the private automobile for movement within cities and between cities and their residential surroundings has led to congestion, particularly at peak hours. In extreme cases it has led to the decay of city central regions and to serious atmospheric pollution. The congestion arises from the limited capacity of the highway for vehicles (about 1800 vehicles per carriageway per hour) and from the fact that, on average, private automobiles usually carry about 1·5 persons. The use of public as opposed to private transport provides an immediate remedy in so far as the vehicles can be much more heavily loaded and capacity up to 30,000 passengers per hour can be achieved on specially segregated bus lanes. (Note: Where the buses are in competition with private vehicles for road space both are equally subject to delay and the bus becomes unattractive to the ordinary citizen who then persists in using his or her own car.)

Greater capacity is provided by rail systems which may utilise trains made up of a number of high-capacity vehicles, which are subject to a greater degree of automation (which improves safety and increases operational capacity) and which are readily adapted to grade separation.

An early advocate of advancement in land-transport technology was President Kennedy, and his successor President Johnson signed the Department of Transportation Bill which set up the department of that name within the Government of the United States of America. The Department has been responsible for extending the use of rail urban systems and much of the cost of improving existing systems and constructing new ones has been assisted by the Federal Government through the agency of the Urban Mass Transportation Administration (known as UMTA) which was set up in accordance with the Urban Mass Transportation Act of 1964.[1]

In addition to rail transit,[2] the Department has supported research on bus and paratransit,[3] and on new systems and automation. (Paratransit is a term used to denote transport resources such as taxis, charter buses, car-pools, etc., which do not operate on a fixed-route fixed-schedule basis.)

16.2. Recently constructed underground railways

Numerous cities have recently constructed "Metros" in order to combat road congestion. These do not differ fundamentally from the "subways" built some 80 years ago. However, construction has tended to be more expensive and to pay more attention to visual amenity than was the case heretofore. For example, the installation at Washington, D.C. has attractive underground stations wherein a semicircular tunnel has a diameter sufficiently large to accommodate two platforms and two tracks. A particularly impressive feature is the point of intersection of two such stations where the routes cross at right angles. Wide passages and galleries at intermediate levels give the impression of spaceousness which should help to win public acceptance. Vehicles are also attractively finished (Fig. 16.1).

FIG. 16.1. Train on Washington Metro.

The ornate stations of the Moscow Underground are well known but more recent construction in the U.S.S.R., notably the Leningrad Underground, maintains a tradition of pleasing visual design of station platforms and approaches. An interesting feature at one terminal station is the elimination of platforms and the provision of ports to provide corridors corresponding to each doorway of the train. It is understood that this feature will not be repeated because, although operation is automatic, the problem of bringing the train to rest at precisely the point where the vehicle doorways correspond with the access ports is very difficult from the control point of view. It can only be solved by reducing the final speed of approach of the train to an extent which negates the saving obtained by eliminating platforms.

It has often been suggested that the loading and unloading of transit vehicles would be greatly expedited if access at stations was provided to both sides of the train, passengers leaving at one side and joining from the other side. This arrangement is applied at an important central station in São Paulo. Perhaps because the system is still incomplete, the arrangement does not appear to be justified in this application. It was noticed that some members of the public ignored the direction signs and entered through the "exit" side. They were induced to do this by the fact that the operation of the doors was so phased that the "exit" doors opened a few seconds before the "entrance" doors. Similar behaviour has been noted on the recently completed Metro in Madrid.

Milan has followed up the successful operation of the "Red Line" with a "Green" line which follows the same design trend but where current is supplied at 1500 V through an overhead catenary system rather than at low-voltage d.c. through a side-contact conductor rail. Although the stations on the tunnel sections are less elaborate than in Washington, the construction of outdoor stations is still substantial.

Table 16.1 shows the cost of some recently constructed urban railways in comparison with that of an inter-urban high-speed link.

TABLE 16.1. THE COST OF MAIN LINE AND UNDERGROUND RAILWAYS (1976 values)

Feature	City					
	Shinkansen Extension	Calcutta	Hong Kong	Atlanta	Lyon	Tyne and Wear
Route length, kilometres	393	16·4	15·6	80·5	12	53·8
Estimated cost, millions (sterling)	1,296	80	490	1000	154	160
Cost per route kilometre (millions sterling)	3·3	4·9	31	12·4	12·8	2·97
Number of stations	10	17	17	37	15	42
Date for first trains	1975	1979	1979	1980	1978	1978
Daily number of passengers (millions)		1·7	1	—	0·1 to 0·15	
Proposed fare (p)		1·66	12	15 average	—	

*The cost of subway construction in Tokyo has been stated to be 500 million yen (£1,000,000) per track kilometre.

These examples confirm that development and construction of Metro's is a worldwide phenomenon and that considerable amounts of capital are being invested. Indeed, the provision of the necessary capital must present a limiting factor to the extension of systems to serve all but the most densely loaded corridors. Apart from automatic operation and some refinement in suspension systems there is little evidence of technical advance over systems built some 80 years ago and the possibility of cost-saving through technical change warrants investigation.

16.3. Intermediate systems

The Metro with capacity of 80,000 passengers in a short time (60,000 passengers per hour in sustained operation) is unrivalled as a means of transporting large numbers of people. However, the cost, several million pounds per single track kilometre, must impose limitations as to the extent to which it can serve the whole of a city. Such high capacity is generally only required on certain corridors which can only be sustained if some other system is available to act as a feeder. This is particularly the case in the outer residential areas where many journeys may be expected to originate by private car. The only other generally accepted mass transport system is the bus. Operating on a two-lane arterial road, the double-decker bus is attributed in the Hong Kong Comprehensive Transport Study with the following capacities: average bus flow 10,800, heavy bus flow 16,200, and maximum bus flow 24,300.[4] It appears therefore that a gap in capacity exists between double-decker bus and the full-scale Metro.

This gap poses a challenge to modern technology. Problems also exist in providing viable transport systems to act as feeders to full-scale bus systems. One solution put forward in the United States was the "dual-mode" system whereby vehicles operated at the periphery of a city under drivers' control but, when they approached the city centre, the driver left the vehicle which was then propelled electrically under automatic control into the city centre. Extensive studies of this system by the G.M.Co. have not resulted in any commercial application.[5, 6, 7, 8]

Another alternative, practised to some extent in both the U.S.A. and Europe, is the "Dial a bus" system. Here, individuals requiring transport telephone a central agency which arranges to route buses to serve the individual house. In one system in Kingston, Ontario, the buses are routed at half-hourly intervals and connect at a central point with other "Dial a bus" routes or with scheduled buses. The economic viability of "Dial a bus" has been much questioned and the only justification would be to reduce dependence on motor cars in city areas having some Metro or other urban transport network.[9]

16.4. Light electric vehicles—Tyne and Wear system

A good example of a light-electric vehicle-based Metro system has been developed by the Tyne and Wear Passenger Transport Executive in and around the City of Newcastle. This system is designed on the basis of an infrastructure of bus routes together with a suburban railway system. This facility has been redeveloped in a co-ordinated plan, provided with modern equipment, and linked to the city by a specially constructed underground route 6·4 km in length. This is connected to some forty-two stations, a former electric suburban railway being re-equipped with light-electric vehicles operated from an overhead 1500 V d.c. with modern cars which are operated as twinned articulated units having seats for 84 passengers and 120 standees (Fig. 16.2).

Acceleration up to 40 km/h takes place at 1 m/s^2 and braking at 1·3 m/s^2. This can be increased to 2·3 m/s^2 by magnetic track brakes. Maximum speed is 80 km/h. Services in the central area run at approximately 3-minute intervals. The units operate either singly or in pairs, according to traffic demand. The capacity is expected to be 20,000 passengers every hour in each direction. The maximum speed of 80 km/h and the performance of the vehicles from the point of view of acceleration and breaking assure an acceptable

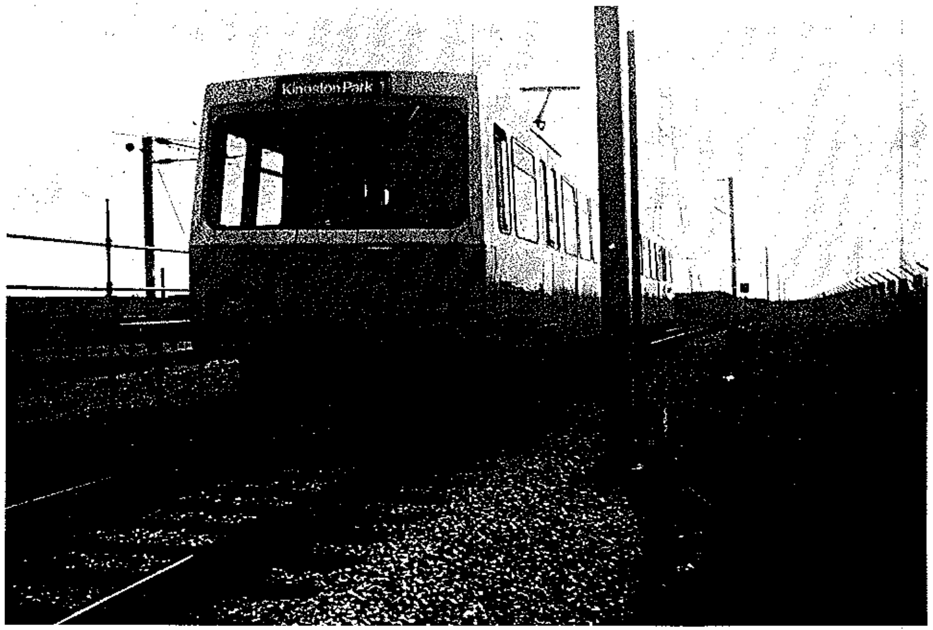


FIG. 16.2. Light electric vehicle on Tyne and Wear Metro.

average speed. Stations are designed to give ready interchange with the bus services which have been rearranged to give an integrated bus and train system. To facilitate public use, a new ticket system provides an opportunity for a journey to be made by bus and Metro using one ticket.

The first concept of the Metro was put forward in 1971, and the first phase of operation was opened in the spring of 1979. The Tyne and Wear Transport system is a good example of how existing assets can be supplemented by major capital investment to provide a cost-effective modern system. Modest increases in train length or reduced headway would enable throughput to be increased to some 30,000 passengers per hour, which would appear to be the desired level for an intermediate system in most cities.[10]

16.5. Pneumatic tyred systems

However well the track is constructed and maintained, a certain amount of noise and vibration will arise when steel wheels run on steel rails due principally to flange contact on curves. This militates against the location of the route above ground notwithstanding the savings in constructional cost which would result from the avoidance of tunnelling. A measure of the importance of this factor may be derived from Fig. 16.3 which shows a special elevated tunnel which has been constructed on the Tsuen Wan extension of the Hong Kong Mass Transit Railway in order to protect the Kwai Fong housing estate from noise impact.

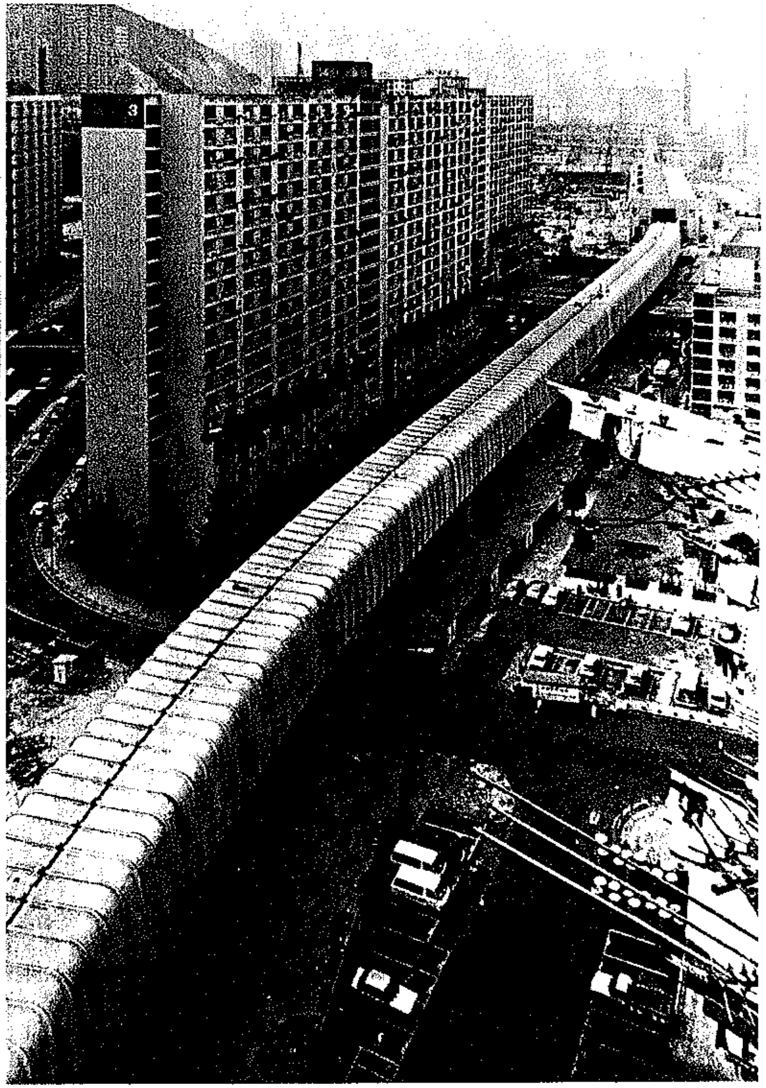


FIG. 16.3. Elevated tunnel to control noise from Mass Transit trains. (Courtesy Mass Transit Railway Corporation, Hong Kong.)

In an effort to diminish environmental impact a number of administrations have ɪopted pneumatic tyres. These also offer promise of improved performance because of ɜtter adhesion. This enables steeper gradients to be negotiated. Climatic factors may, ɔwever, militate against these advantages; thus hydroplaning may occur in heavy rain ɪd track may become iced in winter. Accordingly some administrations have arranged

for all pneumatically-tyred vehicles to act under cover. This was done initially at Mexico City but recent installations have been made in the open air and the tyres used have been manufactured with a specially developed tread pattern. Objections to pneumatic tyres are greater frictional losses, risk of accident due to tyre failure and loss of passenger comfort due to non-circular or eccentrically-mounted tyres.

The most common arrangement for guiding subway vehicles fitted with pneumatic tyres is to provide wheels mounted on vertical spindles attached to the bogies (one at each corner) which interact with vertical guide rails mounted outside the running rails. A general view of this arrangement as adopted at Mexico City is shown in Fig. 16.4. The

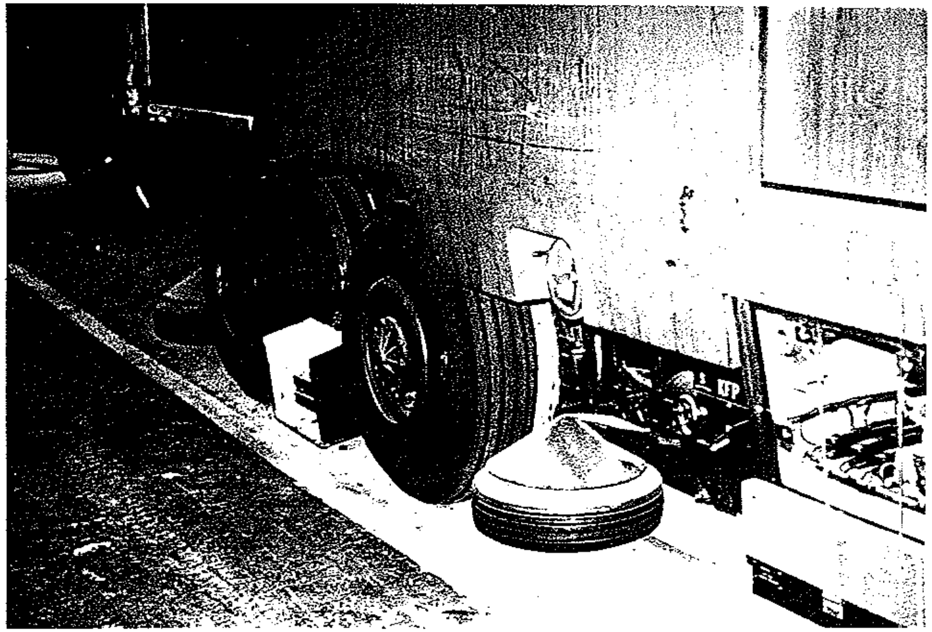


FIG. 16.4. Pneumatic-tyred vehicle for subway in Mexico City. Bogie showing carrying and guiding wheels.

guide rails also act as power-supply rails and the current-collection gear will be seen between the two main carrying wheels. Special features are the driving motors which are mounted in an inclined disposition with their axis at right angles to the main axle, and the steel flanged wheels which are normally clear of the track but which can support the vehicle in the event of a tyre failure. They can also act as drums for clasp brakes and for guiding the vehicles over points provided at terminal stations and depots.

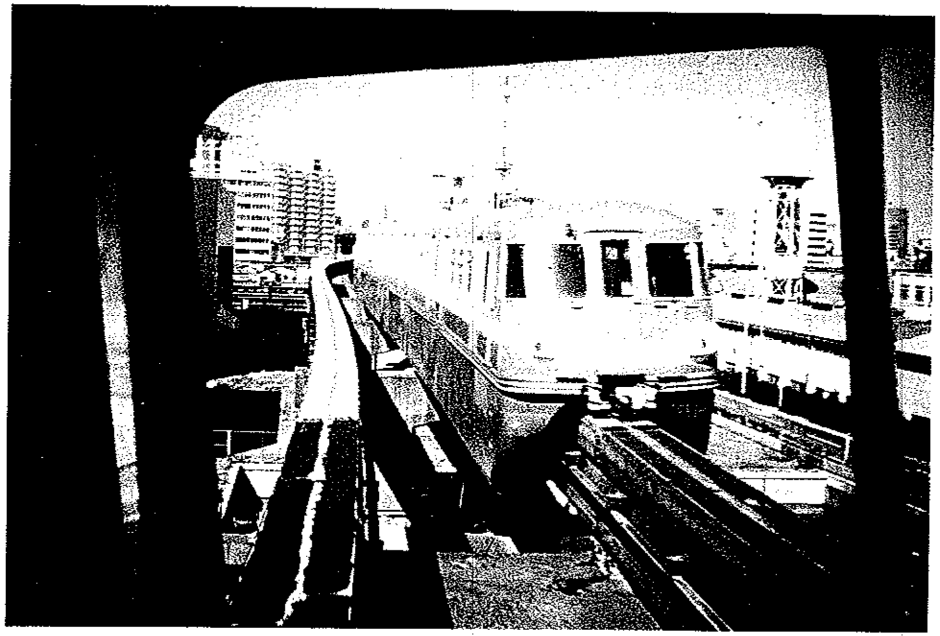
Particulars of various systems are set out in Table 16.2.

TABLE 16.2. COMPARISON OF PNEUMATIC-TYRED GUIDED VEHICLES

	Number of coaches on normal trains	Length and breadth	Weight per car +		Maximum speed (km/hr)	Passengers per car		Rate of acceleration (m/s²)		Material of construction
			Empty	Loaded		Seated	Standing	Driving	Braking	
Tokyo Monorail	3×2 (articulated 2-axle)	(metres) 10 ×3	13·6	—	100	35	45	0·75	1·25	Steel
	2×3	14·5×3	23·5	—	90	60	70	0·75	1·25	Al.
Sapporo	2 to 8 (articulated in pairs)	13·8×3	16·3	21·4	75	42	52	1·27	1·35	Welded Al.
Paris	6	15 ×2·4				24	135			Al.
Mexico City	6 to 9	15 ×2·5	23·3	37·5	80	38	132	—	—	Al.
Lyon	3	18 ×3	26	—	90	55	71	1·2	1·2	Al.
Morgantown	1	4·7×1·8	3·9	5·3	48	8	12	0·61	0·61 and 2·9	Fibreglass
Airtrans	2	7·6×2·1	6·4	—	27	16	24	1·14	1·14	Steel frame Fibreglass
Atlanta Airport (Transit-Expressway)	4	12 ×2·8	13·0	—	112	16	64	1·1	1·1 and 2·7	A.1 + Fibreglass
British Rail study	1 to 3	3 ×2	2·5	—	54	8	4to 15	1·25	1·25	Steel chassis

Tokyo Monorail

A system embodying pneumatic tyres has been commercially operated in Tokyo over a period of 15 years and traffic has continued to increase throughout this period (Fig. 16.5(a)). The system is basically the ALVEG system which has one main disadvantage,



(a) Vehicle for ALVEG system.

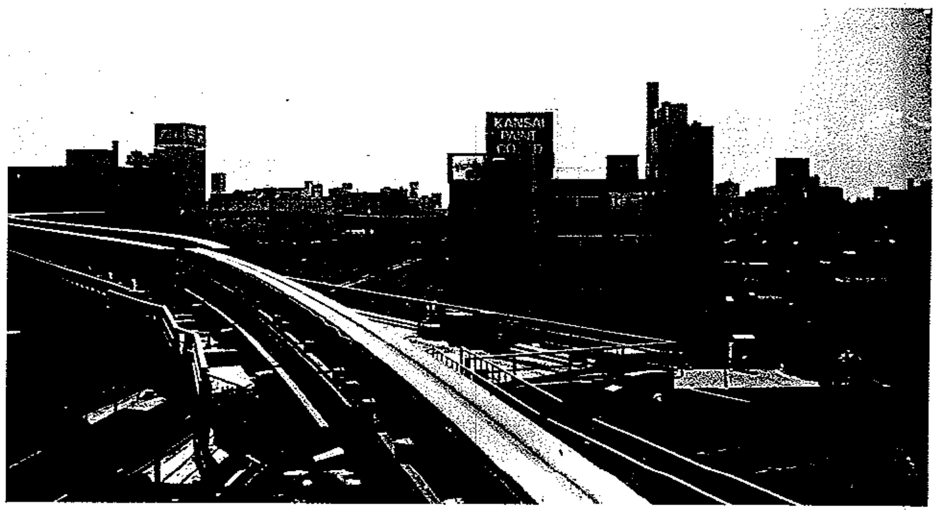


FIG. 16.5. Tokyo Monorail. (b) Points during opeartion.

namely that because of the manner in which the guiding and stabilising wheels impinge on both sides of the supporting structure, the design of points is particularly difficult (Fig. 16.5(b)). This is because the main structure has to be aligned with the chosen route. It is naturally of massive construction and its movement is therefore time-consuming and the construction is expensive.

Apart from this objection, the system has operated very satisfactorily, there being practically no visible wear of the concrete surfaces of the running track. The cost of construction was claimed to be about one-third of that which would have been required for a conventional underground railway. The success of this system indicates that in temperature climates at any rate a combination of pneumatic tyre and concrete is durable and operationally effective.

Lyon

An example of a Metro recently constructed to operate at a moderate capacity is provided by the system at Lyon.[11, 12] Due to waterlogged soil conditions it was necessary to make the system as shallow as possible, and therefore small-wheeled vehicles were adopted which, whilst following on the general principles of the system employed in Paris (notably guidance by lateral guideway and retention of steel wheels and steel rail for use in emergency), nevertheless embodied improved suspension devices (notably air suspensions) which produced a very comfortable ride. A more compact drive arrangement has been provided with a single motor mounted between the two axles, as shown in Fig. 16.6. Particulars of the system are included in Table 16.2 from which it will be seen

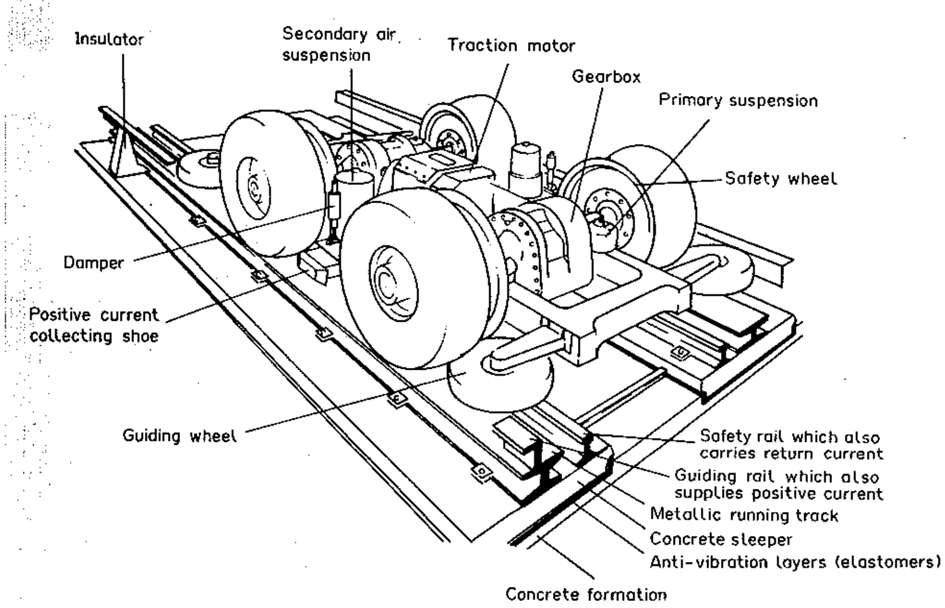
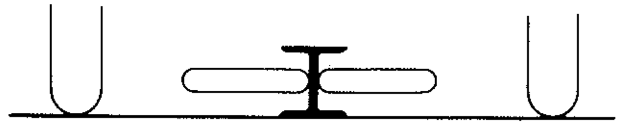


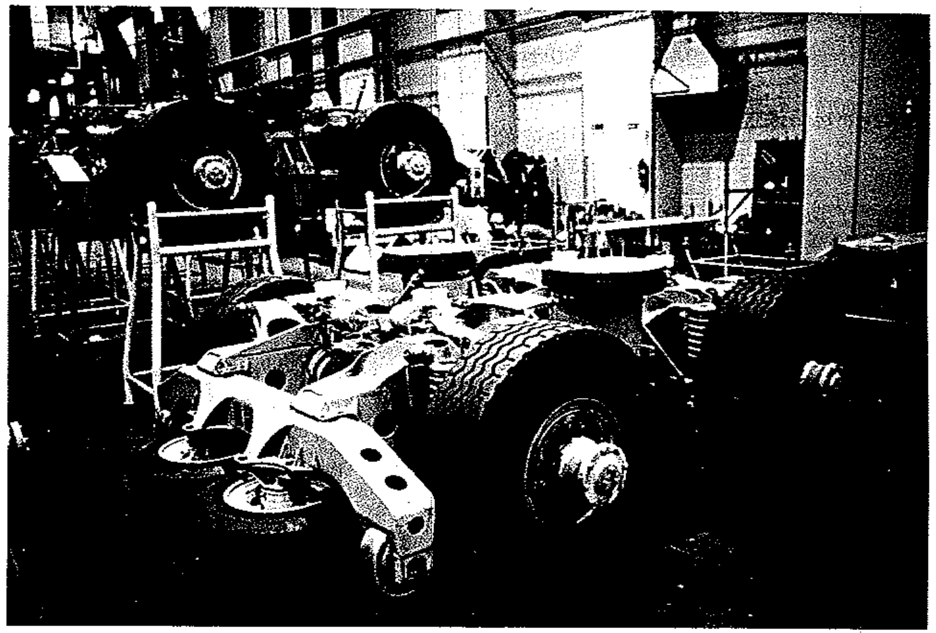
FIG. 16.6. Bogie for Lyon Metro.

that trains will carry 378 passengers each which, at 2-minute headway, leaves capacit well below that associated with an intensively operated bus route.

The initial section to be opened does in fact correspond with a route previously operate with high-capacity articulated buses. Good performance characteristics, notably 1·2 m/s accelerating and braking and the maximum speed of 90 km/h, provide an attractiv alternative to street travel in the central and otherwise congested areas of the town. A interesting feature of the revised arrangements for traffic within the town arises from th fact that during construction of the Metro on a cut-and-cover basis, a central stree



(a) Arrangement of guideway.



(b) General view of bogie.

FIG. 16.7. Sapporo system of guidance.

through the shopping area was taken out of use, traffic management arranging alternative routes. With the completion of the underground, the road concerned was not restored to traffic but kept as a pedestrian precinct which has considerably enhanced the amenity of the central shopping area of the city.

Sapporo

A recent development also in Japan has been the opening of a transport system at Sapporo where pneumatic tyres are also used. However, the supporting wheels are arranged to run on tracks which are permanently in position and the guidance is by means of secondary wheels which operate on a central vertical rail (as sketched in Fig. 16.7(a)) which facilitates the design of points and crossings.

There are two lines, the North and the East–West. The former is operated manually, backed up by an automatic train control (ATC) system but the latter is fully automated although motormen are still carried. The position of a car is detected by an antenna under the frame of the car which interacts with induction loops mounted in the running surface which can be excited at various frequencies corresponding with the command speed operative at any particular time.

Another innovation at Sapporo relative to the systems employed in France, Montreal

16.6. Innovative—fully automated systems

The Airtrans system

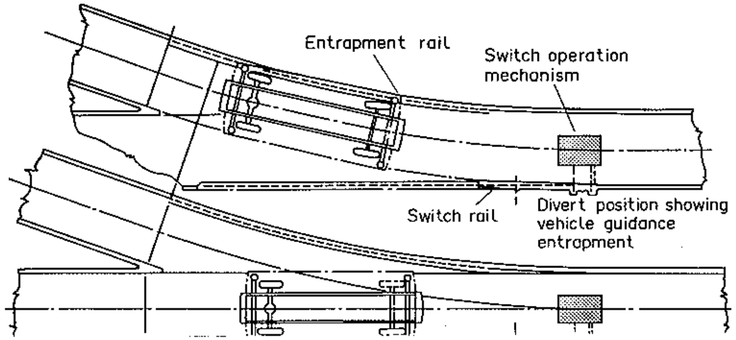
The Airtrans system at Dallas/Fort Worth Airport in Texas[13] was one of the more ambitious schemes to be initiated by the Urban Mass Transportation Administration (UMTA) in the U.S.A. The system was designed to use automatically controlled rubber-tyred vehicles to provide internal transport for a large airport involving interchange between many airlines. In addition to the conveyance of airline passengers, separate provision was made for the transport of airport employees, passengers' luggage, air-freight, mail and garbage. The total length of the system is 21 km and particulars of the passenger-carrying vehicles are given in Table 16.2. There are a total of fifty-three stations.

The guideway consists of an 0·2-m-thick reinforced concrete slab bounded by parapet walls 0·15 m thick and 0·6 metres high. The gauge, as defined by the distance between the guidewalls, is $2{\cdot}4 \pm 0{\cdot}003$ m. The surfaces are finished to inter-state highway standards of smoothness and dimensional tolerance.

The passenger vehicles are steel-framed with external panelling of acrylic-coated fibre-glass. They are provided with double access doors on one side only, together with an emergency exit. The access doors correspond with similar double-doors in a glass partition which encloses the platform. Both sets of doors open automatically when the vehicle has come to a stop and has levelled itself to the station platform. General views of vehicle and guideway are shown in Figs. 16.8(a) and (b).

(a) Vehicle and guideway.

The switches are of the active type as illustrated in Figs. 16.9(a) and (b). The vehicle is guided normally by guide wheels which interact with the parapet walls. At junctions, however, the guide walls are discontinued at the centre of the switch area and separate wheels, known as entrapment wheels, run within entrapment rails to provide guidance. At the area closest to the point of divergence the entrapment rail at one side is discontinued and a corresponding section on the other side is hinged. Thus when the hinged portion is moved to the right the entrapment wheels are directed into engagement with



the entrapment rails on the right-hand side and the vehicle takes the right-hand path. When the hinged portion is moved to the left, the entrapment wheels and rails on the left-hand side are brought into engagement and the vehicle moves to the left.

The above action relates of course to diverging situations (facing points). Where the merge action is required (trailing points) the switch rails are held in the normal position by springs and are moved by the entrapment wheels.

Although the vehicles are supported on rubber tyres, special rails are provided for signalling. One of these rails is mounted above three copper-clad steel rails which supply three-phase power to the vehicle. These are carried on plastic insulators on a concrete parapet at one side of the track (they are duplicated on the other side at junction areas). These rails are divided into insulated sections. A similar rail mounted below the three power rails is earthed. The vehicle carries a conducting member which shunts the upper rail by connecting it to the lower.

The automatic control system is arranged as a hierarchy in three levels as follows: Central Control (CC), Automatic Vehicle Operation (AVO), and Automatic Vehicle Protection (AVP). The AVP consists of a five-block control system, the average block being 90 feet (27 m) long, although they may range in length from 45 to 240 feet (14 to 73 m). The normal operating speed is 17 m.p.h. (7·6 m/s) which requires trains to be separated by five blocks. When the separation becomes less than four blocks a signal is sent to the vehicle to slow to medium speed, which is 9½ m.p.h. (4·27 m/s) and when the separation is less than two blocks the command is to stop. Although the maximum speed is very low this is compensated for to a certain extent by the fact that the minimum headway is only 18 seconds.

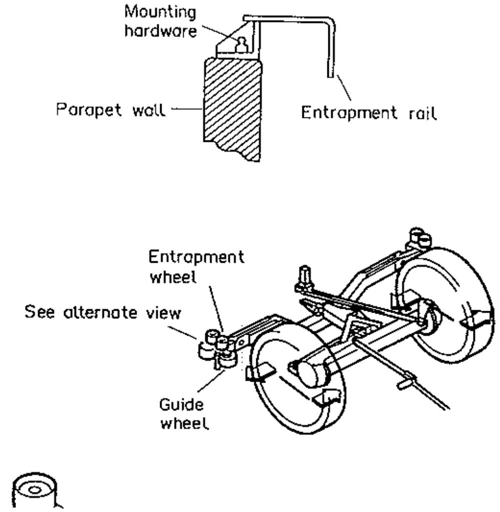


FIG. 16.9(b). Entrapment rail and wheels.

Vehicle route information is stored in an on-board control logic assembly which responds to interrogation from the wayside by providing route information to lineside control stations which decode the route information and set the switches to the required position.

The AVO embodies operational controls which are not fail-safe, notably station stops and door operation.

The CC carries out system status monitoring and embodies supervisory controls.

In addition to the conveyance of passengers, separate provision was made for the transport of airport employees, passengers' luggage and freight, postal traffic and refuse. Particulars of the passenger-carrying vehicles are given in Table 16.2.

The system was constructed at the same time as the airport itself so that delays in finishing airport installation led to insufficient time being available to complete and to fully commission the transport facilities before the commencement of public service. This led to serious operational difficulties and to overspending (the estimated cost was $US 39,651,000 whereas the actual disbursement was $US 53,402,000). The overspending in particular had widespread repercussions and caused an unfortunate reaction against innovative technology in some influential political circles. Shortly after the opening there were 204 men engaged on maintenance and remedial work. However, the passenger-carrying facilities are now working with complete reliability and the staff available has been reduced to 93. A breakdown of the capital cost is shown in Fig. 16.10.

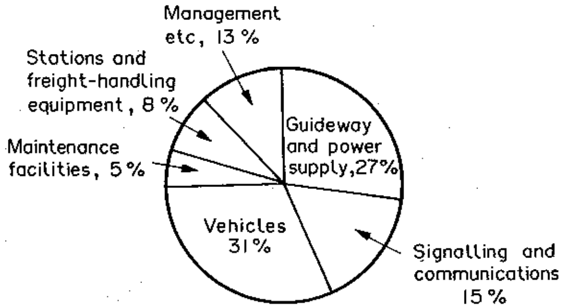


FIG. 16.10. Distribution of capital cost of Automated Guideway System—Airtrans at Dallas/Fort Worth Airport.

Transit Expressway

A system embodying pneumatic tyres, automatic operation and a novel configuration of the vehicle by pneumatic tyres running on a pair of beams is separated from guidance by means of separate wheels running on vertical spindles which interact with a central joist which acts as a guide beam. Propulsion is by electric motors which drive the carrying axles through conventional gears of the type used on heavy road vehicles.

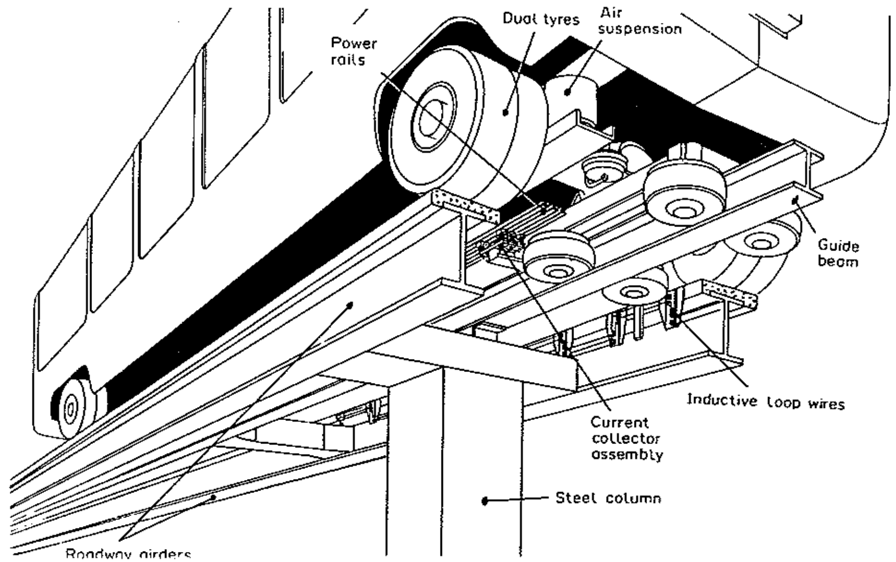
The main axles are pivotted relative to the main vehicle body so that they may set themselves radially on curves under the action of the guidewheels, which are arranged in two sets before and after the main carrying axle.

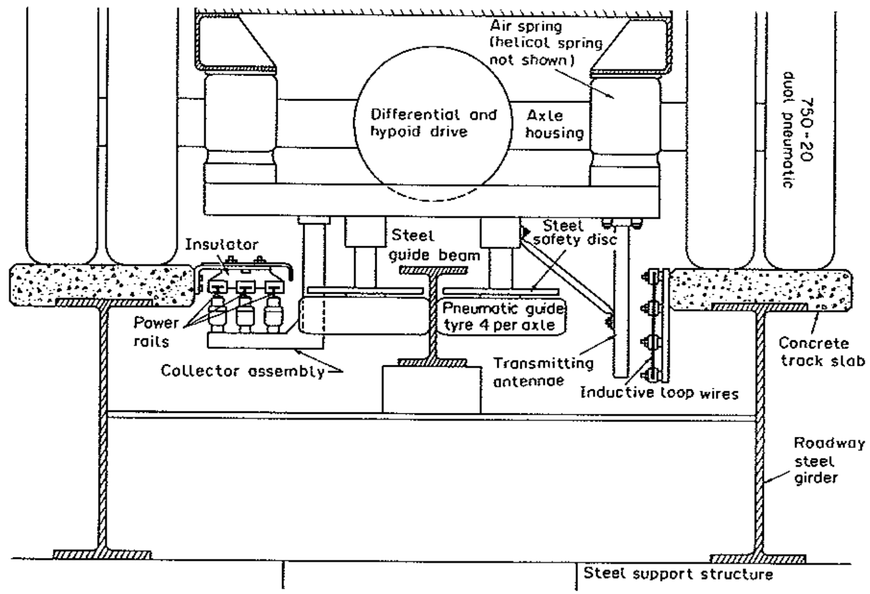
Each car consists of an aluminium shell with fibreglass ends and carries seating for 16 passengers together with accommodation for 64 standees. The length is 12 m (39 ft) and the mass is 12 t. Maximum speed is 3·1 m/s (70 m.p.h., 112 km/h). Normal acceleration and braking is 1·1 m/s and emergency braking is 2·7 m/s.

Power is supplied through three-phase power rails at 600 V and the current is rectified in a thyristor-controlled bridge to provide variable voltage to motors of nominal rating, 74·6 kw (100 h.p.). One motor is mounted on each bogie. Dynamic braking is controlled by regulating the field current of the motors, energy being dissipated in resistances. Air brakes are also provided.

Automatic operation takes the customary form of a hierarchy of three levels. These are designated as the Automatic Train Supervision system (ATS), the Automatic Train Operation system (ATO), and the Automatic Train Protection system (ATP).

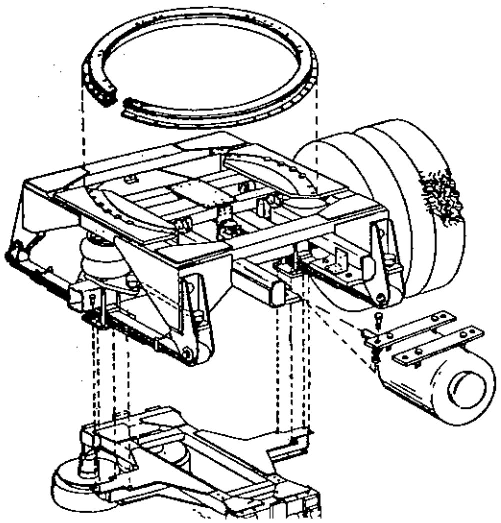
Conventional track circuits not being available because of the absence of steel wheels and rails, two special rails are provided as part of the power-supply system (there are five rails in all) which are divided up by insulated joists to form signal blocks. A code is





(b) Arrangement of vehicle guidance, power and communication equipment

FIG. 16.11. Transit Expressway. General arrangement of vehicle on guideway.



fed to the boundary of each block and its absence at the other end denotes track occupation. This is not a "fail safe" system, so two shunts are used per vehicle and additional safeguards such as confirming the presence of train-shunts, once each round trip, are provided.

The train-detection system is used to generate codes for permissible speed which are transmitted to the train and received inductively by means of antennae mounted on the guidance system. The codes are six-bit binary and provide for five speeds as well as emergency stop and reverse. The signals are interpreted on the vehicle by two microprocessors in a checked redundant fashion to achieve safety. Train speed is measured by tachometer and independently compared with the received code by each microprocessor.

The ATO system controls the stopping of the vehicle at the desired speed, the maintenance of schedule speed (within the limits set by the ATP), the opening and closing of doors, and setting of visual indicators (e.g. destination) on the vehicle.

The programme stop sequence is controlled by a conductor placed along the guiderail which is looped every 0·15 m (6 in.) so that a receiver on the vehicle detects a phase shift for every 0·15 m of travel. The stop routine is initiated by a position signal and a count-down of the requested speed is initiated on the basis of distance travelled.

The door control operates once the vehicle has stopped in proper alignment with the platform.

The ATS system consists of a mini-computer system located in a communications control centre. It monitors all the functions necessary to keep the whole system in operation. Its purpose is purely supervisory insofar as the ATO system can operate independently.

The Morgantown system

The system was introduced to serve an educational purpose. West Virginia has a larg university with 20,000 students and 6000 staff which is spread over two campuses. Thes are only a few miles apart but the distance is too great for easy transfer of students between lectures. In order that the range of choice of each individual student should no be restricted to one or other campus and because his or her programme might necessitat successive lectures in each campus, an inter-campus bus system was introduced. This however, necessitated intervals between classes of some 20 minutes. Eventually this tim became insufficient owing to traffic congestion and the interval between classes reverte to 10 minutes, with the result that students were discouraged from scheduling successiv classes on the different campuses. The Personal Rapid Transit (PRT) system was intende to alleviate this problem and to make it possible for students to choose classes more i keeping with their requirements.

In 1969 the United States Department of Transportation made a grant to study th feasibility of constructing a rapid transport system in Morgantown and, in 1971, contracts to build such a system were awarded by the Department of Transportation. Th system is designed partly as a national transportation research laboratory and partly t serve the immediate needs of the university and city.[14]

and 45 vehicles serving three stations were provided.

Each car can carry 8 seated and 12 standing passengers, as illustrated in Fig. 16.12 and travels at a speed selected from the following:

6·7, 10·0, and 13·4 m/s on the main guideways,
or 0·6, 1·2, and 2·4 m/s in station areas.

Maximum curvature in the vicinity of stations, which are normally taken at a reduce speed, is 9 m (30 ft) radius and the maximum gradient is 1 in 10.

Each vehicle has four wheels and is fitted with steerable axles and tyres which ar air-filled. They have two chambers to minimise the risk of accident in the event of a puncture. The inner air chamber is pressurised to 0·5 MPa (75 psi) and the outer t 0·4 MPa (60 psi).

The vehicle chassis consists of a framework fitted with a front buffer which will withstand any impact up to 1·5 m/s (5 ft/s). The cars are arranged to operate in one directio only, reversing loops being provided at the stations. The axles are developed from lorry type rear axles and embody a heavy-duty differential with a 7·17 : 1 ratio. Suspension i by air springs which are self-inflating and regulated to provide a constant "load-to unload" separation distance and thus to provide a constant floor height for ease of entr and exit at station platforms. Mechanical linkages operate control valves which allow air to flow into or out of the air-springs in order to maintain a constant vehicle floo height as passengers enter or leave the vehicle. (Students have learned to defeat thi system by postponing their entrance to the vehicle to the very last second.) An overloa warning pressure switch operates if the passenger-load exceeds 1·4 tonnes (3150 lb) an the vehicle cannot operate until some passengers decide to leave the vehicle.

FIG. 16.12. Morgantown system. Car in platform showing safety barrier.

The steering system is based on sensing wheels (Fig. 16.13) which impinge on the guiderail and the side of the track. A mechanical bias spring is provided so that a force of up to 800 N (180 lbf) is applied to the guiderail. The power-steering cylinder is arranged to be hydraulically driven so as to provide a force countering the mechanical bias spring. The vehicle will follow the guiderail on one side or the other of the vehicle depending on the setting of a bias-switch cylinder which is responsible for the vehicle switching. This is preset on the vehicle when it passes "switching time" loops which determine whether the vehicle shall go to the right or to the left.

Disc brakes are employed, the calipers containing tandem piston actuators independently supplied. Either piston in the caliper assembly is able to actuate the brakes at full capacity but when both pistons are actuated as is the normal condition, then braking results are not additive. Brake cylinder valves apply pressure to the calipers over the range of 0·17 to 6·2 MPa (25 to 900 lb/in.^{2}). The emergency mode is created by the absence of a 28 volts d.c. signal to the brake amplifier which causes the servo-valve to supply up to 900 lb/in.^{2} to the calipers. The normal braking rate is 0·61 m/s^2 deceleration and emergency brake is 2·94 m/s. Brake energy is provided by the hydraulic system and the electrical control is by battery.

The electrical power system receives three-phase power at 575 volts from the guideway distributed through the vehicle at several different voltages. The main auto-transformer outputs 355 V and 61 V a.c. three-phase for the propulsion and control circuits, respectively. The 575-V input serves as a primary for the transformer as well as providing power

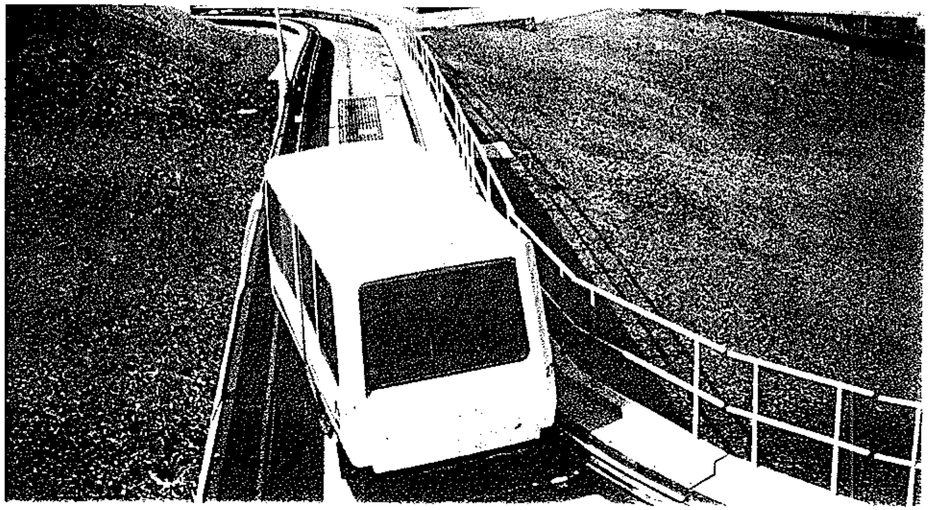


FIG. 16.13. Junction on Morgantown system.

for the hydraulic pump and the electrical control system; 28 V d.c. system charges th emergency batteries, the emergency lights, door actuators, and brake amplifiers as wel as the control relays.

Control system

The signalling system is referred to as the "collision avoidance system" and is used t provide safety even in the event of a failure of the primary vehicle control. It is based o a block system which transmits a tone to the vehicle when it is safe to proceed. This ton signal has a basic frequency of 10·2 kHz and is modulated at 50 Hz. This safe tone i suppressed in the block immediately to the rear of the block occupied by a train.

At junctions in the absence of trains an "off block" state exists at each leg of th merge. As a vehicle approaches the merging point it will be granted priority on the basi of "first come—first served". Thus the normally "off-block" will be turned on. Whe another vehicle approaches the merging point and priority has already been granted t the vehicle on the other leg, the normally "off-block" will remain for the second vehicl and the merge contact will be avoided. Two distinct logic paths are used, one interfacin the station computer, and another logic path utilises special purpose logic circuits t accomplish block control. Both paths must agree on block occupancy or safe tunes are turned off and the system operators notified. The entire guideway is divided into discrete sections; if a disparity between a logic path occurs safe tunes are suppressed only in the affected zone. Other portions are operated normally.

Guideway

Approximately 65% of the guideway is elevated, the remainder being at ground level. The running surface is concrete in which pipes are embedded to permit the circulation of hot water/glycol mixture when required to melt ice or snow. On at-grade sections there are heating pipes across the entire width of the guideway but the elevated sections have pipes confined to the two running pads. There are five heating zones each serviced by three boiler plants. Temperature transfer is accomplished by a 50% water/glycol solution pumped to the guideway at approximately 82°C (180°F) and 0·7 MPa (100 lb/in.^{2}) pressure. This automatically adds or subtracts the number of boilers required to meet load conditions and modulates the boilers with two firing rates. The fuel is natural gas.

Steering and electrical power rails are mounted vertically along the side of the guideway which is provided with emergency walkways for passengers in emergency. Curves

[illegible] each service being programmed to go direct to its destination. Power is distributed by copper busbars inset in a sheet attached to a plastic carrier and secured by an appropriate arrangement of welded studs. A smooth copper surface is intended to reduce brush wear and arcing. Current is picked up from the bus rails by a power collector consisting of sintered copper brushes.

The control and communication system is divided into the following functions:

the central control and communication sub-system;

the station control and communication sub-system;

the guideway control and communication sub-system.

The central computer carries out the automatic system management functions, receiving destination service requests from the station and transmitting commands thereto. Duplex communication with the stations is through asynchronous 2400 bits per second data lines. The station computer receives inputs from the destination selection units, provides passenger instructions and manages vehicle movements. Speed commands, station-stop commands, and steering-switch signals are received by the vehicle through an inductive communication loop buried in the guideway. Redundant computers are provided with automatic switch-over. The central control equipment permits the operator to monitor and control the entire system. The consol includes mimic display which permits the operator to monitor the progress of each vehicle on the system as well as a closed-circuit TV for system security and passenger safety. The station-control communication sub-system controls the vehicle station's operations in response to central supervisory commands. As far as the guideway control and communication sub-system

is concerned this includes the digital data cables, tone signal cables, capacity presence detectors and cable and hardware required to connect stations with the vehicles.

Four sets of loops are provided as follows:

1. The station stop loops transmit a signal to decelerate and stop the vehicles ±0·4 m (±16 in.) from the centre of the station platform access gates. The vehicles enter the "stop loop" at a speed of 1·2 m/s (4 ft/s) and are decelerated to a precise stop by the application of the brakes.
2. The switching tone loop generates the signal to command the vehicle to steer left or steer right. The vehicle receives a "switch command" at every guideway junction, either "merge" or "separate". The vehicle must verify that switching has been accomplished or it is brought to a stop. Calibration loops are provided to transmit a measured-distance reference. This is a non-vital system and it calibrates the vehicle's own measuring device.
3. The speed tone loops transmit brake commands and door commands as well as indicating a safe speed.
4. A further system of loops is used for receiving vehicle identification, door responses and statements. In Phase II the contact rail will be heated electrically.

The system has operated for over 1 million miles and has carried 2 million passengers.

Length of route	actual	8·7 km (5·4 miles)
	authorized	13·0 km (8·2 miles)
No. of stations	actual	3
	authorized	5
Maximum gradient	1 in 10	
Minimum radius of curves	9 metres (30 ft)	
Maximum speed	48 km/h (30 m.p.h.)	

Vehicles

Capacity	seated	8
	standing	12
Mass	empty	3·9 tonnes
	loaded	5·3 tonnes
Length	4·7 m (15·5 ft)	
Width	1·8 m (6·0 ft)	
Power supply	575 V three-phase	
Service braking rate	0·61 m/s^2	
Emergency braking rate	2·94 m/s^2	
Drive motor	52 kW (70 h.p.)	

Control

Collision-avoidance system based on loop antennae in track and on vehicle.
PDP-11 computers (Digital Equipment Corporation).

Mr. Philip H. Morgan of the Urban Mass Transportation Administration of the U.S. Department of Transportation has given the following figures for the cost of building the three-station system at Morgantown, including 45 vehicles. These figures include the cost of research, development, and testing of the system as well as that of building the guide-way and vehicles:

	$
Civil engineering construction	25,800,000
Vehicles	9,000,000
System communication and control	13,600,000
Engineering	11,400,000
Total:	$59,800,000

The V.A.L. rapid transit system

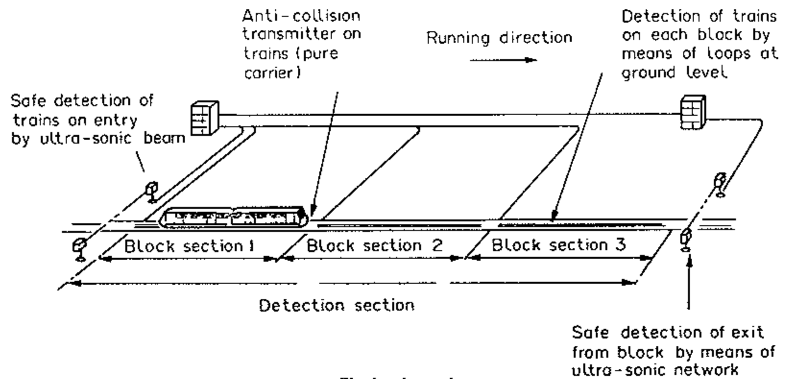
The first section of the Metro system at Lille was opened in the year 1982 and was operated with unmanned trains from the outset.[15] Control requirements are particularly

Fig. 16.14. The negative devices are situated at the entrance and exit of each section. They consist of ultrasonic wave transmitters and receivers mounted on each side of the track. In the absence of a vehicle each receiver receives a constant signal from the transmitter on the other side of the track. The presence of a vehicle leads to the absence of a signal, hence the designation of the devices as "negative" detectors. Vehicles are also detected as they pass over an induction loop located on the track. These loops are the positive detectors and respond to a pure carrier transmitted by the trains.

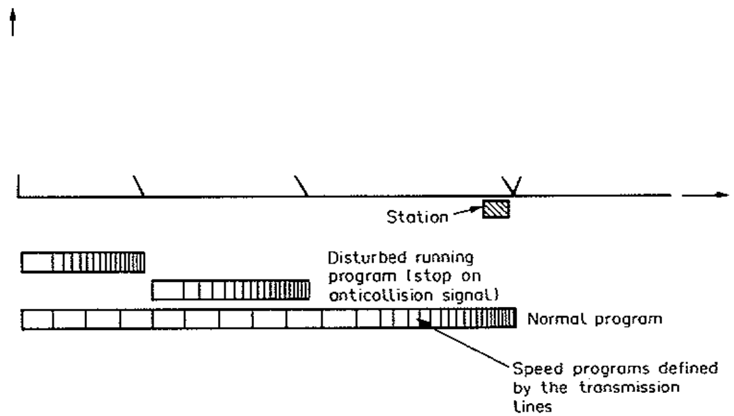
Trains are controlled by the two-wire transmission lines which are each transposed at intervals. One which determines operating speed is crossed at regularly spaced intervals corresponding to a reference time of 0·3 s, and the other embodies a braking programme to bring the train to rest at the end of a block. Trains normally operate up to speeds of 60 km/h (37 m.p.h.), but a higher limit of 80 km/h (50 m.p.h.) or a lower limit of 40 km/h (25 m.p.h.) can be imposed by the central control computer, depending on whether the train is behind or ahead of time, respectively.

The link between the train detection and train control is provided by control logic devices located in the control rooms at the stations. The occupancy of a block is determined from the train-detection devices and the appropriate instructions to the following train are transmitted through the normal transmission line. The transposition of the pair of transmission lines occurs at a reference time interval of 0·35 (corresponding to 5 metres at 60 km/h) which permits a very precise control of speed by onboard security devices. Acceleration is limited to 1·3 m/s² and jerk to ±0·65 m/s³. If the signal is removed from the control transmission line the train is brought to a stand using the emergency braking system.

Under normal circumstances the operating personnel do not take any action except for starting the operation in the morning and stopping it at night.



(a) Train detection.



(b) Control of train speed.

FIG. 16.14. Train detection and control—VAL system.

Although pneumatic tyres are used for support and guidance as in Paris, the V.A.L. system embodies several improvements. Firstly the steel safety-wheels have been discontinued, each bogie is provided with four guiding wheels but only two carrying wheels and a system of points has been developed which can be installed at any location on the route.

Normally vehicles are guided by the steering wheels interacting with vertical side guides. However at junctions sections of these guiderails have to be omitted and at this stage the vehicle is guided by the auxiliary disc shown in Fig. 16.15(a) which engages with the central pair of rails shown in Fig. 16.15(b). These rails are equipped with blades which function in the same manner as conventional railways. These rails are not continued beyond the junction area.

The U.T.D.C. intermediate capacity system

The high cost of tunnelling inhibits the construction of conventional subways except in the most densely used corridors and systems which are more economical in that they are situated above ground pose environmental problems. The use of pneumatic tyres provides one solution but, under unfavourable climatic conditions, expensive precautions are necessary such as electrically heating the guideway or covering it with snow-sheds.

Steel wheels running on steel rails pose problems of noise, particularly on curves, but after much consideration the Urban Transportation Development Corporation of

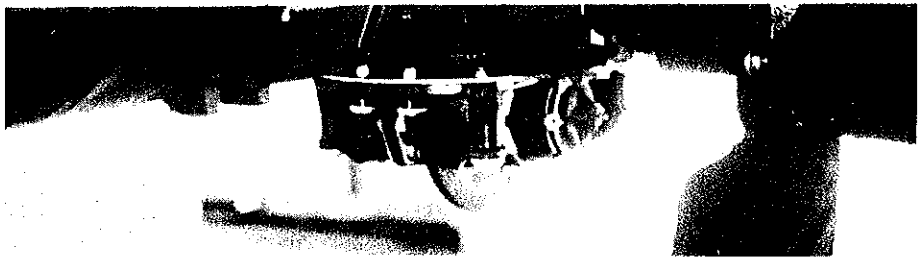
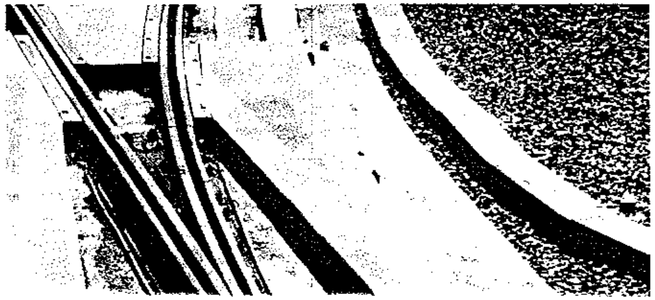


FIG. 16.15(a). V.A.L. System. Undercarriage showing steel disc which interacts with switching system.

Toronto have decided that the necessary technical means exist for overcoming the problems.[16] They ascribe the occurrence of objectionable noise and vibration to two causes: one, the action of guiding the wheels around curves, and the other due to the wheels (and rails) being damaged by skidding and sliding during driving and braking. The first problem is overcome by employing bogies which are so constructed that the axles take up a radial position relative to the curved track (see Chapter 14), and the second by using linear induction motors (see Chapter 17) for propelling and braking the vehicles. These motors operate by electromagnetic forces independently of the wheels which simply function as supports for the vehicles and may thus be expected to retain their shape.

The system has been evaluated on a test track 2·28 km long near to Kingston, Ontario, Canada (Fig. 16.16(a)), and is now being installed on a new Metro system at Vancouver.

It is capable of conveying from 5000 to 20,000 passengers per hour in each direction with trains varying from one to six cars. Each vehicle is 12·7 m long, 2·5 m wide, and weighs 13 tonnes. The linear induction motors are rigidly mounted on the trucks below the axles (see Fig. 16.16(b)) and embody a conventional three-phase six-pole series winding. The gap between the linear motor and the reaction rail is 10 mm, and different types of reaction rail are used along the system. A laminated form of construction is employed where maximum thrust is required as at stations and steep gradients and a cheaper form backed with solid steel is used where thrust requirements are less. The



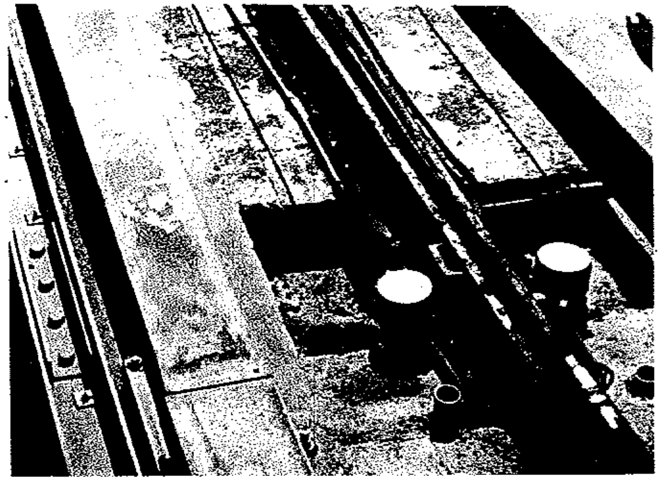
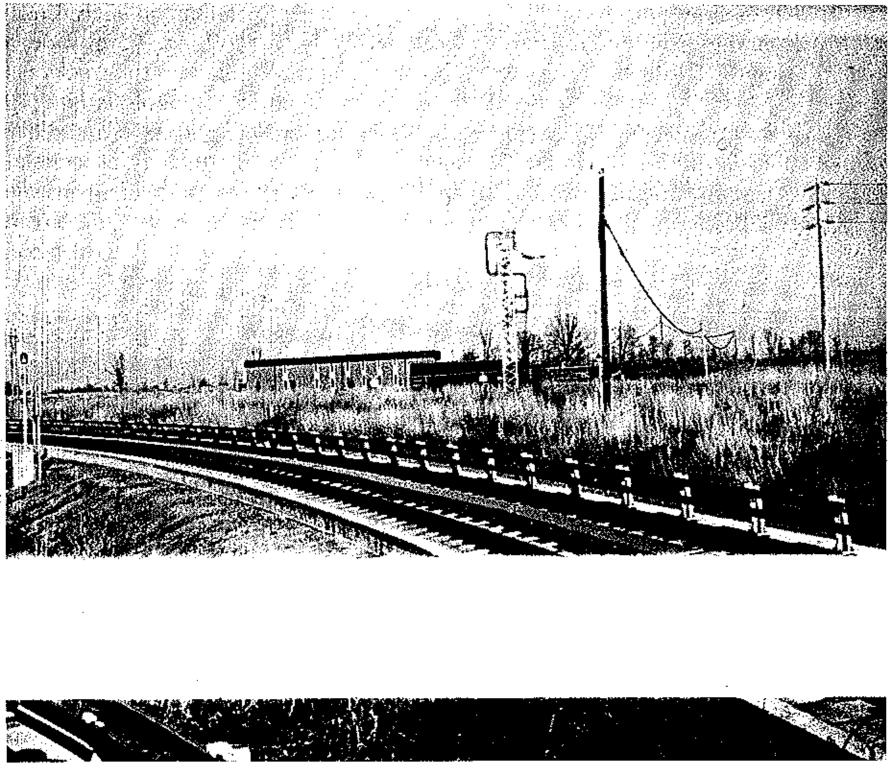


FIG. 16.15(b). Switching system. Steel disc is guided between rails in junction area only.

method of fixing the reaction rail which is composed of iron covered by aluminium is shown in Fig. 16.16(c).

The three-phase a.c. current is derived from 600 V d.c. supplied by a conductor rail situated at the side of the track. Conversion takes place in a transistorised pulse-width



(a) Test track.

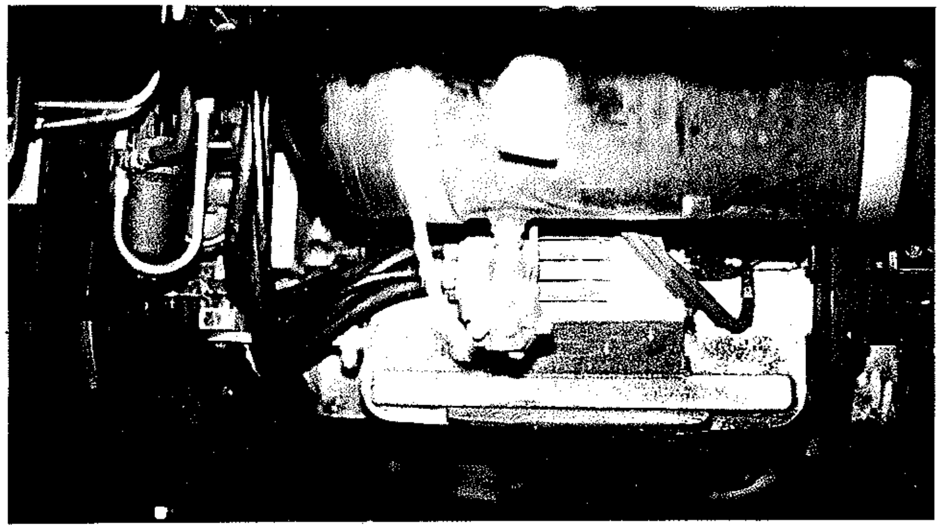


FIG. 16.16. UTDC Intermediate Capacity System. (b) View from below vehicle showing linear motor.

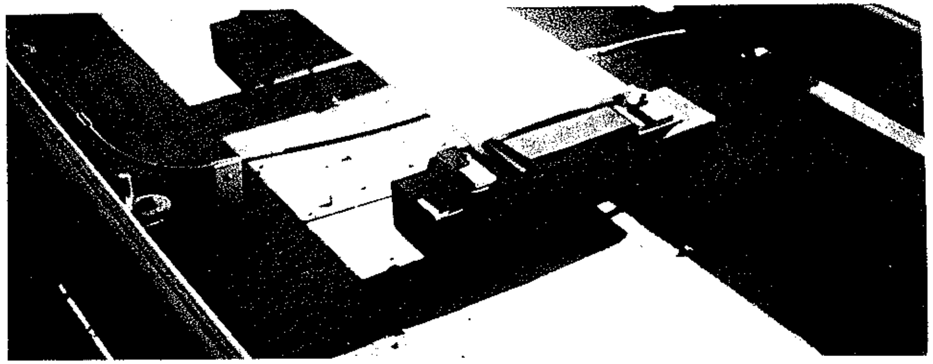


FIG. 16.16. UTDC Intermediate Capacity System. (c) Method of securing reaction rail.

modulated inverter. Current is interrupted at a basic 400 Hz which, when modulated t produce a basic output of 10 Hz, means that there are 40 high-frequency pulses to eacl cycle of output frequency. Those at the peak of the cycle are of maximum duration

varying slip or, as is more common, slip may be held constant as described above, and th inverter can be adjusted to provide the voltage and hence current required to produce th required thrust. System control is provided by the "Seltrac" system.

The SELTRAC and CORECT control systems

The Seltrac system, developed in Germany and applied to their Intermediate Capacit Transit System by the Urban Transportation Development Corporation of Canadւ recognises three levels of control within a hierarchical structure. These are the Systeı Management Centre (SMC), the Vehicle Control Centre (VCC), and the Vehicle Or Board Control (VOBC). Continuous communication between vehicles and linesid equipment, i.e. between the VOBC and the VCC, is provided by interaction of trair borne antennae with a continuous cable arranged as an inductive loop. This loop usually disposed so that one limb is laid at the centre of the track and the return coı ductor is attached to the foot of the right-hand rail. Thus, on single-line routes, each ha of the track can be used for communication with trains travelling in the appropriat direction.

In the case of the Canadian Intermediate system the centre of the track is occupied b the reaction rail of the linear induction motor, so the induction cable is laid close to th running rail and the antennae mounted on the outside edge of the vehicle truck. Th loops are transposed at 100-metre intervals in main-line railway working (COREC System) and at 25-metre intervals for intermediate systems (SELTRAC System). Th maximum length of a loop is 12·7 km and 3·2 km respectively. The cable, which is designed to withstand the hazards of the track environment including normal track maintenance, comprises a twisted copper core surrounded by polythene insulation and a protective outer sheath. The carrier frequency in the forward direction (from VCC to VOBC) is 36 kHz; and in the reverse direction it is 56 kHz. The modulation frequency is 0·4 kHz.

The VCC cyclically monitors the travel direction, position, and velocity of each train. The position of the train is continuously determined from the rotation of the wheels but is calibrated at 25-m intervals using the transposition of the cable. This information is transmitted to the VCC by telegrams having 41 bits. Continuous data transmission is possible at the rate of 600 baud. Using this data the VCC generates a reply telegram including the following:

1. Stopping distance—updated if necessary.
2. Maximum permitted velocity and target velocity at the end of the allowed travel distance. These limits are taken from track-section local speed limits stored in the VCC computers or temporarily imposed by the central operator.
3. Braking characteristics as related to local gradients.

3. a speed restriction; or
4. a station stop.

Similarly, speed adjustment commands are based on the most restrictive of the following:

1. civil engineers speed limit for track section or turn out;
2. a temporary speed restriction for protection of workers on the track; or
3. any speed restriction requested by SMC for schedule adjustment or energy conservation.

The VCC also provides for route-setting by transmitting commands to the appropriate lineside switch-control station.

The VOBC comprises two microprocessor-based units in a checked redundant configuration, together with data transmitter and receiver. It has the following specific functions:

1. Determination of train position, speed, etc.
2. Calculation of braking curve to comply with these commands.
4. Transmission of vehicle status information to the VCC.

Whilst not fail-safe, the system is protected by redundancy. Thus each microprocessor compares its output with that from the other and the signal will be cut off and the brakes applied if there is any discrepancy between the outputs.

Similarly, the VCC uses three computers in a checked redundant configuration. Telegrams from each computer are checked bit-by-bit from the other two computers. If a persistent error develops the faulty computer is isolated from the system and the safe

operation continues on the remaining two computers until maintenance action rectifies the failed unit.

The SMC is the central operating control enabling an operator (often known as a "dispatcher") to supervise the operation of the entire system. Mini-computers are responsible for most automatic train-supervision functions and for the acuation of visual displays. The SMC equipment is not concerned with safety because this is ensured by the VCC and the VOBC equipment.

References

1. HOEL, L. A., Public transportation problems and opportunities. *U.S. Department of Transportation Report No. DOT-TST-77-39*, March 1977.
2. MADIGAN, R. G., *Urban Rail Supporting Technology Program, Fiscal Year 1974*. Year End Summary, U.S. Department of Transportation.
3. PUBLIC TECHNOLOGY IND., *Integration of Para-Transit with Conventional Transit Systems*, October 1976.
4. WILBUR SMITH AND ASSOCIATES, *The Hong Kong Comprehensive Transport Study*, Hong Kong Government, 1976.
5. CAIATI, F. P. and TYSON, H. B., Status Report on General Motors dual mode transit, concept develop-

10. HOWARD, D. F., Tyne and Wear Metro—a modern rapid transit system. *Proc. Instn. Mech. Engrs.*, vol. 190, pp. 121–36 (1976).
11. WALDMANN, M. R., *Le Metro de Lyon. Technica No. 385*, pp. 4–6 (1975).
12. PERNOT, M. J., Materiał roulent et ateliers. *Ibid.*, pp. 42–48.
13. KANGAS, R., LENARD, M., MARINO, J. and HILL, J. H., *Assessment of Operational Automated Guideway System—AIRTRANS (Phase I)*, U.S. Department of Transportation, Transportation Systems Center, Kendall Square, Cambridge MA 02142, U.S.A.
14. *Innovation in Public Transportation, UMTA Fiscal Year 1976*, U.S. Department of Transportation.
15. FERBECK, D., Lille Metro—The VAL ATO System. *Proc. Railway Engineers Forum, 11th March 1982*, pp. 3/1 to 5. The Institution of Electrical Engineers, London.
16. SOBERMAN, R. M., Choices for the future. Summary Report. *Metropolitan Toronto Transportation Plan Review*, Report No. 64 (1975).

CHAPTER 17

Possibilities for the Future

17.1. The need for development of new transport modes

In looking ahead over the period of the rest of this century it is to be expected that demands for all forms of transport will increase and almost certainly will lead to problems of congestion. Professor Kolbuszewski[1] has indicated that these problems will indeed be formidable. Whilst numerous estimates exist of the probable population of England and Wales at the beginning of the next century, it is unlikely that the figures will be less than 56 million and may ind...

... this extra popula-tion require transport at least proportional to the numerical increase but the provision of the necessary goods, services and employment will lead to an even more important demand for freight haulage. It is likely that dispersal will lead to an even greater increase in ton-miles. As the separate centres grow, so more commuter transport will be required and it may be expected that a measure of dispersal will enhance the importance of inter-city transport.

Expansion of existing modes of road and air transport to cater for the increased traffic volume would present formidable problems of finance and land use.

It is shown in Chapter 2 that fundamental limitations exist to the capacity of any transport system and experience relating to motorways as reported by the Ministry of Transport[2] may be of interest in this connection. In the case of the M4 Chiswick Motorway, traffic volumes on the dual two-lane viaduct, which is 3298 yards long, have averaged 50,750 vehicles per 24-hour day with a maximum recorded flow on the west-bound highway of 35,335 vehicles per 24 hours, on Friday, 4th June 1965. This means flow-rates of 1500–1600 vehicles per hour per lane for the peak hours. This was an urban motorway subject to a speed limitation. Capacity limits for motorways are not officially expressed in vehicles per hour but the accepted capacity limit is 50,000 passenger car units per 16-hour day. Such a figure has already been achieved on the M1 and the rate of growth of motorway traffic is such that it is estimated that all our new motorways will be running at over design capacity very soon after their completion. As stated in ref. 2. "Traffic volumes above these limits can be absorbed and carried but only at the expense of free flow and speed."

The continued use of the motor car to and from work in reasonable conditions would require extensive road improvements which would have a detrimental effect on urban environment as is evidenced by the controversy surrounding the abortive proposals of the Greater London Council for the creation of a "Motorway Box". Whilst many improve-ments are necessary, the experience of other countries, notably the United States of America, shows that they do not provide a solution to the problem. Indeed, Buchanan[3] indicates the impossibility of considering cities of the size of Leeds, for example, ever being entirely served by private automobile transport.

The capacity of a right of way has been shown to be dependent on the degree of control applied to the separate mobile units as well as on their capacity. Thus buses and trains are able to provide solutions but only in those circumstances where traffic can be suffi-ciently concentrated to make full use of their capacity. The potential economy in land use and capital investment arising from the full use of railways is indicated by the space requirements in Fig. 17.1.

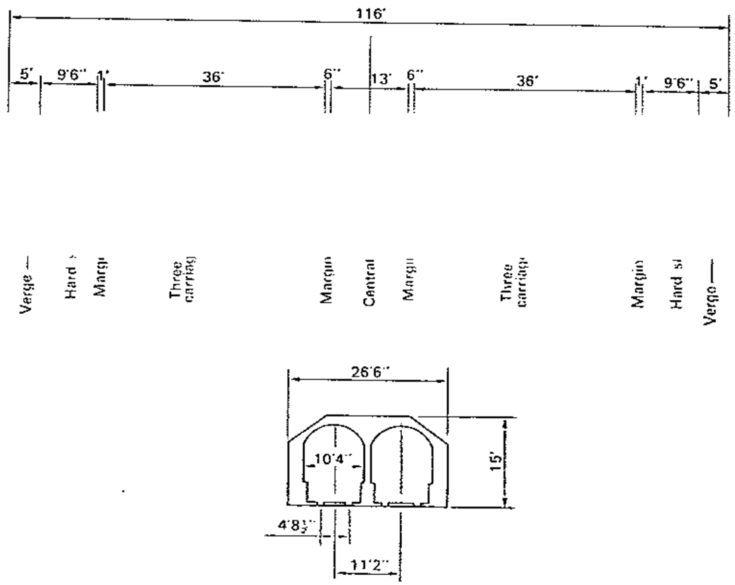


FIG. 17.1. Comparison of space requirements of road and railway.

Unfortunately the number of routes which can provide sufficient traffic fully to utilise an electric railway are few and these are seldom able to offer any individual passenger more than a portion of his total journey from home to workplace. It is not realistic therefore to consider any transport media in isolation. An integrated system is required that is able to concentrate traffic arising in low density areas on to high capacity trunk networks. Present technology offers the private automobile or a bus to feed passengers from their homes to the trunk electric railway.

Theoretical studies can be made in attempts to optimise the system both from the point of view of minimising capital investment and maximising customer satisfaction.

What has the engineer to offer? The basic technology outlined in each case has been with us for the best part of a century. Refinements are constantly being made by a process of evolution. The modern electric train with its high voltage a.c. supply, its solid state rectification, its electric pneumatic brakes and indeed the possibility of complete automatic operation, has reached a high state of development. None of these features will be apparent to the average passenger, although those fortunate to ride in the most modern equipment will be aware of a more comfortable ride than was available in the preceding generation. While the role of a prophet is an unenviable one and simple extrapolation frequently falsified by events, there are some matters about which one can be reasonably sure over the period between now and the year 2000. One is that there will be a continual increase in productivity in industry generally, leading to the labour element in any operation becoming more significant in its costing and the other, that there will be a continuing shortage of capital. Improving standards of living generally will lead to a demand for improved comfort in public transport if people are to consider using it at all. Public transport at present is contrasted with other industries in so far as it is labour intensive and not capital intensive. This results in great difficulty in maintaining economic viability in the face of rising labour costs and a less profitable pattern of usage arising from competition with the automobile, particularly at "off-peak periods".

The possible demands may be considered to be divided into two ranges, one an attempt

17.2. Possible improvements in control on the highway

The question of full automation has been discussed in Chapter 2, Section 3 and, whilst the capacity of manually operated systems is subject to fundamental limitations, the regularity and safety with which these limitations can be approached may be improved by better means of transmitting control information to drivers. It may be salutary to remark that the ordinary traffic signal, when considered on the basis of information content, is hardly more advanced than the disc and crossbar signal of Fig. 3.2.

Whilst control of individual vehicles by signals as in the "block" system is not feasible, safety and capacity may be enhanced by some approach to the railway signalling philosophy, particularly in giving advance warning of the state of the road ahead. It has now been decided that a more comprehensive system of signalling shall be applied to British Motorways.[4] Figure 17.2(a) shows a signal installed on the M4. The system is operated by a signalman, situated in a central office, who is assisted by an automatic stepdown procedure which serves to prevent abrupt or possibly confusing changes of signal aspect. Thus if a low speed restriction, say 10 m.p.h. (4·5 m/s), is selected by the signalman, the motorway signals first show "60" and then change to "40", "20" and "10" at 5-second intervals.

The system offers the advantage that, under unfavourable conditions of weather or congestion, the overall speed of traffic can be regulated to safe limits and free flow

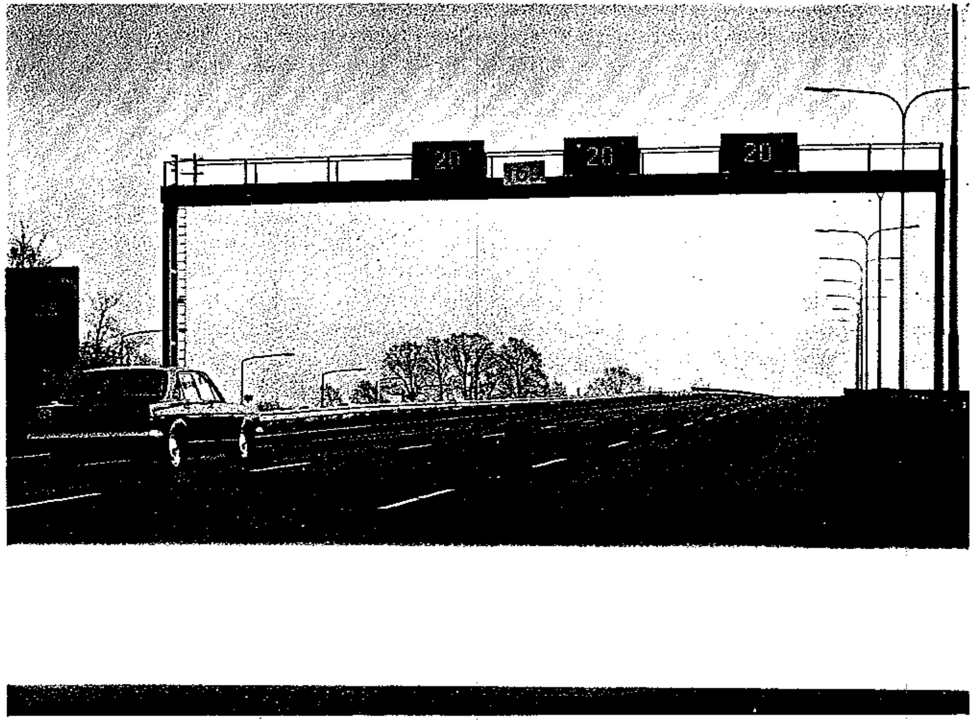


FIG. 17.2(a). Signal gantry. Signalling system of M4 motorway.

maintained. The signal shown in Fig. 17.2(b) is provided with amber lamps at each corner which flash alternately in horizontal pairs whenever an aspect is being displayed. These are intended to attract the attention of drivers at a sighting distance of 400 metres. In modern installations computers are provided to assist the operator in maintaining consistent signalling sequences and in monitoring the information available to him. Some possible legends for motorway signals are shown in Fig. 17.3.

The information available at present to a road driver is almost exclusively visual in character but the availability of cheap, short-range radio transmission, known in the U.S.A. as "Citizens Band" radio, enables supplementary information to be presented to him in audio form. In the DAIR concept[6] a visual sign minder reproduces roadside traffic signs on a display panel in the car and a "route minder" directs a driver along a predetermined route to his destination. The vehicle is triggered to receive an appropriate signal by passing over a "trap" or beacon situated at an appropriate position in the surface of the roadway. This consists of an array of permanent magnets arranged in sequence in accordance with a binary code. The first three bits contain the sign information and route information, where necessary, is provided by a further three bits as shown in Table 17.1.

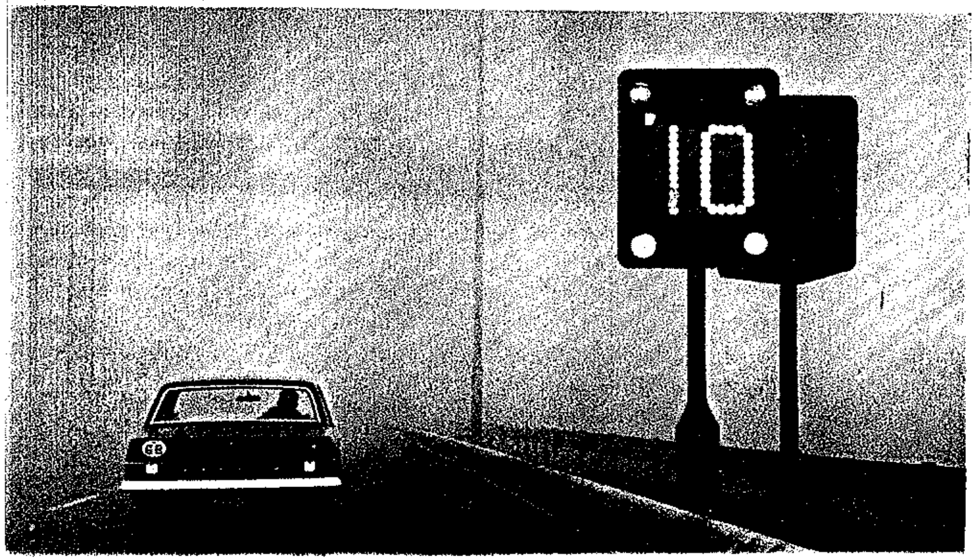
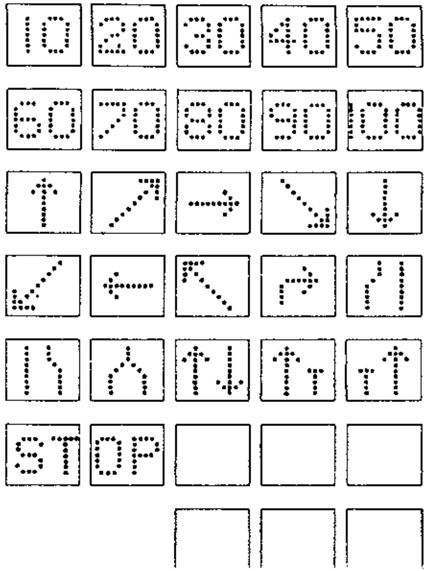


FIG. 17.2(b). Roadside signal. Signalling system of M4 motorway.

TABLE 17.1.
CODING PROPOSED FOR DAIR MAGNETIC ROAD BEACONS

Arrangement of code magnets	Message indicated by sign code	Intersection code number indicated by route code
S S S	Not used	1
W S S	Stop	2
S N S	Yield	3
N N S	Level crossing	4
S S N	Curve	5
N S N	Speed 15 m.p.h.	6
S N N	Speed 25 m.p.h.	7
N N N	Listen to recorded audio message	8



displayed on a display unit mounted over the dashboard. A 2·5 kHz audible "bleep" was sounded to alert the driver to the display.

When the code NNN is employed the CB radio receiver on the car is used to receive taped information from a low-powered transmitter placed on the road adjacent to the magnets. Range is limited to 300 metres which allows at least two 4-second audio messages to be received at 33 m/s (75 m.p.h.).

In addition to the roadside equipment, provision is made for the motorist to receive verbal messages of a general character relating to the weather or traffic conditions on the road ahead and for direct voice and coded communications between the vehicle and "Aid and Information" centres.

Buses and trams, being confined to set routes, may be equipped with means for exchanging information with a central control so as to provide continuous and comprehensive supervision of their operations. Data can be interchanged between vehicles and central control by radio supplemented by position transponders and bus scanning devices of the type shown in Fig. 8.2.

Schedules are computerised and a visual display of the actual position of each bus is provided at the control centre together with the scheduled position. This enables the controller to take effective action in the event of a delay due to traffic congestion or equipment fault. The sequence of traffic signals may be modified to reduce congestion or delay and passengers may be informed of the situation by loudspeakers or visual indicators within vehicles or at stopping places. Data can be automatically recorded for use in planning future service provision.

17.3. High capacity systems—effect of station stops

In the case of urban railways, the duration of station stops may limit track capacity. Figure 17.4 illustrates a situation governed by full automatic moving-block operation and wherein maximum acceleration is immediately followed by maximum braking. The optimum condition applies where braking time is measured from the instant where the previous train clears the platform. Equation (17.1) shows the capacity of such a railway.

$$\text{Capacity} = \frac{\text{payload metres}}{\text{seconds}} = \frac{l}{t_s + \sqrt{2l/a} + \sqrt{L/a}} \qquad (17.1)$$

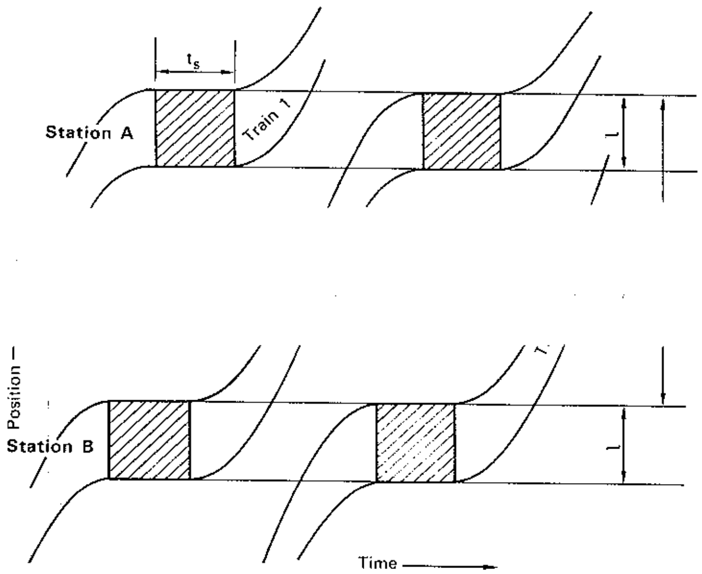


FIG. 17.4. Effect of station stops—optimum throughput with short platforms.

where l = length of train = length of platform,
a = both acceleration and braking rate,
L = distance between stops and t_s = duration of station stop.

The unit payload m/s is used to allow for different standards of passenger accommodation and different widths of vehicle, factors which are irrelevant to the present argument.

One method of increasing capacity would be to eliminate station stops altogether and to devise some method whereby passengers could be accelerated to join the train and decelerated after leaving it by some auxiliary equipment. An ingenious proposal for effecting this has been put forward by M. Bouladon of the Institute Battelle, Geneva.[7] He envisaged an integrator which collected a limited number of passengers, perhaps four, at walking pace and then accelerated them and aligned them longitudinally as

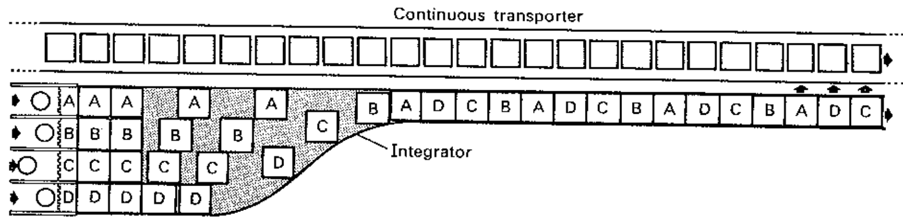


FIG. 17.5. Bouladon system.

shown in Fig. 17.5. By this time they would have reached a speed corresponding to t[h]e main train which was continuous and to which they could transfer by walking. The limit[a]tions to the Bouladon system were, of course, in the speed at which the main syste[m] operated and this would be determined by the maximum speed of the accelerating syste[m] plus a small differential determined by the amount of shock a person can tolerate i[n] stepping from one surface to another when they are in relative motion. This puts th[e] speed of the main conveyor at between 6 and 10 m.p.h. (2·8 and 4·5 m/s). Clearly th[e] accelerating unit must be considered as an increase in the effective area of the main systen

[illegible] per hour.

If an attempt is made to seek an optimum value of the ratio of station length to sectio[n] length (l/L) in equation (17.1) a limit is reached when $L = l$ so that platforms becom[e] continuous as in Fig. 17.6 from which it will be noted that the operation of succeedin[g] trains is now completely in phase.[7] This permits the revolutionary step to be taken o[f] lengthening the trains so that they become continuous. Capacity is now given by equatio[n] (17.2).

$$\text{Capacity} = \frac{L}{t_s + 2\sqrt{L/a}} \quad \text{or} \quad \frac{1}{t_s/L + 2\sqrt{1/aL}} \qquad (17.2)$$

If the comparatively moderate value for average acceleration and deceleration of 1 m/s a unit journey length of 800 metres and passenger loading (allowing for seats and gang ways) of six persons per metre is assumed, a throughput per hour of some 200,000 peopl[e] results, which is unmatched by any other form of transport. The average speed is compara tively low at 10 m/s but this compares well with present-day practice where overall speed of 7 m/s are tolerated. Increase of both capacity and average speed will follow fro[m] reduction in time of stopping, increase in rate of acceleration and increase in uni[t] journey length. Further improvement might be made by making the platform to mov[e] at walking pace as in an escalator so that the trains would not stop but only slow t[o] walking pace.

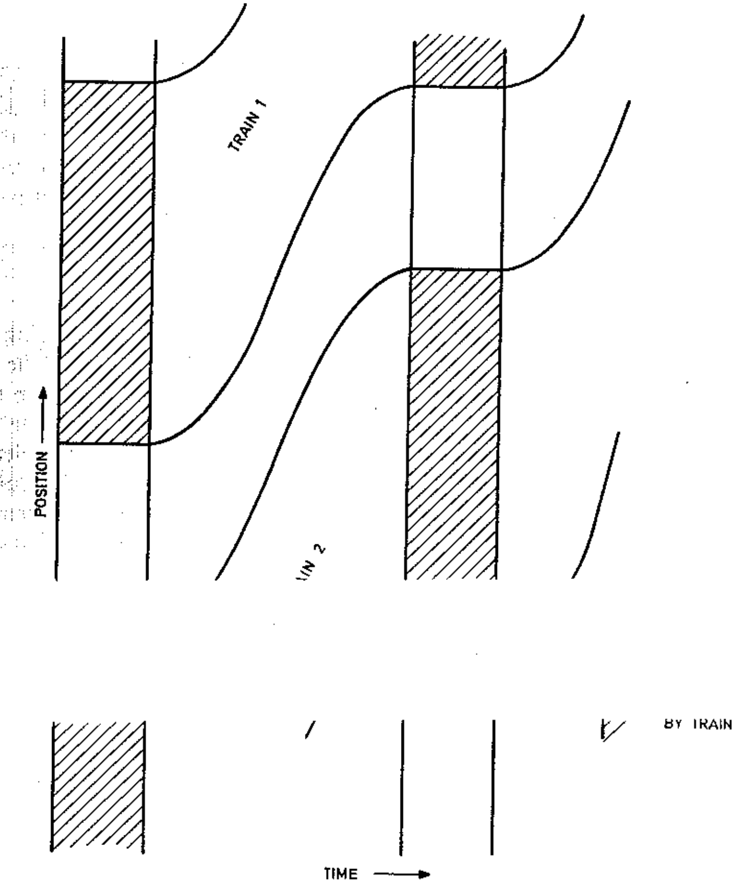


FIG. 17.6. Effect of station stops—continuous platforms.

Vast capital expenditure might be avoided by grouping routes so that existing double-track routes could be converted to single-track routes, the continuous platform occupying the space at present taken up by one of the tracks. Such a system would integrate closely with modern aspects of town planning in so far as there would be a much better distribution of entrances and exits throughout areas of congestion. As regards distance to be walked, passengers would be no worse off and in some cases better off than they are at present because of the wide choice of exits available.

The fact that trains are continuous considerably simplifies the provision of motive power. The trains themselves would have no motive power or control equipment, this being mounted in a fixed position. The linear motor is particularly appropriate in this connection (see Section 17.9).

Power supply presents a problem because of the absence of any diversity factor and it is likely that it will be necessary to integrate power-generating capacity with the railway in order to accommodate regular but extensive variations in demand.

The continuous train represents a hypothesis regarding the ultimate that might be achieved. The number of routes, however, that can justify such a generous provision of capacity must be limited, although this is not to say that the system might still be the cheapest solution even in circumstances where the capacity was considerably in excess of demand. It is possible that the next stage in a radiating network would be for conventional electric trains to feed the continuous train. These would, however, have to operate at a frequent headway in order to provide acceptable overall journey times and there would seem to be some case for examining this concept and providing some alternative technically. Since proposing the above system the author notes[18] that a similar scheme was put forward by Dalifol in 1880.

17.4. Small vehicles under computer control (CVS)

If it is assumed that the city system consists of a continuous train/platform central spine with laterals of lightweight, guided trains, then feeding points for such a system have to be considered. It is apparent that both the former elements are best suited for journeys of some distance. For the final distribution within the city a low capacity system is wanted but one that does have a high frequency to avoid the objection of passengers to long waiting periods and which can be joined or left at many points. One such system has been proposed by Blake[19] and is illustrated in Fig. 17.7. Small vehicles, electrically driven by

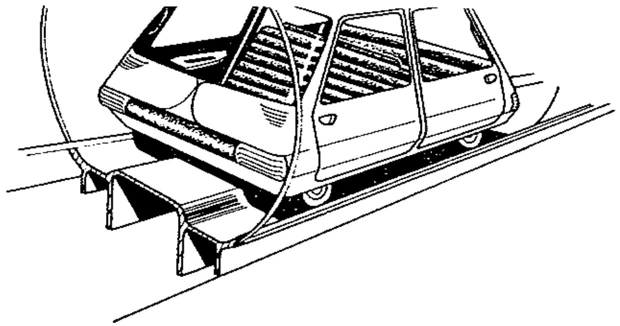


FIG. 17.7. Blake system.

a 50-Hz a.c. induction motor, are proposed to carry not more than four adults plus two children. An average figure of 1·5 adults can be taken for estimation of the effective capacity. The cars are carried on rubber wheels of about 15-inch diameter and would be driven automatically, being self-steering and self-routing. They would travel in enclosed tubeways on which would be mounted immediately above the vehicle path a longitudinal member that combines the function of guidance input as well as electrical power supply. Cars would travel at a constant speed of 35 m.p.h. (16 m/s) requiring 215 feet (65 metres) to stop and some 170 feet (55 metres) to accelerate to full speed. Power requirements would be 3·7 kW at 35 m.p.h. (15·6 m/s).

If the concept that vehicles must be separated by their braking distance is retained, then the capacity of the Blake system will be only marginally greater than that of automobiles on the roadway. The improvements will be attributable to the reduction in effective braking distance due to the elimination of driver's time reactions because of automatic control and because of the short vehicles used. However, because of the induction-motor drive would effectively limit the speed of all vehicles to a constant value, cars could be allowed to bunch together into trains so that a figure of 4000 cars per hour can be put forward as a conservative estimate. If the previous estimate of 1·5 passengers per car is used, this leads to a capacity of 6000 passengers per hour. The system cannot therefore compete economically with a fully loaded electric railway but constructional costs should be rather lower and the "taxi-type" service provided might command a proportionately higher level of fares.

TABLE 17.2.
CAPACITY OF VARIOUS TRANSPORT MODES

Mode	Capacity	Probable speed (m/s)
Blake system	6000	10
Continuous belt	12,000[2]	0·6
Continuous train	200,000[2]	10

Notes: (1) Occupancy taken as 1·5 persons per vehicle.
(2) Width taken so as to accommodate 6 persons/m length.

Professor Ishi and his colleagues[11] have also studied the possibilities of systems based on small vehicles but with particular emphasis on the application of computerised control. They propose a system based on coded induction loops of half-metre pitch laid in the track. The computer calls each car in turn every 500 milliseconds. In recent tests one second headway was maintained on four-passenger vehicles operating at a speed of 60 km/h. This implied a very high emergency braking rate, so that all seats were required to be backward-facing. A test track 4·8 km long has been constructed together with sixty vehicles.

Even at 1-second headway, the capacity of the system is only 14,400 passengers per hour and it is unlikely that the high braking rates which this would require would be feasible or acceptable to the public.

The control system departs from the "car following" method and embodies a new concept, the "Moving Target" method. In this method the guideway image is created in the computer along which moves a virtual point called the "Target".[12] The time-distance pattern of the motion of the target is determined in advance from dynamical

considerations. The motion of the vehicle is controlled so that its position on the rea guideway corresponds with the position of the target on the guideway image in the computer. The interval between two adjacent targets is made sufficiently long that the tw vehicles which are tracking them do not collide. Position error is therefore critical.

Another concept of automatic taxi relies on battery-powered electric vehicles that ar guided electronically. Control of forward motion and direction is intended to be effecte electronically as on the automatic highway described in Chapter 2. If the guidance rail i simply to be regarded as a datum against which the deviation of the vehicle from th required path is measured in order to provide an error function for input to a contro system, there is no need for it to be mechanical at all and a wire laid into the surface o the track could provide all the information required.

The capacity of various transport modes is estimated in Table 17.2.[12]

17.5. Inter-city transport—high-speed rail

The Shinkansen system

Growing congestion of roads and airports has led a number of countries to reconsider

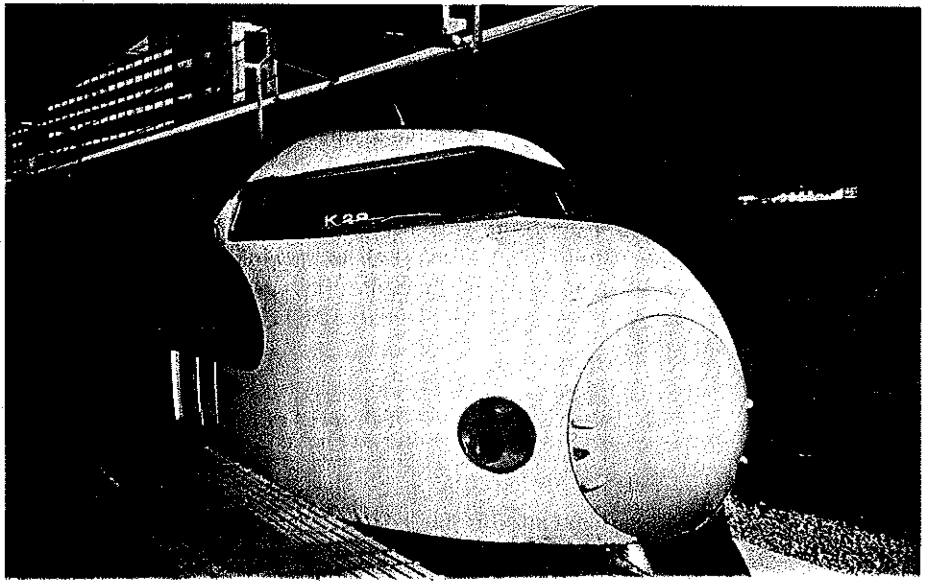
This was an entirely new system having a track gauge of 1·435 m (4 ft 8½ in.) which was superimposed onto an existing dense national system built to a gauge of 1·067 m (3 ft 6 in.).

Whilst the existence of the Shinkansen system with its maximum operational speed of 210 km/h is common knowledge, it is often assumed that this consists of prestige trains analogous to such European trains as the "Rheingold" or the "Mistral" which are classified as T.E.E. (first-class only, plus supplementary fares). Far from being exceptional, the high-speed trains were playing a major part in the national transport plan. A second extension to the Shinkansen route has now been brought into service, giving a continuous run of 1176 km (730 miles) from Tokyo to Hakata. The frequency of the service is a matter of comment. In a typical hour, trains leave Tokyo at the following minutes after the hour: 04, 12, 16, 24, 28, 40, 48, and 52. Only two of these make the full journey to Hakata which takes 6 hours 56 minutes with six intermediate halts.

There are two types of train, fast and semi-fast, known as the "Hikari" and "Kodoma" respectively. These trains, one of which is illustrated in Fig. 17.8, comprise 16 cars. In the case of "Hikari" trains two vehicles known as "green cars" are provided and in the "Kodoma", one. These correspond to first-class in Europe.

The small proportion of first-class accommodation is instanced as an indication that the service is supported by the general public as distinct from the more affluent business community. Most coaches carry 110 passengers so that the capacity of the line operating south from Tokyo is 14,000 passengers per hour in each direction. A corresponding figure for passengers in private cars on a six-lane motorway is 8000.

The development of the Shinkansen system is part of a national plan to link the islands



which make up Japan into a single entity. The extension to Hakata has already provided a link with the island of Kyushu and a tunnel on the projected extension to serve the northern island of Hokaido will have a maximum length of 53·86 km (33 miles) of which 23·3 km (15 miles) are under the sea. Several bridges are planned to join the island of Shikoku to the main island of Honshu. Figure 17.9 shows a schematic map of the effect of the provision of high-speed travel on the isolation of remote communities.

It is intended to increase the maximum speed of the trains from 210 to 260 km/h (130 to 143 m.p.h.) and a trial run completed in 1972 reached a speed of 286 km/h (178 m.p.h.).

The cost of the 393 km (244 miles) extension of the Shinkansen line from Okayama to Hakata during the years 1970 to 1975 was 648,400 million yen. This includes 222 km (138 miles) of tunnel. This yields an average cost of 1·65 million pound sterling per single track kilometre.

Table 17.3 presents particulars of the Shinkansen trains in comparison with the High Speed and Advanced Passenger trains of British Rail and the T.G.V. of the S.N.C.F.

Extension of the Shinkansen route north of Tokyo was opposed on environmental grounds; notably on account of noise and vibration affecting residential property. A special track was built in a non-sensitive part of the proposed route in order to experiment on various forms of noise and vibration insulation and it is understood that satisfactory solutions have been found.

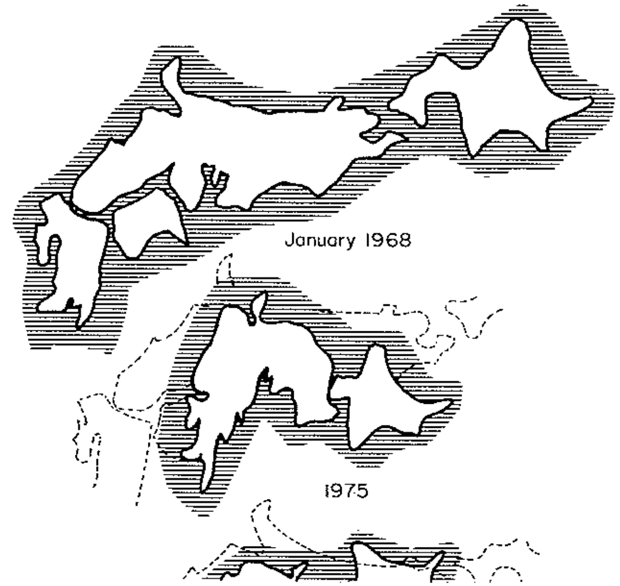


FIG. 17.9. Effect of Shinkansen trains on isolation of remote areas in Japan.

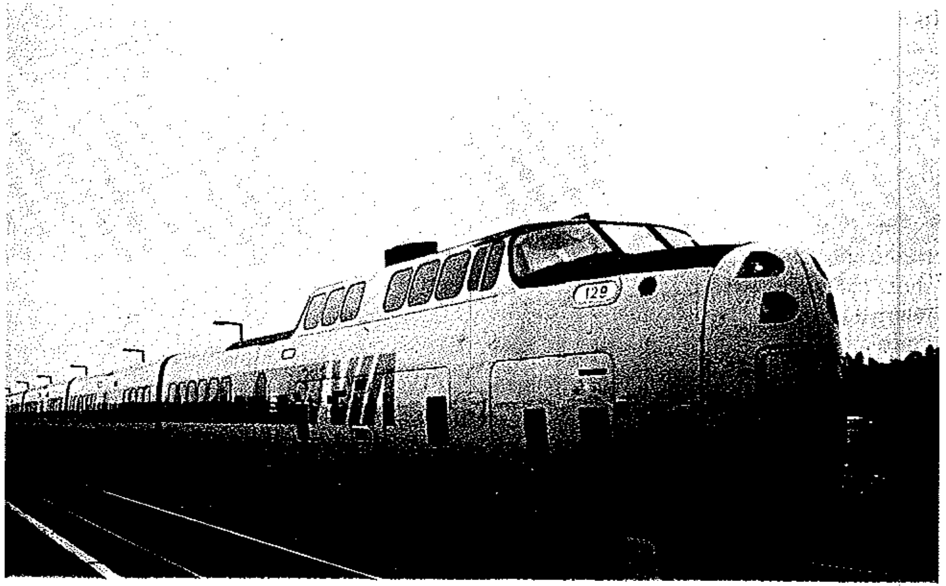
Canadian and U.S. turbo trains

The first of a series of turbo-trains was completed in 1967. One version which was extensively tested in the U.S.A., particularly on the service from New York to Boston, was fitted with a pendular suspension which permitted an increase of 30% over the maximum speed allowable on curves for conventional vehicles. Although this was a passive system, as opposed to the active system of the A.P.T. referred to later, it was extremely comfortable. Some sensation could be experienced by a critical observer on entering and leaving curves but in general passengers were not affected by negotiation of curves at speed. There were three cars having a total weight of 128 tons which were propelled through mechanical transmissions by five gas turbine engines of the aircraft type. The total power was 1490 kw (2000 h.p.). Maximum speed in traffic was 200 km/h (125 m.p.h.) although 274 km/h (170 m.p.h.) had been obtained in test when additional power was installed to give a total of 2460 kw (3300 h.p.).

The Canadian version which operated between Toronto and Montreal had seven cars and was not fitted with pendular suspension (Fig. 17.10(a)). In both versions a single axle was provided between each car which are guided by links from each of the adjacent

TABLE 17.3. H... NS

Designation	High-speed train	Advanced passe...	Shinkansen	Paris/Lyon
Code name	H.S.T.	A.P.T.-P/E	—	T.G.V.
Speed (km-h)	200	200 up to 250 la...	210	260
Motive power	Diesel-electric 2×2250 h.p.	Electric 25 kV. ... 8×1000	Electric 25 kV. 60 Hz 64×250	Electric 25 kV. 50 Hz (also 1·5 kV d.c.) 2×3000 kW
No. of passenger cars	5 to 8	12	16	8
Length of cars (m)	23	21 articulated	25	18·7 articulated
Suspension	Swing links -damped-air secondary springs	Special dynamic hydro-kinetic mechanism	Conventional with air secondary springs	Articulated bogies with top of secondary suspension close to c.g. of body
Power drive	Bogie-mounted motors transverse cardon shaft	Body-mounted n... longitudinal c...	Transverse cardon shaft	Body-mounted motors and helical gear reduction unit
Brakes	Disc + scrubbers	Hydro-kinetic ar... blocks	Electro-dynamic and pneumatic	Electrodynamic, discs and shoes
Start of project	1970	1967	1959	1968
Date in service	1975	?	1964	1982
Speed in trials (km/h)	230	245 (Gas turbine prototype A.P...	286	380
Braking	Both trains can stop from 200 km/h in same dist... conventional trains from 160 km/h		Electro-dynamic brake applied to all axles down to 50 km/h. Pneumatic brakes only at low speed.	Rheostatic brake on motored axles. Disc and shoe, brakes on carriage axles



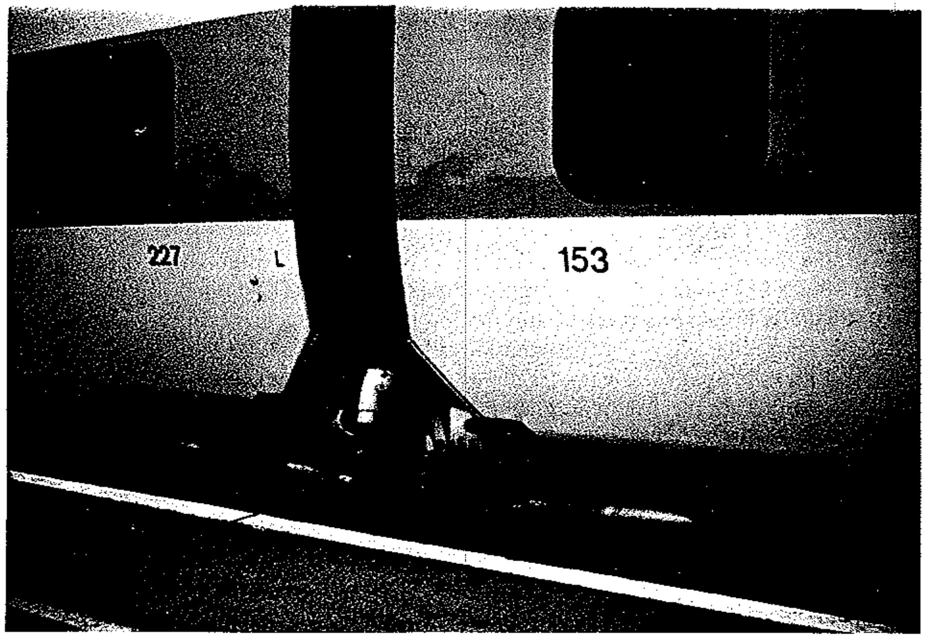


FIG. 17.10(b). Suspension of Canadian turbo train.

bodies and which are loaded through air springs (Fig. 17.10(b)). The power installed in the Canadian version was 1200 kw (1600 h.p.) and the maximum speed was stated to be 200 km/h (125 m.p.h.) although this was not normally required by the schedules actually operated.

The Talgo trains

The principle of pendular suspension has also been applied to a series of high-speed trains operating in Spain. These trains are unique in so far as the wheels are mounted on short axles independently of each other. They are caused to take up a radial position relative to the track by linkwork as in the case of the turbo trains. Pivotting of the car body under the action of centrifugal force occurs due to the differential compression of air-springs, as shown in Fig. 17.10(c). The vehicles of the Talgo train are constructed to

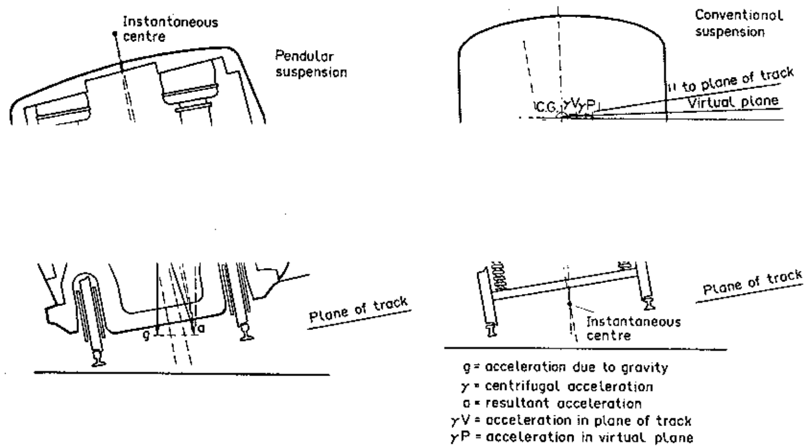


FIG. 17.10(c). Suspension of "Talgo" train compared with conventional (primary) suspension.

a low profile (maximum height above track level of 3·28 m (10 feet)) and are all trailers. They are hauled by locomotives at speeds of up to 200 km/h (125 m.p.h.).

A particularly interesting piece of automation is the arrangement on certain trains for an automatic change of gauge between the standard international gauge of 1·435 m (4ft 8½ in.) and the Spanish gauge of 1·668 m (5 ft 3 in.). The train is pushed through the small hut containing the fixed equipment situated between one railway system and the other, as shown in Fig. 17.11, and coupled to a waiting locomotive which then pulls the remainder of the train through the apparatus. During the gauge-changing operation the weight of the vehicle is carried on fixed rails at each side of the track area through brackets attached to the axle boxes (Fig. 17.12). The wheels are constructed so that they are free to move along the axles unless locked in one of two positions by bolts. These bolts can be withdrawn in a downward direction by specially shaped rails which are installed on each side of each axle. Once the bolts have been withdrawn the wheels are free to move axially and roll along guide rails which are no longer parallel to each other

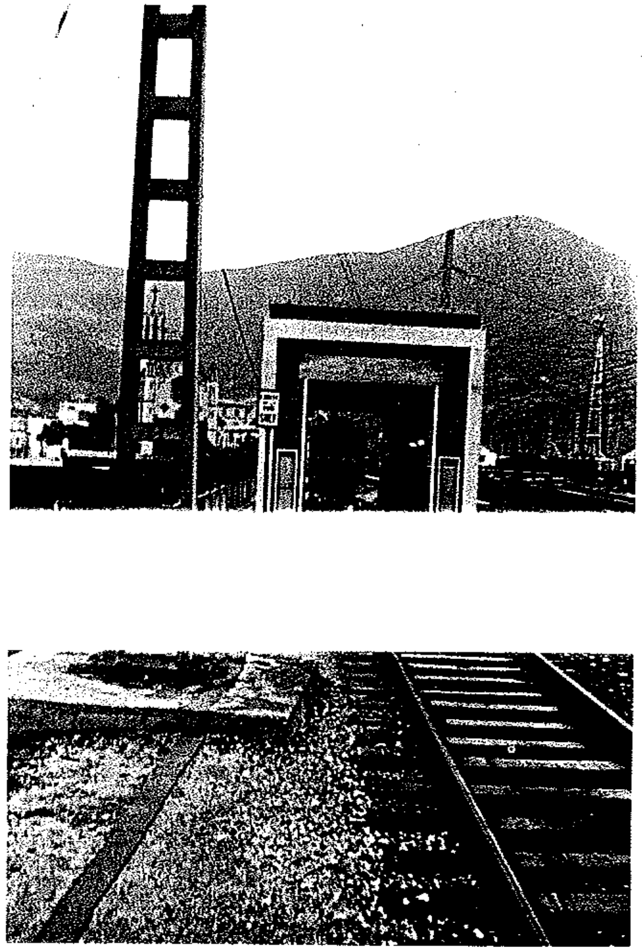


FIG. 17.11. Installation at French-Spanish border for automatic gauge changing.

but which lead from one gauge to the other. Once the wheels are rolling on parallel rails corresponding to the new gauge, the bolts are restored to the locking position and weight restored to the wheels as the vehicle passes beyond the support rails. It will be noted that all the ground-mounted equipment is entirely static. The operation of changing gauge is carried out at speeds ranging from 5 to 15 km/h (3 to 10 m.p.h.).

The Advanced Passenger Train (A.P.T.) and High-speed Train (H.S.T.)

British development has been based on two projects which were carried out in parallel as shown in Table 17.3. These were known respectively as "Advanced Passenger Train"

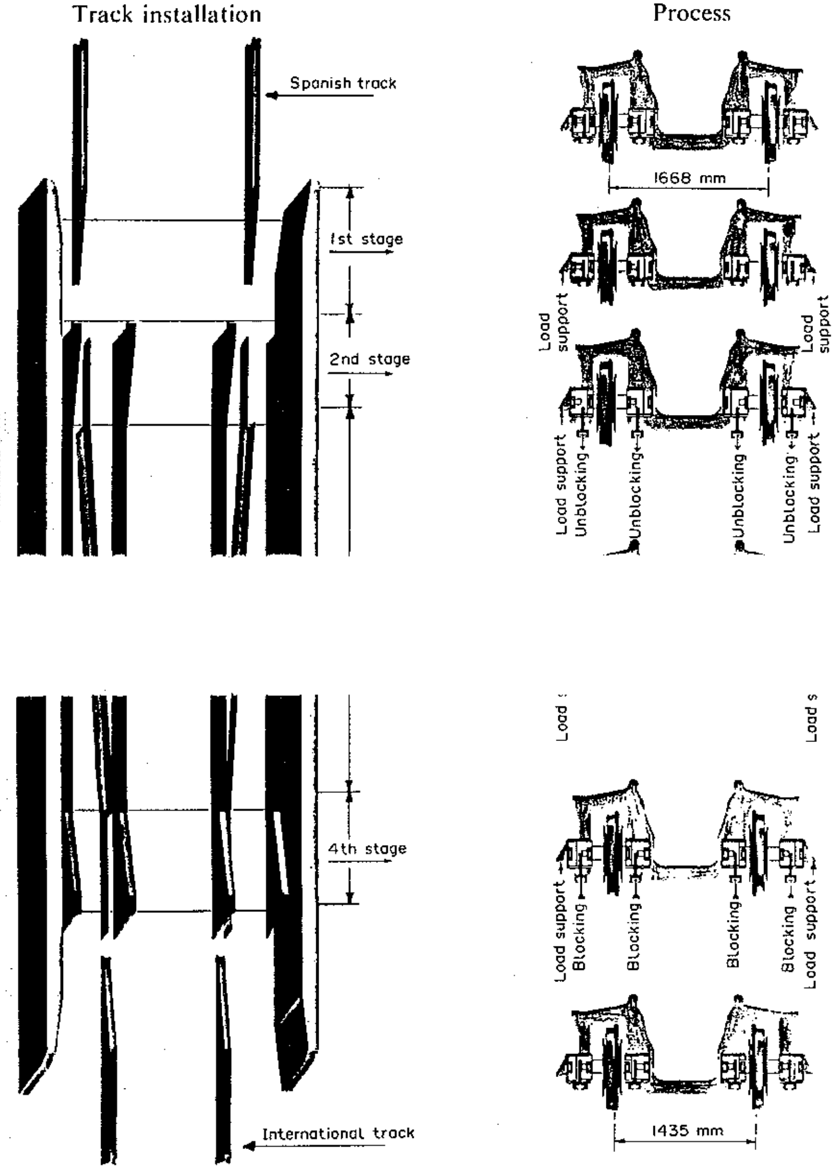


FIG. 17.12. Method for changing gauge automatically.

and "High-speed Train". As will be seen from the table the Advanced Passenger Train project was commenced in 1967 and embodied advanced concepts such as the hydro-kinetic tilt mechanism to facilitate more rapid negotiation of curves without detriment to passenger comfort and special dynamic design which would minimise the adverse

interactions of wheel and rail. The initial trials employed gas-turbines for propulsion but the first commercial application will be to an electrified railway.

The high-speed train project was commenced some three years later and the design was based on elements which had been well tried in existing practice, notably the diesel-engines and air-conditioned passenger coaches which had operated extensively at speeds exceeding 160 km/h (100 m.p.h.). These trains have been widely applied to the non-electrified main lines of British Rail and have been well received by the travelling public.

The A.P.T. has been redesigned to operate on the electrified main lines and, at the time of writing, is still under development. Its main features are as follows:

Streamlined train profile—to reduce power consumption, pressure pulses and air turbulence noise.

Articulated train configuration—weight saving and better track stability on curves.

Active vehicle tilt mechanism—each vehicle can tilt independently of its neighbours up to an angle of 9° under the control of a closed-loop hydraulic system. In addition to the benefits to passenger comfort on curves, the tilt mechanism retains vehicles centrally on their suspensions. By this means the transfer of weight from inner to outer wheels is limited.

Advanced bogie suspensions. Parameters are selected to give dynamically stable operation

Low centre of gravity. Maintains safety margins during high-speed curving and under extremely high winds.

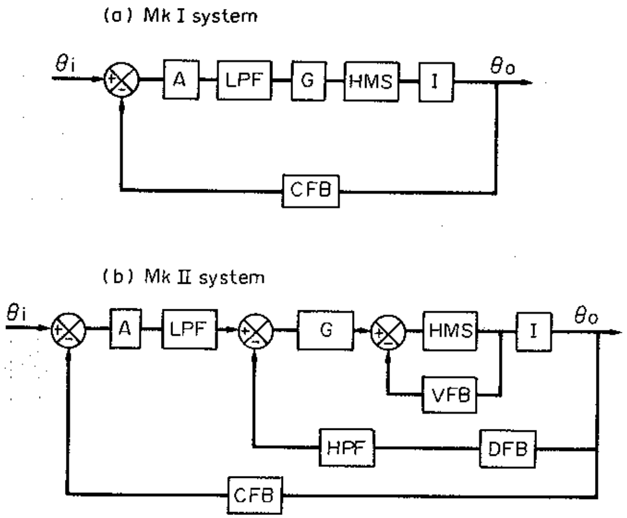
Hydro-kinetic braking. These deal with the very high levels of power dissipation when braking from high speed and comparatively light compact apparatus. Conventional braking is applicable at low speeds.

Transmission. 750 kw is supplied to each power axle from body-mounted motors through a body-mounted gearbox, a cardon shaft, a bogie-mounted gearbox and finally a flexible quill to the axles. This allows unsprung mass to be minimised.

The most novel feature, the tilt mechanism, has been the subject of much development. Three successive developments of the tilt-control system are shown in Fig. 17.13.[13] The specification required that the vehicles should tilt by up to 9° in response to a 5°/s rate of change of cant deficiency and an 8·3°/s rate of change of cant plus cant deficiency without passengers experiencing more than 10% g acceleration (0·98 m/s²). An electro-hydraulic tilt mechanism was fitted to achieve the desired tilt rates. Originally an accelerometer was mounted on a tilting bolster within the bogie. This provided a mechanical feedback loop with a low-pass filter to eliminate the effect of track irregularities. Stability margins were shown to be inadequate in relation to the required response capability.

A second loop was added to the system embodying a tilt-displacement transducer coupled with a high-pass filter. This produced some improvement but stability margins were still inadequate in relation to the required tilt response rates.

The problem was solved using an open-loop feature in which the accelerometer is mounted on the bogie frame so that its reading represents the input to the system. Tilt



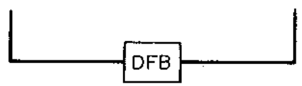


FIG. 17.13. Tilt control circuit for A.P.T.

is measured within the closed loop by a displacement transducer so that tilting of the body occurs by a definite amount for a given intensity of acceleration. Advantage of the "open-loop" situation of the accelerometer has been taken to introduce an anticipation feature. The accelerometer is mounted on the leading bogie of the preceding vehicle which gives an approximately half-second anticipation.

Although no exceptional signalling system is envisaged on the proposed introduction of the A.P.T. on the London–Glasgow route where the maximum speed will be 200 km/h (125 m.p.h.), a special speed advisory system will be introduced to ensure observance of the many speed limits on the line. (These limits are of course set at a value which is

higher than that applicable to conventional trains.) A continuous display of permitted speed is provided in the cab with automatic brake application if this is ignored. The system is based on 2000 transponders which have been installed at appropriate places which are coded to represent the speed restrictions.

Because it was feared that the disturbance of the catenary system by one pantograph would lead to unstable operation of any other pantograph on the same train, the two power cars of the prototype train are placed together at the centre of the fourteen-vehicle unit. Proposals for the production version envisage that only one power-unit will be provided at one end of the train. Ten trailing vehicles will be hauled or propelled and the maximum speed potential will be limited to 225 km/h (140 m.p.h.).

European experience—the T.G.V.

Early development in Germany may be typified by the operation during the 1965 International Transport Exhibition at Munich of a regular service between Munich and Augsburg which attained a maximum speed of 200 km/h.

In France, as early as the 1950s the S.N.C.F. carried out trials with electric trains during which maximum speeds of over 300 km/h (148 m.p.h.) were attained. This has been [illegible]

specially constructed high-speed railways. In contrast to the policy underlying the A.P.T. which was based on the design of the train to suit existing track, new tracks have been built especially adapted for high-speed running. The minimum curvature adopted has been 4000 m (13,000 ft). Because of the high power-weight ratio characteristic of trains designed for high speed, steeper gradients than usual for main line railways have been incorporated. The maximum is 35% (1 in 28).

During high-speed operation only one pantograph will be used on each ten-vehicle train unit necessitating the installation of 25 kV roof line connecting the two pantographs. Another pantograph will be provided for use on existing railways which enter the existing terminals at Paris or Lyon and which permit the trains to operate (at reduced speed) over contiguous routes. The length of the train unit, 200 m (657 ft), is such that two units may be coupled together without serious interaction occurring between the two pantographs.

Very serious consideration was given to the introduction of a high-speed link connecting Paris with Lyon based upon the Aerotrain concept of Berlin. An important consideration leading to the selection of the conventional system was that with this system it was possible for the T.G.V. trains to continue their journeys over existing railways to important destinations in Switzerland and on the French Riviera.

Current-collection at high speed—the servo pantograph

Power is supplied to electric vehicles either by rigid rails, usually energised to about 800 V d.c., or by flexible overhead conductors which are supplied with current at voltages which vary from 1500 (d.c.) to 50,000 V (a.c.). The most common system employed for new construction is now 25 kV single-phase at either 50 or 60 Hz.

Where speeds are low, street railways for example, a single copper or bronze wire is supported under tension above the track. For higher speeds a catenary is provided to support the weight of the overhead system and the contact wire is supported therefrom by various systems of droppers as described in reference 14. Because of geographical constraints on the railway the "wire" must be erected at different heights from place to place. At level-crossings (grade-crossings) for example, the wire must be high to provide adequate clearance for road vehicles to pass underneath although when passing through tunnels it will be mounted as low as possible to minimise constructional cost of the railway.

On straight track the route of the wire is zigzagged from side to side between the supports to distribute wear on the current collector carried by the vehicle. On curved track it forms a series of chords so arranged that the wire is always within a prescribed distance from the centre line of the track.

The current collection shoe, usually composed of graphite impregnated with metal, is held against the underside of the contact wire by a linkwork known as a pantograph which is intended to exert a constant pressure as the head rises and falls in consonence with changes in the height of the wire. It must also be noted that the pantograph is itself

gravitational force acting on the wire and fittings together with the negligible vertical components of the horizontal tension in the wire (this tension is usually about 10,000 N). Naturally the system will be more compliant near the centre of a span than near the

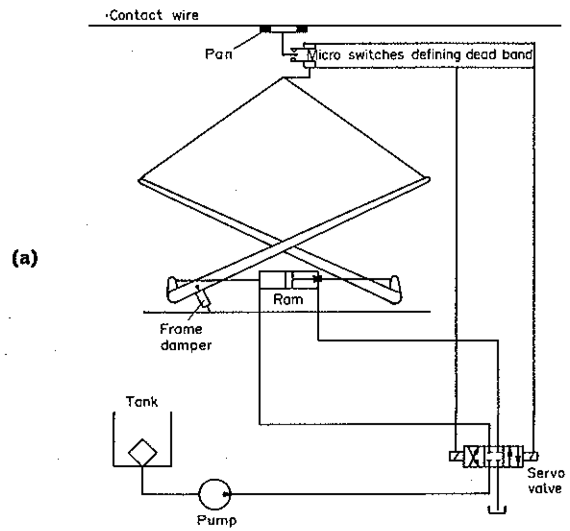


FIG. 17.14(a). Arrangement of power drive Servo actuated pantograph.

supports so that, upward force being constant, the collector will rise and fall as the vehicle passes along the track. As speed increases the inertia forces arising from the acceleration of the head in the vertical direction will be added to the static forces.

The object of design of the overhead system in order to optimise the quality of current collection (to minimise loss of contact) was therefore to secure as uniform a value of compliance as possible throughout the span. This was achieved by placing the droppers more closely together towards the centre of a span. The effect of this was augmented by artificially "sagging" the line where the compliance was lowest so that the path of the head was as nearly as possible in a straight line.

The pantograph frame was usually of massive construction which could only follow variations in the wire at a sluggish pace and it was usual to make the current collection assembly (the pan) as light as possible and to connect it to the pantograph proper through a resilient suspension.

Application of control theory to pantograph design has improved performance so that higher speeds (up to 50 km/h increase) can be run with the same quality of current collection. Referring to Fig. 7.14(a), the current-collecting head or "pan" is forced against the contact wire by a secondary suspension consisting of springs arranged to give a substantially constant static force over the whole range of secondary suspension travel. This suspension was made as frictionless as possible and the mass of the pan was limited

pantograph frame as shown schematically in Fig. 17.14(b). A damper was arranged to act on the frame. Because of the existence of the "dead zone" the system was non-linear but could be studied using the "phase-plane" diagram described in Section 1.4.

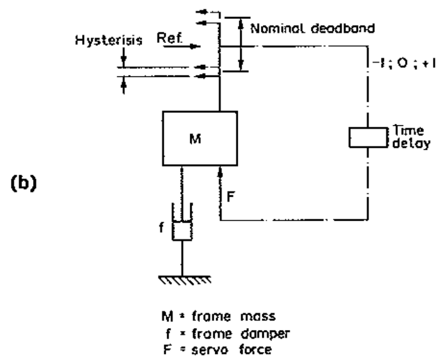


FIG. 17.14(b). Model for phase-plane treatment.

Referring to equation (1.13) M can be assigned to the mass of the pantograph frame, f to be the coefficient of the damper and F is constant in this case. The equation then becomes

$$M\frac{d^2x}{dt^2} + f\left(\frac{dx}{dt}\right) = F. \tag{17.3}$$

Representative values from reference[15] are as follows: $M = 25$ kg, $f = 200$ N/m/s and $F = 150$N. Writing $y = dx/dt$

$$M\dot{y} + fy = F \tag{17.4}$$

Within the dead band $F = 0$ and $\dot{y} = -f/My$ ($= -8y$ ms^{-2}) and outside the dead band

$$\dot{y} = \frac{F - fy}{M} \quad (= [6 - 8y] \text{ ms}^{-2})$$

A phase-plane plot based upon these expressions characterises a stable system which speedily comes to rest. In practice, however, the electro-hydraulic servo-valve possesses a finite operating time and the switches which limit the dead zone are subject to hysteresis. (Values reported by B.R. are as follows: servo-valve operating time 0·020s, limit switch hysteresis 0·005 m.) Figure 17.14(c) shows a phase-plane diagram taking these two factors

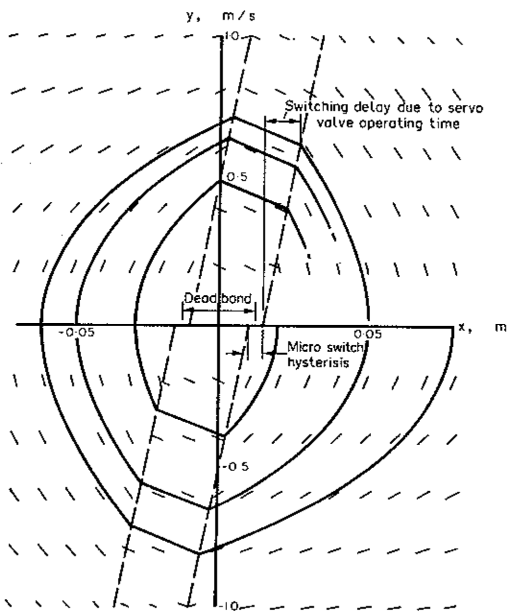


FIG. 17.14(c). Phase plane diagram for pantograph system.

17.6. Studies of new forms of urban and inter-city transport

As mentioned in Chapter 16 the U.S. Department of Transportation has engaged extensively in research and development in transport technology, particularly in establishing the Transportation Systems Centre at Cambridge, Massachusetts, and the Transportation Test Centre at Pueblo, Colorado.

A report from the M.I.T. published in 1969 recommended that new technology was necessary[14] including the application to traction of linear induction motors.

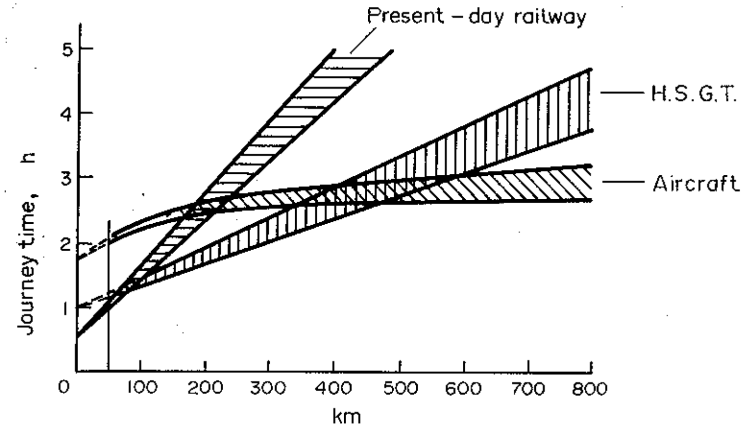
In Great Britain an interdepartmental working party[15] based on the Department of the Environment and the Department of Trade and Industry was set up in 1969 and issued its final report in 1971. It recognised that a London to Manchester hovercraft system could be profitable by the year 1985. For other routes the A.P.T. was recommended because it avoided the economic burden of new tracks. With the benefit of hindsight it may be suggested that the working party underestimated the difficulty of developing new trains within the constraints imposed by the existing system.

A major study was commissioned in the Federal Republic of Germany in 1969 by the Hochleistungs-Schnellbahn Studiensgesellschaft B.H. which is referred to by the initials

mented with estimates of traffic volume which were related to general forecasts of increasing population, of increasing gross national product (which has been shown to be a very sensitive indicator of the demand for transport) and for the continuing tendency of people to move towards larger concentrations of population. As a result of these estimates of traffic it was decided that a route of the configuration shown in Fig. 17.15(b) could be justified by the traffic offering. A more complete and more complex system in the form of a figure-of-eight was rejected because it was not considered that this would attract sufficient traffic during the remainder of this century.

A number of alternative technical configurations were studied including joint use of some tracks with the D.B. However, it became clear that these would prevent either system from operating to peak efficiency. Out of a large number of configurations which were studied in a preliminary manner, four were finally selected for detailed comparison as shown diagrammatically in Fig. 17.16. In Mode 1 all traffic was carried in one type of train. These trains would either be conventional, that is running on wheels, or might be magnetically levitated or air-cushion supported depending on the speed range which, for the purpose of the study, was varied from 225 to 375 km/h.

Under System 2 there would be one set of tracks but the trains would be separated in so far as the vehicle-carrying trains would go at 275 km/h and the passenger-only trains at 375. Under Scheme 3, which would appear to be the one which most countries are planning, i.e. the carriage of vehicles on conventional trains at a speed of 225 km/h, was supplemented by an entirely different system of high-speed ground transport (either magnetically or air-cushion supported) with a speed of 500 km/h. The fourth system, which was introduced late in the study, is equivalent to System 1 but operating over the lower speed range of 200–300 km/h using entirely conventional technology.



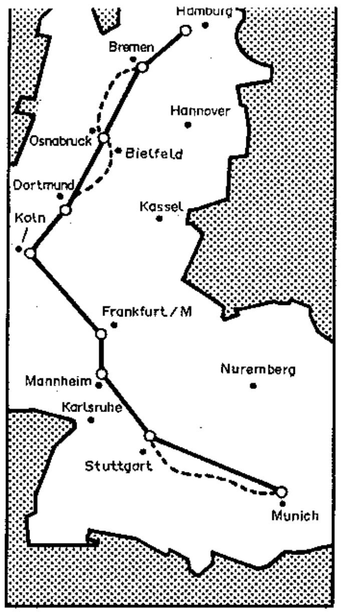


FIG. 17.15. (b) H.S.B. studies proposed route.

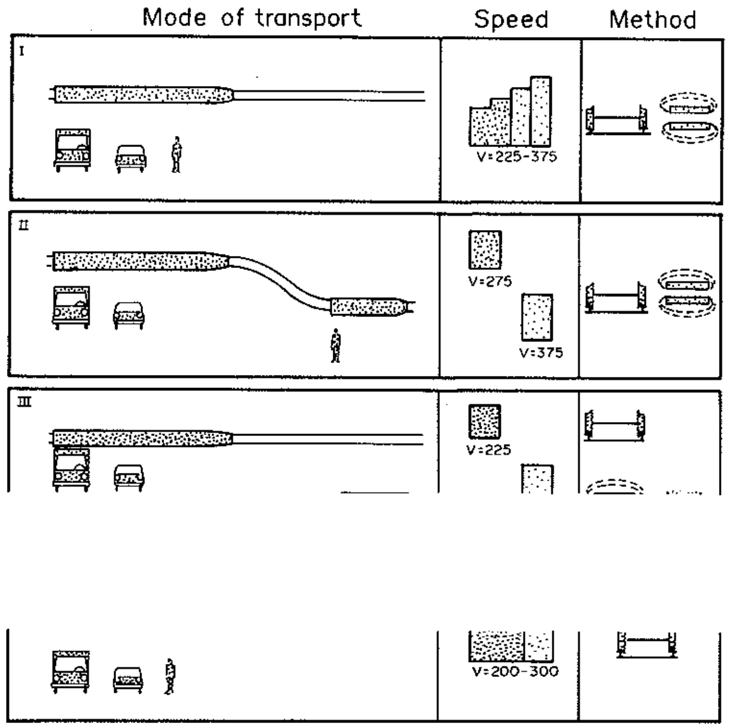
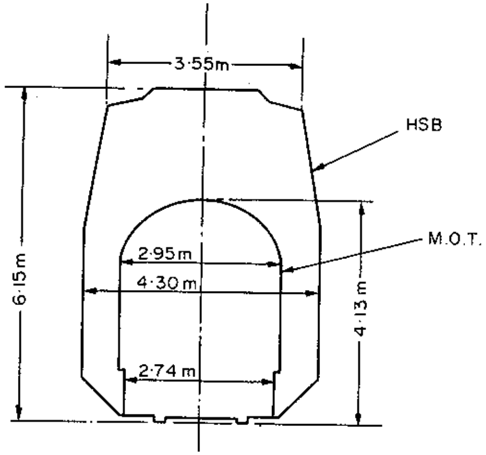


FIG. 17.16. Different modes and combinations for H.S.B.

- I Single route—wheeled—linear motor.
- II Single route—wheeled—linear motor.
- III Separate systems. Conventional for freight—Maglev for passenger.
- IV Single conventional wheeled system.

Professor Baseler had worked out a scheme for the German Federal Ministry of Transport in 1962 in which very large vehicles were envisaged which operated on a broad gauge (3 m). However, the H.S.B. studies finally concluded that the most economic system would be to retain standard rail gauge but to adopt an extended loading gauge so as to permit lorries to be carried in covered wagons. A number of ways of fitting lorries and cars into wagons were considered but it was finally decided that the loading gauge indicated in Fig. 17.17 would be suitable. In this case a single lorry would be used but there would be room for two cars to be carried one above the other. It was envisaged that the power-equipment would be located under the floor, irrespective of the system of propulsion and support finally chosen, and that the vehicles would be loaded from the rear whilst being unloaded from side doors at the front of the train.



HSB and BR loading gauges

of trains which operate between two stations only. Because there are nine stations, this means that seventy-two trains would be required. An operational system was then proposed embodying a train service of 1-hour intervals; trains would be up to 400 m long and as many as thirty lorries would be carried at one time together with cars and passengers. In the system finally selected there was only one class of the train operating at a speed of 300 km/h. The system was fully costed on the basis of detailed estimates of track and operating costs for each system studied. The capital cost varied from DM 15.5×10^9 (£2800 million) for conventional tracks at 225 km/h to DM 24.5×10^9 (£4400 million) for air-cushion or magnetic track suitable for a speed of 500 km/h. In addition to the direct financial implications, cost–benefit analysis was carried out from the point of view of the economy as a whole. The method adopted was to produce two estimates of the total expenditure on transportation, one if the scheme were carried out, and the other if it were not. The difference between the two therefore was indicative of the benefit as shown in Fig. 17.18. The factors taken into account in the cost–benefit analysis were the advantages due to the time-saving of the commercial and private users, saving in operating costs, saving in road costs, saving in accident costs. The effects of the existence or otherwise of the H.S.B. on rail and air transport was also quantified.

Factors which were not included in the comparison were improvement of the environment, greater ease of meeting transportation needs, repercussions on other transport systems, effect on location of industry, effect on city planning, regional development and saving in occupation of land. It will be seen from Fig. 17.18 that the benefit continued to improve throughout the speed range. Nevertheless, costs also increased with speed which

leaves an optimum value at about 300 km/h. This is of course well within the range of conventional rail technology (A.P.T.). Moreover, this speed was considered by the authors of the report to provide a sufficient margin over the speeds of the road and to be sufficiently competitive with air to meet the requirements of the region chosen for study.

A report on Tracked Air Cushion Vehicles issued in September 1970[16] is evidence of early Canadian interest in innovative guided transport. In contrast to the view of the British interdepartmental working party, a new infrastructure was envisaged because it was predicted that the growth of freight traffic on the critical corridors would be so great that eventually there would be no room for passengers. A three-volume report issued in 1980[17] by the Canadian Institute of Guided Ground Transport compared three solutions for operating the Toronto–Ottawa–Montreal Corridor as follows: a 450 km/h magnetically-levitated (Maglev) system, a 260 km/h electrified high-speed railway (H.S.R.) operating on a dedicated double-track in a new right-of-way (cf. the T.G.V.) and a 200 km/h diesel-electric intermediate-speed railway (I.S.R.) operating on partial double-track in a combination of new and existing rights-of-way (cf. the H.S.T.). Conventional aircraft (CTOL) and short take-off and landing (STOL) were also evaluated in comparison.

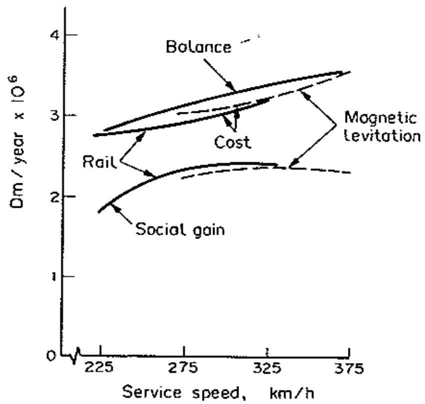


FIG. 17.18. Cost–benefit comparisons, H.S.B.

After much detailed planning of routes and various assumptions of travel demand, energy cost, etc., it was concluded that either the Maglev system or the H.S.R. would capture much of the existing air patronage. On the basis of the 1978 Canadian dollar, a ticket between Montreal and Toronto would cost $22 for H.S.R. and $28 for Maglev, which would compare with $57 computed for air by the same procedure. At 200 km/h

(125 m.p.h.) the diesel-powered system (I.S.R.) would not penetrate the air market to the same degree but would cost $2 more than the H.S.B. Because of the lower infrastructure requirements it might however be useful as an intermediate application.

The cost for the 540-km (335-mile) system was £3 billion for Maglev, $1.5 billion for H.S.R., and $1 billion for I.S.R. The estimated net benefits from the various systems exceeded the investment costs by impressive amounts as follows: $1.26 billion for Maglev, $1.5 billion for H.S.R., and $500 million for I.S.R.

17.7. Air-cushion support

There is no evidence that the conventional railway has reached the limit of speed of which it is capable. Nevertheless, higher speeds necessitate precision of rolling elements and above all, make particular demands on the alignment of curves. Many of the problems of the track and suspension arise from the need for a massive wheel to make contact with the rail, which is never entirely free from imperfections. The basic payloads are usually distributed over the floor of the vehicle and, at present, this loading is concen-

distribute the load over a wide area of track surface thereby avoiding the constructional expense arising from the use of wheeled vehicles. The sensitivity to small-scale track irregularities is eliminated because it is no longer necessary for a massive component to follow the track intimately and friction is virtually eliminated albeit at a price of the energy required to maintain the air cushion. Air-cushion systems are attractive, not only because they permit surfaces to be constructed to wider tolerances but because they need only be sufficiently well finished to withstand aerodynamic forces. Most important of all, they would not be subject to any form of wear.

The basic principle, of course, is that the air beneath the vehicle is raised to a specific pressure to withstand its weight by being fed from an external source. Because of the need to arrange a clearance between the vehicle and the road bed during high-speed operation, some air will escape and will actually be replaced by the power system, so that the power consumed must represent a direct loss. Because this power is independent of speed, the loss per ton-mile is inversely proportional to speed so that these systems become more economic the higher the speeds involved.

There are three possible applications of the air-support system, the difference being the manner in which the flow of air from the surface is restricted. These are illustrated in Fig. 17.19. The simplest type, based on the conventional aerodynamic bearing (Fig. 17.19(a)), relies on a viscous restraint of the air as it passes between two closely aligned parallel surfaces. Such a system requires close conformity with fixed and moving surfaces in order to provide maximum constraint for a minimum quantity of air. It is considered that such a system is unlikely to be competitive with other systems because of the fine tolerance

required, but the introduction of compliant material as in Fig. 17.19(b) may overcom this difficulty.

Two other systems do not rely on a close alignment but derive their resistance to th outflow of air from the transfer of pressure into momentum in passing through th restriction between the pressure area and the external atmosphere. The simplest arrang ment is the plenum chamber (Fig. 17.19(c)), which can be fitted with a flexible skirt t enable the clearance gap to be reduced to a minimum. The "hovercraft" (Fig. 17.19(d) is also based on the momentum to pressure relation and employs high-speed jets t impinge on the pressure area in an inwards direction so that the initial momentum ca be converted into pressure.

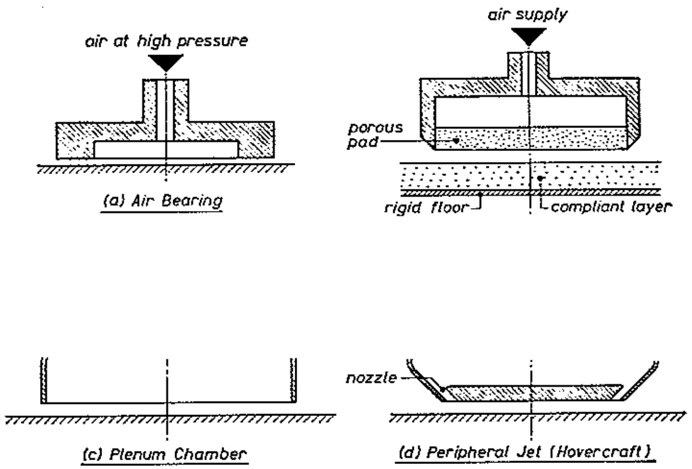


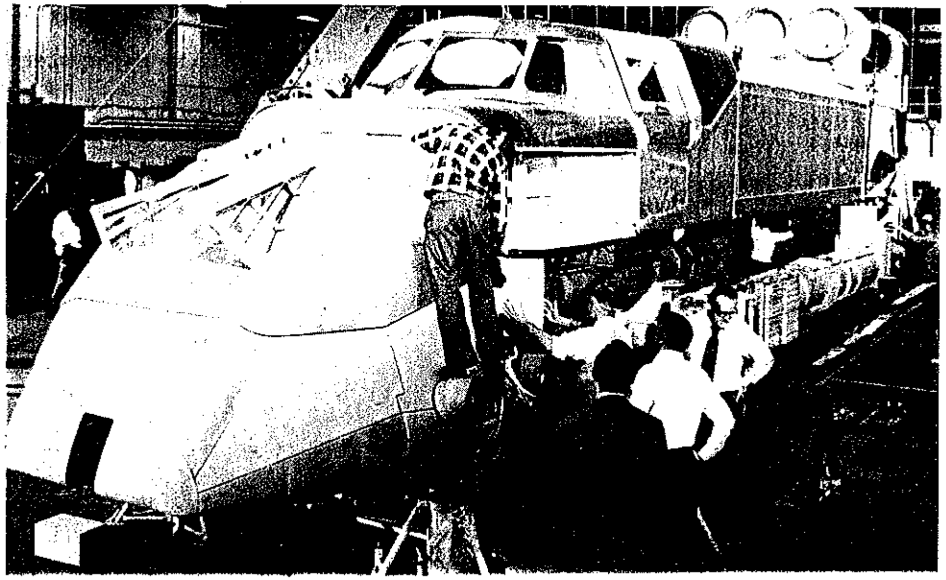
FIG. 17.19. Forms of air suspension.

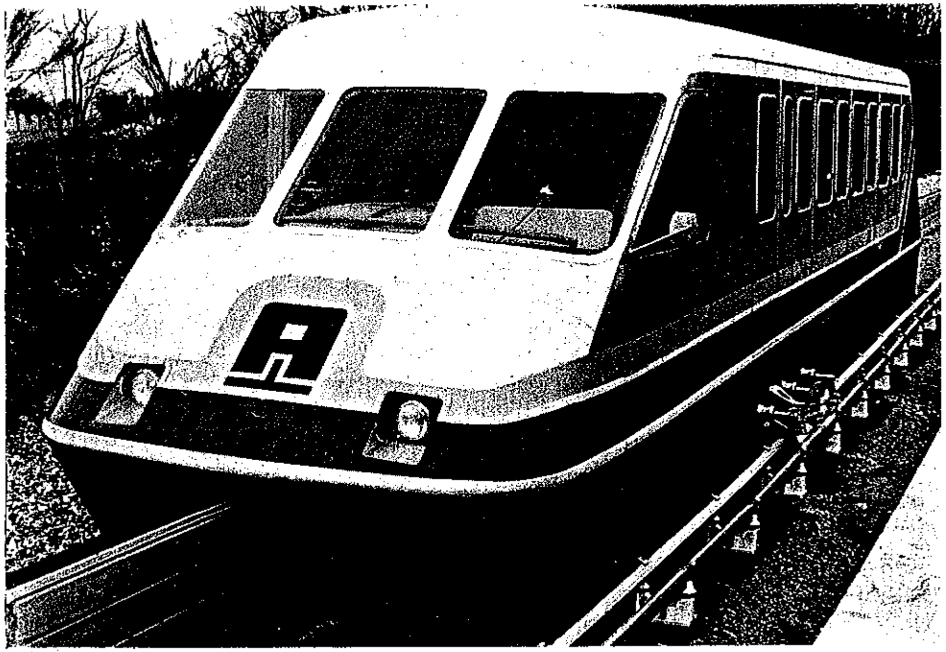
A number of configurations for air cushion support and guidance are possible an developments of hovertrains based on the principles of Fig. 17.19(c) and (d) respective are shown in Fig. 17.20. The "Aerotrain" developed by the late M. Bertin in France ru on an inverted T-shaped track. The centre spine provides guidance and embodies a aluminium reaction rail for the linear motor (see below).

The research vehicle shown in Fig. 17.20(a) runs within a trough-shaped track. It shown under construction at the works of the Grumman Aerospace Corporation Hicksville, New Jersey, for testing on the High Speed Ground Test Centre of the Depa ment of Transportation, Pueblo, Colorado.

17.8. Linear induction motors

The function to be fulfilled on a vehicle by its power unit is the provision of linear thru rather than the rotary torque provided by the conventional electrical motor. For norm





(b) Suburban "Aerotrain" developed in France. (Courtesy of Sté Bertin.)

FIG. 17.20. Types of air-cushion vehicles using double-sided linear motors.

speeds it is customary to gear such motors to the axles with a step-down ratio of something like 3:1. If, however, we envisage higher speeds this gear ratio will be reduced as a result of the limitations imposed on the conventional motor by centrifugal force and even more by commutation. Therefore as speeds increase, the attraction of a system of providing direct thrust increases and, moreover, the substitution of electro magnetic for friction forces should ease the problem of the design of the motive power unit from the point of view of acceleration and traction as well as removing one of the main hazards of braking.[18, 19]

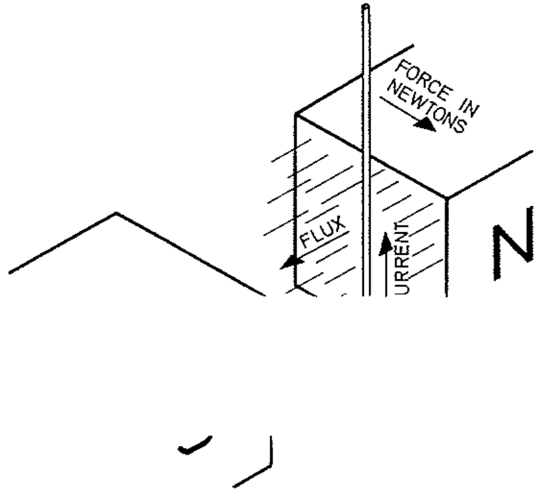


FIG. 17.21. Faraday's law.

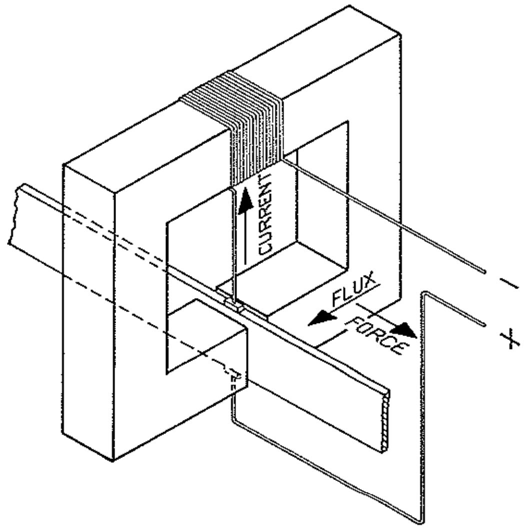


FIG. 17.22. Homopolar linear motor.

The basic method used in converting electrical to mechanical power is shown on Fig. 17.21. The output is a force which is mutually at right angles with the flux and the current flowing. If the flux were stationary and the conductor mounted on the train, a very simple method of propulsion would result. Figure 17.22 shows a feasible mechanical system in which the situation is reversed with the electromagnet being carried on the vehicle and current directed by brushes through a plate which is fixed relative to the earth.

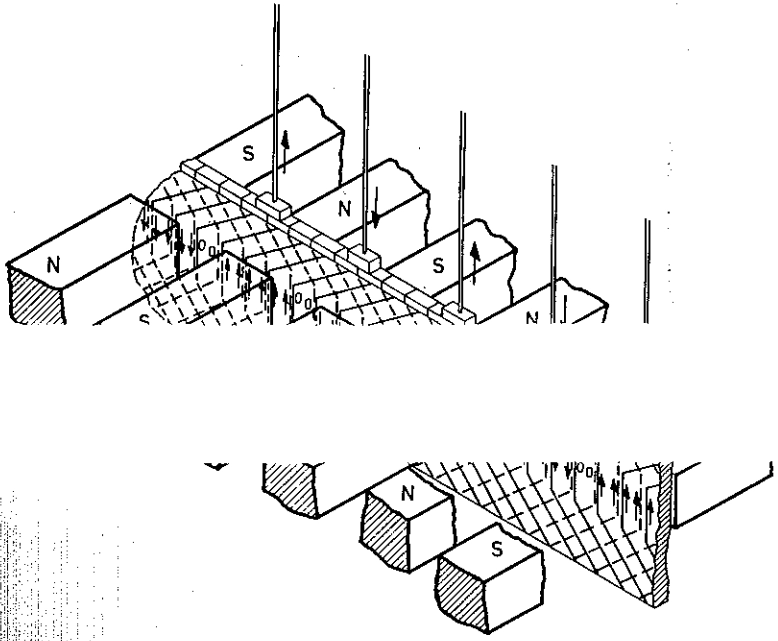


FIG. 17.23. Linear motor with commutator.

Such a system, however, would be extremely uneconomic unless very great currents were used. A series action could be achieved by winding the fixed plate with either a lap or wave winding and feeding it with a linear commutator as shown in Fig. 17.23. Of course such a unit would be far too costly and complex for a transport system and the induction principle, sketched in Fig. 17.24, must be employed.

A three-phase winding, in which a sinusoidal wave of magnetic force in space progresses at the rate of $2pf$ (where p is the pole pitch and f is the frequency of supply), is applied, and a force on the plate will be generated which will tend to accelerate the vehicle until a speed of $-2pf$ is reached (Fig. 17.24). At this speed there will no longer be any relative motion between the moving field and the reaction rail. In the absence of interaction of current and flux the force would fall to zero. This is known as the "synchronous" speed. If as a result of some external force the vehicle is propelled leftwards at a greater speed than this, the relative motion between field and reactor will be

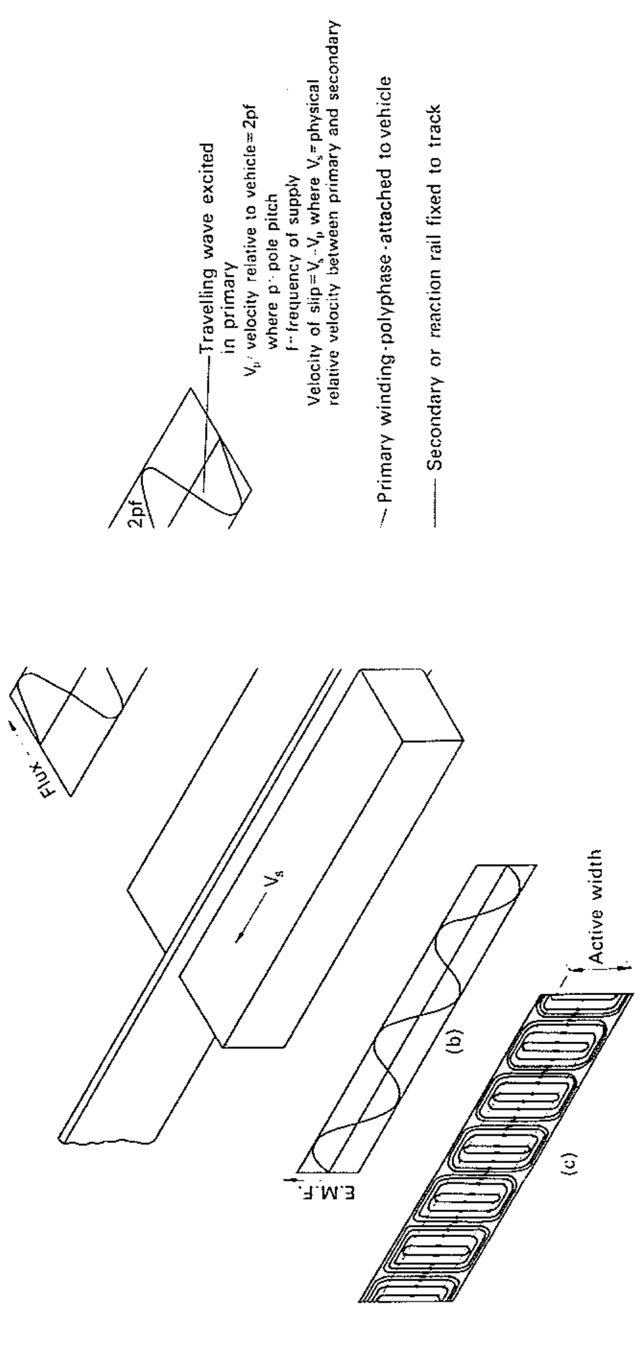


FIG. 17.24. Princip

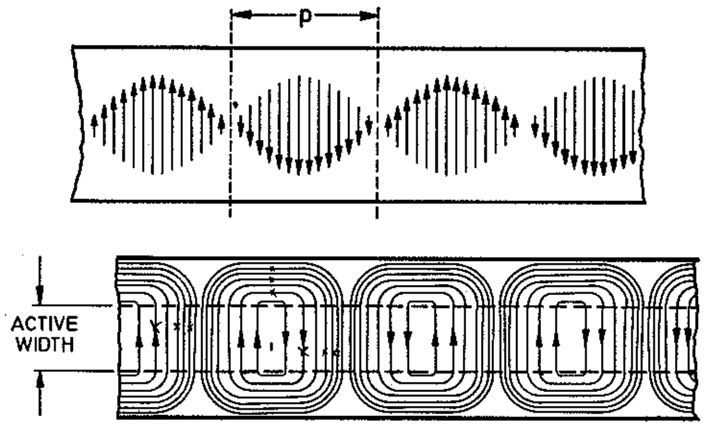


Fig. 17.24. Circulation of current.

be achieved by reducing the supply frequency in consonance with the speed of the train so that the requisite amount of negative slip exists. Such braking is, of course, regenerative.

Application of a linear motor to a transport system may increase capacity because of the independence of braking from adhesion limitations and, where high speeds are involved, can offer promise of a compact, simple, light and inexpensive propulsion unit. Set against this, additional investment is required to provide a reaction rail throughout the entire length of the route and the problem of the design of points and crossings and other intersections becomes very complex. The conversion of existing railway systems with today's speeds and traffic pattern to linear propulsion is therefore unlikely to be economic, particularly bearing in mind that such systems have been designed to take advantage only of the limited propulsive characteristics of the steam locomotive. However, in a new system, the freedom from the need to provide adhesion between wheel and rail for propulsion, braking and indeed guidance, opens the field to novel systems of suspension as well as offering economies in constructional costs by the avoidance of the need for heavy axle loads and moderate gradients. Thus the linear motor forms a natural ally to the "hovercraft" system and indeed enables many of the objections to pneumatic tyres for urban services to be overcome.

To be efficient it is necessary that the air gap in a linear motor should be kept down to the smallest possible value. However, for smooth operation most systems of guidance require a reasonable tolerance on lateral position for the vehicle. Fortunately, the mass of the linear motor element is relatively low so that the solution to the problem arises from the possibility of separating it physically from the vehicle itself and allowing it to be guided relative to the reaction rail, the main vehicle being separately guided with

greater lateral freedom. The method is illustrated in the photograph of one of the early experiments on the linear motor (Fig. 17.26). In a high-speed application the guidance would be pneumatic. The connection with the main vehicle simply has to transmit longitudinal forces, lateral freedom being permissible.

Some apprehension may be felt in relying on the linear motor for braking in the event of a power cut. However, it must not be forgotten that the induction motor can also form a generator and, if suitably excited from a vehicle-borne source, it may act as a brake independently of the line. Thus the numerous methods available for braking induction motors are directly applicable.

It is necessary to consider whether or not the linear motor can be regarded as an efficient and effective propulsion device. In comparison with conventional electrical machines it presents two main disadvantages. Firstly, considerations of alignment and the use of a non-magnetic secondary will necessitate the use of wider effective air gaps than in conventional machines and, secondly, the unwound metal rotor will be less effective in utilising the currents generated therein than would a conventional slotted construction. The subsidiary disadvantage arises from the finite length of a linear field winding as opposed to the continuous fields of the rotary machine. This leads to "end effects" which can be detrimental to performance.

Regarding the effective air gap, the disadvantage of this is that it requires a greater mmf

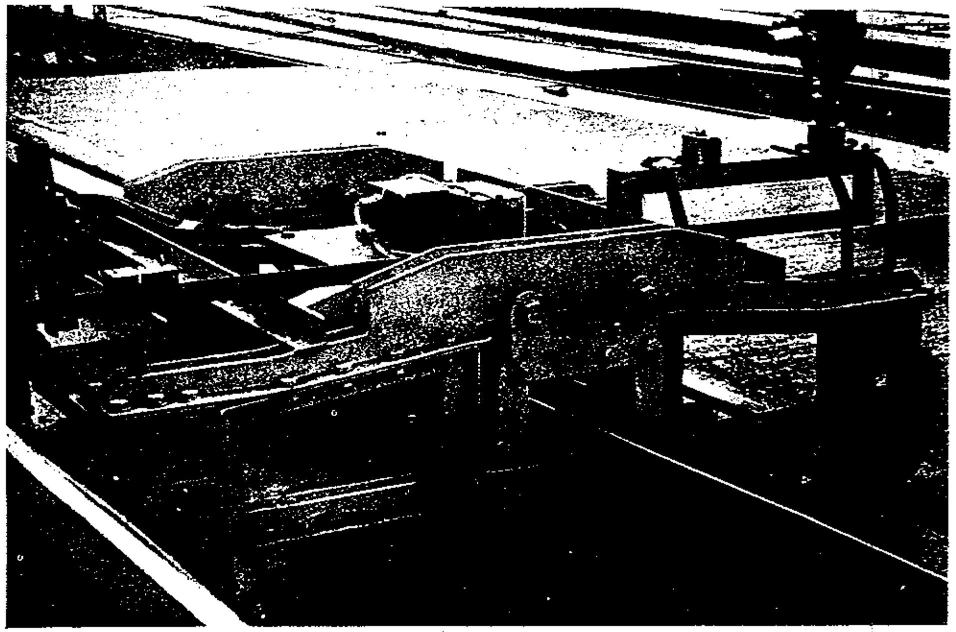


Fig. 17.26. Early experiment with linear motor.

effective resistance of the rotor. However, when these factors were taken into account in the actual design of the machine required to operate under conditions as expressed in equation

$$G = V^2\mu_0/2\pi f\rho_0 g \qquad (17.5)$$

the great importance of velocity in utilising the potential of an electrical machine will be realised.

The significance of the terms is as follows: V is the synchronous velocity, p is the pole pitch $= V/2f$, μ_0 is the permeability of free space, f is the supply frequency, ρ_0 is the surface resistivity of the rotor and g is the effective air gap. Because of the simple construction of the secondary, the effective value of ρ_0 will be greater than in a conventional machine. In the linear motor, however, there is no limit placed on p which can be increased without restriction by adjusting the frequency with which the machine is to be supplied. Also the equation can be rewritten in the form shown in equation (17.6) and regrouping terms it will be seen that the first term represents velocity which is the design requirement.

$$G = V_s \frac{\mu_0}{\pi}\left[\frac{1}{\rho_0}\left(\frac{p}{g}\right)\right]. \qquad (17.6)$$

The central group represents natural constants largely beyond our control and the final

suspension is very necessary to avoid the allowance of an air gap of the same magnitude as is permitted by the normal flange contours and because it would be unwise to insist on a high precision in relating the reaction rail to the carrying rails. In the case of air suspended vehicles it is very likely that the reaction rail would itself provide the guidance of the main vehicle. Nevertheless, freedom of design will arise from separate suspension characteristics for the linear motor and the main vehicle which might be worth preserving.

In consideration of the best way of disposing the reaction rail relative to the geometry of the system as a whole, electrical considerations show that a non-ferrous material is very much preferred to steel as a reaction rail and cost factors probably point to the use of some aluminium alloy. The dimensions would have to be determined by the mechanical forces likely to be imposed as well as by necessity of providing an effective area for the electromagnetic action. At the present state of development it appears that tractive effort available from a double-sided linear motor of reasonably large size may be estimated from the fact that 44,500 newtons may be developed per square metre of pole face. This is equivalent to 6·5 lb/in². Thus to accelerate a vehicle weighing 1650 kgm at a rate of 2 m/s would require an effective depth of 7·2 cm. To this must be added the provision for conducting circulating currents longitudinally at the edges of the reaction rail.

End effects which can be regarded as a flux wave travelling at the speed of the vehicle can be minimised by suitably shaping the poles and graduating the ends of the windings and, of course, by making the linear secondary elements as long as possible so as to reduce the proportion of end poles to the total number. This leads to the consideration of the best shape for a linear motor. Because the reaction rail must be provided throughout the full length of the track it is desirable to make this as small as possible, particularly bearing in mind that it is essential to allow for the plate to be wider than the effective width of the poles because of the need to allow for circulation of induced current as shown in Fig. 17.25. This, however, reduces the effectiveness of the field windings from the point of view of the utilisation of copper because it is necessary to connect the poles through end turns, thus the wider the field coils are made, the smaller will be the proportion of copper which is ineffective.

The actual designs, however, will be conditioned by two factors. That is, the longest individual linear motor which can be accommodated to provide reasonably uniform clearance during their negotiation of curves and the necessity to provide adequate size of machine to exert the required thrusts. It is to be expected that the linear motor element will be considerably shorter than the main vehicle because, in order to reduce the air gap to the minimum, it is desirable that the motor should be suspended independently therefrom so as to follow more closely the irregularities of the reaction rail. This is because its mass is considerably less than the vehicle and because, due to its robust construction, it can be expected to tolerate greater transient lateral accelerations than would be acceptable in the passenger space. In the case of conventional passenger vehicles, independent

rates of acceleration at medium speed may suffice to provide the power required at maximum speed.

The first linear induction motor to be fitted to a standard gauge rail vehicle was exhibited to the press at the Gorton Works of British Railways on 14th November 1962 (Fig. 17.26). This was constructed to the design of Dr. E. Laithwaite of the University of Manchester (now Professor at Imperial College, London) at the request of the Chief Electrical Engineer of the British Transport Commission. It was later removed to Derby where it was inspected by H.R.H. The Duke of Edinburgh on 14th May 1964 (Fig. 17.27).

Whilst the double-sided linear motor is the most attractive from the mechanical point of view, considerable advantages from the point of view of the overall construction of the system arise from placing the reaction rail flush with the surface of the track. Disadvantages are that a continuous ferrous member must be provided to complete the magnetic circuit and that unbalanced forces exist normal to the track surface. These are usually attractive at low currents, i.e. near synchronism and repulsive when slip is considerable.

At the time of writing, political support for high-speed ground transport is lacking in Britain and the U.S.A. where financial support is no longer available for effective progress. Nevertheless, a good deal of successful testing has taken place at the Transportation Test Centre at Pueblo, California. Some results of these trials are summarised in Table 17.4 which also includes reference to the Tracked Hovercraft illustrated in Fig. 17.28. The LIMRV was devised for testing linear motors on conventional track. To simplify the test arrangement this vehicle was fitted with a 2240-kw two-shaft gas turbine

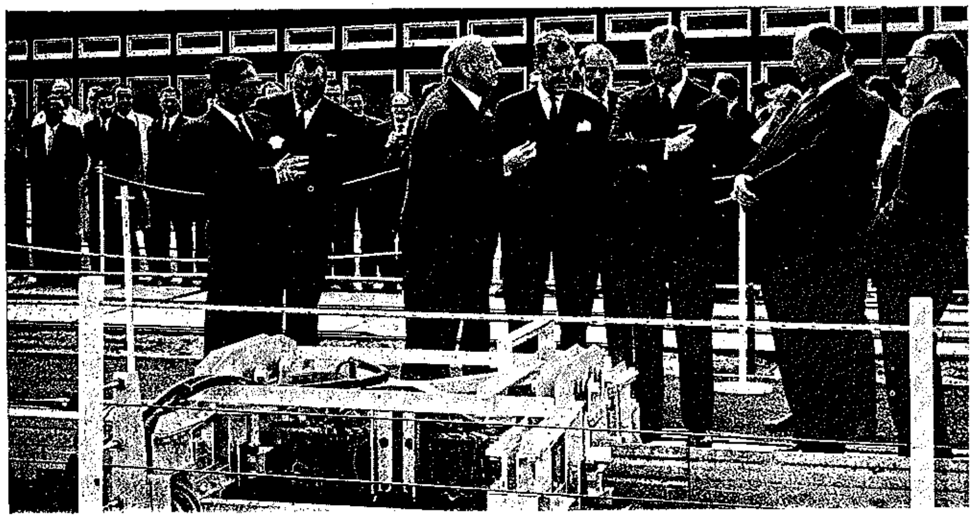


TABLE 17.4. H.S.G.T. TRIALS AT EARITH AND PUEBLO

Feature	System			
	LIMRV	TLRV	PTACV	RTV 13
Projected speed (km/h)	402	483	241	483
Means of suspension and guidance	Wheel on rail	Air cushion with peripheral jets. Trough section.	Air-cushion plenum chamber with central reactor rail	Air cushion with peripheral jets, box girder
Power system	Double-sided L.I.M. augmented by jet	Turbine exhaust reaction only	Double-sided L.I.M. (start by rollers)	Single-sided L.I.M.
Length of track (km)				
Projected	9·66	35·4	9·66	11·2
Actual	9·66	8	9·66	1·93
Maximum speed attained (km/h)	410	151	164	172
Principal contractor	Garrett	Grumman	Rohr	Tracked hovercraft

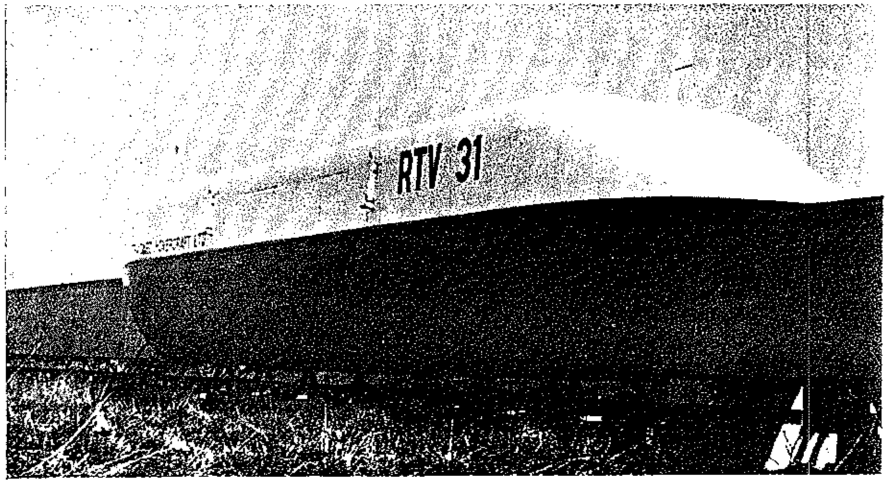


FIG. 17.28. Test vehicle embodying single-sided linear motor. (Courtesy of Tracked Hovercraft Limited.)

drag parachutes. Successful test running has been recorded up to speeds of 400

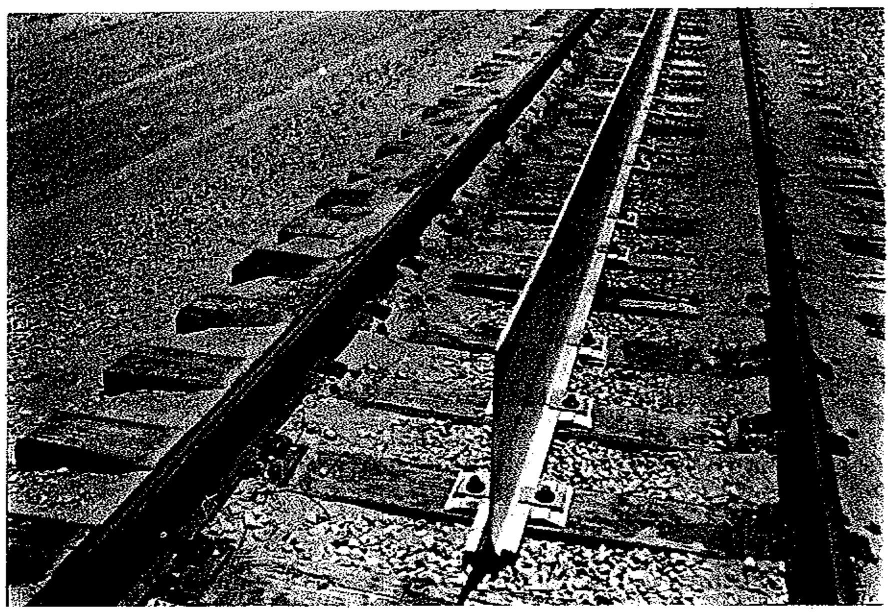


FIG. 17.29. Track for LIMRV showing end of reaction rail.

(255 m.p.h.).[20] The running of the vehicle, which was supported on Budd "Pioneer" tracks embodying an air-suspension system, was particularly good, cinematographic records indicating the virtual absence of "hunting" even at the highest speed. Figure 17.29 shows an arrangement of track including the central reaction rail.

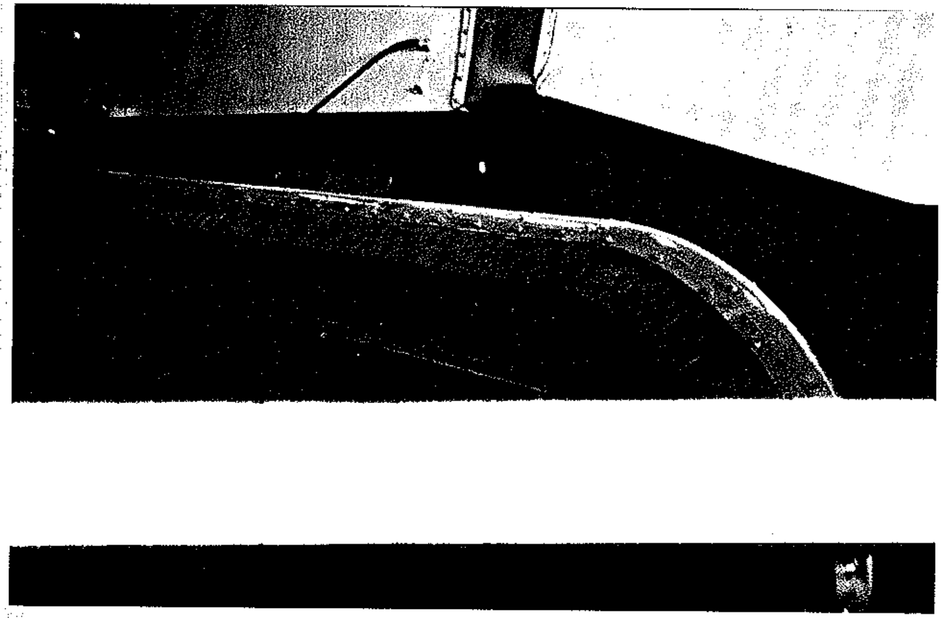


FIG. 17.30. Guidance cushion for TLRV.

Combination of linear induction motors with air cushion support

It seemed natural that the two British inventions, the linear induction motor and the hovercraft, should come together as a solution to the problem of providing high-speed guided transport. The British Hovercraft Corporation exhibited a model at Gosport in 1965 which embodied all the features of subsequent construction excepting that the reaction rail was mounted vertically.

The next stage was the construction of a full-scale vehicle and track by a company specially set up for the purpose called "Tracked Hovercraft Limited" under the aegis of the National Research Development Corporation. Chiefly because of the difficulty of devising points and crossings but also because the elevated reaction rail was regarded as being hazardous should it become broken or detached, single-sided linear motors were introduced in which a continuous iron member was laid in the track to complete the magnetic circuit. This was covered by a thin aluminium plate which afforded a low-resistance path for the circulating currents.

Air cushions may take two forms, the hovercraft which embodies jets which are directed inwardly at the periphery of support cushions so as to induce an element

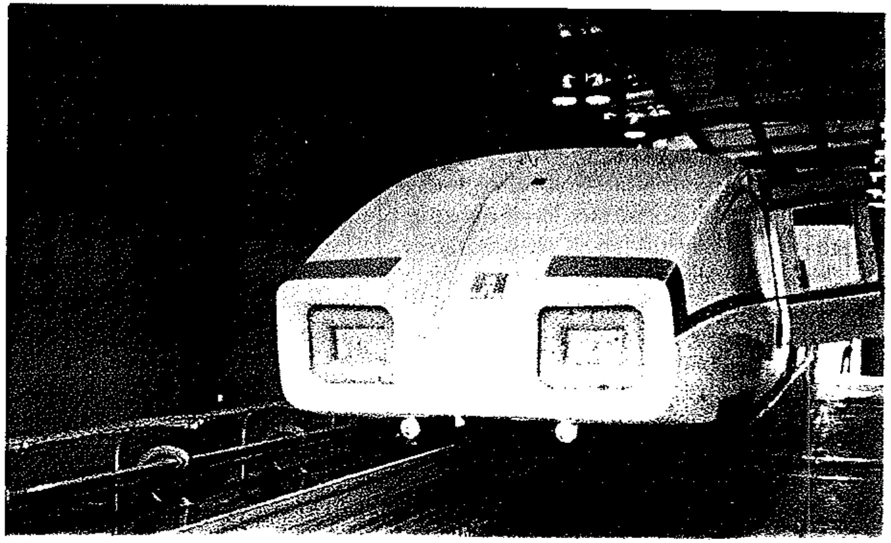


FIG. 17.31. PTACV.

of momentum to pressure conversion, or the plenum chamber introduced by late M. Bertin in France and tested in the U.S.A. as the PTACV System (see 17.31).

Considerable experience in the use of hover-pad suspension was accumulated by Tracked Hovercraft Company at its test track at Earith, Hunts., during the per 1969–72. A research test vehicle RTV 31 (Fig. 17.28) was constructed being suppoı and guided by separate air cushions. Eight hover-pads were provided, two for supp and six for guidance, each being provided with its own supply. The lift pads had tl fans each and the guide pads one only. The cushion pressure was 14·360 kN/m² wit gap of 10 mm in the case of the support pads and 5220 kN/m² and 20 mm respecti for the guidance pads. At speeds of 45 m/s (163 km/h) the ride was satisfactory, vert acceleration being 0·03 g and 0·04 g for lateral vibrations. A speed of 170 km/h (48 n was achieved on a track just over 1·6 km in length[21] (see Table 17.4).

Similar experience was obtained by the U.S. Department of Transportation at Transportation Systems Centre at Pueblo, Colorado, on a vehicle known as TL (Track Levitated Research Vehicle) (Figs. 17.20(a) and 17.30). At the time of the writ visit it had only operated as fast as the reaction from a gas turbine exhaust would peri However, at over 145 km/h (40 m/s) it rode particularly well even when artificial irre larities were on the track.

The plenum chamber system was also tested at Pueblo in a vehicle known by initials PTACV (Prototype Tracked Air Cushion Vehicle) (see Fig. 17.31) wl operated satisfactorily at a speed of 164 km/h. The ride was described as "extremely comfortable".

These experiences coupled with the demonstration provided in France by the late M. Bertin provide ample evidence of the effectiveness of the air-cushion system as a means for the support of guided vehicles.

Although a Select Committee of the House of Commons recommended retention of the track for continued trials, the government of the day decided on demolition.[22, 23]

Figure 17.32 shows a linear motor which has been used for towing test vehicles at the Motor Industry Research Association Laboratory at Nuneaton.

In the MIRA installation (Fig. 17.33) the reactor or secondary is an aluminium plate 1 inch thick and 30 inches deep with steel guide plates at the top and bottom extending the whole length of the building in a duct or trough in the floor and acting as a track for the moving primary part of the motor. This moving primary is in effect a trolley which tows the test vehicle and by means of guiding wheels runs on the reactor track. A quick-release connecting device on the trolley protruding above the floor is used to tow the test vehicle.

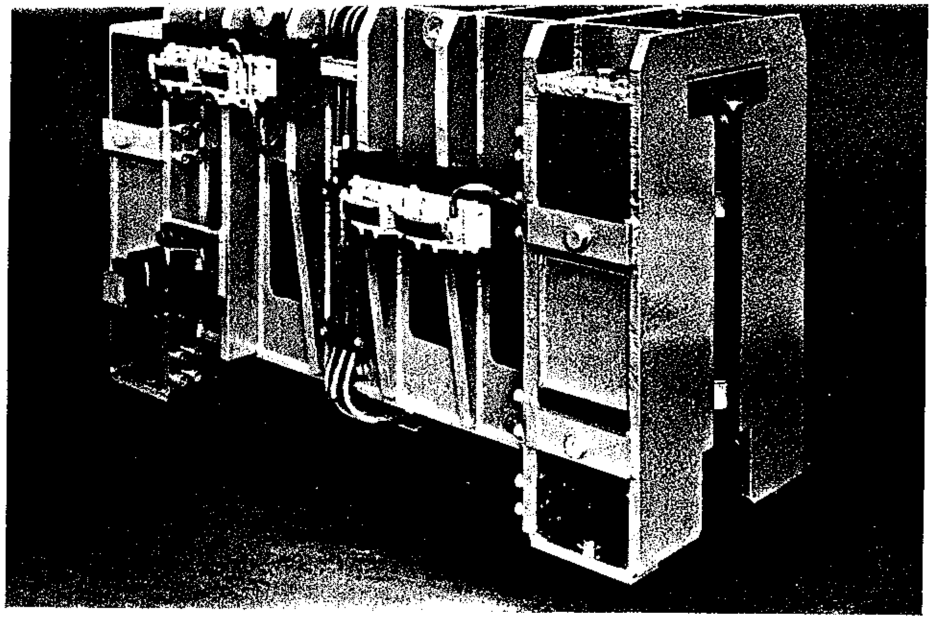


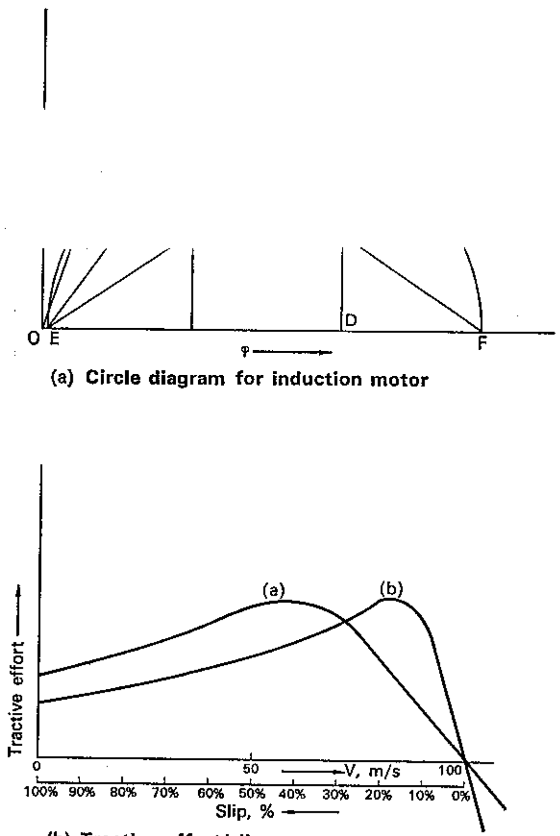
Fig. 17.32. Linear motor used for automobile testing. (Courtesy of Motor Industry Research Association.)

The moving primary trolley picks up current through collector brushes and rails mou
on the sides of the duct. The installation has been so designed that the test ve
acceleration does not exceed 1·5 g. To ensure that the dummy occupants of the ve
return to their normal positions after the applications of high accelerating force, the n
is designed to run for a distance before impact at a steady speed. Accurate contr
impact velocity is provided between the limits of 13·4 and 13·8 m/s (30 and 31 m.
At a point 10 feet (3·05 metres) before the impact barrier the current supply is dis
nected and the test vehicle is automatically detached from the trolley. Whilst the ve
travels freely on and crashes into the barrier, the trolley passes under the barrier a
arrested by the extension of undrawn nylon ropes. These ropes have the characterist
absorbing the energy from the trolley without excessive rebound. Provision is mad
emergency stopping in the unlikely event of rope failure. The barrier into which
vehicle is impacted is a steel-faced concrete block 20 feet long × 12 feet wide × 6
high and weighs 100 tons. It is keyed to a further 100 tons of concrete foundat
Provision can be made to angle the front of the block and the safety rails on each si
the track have been made adjustable to permit cars to be propelled in a sideways posi
The use of a sled (an "indestructible" car with buffers) in place of a vehicle will en
components to be tested.

Fig. 17.33. Crash test on automobile propelled by linear motor.

The specification of the linear motor is as follows:

Nominal supply voltage	3·3 kV, 3 phase, 50 Hz
Number of poles	6
Winding	2 slots/pole/phase
Reactor plate material	aluminium alloy
Reactor plate thickness	1 inch (0·0244 metre)
Reactor plate resistivity	3·4 microhms/cm³
Air gap	⅛ inch (0·0032 metre)
Maximum thrust	6000 lb (approx.) (27200 N)
Maximum stall thrust	3700 lb (approx.) (16800 N)
Synchronous speed	32·6 m.p.h. (14·8 m/s)
Maximum stall current	290 amps
Current collection gear	copper/carbon slippers on Mn bronze rails
Mass of motor	3150 lb (1400 kg)



(a) **Circle diagram for induction motor**

(b) **Tractive effort/slip curves**

Fig. 17.34. Tractive effort/speed relationship of linear induction motor.

Control of linear motors

The linear induction motor behaves in a similar manner to its rotary counterpart with the exception of end effects arising from the need to set up circulating currents at the entry region. This effect may be minimised by good design and does not affect the control problem.

Unlike the rotary machine no means are available for varying secondary resistance which is determined by the shape and material of the reaction rail. The tractive effort exerted at any speed for a given secondary rail will depend on the amount of slip and can be derived from the conventional circle diagram (Fig. 17.34). Thus if a motor is subjected to an applied alternating voltage when stationary in the absence of the secondary rail, the input will be represented by the line OA. If it is now clamped with the rail in position the input will be shown as OB. The total equivalent flux relative to the primary windings will be in quadrature with the applied e.m.f. and can be represented by a horizontal line. A semicircle with diameter coincident with this line and passing through A

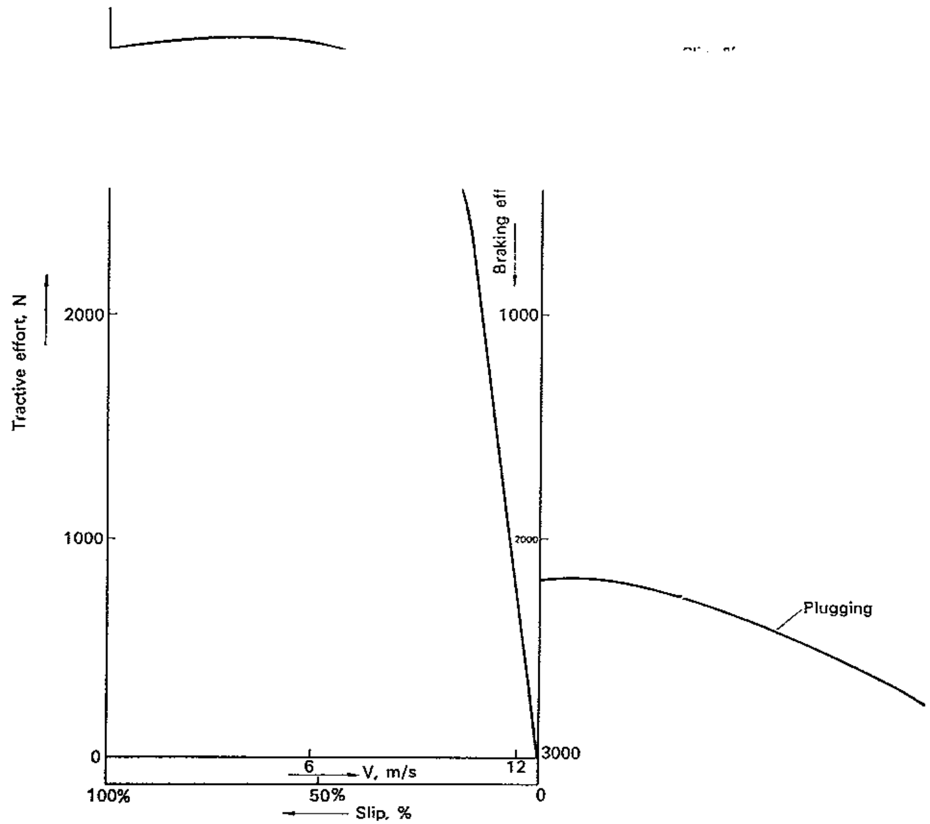


Fig. 17.35. Acceleration and braking of linear motor at constant supply frequency.

and B will complete the circle diagram. A horizontal line drawn through A will represent the constant stator losses. BD represents total losses = total input in the stalled condition and AB the losses for intermediate conditions such as those represented by C. EC represents secondary current and CF secondary flux. These are always in quadrature as consistent with the geometrical construction of the circle diagram. The construction is completed by AG which apportions the copper losses between rotor and stator, BG representing copper losses in the stalled condition. Then for any intermediate condition C, tractive effort is proportional to CH and slip is proportional

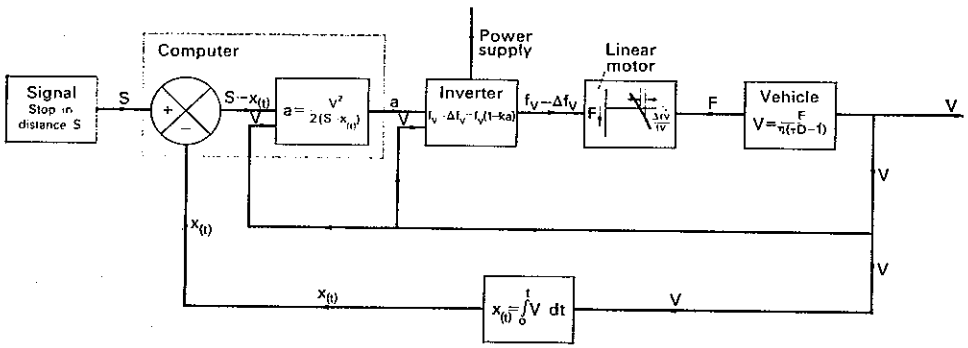
$$\frac{JH}{CJ}, \text{ i.e. } \frac{\text{secondary losses}}{\text{secondary input}}.$$

Current drawn from the line is represented by CO and power factor by $\cos \Phi$. This data can be replotted as a tractive effort/speed curve as in Fig. 17.34(b) and Fig. 17.35 shows some experimental results for a small linear motor operating under somewhat unfavourable conditions.

If rotor resistance is reduced, point D of Fig. 17.34 will be moved to the right which will have the effect of reducing starting effort but of increasing tractive effort at high speeds where it is required to overcome aerodynamic

intersection of curves a and b of Fig. 17.34(b).

Whatever the value of secondary resistance, the tractive effort/speed curve in the vicinity of the synchronous speed is practically linear. This continues if synchronous speed is exceeded for any reason so as to produce a braking force. The motor then acts as an induction generator. Thus the motor characteristics themselves are conducive to stability in the system and are adequate for its control if operational flexibility is not required.



S = distance from desired stopping point when t = 0 and $x = x_{(0)}$

$x_{(t)}$ = position of vehicle at time t

FIG. 17.36. Variable frequency control of linear motor.

When the direction of current in one of the phases is reversed so as to produce a tra ling field in the direction of motion relative to the vehicle, slip exceeds 100% in the ne tive sense and a powerful braking action results. This is shown in Fig. 17.35 and is kn as "plugging".

A more flexible control system both for traction and braking can be achieved by vary the frequency of supply, possibly by the means described in Chapter 12, Section 10.

At any given speed the synchronous frequency corresponding to the actual spee could be determined and then a number Δf_v could be added therefrom so that the mo was supplied at frequency $f_v + \Delta f_v$. Tractive effort would then be proportional to When braking was required the frequency of supply would be $f_v - \Delta f_v$, the magnitud Δf_v determining the intensity of braking. A system for braking within a certain dista specified by some external control is shown in Fig. 17.36.

Track-mounted stators

In the most economical arrangements for linear-induction-motor propulsion vehicles, the stator which contains the field windings is carried on the vehicle and reaction element (or rotor) is fixed to the track. This requires that the vehicle must

power.

The tests on the Maglev vehicles at Emsland (see page 340) use a continuous polyph winding on the track.

Transverse flux; When the relationship between the highest available frequency a vehicle speed is such that the pole pitch of a conventional (longitudinal flux) machin large, it is more economical to arrange for the flux to flow transversely. This supplies "single-sided" configuration with two or more air gaps. The merits of transverse-f machines are described by Laithwaite *et al.*[20]

The magnetic river

The radial forces in rotary machines and the normal forces in linear machines usually large (as much as 20 times the tangential force) and are regarded as parasitic. the case of the linear induction motor the normal forces are usually repulsive at l speed, becoming tractive at higher values.

If a motor can be designed to exercise repulsive forces over the full range of s utilised for propulsion then the otherwise parasitic forces can be used to replace t levitation and guidance forces necessary in the electro-magnetic and electro-dynar systems. Thus if a linear motor is used for propulsion these secondary functions can fulfilled without additional capital investment or energy consumption.[24]

In addition to the provision of normal forces, stabilising moments can result fr suitably designed magnetic cores.

The simplicity, economy and versatility of the repulsion magnetic river offers t greatest promise of all the electrical systems, and trials on a technical scale are bei organised in Britain by Landspeed Limited.

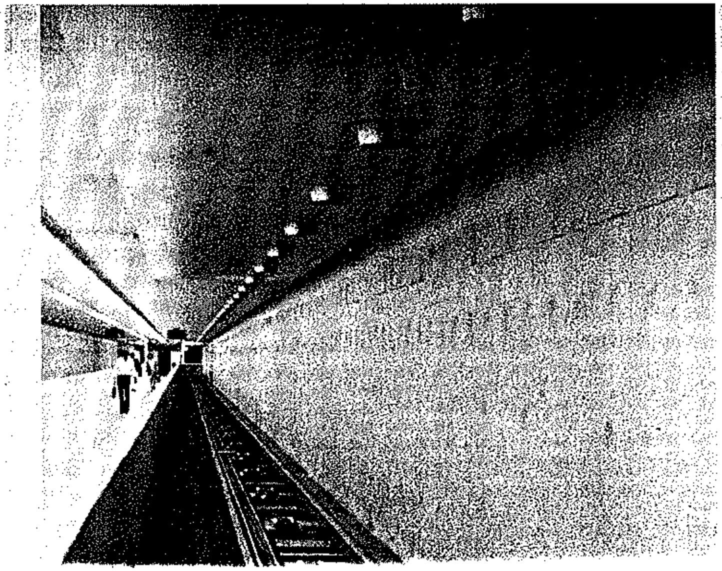


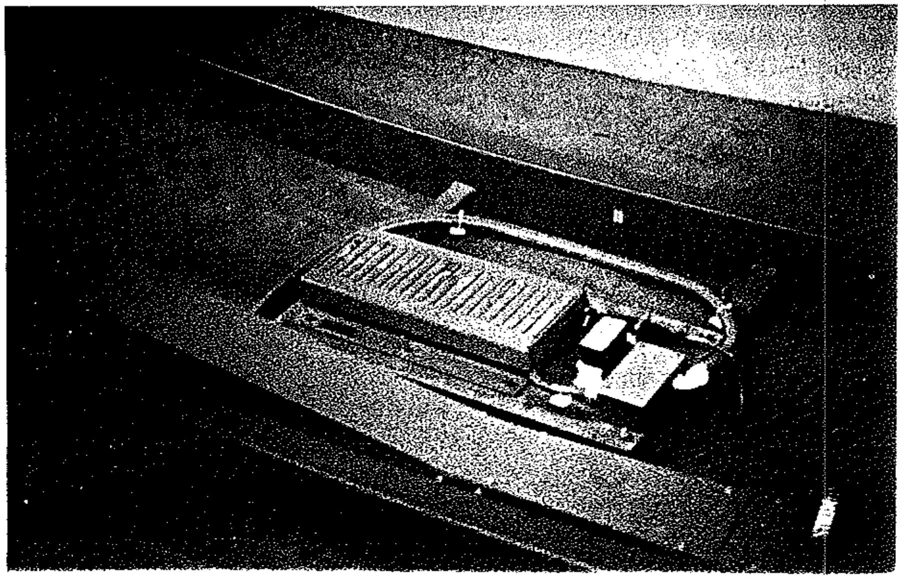
FIG. 17.37. Linear motor drive for People Mover. (a) Track.

Active and passive switching

Repulsion forces are stabilising in so far as they increase as air gap is diminished and thus the magnetic river is ideally adapted for guidance.

The conventional turnout used on railways can be described as "active" in so far as the route to be taken by the vehicle is determined by the movement of switch blades connected with the track. The route has therefore to be "set" by an external agency, usually a signalman.

The alternative method, Passive Switching (suggested by Perrot), requires that the



track features remain unchanged and that the route should be selected by the veh itself. The Morgantown installation already described provides an example of this syst

Active switching has been held to present the disadvantage that when a route is bε set the movement of the switch blades must be complete and proved to be comp before the next vehicle may be allowed to pass at speed, a point which is situated at l one braking distance from the points.

It is generally held that no such requirement exists for passive switching altho British Rail point out that, if as is usual, one of the routes controlled by a switch is s ject to a speed restriction, the speed of approach of the vehicle should be limited to t speed until the route-selection apparatus on the vehicle has been shown to have fι tioned correctly.

A proposal for an active system is illustrated in Fig. 17.38(a) and a passive systen Fig. 17.38(b).

In the latter case the vehicle would be fitted with sets of rollers which would inte with the check rails. Route selection would be effected by lowering the set of rollers the side of the vehicle corresponding to the desired direction of divergence.

Attraction method of magnetic river

A series of tests has been carried out at the Canadian Institute for Guided Gro Transport on a linear induction machine for integrated magnetic suspension and pro sion.[25] These were carried out using a 7·6 m (24 ft) diameter test wheel giving peripheral speeds up to 80 km/h (50 m.p.h.). The motor/suspension unit tested was 1·73 m (9 ft) long, having six poles. The pole pitch was 0·25 m (9·6 in.). Two forms of rail were applied to the drum, one a cage with solid aluminium bars let into a laminated core and the other a solid steel slab.

The tests have demonstrated the feasibility of an integrated magnetic support and propulsion system for a track guided vehicle. Guidance could also be obtained using offset linear motor pairs. The performance of the solid rail was only marginally inferior to the cage but introduced non-linear behaviour which required further investigation.[26]

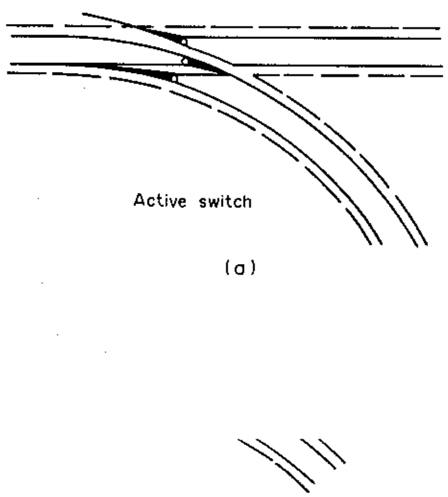


FIG. 17.38. Switching of Magnetic River System. (a) Active system. (b) Passive.

17.9. Linear synchronous motors

The linear induction motor has the advantage that power need only be supplied to one member of the interacting electro-magnetic circuits, reactive current being generated in the other by induction whereas the synchronous motor requires supplies to both rotor and stator. One proposal for overcoming this difficulty has been put forward by Levi[27] in the "claw motor". This is illustrated in Fig. 17.39 and involves two windings carried on the vehicle. The field winding is fed with d.c. to produce a unidirectional flux which is arranged so that the "north" pole is on the left and the "south" pole on the right. Flux is transferred through an air gap from the moving to the stationary parts of the motor. The stationary flux path is made up of interlocking "saw tooth" or claw

components so that, in the control region, there is an alternating polarity. Polyphase a.c. is fed to the armature winding so as to provide a travelling field which interacts with the saw tooth fluxes.

The Japanese National Railways have developed a test facility in which the field magnet (d.c.) is carried on the vehicle and which interacts with coils mounted on the track which are fed intermittently by d.c. currents switched on and off by thyristors so as to produce a travelling field.[28] Test results yielded a propulsion force of 20,000 N (4500 lb) at 165 km/h (102 m.p.h.). Figure 17.40 shows equipment for testing pantographs at speed using this type of motor for propulsion.

An alternative system of electro-magnetic linear propulsion is analogous to a rotary synchronous motor wherein a direct current excitation winding is driven by a polyphase stator winding. When the repulsion system of levitation is used powerful direct current magnets are available on the vehicle and these can be arranged to interact with polyphase windings mounted on the track. One advantage of the synchronous linear motor is that no power has to be supplied to the vehicle (other than that required to excite the cryogenic magnet) the track-mounted coils being connected to the supply through a power conditioner which adjusts the frequency to correspond with the desired speed of the vehicle.

Tests have been carried out on track in Japan and on a rotating drum in Canada (Fig. 17.41).

It appears that cryogenic synchronous linear motors are especially associated with

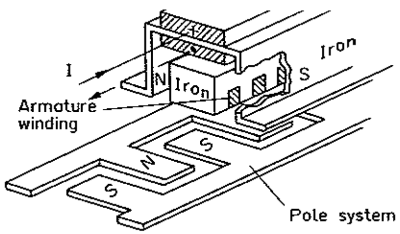


FIG. 17.39. "Claw" motor.

17.10. Electro-magnetic support and guidance (attraction mode)

The electro-magnetic system of support and guidance embodies magnets mounted on the vehicle which attracts steel rails embodied in the guideway. The system is therefore referred to as the "attraction" system to distinguish it from the "repulsion" system discussed in Section 17.11.

Support is obtained by magnets mounted on brackets below the vehicle which pass underneath the running rail. Guidance may be obtained either by separate magnets acting

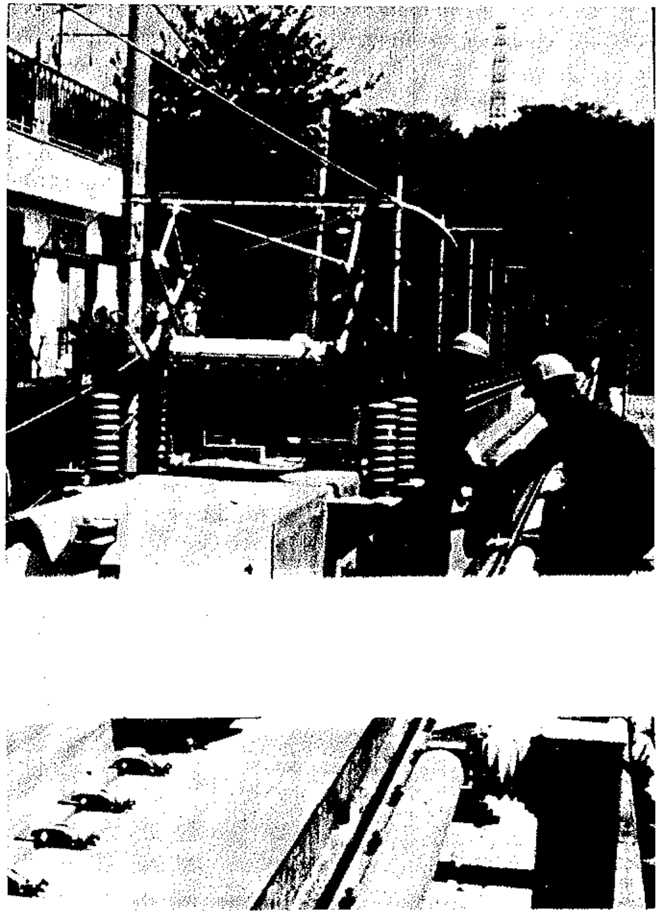


Fig. 17.40. Apparatus for testing pantographs propelled by linear synchronous motor.

horizontally or by the exercise of transverse forces on the support magnets. The attraction system is inherently unstable. Suppose that a certain definite force acts between vehicle and track; if this force were less than that due to the weight of the vehicle, the gap would widen, the magnetic circuit would extend, the magnetic flux would fall off and the vehicle would fall to the ground. Conversely, if the magnetic force were greater than the gravitational force the vehicle would lift, thereby closing the air gap enabling the force to become still greater. Thus there would be a tendency either for the vehicle to fall away or for it to rise to the greatest height thereby eliminating the gap between the magnet and the track. If a solenoid were used to provide the whole of the magnetic flux or to moderate the flux provided by permanent magnets, a control system would be required which measured the value of the air gap at any instant of time and moderated the applied

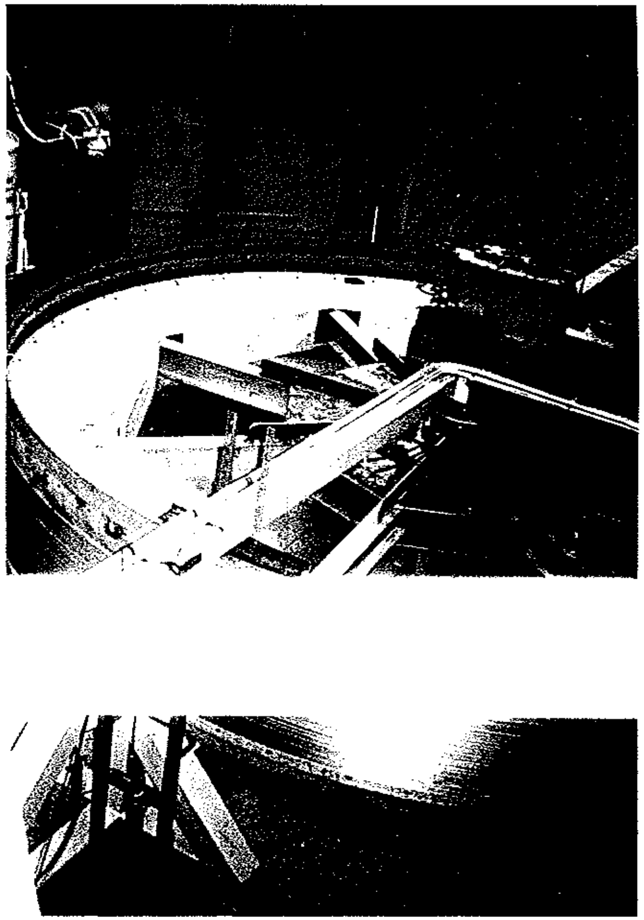


Fig. 17.41. Drum for testing electromagnetic propulsion and levitation systems, Canadian Institute for Guided Ground Transport, Queen's University, Kingston, Ontario.

current in order to keep the separation within acceptable limits. When this was considered by British Rail some 20 years ago it was decided that the power required for such a control system would be so great that the necessary equipment presented unacceptable weight and volume penalties on any vehicle. It is a tribute to the development of control engineering during the intervening period, particularly in the application of solid-state devices, that the system can now be regarded as being feasible.

The City of Toronto which is one of the most successful in developing a co-ordinated transport system so as to avoid the impact on the centre of the city of the motor car, having successfully developed underground systems as well as bus, trolley-bus, and tram systems, felt the need for something intermediate between them—something which was environmentally compatible with city conditions but cheaper and more flexible than conventional undergrounds. As the result of a world-wide study and a comparison of many tenders, they finally chose that of the German firm Krauss-Maffei which proposed vehicles embodying the magnetic attraction system.

Two horizontal coils were provided which could be selected so as to provide a directional control at junctions which was controlled entirely from the vehicle. In other words, passive switching would be employed which has important connotations from the point of route capacity. The German Government was supporting the firm of Krauss-Maffei in this development and it was a great shock at the beginning of December 1974 when it decided not to continue its support into 1975 on the grounds that the application to urban work was not necessary and that development should be confined to high speeds. This led to the firm withdrawing from the contract at Toronto.

Nevertheless work on application at high speeds continues in Germany under the aegis of the Abeitsgemenschaft-EMS, Munich. This joint venture includes the Krauss-Maffei company of locomotive builders, and Messerschmitt-Bolkow-Blohm, an aero-space company. An experimental vehicle manufactured by Krauss-Maffei known as "Transrapid 02" achieved a maximum speed of 163 km/h (101 m.p.h.) in 1971 on a track which was 930 m long and which embodied a curve of 800 m radius.

A further vehicle, known as "Transrapid 04", reached a top speed of 220 km/h

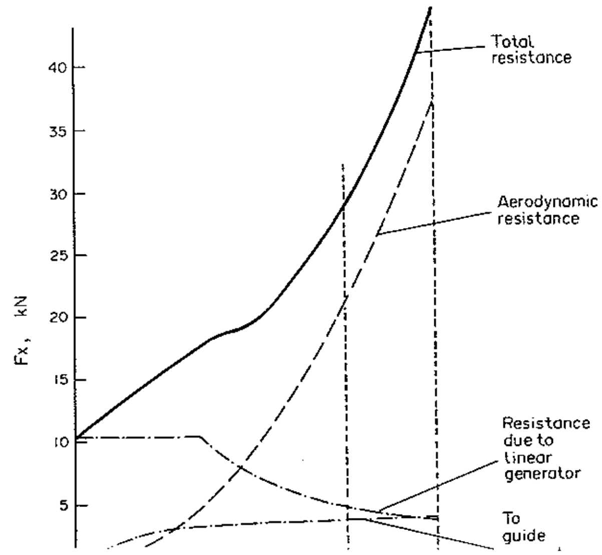
which acts as a test vehicle for various components. It is 8·5 m long, weighs 8 tons, is supported by ten magnets and guided by four magnets.

The next stage in the development of magnetic levitation in Germany was the demonstration of Transrapid 05 at the IVA transport exhibition in Hamburg in 1979. At the time of writing the Transrapid 06 magnetic levitation vehicle is being developed by Krauss-Maffei with AEC-Telefunken and Messerschmitt-Bolkow-Blohm making important contributions. The tests will take place on a track which is being constructed at Emsland in North Germany. This will consist of a straight track some 30 km long (19 miles) with turning circles of 1690 m (1 mile) radius at the north end and 1000 m (0·62 miles) at the south.

The new vehicle is designed for a speed of 400 km/h (250 m.p.h.) and the expected speed/tractive effort curve is reproduced in Fig. 17.42(a). It is constructed in two sections each 27·1 m (90 ft) long and weighing 61 tons. Seats are provided for 96 passengers.

The two-car train is supported by four air springs on four levitation bogies, each being fitted with eight carrying magnets and seven guiding magnets. The guide magnets interact with vertical rails and the support magnets with a horizontal member which also acts as the stator of the propulsion system. Guide rails and stator are cantilevered from each side of the track structure and the magnets are supported through a mechanical suspension system by brackets which project from the levitation bogies.

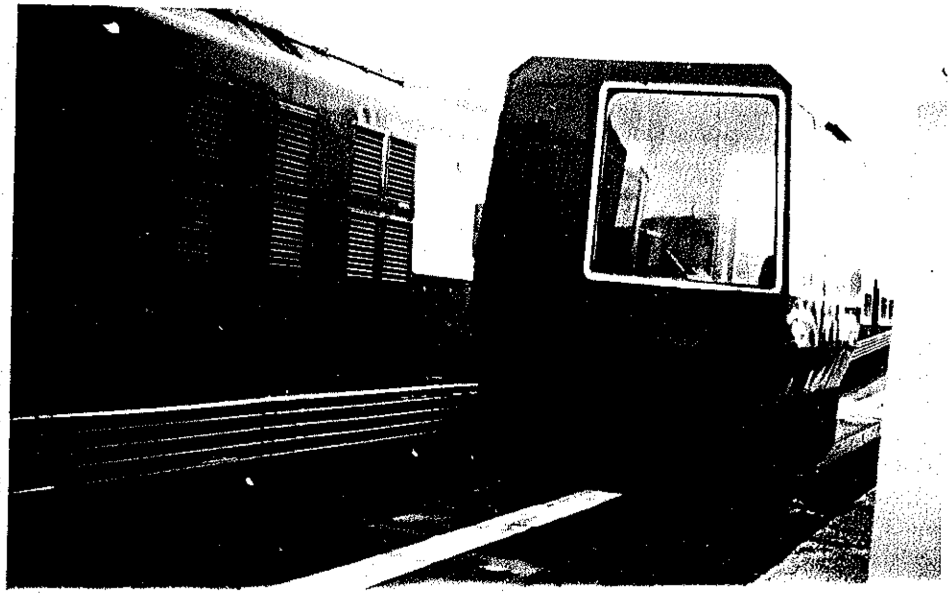
Power for the support and guidance systems and air conditioning is transmitted from the stator to the vehicle by linear generators. Batteries are provided to feed the supporting magnets in the event of power failure. Conductor rails are provided to feed energy to the vehicle when it is stopped at a station or proceeding at low speed (up to 90 km/h).

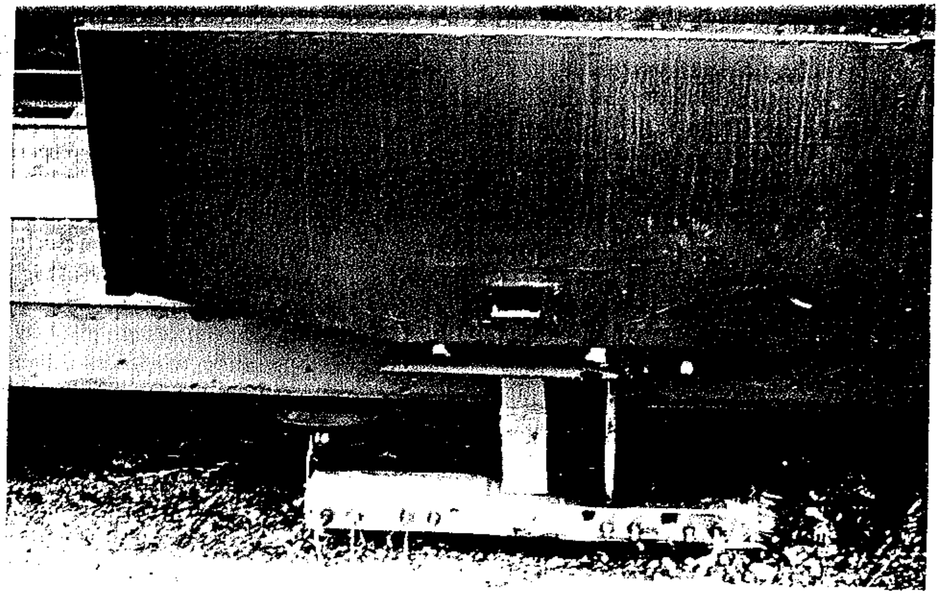


The stator winding on the track is divided into sections, each connected to the pow supply by vacuum circuit-breakers. Voltage and frequency are varied to suit the speed the train. Below 120 Hz voltage is varied by pulse-width modulation. In the upp frequency range (120 to 215 Hz) two separate units are used to control voltage by i phase addition of two voltage components. Voltage at the highest value (4·25 kV p phase) is obtained by means of a transformer.

In September 1975 British Rail exhibited a magnetically levitated vehicle which h been requested by the British Department of the Environment (Fig. 17.42(b). The vehic weighed 2·7 Mg including driver and four passengers. The support magnets (Fi 17.42(c) were of the d.c. attraction type and were controlled to hover suspended from steel rail within the limits ±12 mm. Propulsion was by three linear motors supplied three-phase contact rails at 415 V 50 Hz. The particular demonstration was for an urb miniature passenger system and 8-m radius curves, and 1 in 20 gradients were includ in a short test track. However, British Rail consider the system applicable to very hig speed ground transport running at perhaps 500 km/h (310 m.p.h.). The control proble necessary to maintain constant hover-height have been satisfactorily solved. By duplic ing the magnets at each corner and arranging them to be slightly offset to each side of t rail centre line, it is possible to exercise lateral damping by varying the power different between the two sets. The arrangement is such that the magnets automatically follow t guide-rail without any separate steering device.

The British Rail system is at present being projected to provide a shuttle service betwe

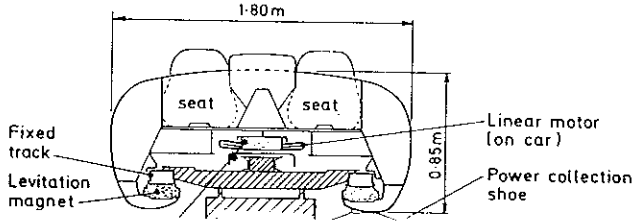


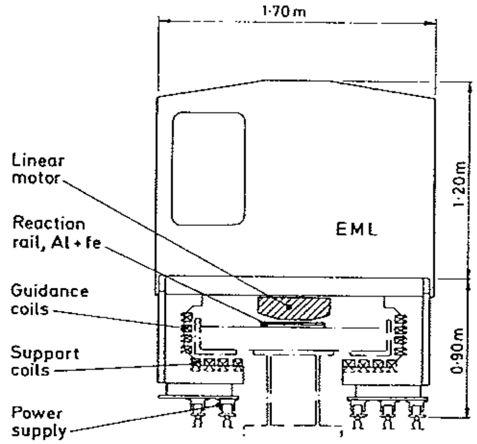


(c) Support magnets for B.R. vehicle.
FIG. 17.42.

the air terminus and the National Exhibition Centre at Birmingham. The propc guideway has a maximum gradient of 1 in 10 and a minimum radius of curvature of 1: The vehicles will be 20 m long and 2·6 metres wide. They will weigh 4 tons loa Duration of stops at the terminals will be 30 seconds and the journey time will b seconds.

The Japan Railway Technology Association has selected the attraction system development in preference to the repulsion method because it was thought to be suit for low-speed urban systems as well as high-speed inter-city applications. Speed 40 km/h have been achieved on a 200-m-long test track using the vehicle illustrate Fig. 17.43. It will be noted that separate magnets are used for support and guidance.





(b) EML system.

FIG. 17.43. Magnetic levitation.

A large international airport constructed at Narita has been the subject of much controversy from an environmental point of view and suffers from remoteness from the centre of Tokyo, the distance being 65 km. The plans for extending the Shinkansen line from central Tokyo to this airport are at present in abeyance.

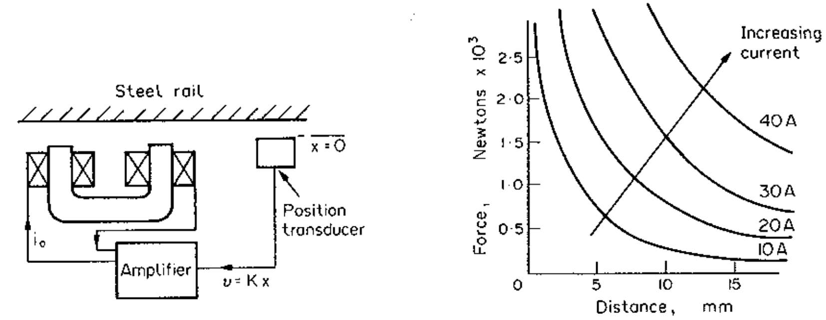
The non-use of the airport is of considerable embarrassment to Japanese airlines who have embarked on the development of high-speed ground transport with the intention that the application of the designs and the operation of the airport link would be handled by a new organisation which was a government-backed corporation, perhaps including private investment. The Japanese airlines (JAL) have also selected an attraction system similar to that developed by Krauss-Maffei AG. The main difference between the systems is that in the functions of support and guidance they are separated, whereas in the FNL scheme they are combined. This has important control implications as well as affecting the design of points and crossings. A demonstration run employed a test vehicle (illustrated in Fig. 17.43 which was 4 metres long, 1·7 metres wide, and 0·85 metre high and which was fitted with a 2000 kVA linear induction motor fed on a frequency which varied from 01 to 50 Hz and a voltage of 0–600 V. The reaction plate consisted of aluminium-coated iron, the aluminium coating being thicker at deceleration and acceleration zones of the track. Suspension was by direct-current electro-magnets which were attracted to

for the full-scale are as follows:

Cruising speed	300 km/h
Acceleration	0·1 g
Power required at cruising speed	5 kW per passenger
Radius of track at 300/km/h	3000 metres

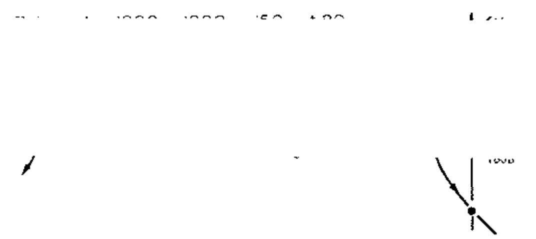
A study in depth of the economic possibilities of high-speed ground transport has been carried out in the U.S.S.R. and it was considered that technical questions were worth investigation. As the result of consideration of various systems the electro-magnetic (attraction) system had been selected on the grounds of simplicity. Tests will be made at Noverchertask which is in the south of the U.S.S.R., near Rostov.

The nature of the control problem is clarified by Jayawant *et al.*[29] who discuss both single magnets and the multi-magnet systems necessary to control the position of a vehicle in three dimensions. Considering the single-magnet system by way of illustration, the physical configuration of the system is shown in Fig. 17.44(a). The system's non-linear typical force distance characteristics for different values of current are shown in Fig. 17.44(b). For a given load in the steady state there will be a particular value of current necessary to maintain the air gap at the prescribed distance. If the vehicle were to move over a track irregularly which altered the gap there would be a discrepancy between magnetic force and gravitational force and the mass would accelerate until the two forces were in equilibrium and then oscillate.

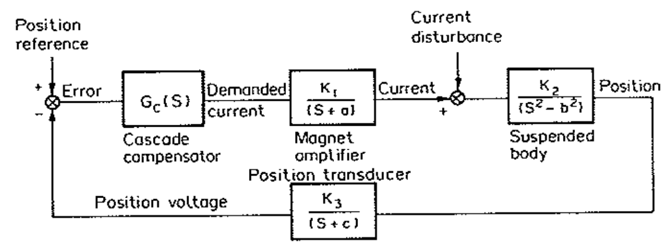


(a) Suspension of single magnet using a position transducer

(b) Force/distance characteristic



(c) Root locus



(d) Schematic

FIG. 17.44. Magnetic levitation control system. (a) Physical configuration of system. (b) Force-distance characteristics. (c) Root locus. (d) Control system.

If f is the force acting on the suspended body and m is its mass

$$\frac{md^2x}{dt^2} = -f. \tag{17.7}$$

If force is related to gap by $-Kd$ and to instantaneous current by K_i, then

$$m\ddot{x} = -(-K_d x + K_i). \tag{17.8}$$

Writing $H(s)$ as

$$\frac{-K}{m(S^2 - K_d/m)}$$

If $G(s)$ is the transfer function of the magnet-amplifier combination

$$= \frac{1}{(1+sT_m)}$$

then

$$G(s)\,H(s) = \frac{-K}{m(1+sT_m)\,(S^2 - K_{dm})} \tag{17.9}$$

Design of multi-magnet systems is complicated by two factors. Firstly, under practical conditions of air gap the magnet is operating in the highly non-linear part of its force–distance characteristic, and secondly, because of the mechanical coupling of several magnets mounted on a single vehicle body.

Suspension with several degrees of freedom requires at least one magnet for each degree of freedom. The action of these magnets must be co-ordinated requiring a multi-variable control system. Hazlerigg and Sinka[30] describe the design of such systems and identify two basic approaches. One of these is to control the motion in each degree of freedom so as to maintain roll and pitch angles and the height of the vehicle constant in a four-input/four-output system. Change in any one of the forces will cause a response in all the output variables to compensate for cross-coupling between them. An alternative design procedure is to admit the presence of cross-coupling but to base control on a direct association between a magnet and its nearest transducer.

Considerable improvement in the dynamic characteristics of multi-magnet systems can be obtained using acceleration and velocity transducers in the feedback system to supplement position transducers.

17.11. Magnetic support and guidance—the repulsion (electrodynamic) system

If a constant magnetic field is moved relative to a conducting sheet, the interaction induces currents within the sheet. These currents flow in such a path that they become

equivalent to a magnet of the same polarity as the real magnet and positioned as though it were the mirror image thereof. A repulsion effect exists therefore which can be regarded as the reaction between the real and virtual magnets and which can be used to support a vehicle.

The number of ampere-turns necessary to provide support at a reasonable height is very large indeed and is made possible by using a cryogenic system wherein an electromagnet is kept at an extremely low temperature so that many thousands of amperes can flow with negligible I^2R loss. A demerit of the system is that the action of inducing levitation produces a force which opposes vehicle motion. This increases rapidly to peak at about 45 m/s (160 km/h) and then falls off as speed is further raised. There is no lift as the vehicle starts from rest so that an auxiliary support system must be provided for starting and slow-speed running.

Interest in the system is based on the fact that the air gap may be made very much larger than that of the attraction system which may reduce the degree of precision demanded in the construction of the guideway.

Investigation is proceeding in a number of countries. In Britain at the University of Warwick, in Germany by a consortium of firms (embodying principally A.E.G., Brown Boveri, and Siemens), in Canada where a drum-type test-rig has been installed at the Canadian Institute of Guided Transport at Kingston, Ontario. System parameters for an

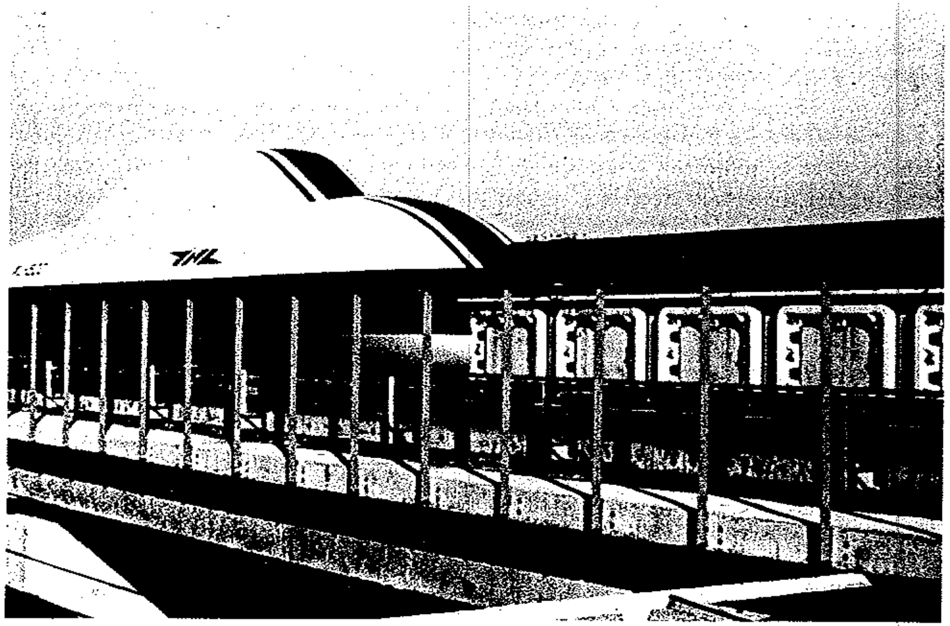


FIG. 17.45. High-speed test track, Japan.

The transverse configuration of the track used in these trials was an inverted "tee". The levitation system was based on coils mounted on the horizontal section whilst the vertical central spine carried coils which were fed with polyphase a.c. to act as the secondary of the linear synchronous motor and to provide guidance by the "no-flux" method.

The maximum speed achieved by the inverted tee system is 517 km/h (321 m.p.h.). Some concern has been felt regarding the stability of ride of the inverted T section and an inverted U has been adopted for future trials. Refrigeration will now be provided on-board to produce liquid helium at a temperature of −269°C to a more powerful I-shaped magnet capable of combining the functions of suspension, propulsion and guidance. The power of this magnet will be 700,000 ampere-turns.

17.12. Unexplored possibilities

The majority of the energy which must be applied to propel a vehicle at speed is devoted to the removal of air from its path and to overcome air resistance on its surface. It appears that, given a suitable configuration of the space between vehicle and track, much of this energy could provide lift, thereby reducing the power necessary to supply

... [illegible] that the system was stable with little rolling or pitching.[(33)]

Support for continued research on the "ram jet" approach has not been forthcoming but it is the writer's view that in certain respects it has advantages over the magnetic systems for operation at high speeds.

A further attempt to eliminate aero-dynamic losses on ground-based vehicles might be to contain them entirely within evacuated tubes. Thus Brunel's atmospheric system may have its place in tomorrow's technology but with the pipe made sufficiently large to fully embrace the vehicle.

References

1. Kolbuszewski, J., *Environment—Quantitative Approach*, Town and Country Planning Association, National Conference, 1966 (see also: Human environment of the future, *J. Soc. Engrs.*, April–June 1967, p. 39).
2. Ginns, H. N., English motorways—development and progress. *Proc. Inst. Civ. Engrs.*, vol. 37, p. 449 (1967).
3. Buchanan, C., *Traffic in Towns*, Penguin Books in association with H.M.S.O. (1963).
4. Austin, W. T. F. and Nash, H. E. C., Thoughts on vehicle design for safety. *Proc. Instn. Mech. Engrs.*, vol. 183, pt. 3A, p. 62 (1968).
5. Brockman, M. J. and Michaels, D., Computer-aided signal systems for motorways. *J. Sci. Technol.*, vol. 36, p. 80 (1969).
6. Han, E. A., Quinn, C. E., Stevens, J. E. and Trabold, W. G., DAIR—a new concept in highway communication for added safety and driving convenience. *Instn. Elect. Electron. Engrs. Trans. on Vehicular Technology*, vol. VT-16, no. 1, p. 33 (1967).
7. Bouladon, G., Transport. *Science Journal*, October 1967, p. 93.

8. Touch, J. M. and O'Flaherty, C. A., *Passenger Conveyors*, Ian Allan, London, 1971.
9. Blake, L. R., A public transport system using four passenger self-routing cars. *Proc. Instn. Mec Engrs.*, vol. 181, pt. 3G, p. 64 (1966).
10. Barwell, F. T. (in discussion), *Proc. Instn. Civ. Engrs., Transportation Engineering Conferenc* April 1968, p. 113.
11. Ishi, T., Masakaya, Nakaham, T., Kohsaka, Y. and Yasatrsugu, D., Computer controlled min car systems in Expo '76. An experiment in a new personal urban transport system. *Instn. Elec Electron. Engrs. Trans. on Vehicle Technology*, vol. VI-21, p. 77 (1972).
12. Ishi, T., Kashida, K., Kinoshita, K. and Takaoha, H., The control system of CVS using the tw target tracking scheme. *Proc. 1973 International Conference on Personal Rapid Transit, Minneapoli Minnesota, U.S.A.*
13. Bocock, D. and King, B. L., The development of the advanced passenger train. *Railway Engine* 1982/issue, p. 20.
14. Sell, R. G., Prince, G. E. and Twine, D., An experimental study of the overhead contact system f electric traction at 25 kV. *Proc. Instn. Mech. Engrs.*, vol. 179, pt. 1, p. 765 (1964).
15. Beadle, A. R. and Adams, C. J., The development and scaled line testing of a servo-pantograp *British Railways Board R. and D. Division tech. Mem. TMETR 60* (1975).
16. Seifert, W. W., *Survey of Technology for High Speed Ground Transport*, Part 1, Massachuset Institute of Technology, Cambridge, Massachusetts. June 1965.
17. Comparative assessment of new form of inter-city transport. *TRRL Report SR 1*, vols. 1 to December 1970. Reissued April 1973.
18. Parkinson, T. E. (Ed.), Tracked air-cushion vehicles in the Canadian corridor, *Canadian Transpo Commission Research Branch, Report 07*, September 1970.
19. Lake, R. W., Boon, C. J., English, G. W., Schwier, C., Fitzpatrick, C., Bunting, P. M. an Eastham, P. R., Alternatives to air, a feasible concept for the Toronto Ottawa Montreal corrid

22. Chirgwin, K. M., Test results from the U.S. linear motor research program. *I.E.E. Conferen Publication No. 120*, p. 236, together with addendum, October 1974.
23. Fellows, T. G., High speed surface transport. *Inst. Mech. Engrs. Railway Engineering Journa* vol. 3, no. 2, p. 4 (November 1973).
24. Select Committee on Science and Technology, *Second Report, Session 1975–76, Advanced Groun Transport*, H.M.S.O., London.
25. Select Committee on Science and Technology, *Report, Session 1972–73, Tracked Hovercra* H.M.S.O., London.
26. Laithwaite, E. R., Linear propulsion by electro-magnetic river. *International Conference on Hoverin Craft, Hydrofoils and Advanced Transit Systems*, Brighton (1974).
27. Katz, R. M. and Eastham, T. R., Single-sided linear induction motor with cage and solid ste reaction rails for integrated magnetic suspension and propulsion of guided ground transport. *Pro Instn. Elect. Engrs.* CH 1575-0/80/000-268 (1980).
28. Katz, R. M., Linear induction machine with solid iron reaction rail for integrated magnetic suspen sion and propulsion. *UMTA-VA-06-0068-80-1*. U.S. Department of Transportation (1981).
29. Levi, E., High speed, iron-cored synchronously operating linear motors. *I.E.E. Conference Publica tion No. 120*, p. 155 (1974).
30. Matsui, K., Umemori, T., Takeтsusi, Y. and Hosoda, Y., D.C. linear motor controlled by thyristo and testing equipment for its high-speed characteristics. *Ibid.*, p. 149.
31. Jayawant, B. W., Hodkinson, R. L., Wheler, A. R. and Whorlow, R. J., Transducers and the influence in the design of magnetically suspended vehicles. *I.E.E. Conference Publication No. 11* p. 200 (1974).
32. Hazlerigg, A. D. C. and Sinha, P. K., Design of a multi-variable controller for a magneticall suspended vehicle. *Ibid.*, p. 233.
33. U.S. Department of Transportation—Office of Research and Development, *Federal Railroa Administration Tenth and Final Report on the High Speed Ground Transportation Act of 1965.*

APPENDIX I

Control Theory—Guide to Literature

A1. Equivalence of dynamic and active systems

Consider the simple linear differential equation describing a system comprising one mass, one dashpot and one spring acted upon by a force which varies with time.

Then
$$M\frac{d^2x}{dt^2} + \eta\frac{dx}{dt} + kx = F(t) \tag{A1}$$

where M = mass,
η = coefficient of viscous friction,

of momentum which will immediately be seen to be equivalent to

$$L\frac{di}{dt} + Ri + \frac{1}{C}\int_0^t idt = V \tag{A3}$$

representing the electrical circuit of Fig. A1 where I = inductance, R = resistance, C = capacitance and i and v represent the instantaneous values of current and voltage respectively. Here the circuit components are in series, the current being constant throughout the circuit but the voltages additive.

Where the elements are in parallel as in Fig. A2 the voltage is constant throughout the circuit but the currents are additive, thus

$$\frac{1}{L}\int vdt + \frac{v}{R} + C\frac{dv}{dt} = i. \tag{A4}$$

Reverting to the mechanical system of equation (A1) and representing this in Fig. A3, in the absence of an externally applied force and neglecting damping

$$M\frac{d^2x}{dt^2} + kx = 0 \tag{A5}$$

or
$$M\frac{dx}{dt} + k\int_0^t xdt = 0 \tag{A6}$$

which will be equivalent to the electric circuit of Fig. A4 which is represented by the equation

$$L\frac{di}{dt} + \frac{1}{C}\int_0^t idt = 0. \qquad \text{(A7}$$

This will be recognised as an oscillatory circuit having a resonant frequency of $1/2\pi\sqrt{LC}$ The amplitude of the oscillation will depend on the initial conditions and will remair unchanged.

Let $i = a\cos\omega t$ then $di/dt = -a\omega\sin\omega t$

and $\int idt = a/\omega\sin\omega t$ (A8

thus the general solution is $\omega = \dfrac{1}{\sqrt{LC}}$ and $i = a\cos\dfrac{t}{\sqrt{LC}}$. (A9

There will be a value of frequency

$$f = \frac{\omega}{2\pi} = \frac{1}{2\pi\sqrt{LC}}$$

at any time t, the situation can be described completely by a "particular" solution. Thu if amplitude is A units of length when $t = 2\pi\sqrt{M/k}(n + 1)$ where n is zero or an positive integer

$$x = A\cos\omega t \qquad \text{(A10}$$

where
$$\omega = \sqrt{\frac{k}{M}}.$$

Road and rail vehicles are generally fitted with springs giving a static deflection o between 0·1 and 0·2 metre (4 and 8 inches) when unloaded. The stiffness to mass ratio i therefore determined within fairly close limits as follows:

$$k/M = \text{g/static deflection}. \qquad \text{(A11}$$

Therefore the natural frequency of bounce $= \dfrac{1}{2\pi}\sqrt{\dfrac{g}{\text{static deflection}}}$ (A12

= 1·11 Hz for 8 inches static deflection or
1·56 Hz for 4 inches static deflection. (A13

Equation (A10) can be represented by a rotating vector tracing out a circle as in Fig. A5 If A represents the situation at any point in time, the intercept of the perpendicula through A with the X-axis represents displacement and that of the horizontal with th

Y-axis represents velocity to a scale which includes the factor ω arising from the differentiation.

If set in motion, the system of Fig. A3 would continue to operate on one of a series of concentric circles. These circles, as in Fig. A6, represent a constant frequency with amplitude determined by initial conditions.

If the effect of damping in the system is taken into account, it can be shown from equation (A1) that action of the system after force has been removed is represented by

$$x = \epsilon^{-\eta t/2M}(A \cos \omega_1 t + iB \sin \omega_1 t) \quad \text{where} \quad \omega_1 = \sqrt{\omega^2 - \frac{\eta^2}{4M^2}} \qquad \text{(A14)}$$

and the motion can be represented on the phase-plane diagram of Fig. A7 as a trajectory forming part of a stable node or focus, depending on the value of $\frac{\eta}{M}$.

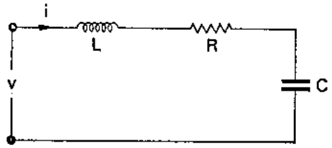


FIG. A1. Series circuit.

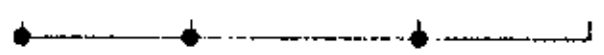


FIG. A2. Shunt circuit.

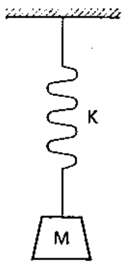


FIG. A3. Mechanical system.

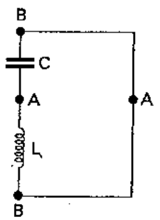


FIG. A4. Diagram illustrating notation for simple system.

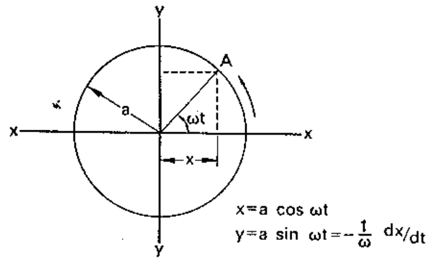


FIG. A5. Harmonic motion.

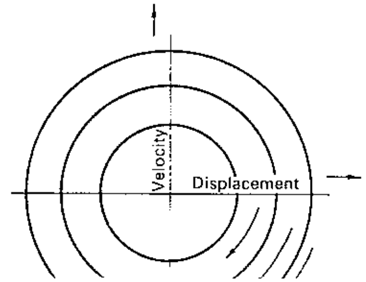


FIG. A6. Phase-plane representation—undamped.

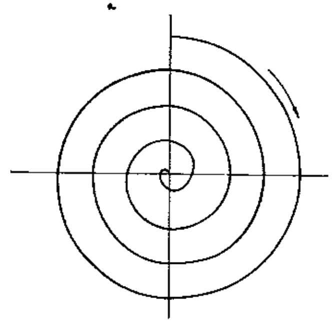
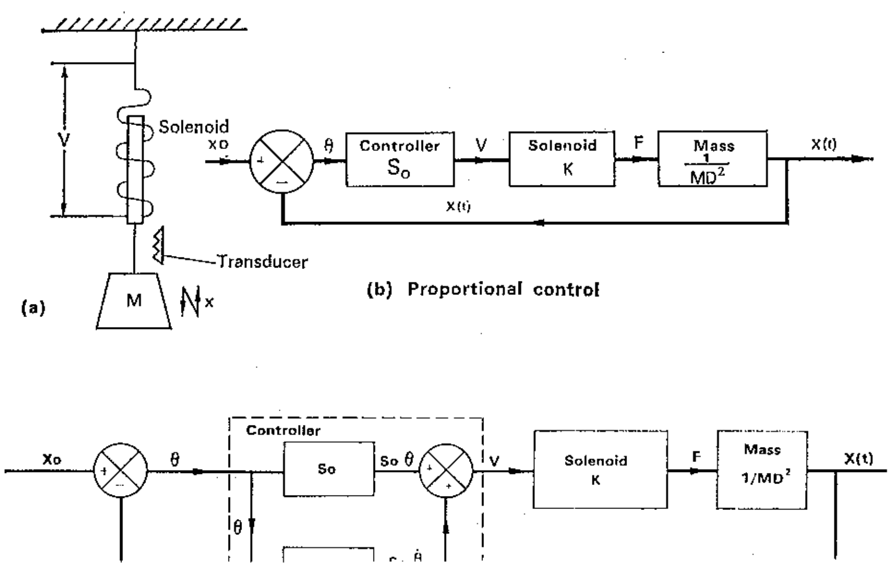


FIG. A7. Phase-plane representation of damped system.

Supposing instead of the spring in Fig. A3, a transducer were fitted to measure the position of the mass as in Fig. A8 and some form of actuator such as a pneumatic cylinder or solenoid was excited so as to apply a force proportional to displacement acting in a direction tending to oppose deviation from the initial condition, this would be a case of negative feedback. Such a system is known as an "active" system. Then in place of equation (A5), equation (A15) can be written

$$M\frac{d^2x}{dt^2} = -S_0(x(t) - x_0) \qquad \text{(A15)}$$



(c) Proportional + derivative control

FIG. A8. Active control circuits.

where $x(t)$ equals displacement at time t and x_0 the initial position. S_0 equals the gain in the feedback circuit as defined in Table 12.2. Writing θ for $x(t) - x_0$ the solution becomes

$$\theta = \cos \omega t \quad \text{where} \quad \omega = \sqrt{\frac{S_0}{M}} \qquad \text{(A16)}$$

which corresponds to equation (A10).

If displaced, the system will continue to oscillate with period $2\pi\omega$.

Suppose a further loop were provided to add the term $-S_1\, d\theta/dt$

then
$$M\frac{d^2x}{dt^2} = -S_0\theta - S_1\frac{d\theta}{dt}. \qquad \text{(A17)}$$

and the solution becomes

$$\theta = \epsilon^{-(S_1/2M)t}\,(A\cos\omega_1 t + iB\sin\omega_1 t) \quad \text{where} \quad \omega_1 = \sqrt{\omega^2 - \frac{S_1^{\,2}}{4M^2}} \qquad \text{(A18)}$$

which corresponds to equation (A14).

Thus the incorporation of derivative control in an active system is precisely equivalent to the provision of damping in a *passive* or dynamic system.

Were an additional mass placed on M the system would settle down to a new position in exactly the same way as the system of Fig. A3 where the spring would be permanently

extended. This corresponds to the "droop" of a governor and can be corrected in cisely the same manner, that is by the addition of an element of "feedback" proporti to the integral of error. Thus a "three-term" controller is needed to bring mass M I to its initial position under all conditions of externally applied load. Modern a automobile suspensions embody this principle.

Control system design has become increasingly dependent on digital treatm Nevertheless it is still important for a designer to possess skill in manipulating the established graphical aids to system design (such as the root-locus and Bode diagr to aid him in formulating preliminary designs of control systems although it ma unnecessary for him to achieve great accuracy in plotting because computer prog for calculating and tracing these plots are now available.

Where computer methods are particularly advantageous is in the treatment of "m variable" systems, particularly when different outputs are coupled to a greater or le degree—the steering and levitational behaviour of a Maglev vehicle for example.

It is of course desirable to limit the degree of interaction or cross-talk in a system to reduce the design problem to a number of uncoupled loops in order that the well-methods used to design one-input/one-output systems may be applied to each such l

Multi-variable control theory has been treated by Layton and Fossard.*

principles are mentioned below for ease of reference but for derivations and applicati the reader is referred to one of the specialist texts now available, some of which are li below. Each entry is preceded by a roman numeral which is used in the subsequent n together with page numbers to aid the reader in referring to the chosen text.

References

I. DAVIES, P. D. A., *An Introduction to Dynamic Analysis and Automatic Control*, Wiley.
II. CHESTNUT, H. and MAYER, R. W., *Servo-mechanisms and Regulating System Design*, Wiley
III. DORF, R. C., *Modern Control Systems*, Addison-Wesley Publishing Company, Readi Massachusetts.
IV. HARRISON, H. L. and BOLLINGER, I. G., *Introduction to Automatic Controls*, Intertext Bo London.
V. MACMILLAN, R. H., *Non-linear Control Systems Analysis*, Pergamon Press.
VI. MARTIN, H. R., *Introduction to Feedback Systems*, McGraw-Hill, London.
VII. NEWCOMB, R. W., *Concepts of Linear Systems and Controls*, Brooks/Cole, Belmont, Califo
VIII. NASLIN, P., *The Dynamics of Linear and Non-Linear Systems*, Blackie.
IX. PITMAN, R. J. G., *Automatic Control Systems Explained*, Macmillan.
X. PORTER, A., *Introduction to Servo-mechanisms*, Methuen.
XI. ROCARD, Y., *Dynamic Instability*, Crosby Lockwood.
XII. SAGE, A. D., *Linear Systems Control*, Pitman, London.
XIII. SAVANT, C. J., *Control System Design*, McGraw-Hill.
XIV. SHEARER, J. L., MURPHY, A. T. and RICHARDSON, H. H., *Introduction to System Dyna* Addison-Wesley.
XV. SHINNERS, S. M., *Control System Design*, Wiley.
XVI. TAYLOR, P. L., *Servo-mechanisms*, Longmans.

* Layton, J. M., *Multi-variable Control Theory*, Peter Peregrinus, London (1976). Fossard, A *Multi-variable System Control*, North Holland Publishing Company, Amsterdam.

XVII. Thaler, G. J. and Brown, R. G., *Analysis and Design of Feedback Control Systems*, McGraw-Hill.
XVIII. Webb, C. R., *Automatic Control*, McGraw-Hill.
XIX. Welbourne, D. B., *Essentials of Control Theory for Mechanical Engineers*, Arnold.
XX. West, J. C., *Servo-mechanisms*, English Universities Press.

A3. The Laplace transform–transfer function

Harmonic response

Recalling equation (1.5) $V(\tau D + 1) = F/\eta$ (p is written for D in some texts). Supposing instead of the forcing function F/η a sinusoidal input $\sin \omega t$ is applied. Then

$$V(\tau D + 1) = \sin \omega t. \qquad \text{(A19)}$$

For the particular integral let the output be written

$$V = A \sin (\omega t - \psi). \qquad \text{(A20)}$$

Substituting in (A19)

Equating coefficients of $\sin \omega t$ and $\cos \omega t$

$$\tan \psi = \omega\tau \qquad \text{(A21)}$$

and

$$A = \frac{1}{\sqrt{1 + \omega^2\tau^2}}.$$

The complementary function is of the form $R\epsilon^{-t/\tau}$ as before. If when $t = 0$, $V = 0$

$$0 = \frac{1}{\sqrt{1 + \omega^2\tau^2}} \sin \psi + R,$$

$$R = \frac{\omega\tau}{1 + \omega^2\tau^2},$$

$\therefore$ the complete solution is

$$V = \frac{1}{\sqrt{1 + \omega^2\tau^2}} \left\{ \sin (\omega t - \tan^{-1} \omega\tau) \right\} + \frac{\omega\tau}{1 + \omega^2\tau^2} \epsilon^{-t/\tau} \qquad \text{(A22)}$$

Recalling that $\psi = \tan^{-1}\omega\tau$, when $\psi = \pi/4$, $\omega\tau = 1$, therefore delay $= \psi/\omega = \pi\tau/4$ seconds which gives further significance to the description of τ as the time constant.

In the steady state the final term of equation (A22) becomes zero.

It is convenient to represent the response to any input of period $2\pi\omega$ as a vector or *phasor* on an Argand diagram as in Fig. A9. A is referred to as the "modulus" and ψ as the "phase".

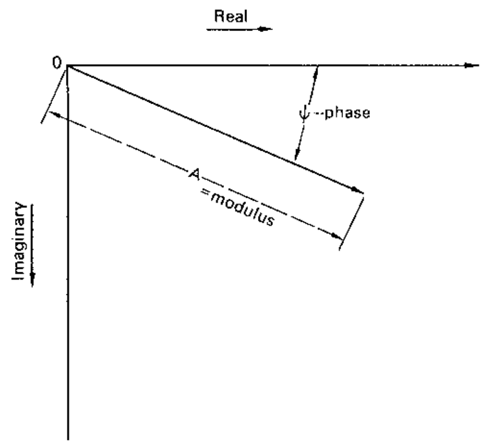


Fig. A9. Harmonic response phasor.

Real and complex roots

The quantity $1/(i\omega\tau + 1)$ is usually written with modulus as $|\, T(i\omega) \,|$ and phase as $\angle T(i\omega)$.

Being concerned solely with steady state conditions, harmonic analysis deals only with the particular integral. The complete solution contains the term $\epsilon^{-t/\tau}$ which is real and it can be shown that the function $\sigma + i\omega$, known as the *complex pulsatance*, is required for any completely general treatment. This is generally known by the symbol "s" (although p is sometimes used here also) and can be represented on another Argand diagram known as the "s-plane" (see Fig. A10). The quantity written $T(s)$ is known as the "Transfer function".

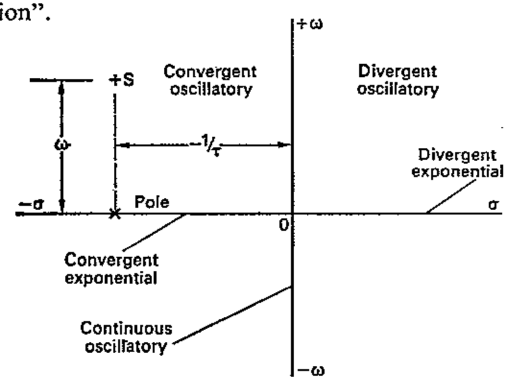


Fig. A10. The "s-plane".

The values of s making $T(s) = \infty$ are known as *poles* and those making $T(s) = 0$ as *zeros*. These are usually shown as crosses and circles respectively.

I. 20–67; II. 20–32; IV. 74–97; VI. 63–107; VII. 1–10; VIII. 137–61; IX. 92–95; XIII. 17; XIV. 252–63, 323–330; XVIII. 29–40; XIX. 4–9; XX. 100–8.

The Laplace transform

This function is valuable in the manipulation of control equations in the same manner as logarithms facilitate numerical calculations and tables are available which can be used in a similar manner.

A function of t may be transformed into a function of the complex variable s by multiplying it by ϵ^{-st} and integrating from 0 to ∞ thus

$$F(s) = \mathscr{L}[f(t)] = \int_0^\infty f(t)\epsilon^{-st}dt. \qquad \text{(A24)}$$

Some useful examples are given in Table A1.

Reverting to equation (1.5) this may be written $V_0(\tau D + 1) = V_i$ where V_i represents the step function input denoting the driver's sudden decision to increase speed to the new value. The Laplace transform of this

TABLE A1. SOME USEFUL TRANSFORMS

$f(t)$	$\mathscr{L}[f(t)] = F(s)$	$F(z) = e^{sT} = Z[f(kT)]$
1. 1 (unit step)	$1/s$	$1/(1 - z^{-1})$
2. t	$1/s^2$	$Tz^{-1}/(1 - z^{-1})^2$
3. e^{-at}	$\dfrac{1}{s + a}$	$\dfrac{1}{(1 - e^{-aT}z^{-1})}$
4. $t\epsilon^{-at}$	$\dfrac{1}{(s + a)^2}$	$\dfrac{Te^{-aT}z - 1}{(1 - e^{-aT}z^{-1})^2}$
5. $1 - \epsilon^{-at}$	$\dfrac{a}{s(s + a)}$	$\dfrac{(1 - e^{-aT})z^{-1}}{(1 - z^{-1})(1 - e^{-aT}z^{-1})}$
6. $e^{-at} - e^{-bt}$	$\dfrac{(b - a)}{(s + a)(s + b)}$	$\dfrac{(e^{-aT} - e^{-bT})z^{-1}}{(1 - e^{-aT}z^{-1})(1 - e^{-bT}z^{-2})}$
	$(a - b)s$	$(e^{-aT} - e^{-bT})z - 1$

Referring to Table A1, line 5, and writing a for $1/\tau$ this gives the inverse transform

$$1 - \epsilon^{-at} \quad \text{or} \quad 1 - \epsilon^{-t/\tau}. \qquad \text{(A25)}$$

This is identical in form to equation (1.7).

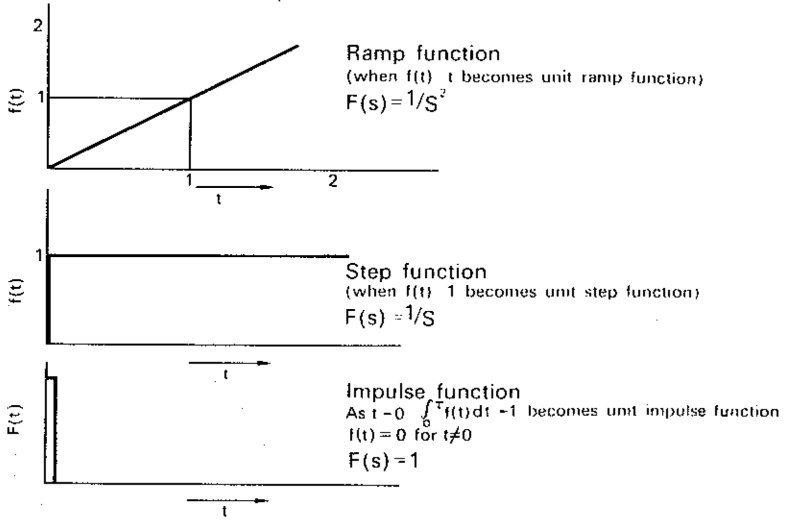


FIG. A11. Laplace transforms of input functions.

9. $\cos \omega t$	$\dfrac{s}{s^2 + \omega^2}$	$\dfrac{1 - (\cos \omega T)z^{-1}}{1 - 2(\cos \omega T)z^{-1} + z^{-2}}$
10. $\epsilon^{-at} \sin \omega t$	$\dfrac{\omega}{(s + a)^2 + \omega^2}$	$\dfrac{(e^{-at} \sin \omega T)z^{-1}}{1 - 2(e^{-aT} \cos \omega T)z^{-1} + e^{-2aT}z^{-2}}$
11. $\epsilon^{-at} \cos \omega t$	$\dfrac{s + a}{(s + a)^2 + \omega^2}$	$\dfrac{1 - (e^{-aT} \cos \omega T)z^{-1}}{1 - 2(e^{-aT} \cos \omega T)z^{-1} + e^{-2aT}z^{-2}}$
12. $\sinh \omega t$	$\dfrac{\omega}{s^2 - \omega^2}$	$\dfrac{(\sinh \omega T)z^{-1}}{1 - 2(\cosh \omega T)z^{-1} + z^{-2}}$
13. $\cosh \omega t$	$\dfrac{s}{s^2 - \omega^2}$	$\dfrac{1 - (\cosh \omega T)z^{-1}}{1 - 2(\cosh \omega T)z^{-1} + z^{-2}}$

This example which illustrates the method of using Laplace transforms for the solution of system equations is, of course, trivial. However, much labour can be saved in more complex cases.

II. 70–102; III. 31–48; IV. 98–122; VI. 313–16; VII. 25–36; VIII. 552–71; XII. 21–49; XIII. 385–98; XIV. 391–405; XV. 9–18; XVI. 279–93; XVII. 9–29; XVIII. 40–42, 288–9; XIX. 176–81.

A4. Effect of time delay

In many systems, for example that represented by the car-driving equation (2.9), a time lag exists between stimulus and response. This sometimes occurs in process control where a quantity is measured at a point displaced by some distance from the agency in the process which determines that quantity. Thus some time must elapse during which the material under treatment travels from the operating agency to the measuring point. The effect is therefore sometimes known as transportation lag. Other terms are "finite time delay", "dead time", "pure lag" (in contrast to exponential lag, see Section 1.2) and "distance velocity lag".

Supposing $f_0 t = f_1(t)(t - T)$ where $f_0(t)$ represents the output of the system for input $f_1(t)$.

The right-hand side can be expanded by Taylor's series as

$$f_i(t) - T\dot{f}_i(t) - \frac{T^2}{2!}\ddot{f}(t) \quad \dots \quad \frac{T^n}{n! - 1}\overset{n}{f}_i(t), \text{ etc.}$$

$$= \epsilon^{-TD} f_i(t).$$

A5. Criteria for stability

The quality of a system is determined by two criteria, performance and stability. Performance is determined by how closely output follows input. Stability ensures that the system will not oscillate violently or allow the output to increase indefinitely until either damage is caused to the equipment or some limiting factor intervenes. This implies that whenever the complementary function which governs transient behaviour contains exponential terms with real or complex roots, the real roots or the real part of complex roots must be negative.

A number of methods are available to determine compliance with this requirement without it being necessary to solve the actual equations.

A method put forward by Routh and Hurwitz (1905 and 1895 respectively) consists of examining the characteristic equation of the system without actually solving it.

Thus if this equation be expressed in the following forms,

$$a_0 D^n + a_1 D^{n-1} + a_2 D^{n-2}, \cdots, a_{n-1} D + a_n = 0,$$

the coefficients are arranged as follows:

$$\left.\begin{array}{lllll} a_0 & a_2 & a_4 & \dots & a_{(n-1)} \\ a_1 & a_3 & a_5 & & a_{(n)} \\ x_1 & x_2 & & & \\ y_1 & x_3 & & & \end{array}\right.$$

where

$$\left.\begin{aligned} x_1 &= \frac{a_1 a_2 - a_0 a_3}{a_1} \\ x_2 &= \frac{a_1 a_4 - a_0 a_5}{a_1} \\ y_1 &= \frac{x_1 a_3 - x_2 a_1}{x_1} \\ y_2 &= \frac{x_1 a_5 - x_3 a_1}{x_1} \end{aligned}\right\} \qquad \text{(A27)}$$

The cross multiplication is continued until a row of zeros appears.

If there are no positive real roots all the terms in the first column will be positive and therefore the system will be stable. If there exist roots with real positive parts their number will be indicated by the number of sign changes in the first column.

I. 193–248; II. 6, 264; III. 145–9; IV. 87–95; VI. 128–33; VIII. 238–303; IX. 53–62; X. 48–109 XI. 37–82; XIII. 406–8; XIV. 131–2; XV. 107–10; XVI. 302–9; XIX. 26–30; XX. 79–83.

The Routh–Hurwitz criterion, whilst providing a rapid method for determining whether or not a system is stable, provides no indication of the margin of stability or the quality of performance, neither does it provide direct guidance as to desirable modifications to an unstable system.

Consider the system

$$G(s)H(s) = \frac{1}{(\tau_1 s + 1)} \times \frac{K}{s(\tau_2 s + 1)} \qquad \text{(A28)}$$

where τ_2 might be the time constant of a vehicle and τ_1 that of some element in a control system which also embodies an integration feature represented by $1/s$.

$G(s)$ represents the transfer function of the forward elements and $H(s)$ that of the feedback elements. Such a system can be studied using the root-locus method introduced by Evans in 1950.

Taking first the simpler case where

$$G(s) = \frac{K}{\tau s + 1}$$

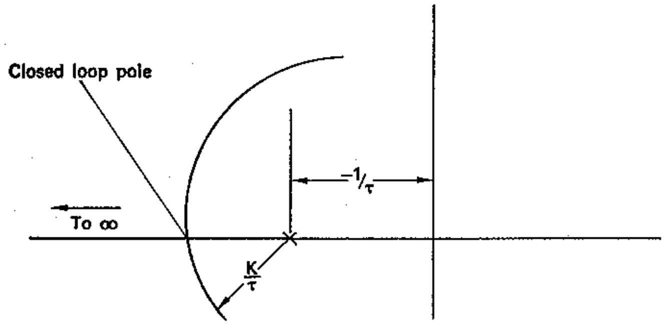
representing an automobile in which the driver is following a strategy of proportional control with "gain" or amplification K.

Writing $$A = \frac{K}{\tau}, \quad F(s) = \frac{1}{s + 1/\tau}$$

and referring to Fig. A10 there is one finite pole situated on the real axis at $-1/\tau$ and one zero at infinity.

Now $1 + G(s) = 0$, $\therefore G(s) = AF(s) = -1$,

$$\therefore |F(s)| = \frac{\tau}{K}; \quad \angle F(s) = 180°, \tag{A29}$$



therefore the locus of the root starts at the pole and continues leftwards along the real axis to infinity. The closed-loop pole will be found on the intersection of a circular arc of radius K/τ centred at the pole with the real axis, Fig. A12. Thus whatever the value of K adopted by the driver, the root will always lie to the left of the origin, indeed on the left of $-1/\tau$ so that the system will always be stable. Thus the main factor in τ, namely the increase in air resistance, ensures stability in the same way as a damping device might improve the stability of a governor.

Reverting to the more complex system of equation (A27)

$$G(s)H(s) = A \times \frac{1}{s} \times \frac{1}{s + 1/\tau_1} \times \frac{1}{s + 1/\tau_2} \tag{A30}$$

where $A = K/\tau_1\tau_2$.

There will now be three loci on the real axis starting at 0, $-1/\tau_1$ and $-1/\tau_2$ and ending at infinity. The root locus plot for varying values of A will be as shown in Fig. A13, curve (a). The system becomes unstable at the values of A corresponding to the points where this curve crosses the imaginary axis.

Rules for plotting the root locus are as follows:

1. Plot position of poles and zeros.
2. Each branch of a root locus diagram starts at a pole and terminates at a zero.
3. There are as many branches as there are roots of the characteristic equation of the system.
4. Loci start at poles when $A = 0$ and end at zeros when $A = \infty$.

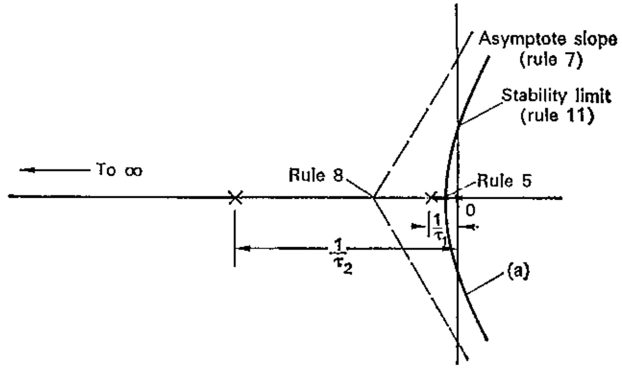


FIG. A13. Root locus plot for second order system with integration.

5. The root locus crosses the real axis to the left of an odd number of poles or ze on that axis.
6. When zeros are at infinity the asymptote of the locus can be obtained from tl characteristic equation slope and intercept being determined separately as belo

for $n = \pm 1, \pm 3$, etc.

8. The asymptotes leave the real axis at the centre of gravity of the poles and zero Real parts need only be considered because complex poles and zeros only occur i conjugate pairs. The displacement of the intersection of the asymptotes with tl real axis is then given by
$$\frac{\text{sum of poles} - \text{sum of zeros}}{\text{number of poles} - \text{number of zeros}}.$$
9. A root locus based on two poles on the real axis consists of two branches whic approach each other on the real axis until they meet and then turn sharply rigl and left of that axis. This occurs when $|F(s)|$ is a minimum and is determined t differentiating and equating to zero.
10. The inclination in radians of a locus at the point of origin from a complex pole (terminating at a complex zero is obtained by summing the angles from all the oth poles and zeros and subtracting 2π.
11. It is seldom necessary to plot the whole locus with accuracy because the onset (instability can be calculated by substituting $i\omega$ for s in $H(s)G(s)$ and equating t zero. Amplification factor K is usually the easiest feature of a system to manipulat For low values, systems are generally stable but very sluggish. As A is increased s does speed of response with increasing tendency to overshoot and finally the syste may become unstable.

I. 202–9; II. 161–78; III. 165–202; IV. 263–88; VI. 198–211; VII. 137–65; XII. 296–350; XII 99–144; XIV. 322; XV. 143–53; XVI. 320–39; XVII. 176–97; XVIII. 76–117; XIX. 182–4.

A7. Nyquist and Bode representation

It is frequently sufficient to study the steady state response of a system to sinusoidal excitation in order to evaluate its stability. Thus only the imaginary part of the roots of the system equation are involved and s may be replaced by $i\omega$. The response of the system

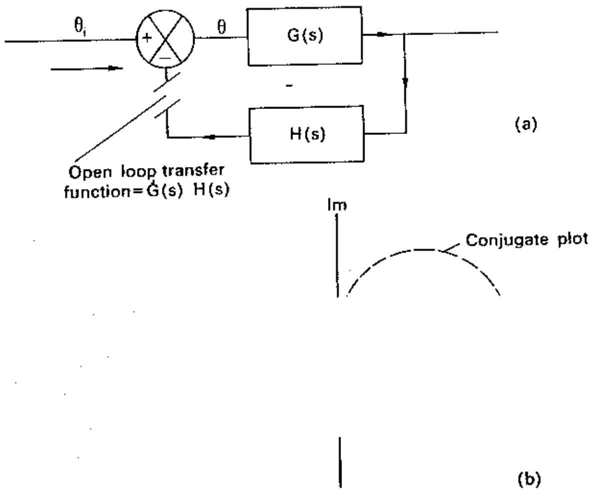


FIG. A14. Nyquist plot.

to unity input may be represented by a vector whose magnitude represents the amplification or attenuation and whose angle represents the phase difference between input and output .This again makes use of an Argand diagram but the actual functions are plotted and not the roots as previously.

In the system shown in Fig. A14(a), the "open-loop" transfer function is defined as $KG(s)H(s)$ as though the feedback loop has been interrupted at the summing point. The closed-loop function would be

$$\frac{KG(s)}{1 + KG(s)H(s)}.$$

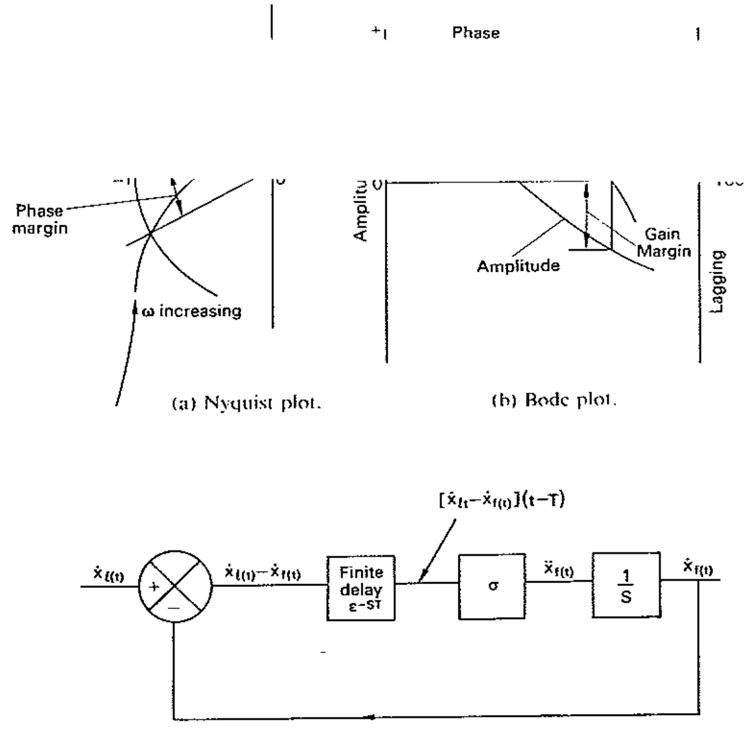
The polar plot of the function $KG(s)H(s)$ must not encircle the -1, $i0$ point on the polar diagram if the system is to be stable. This means that if and when the phase difference between input and output is π the denominator must be positive. In the case of equation (A18), $H(s)$ is unity and the response of the system to variable frequency and constant K is shown in Fig. A14(b). It will be noted that the diagram is situated wholly to the right of the imaginary axis so that whatever the value of K it cannot encircle the -1, $i0$ point. The system is therefore stable.

One or more clockwise encirclements of the -1, $i0$ point may occur. If there is net encirclement, i.e. one encirclement is not cancelled out by one in the opposite sense, the system will be unstable. If there are no encirclements or net encirclements in the anticlockwise direction, further tests may be necessary to establish stability. In most cases, however, the locus takes a simple form and a useful working rule is that a hypothetical individual travelling along the locus from $\omega = 0$ to $\omega = \infty$ would pass the -1 $i0$ point on his left-hand side if the system were stable.

The diagram allows an estimate to be made of the margin of stability of the system This is defined in two ways; the gain margin which indicates by how much the amplification may be increased before instability occurs, and the phase margin. These are illustrate in Fig. A15(a).

An alternative method of presenting the frequency response of a system is known as th Bode diagram. Amplitude is plotted vertically, usually in units of decibels, and frequenc horizontally also on a logarithmic scale. Phase, lagging or leading is also plotted verticall against the same horizontal scale. Phase and gain margins are indicated in Fig. A15(b

Because the moduli of successive elements of a control circuit have to be multiplied t obtain the system modulus, logarithmic plotting is particularly convenient because mult



(a) Nyquist plot. (b) Bode plot.

(c) Block diagram.

FIG. A15. Phase and gain margins.

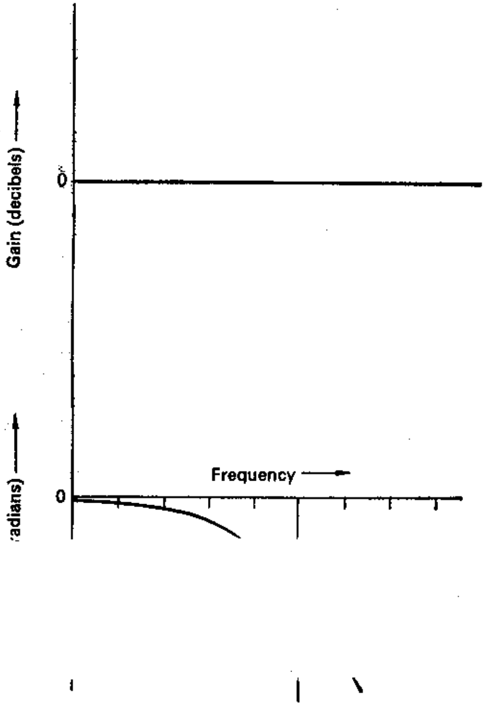


FIG. A 15 (d).

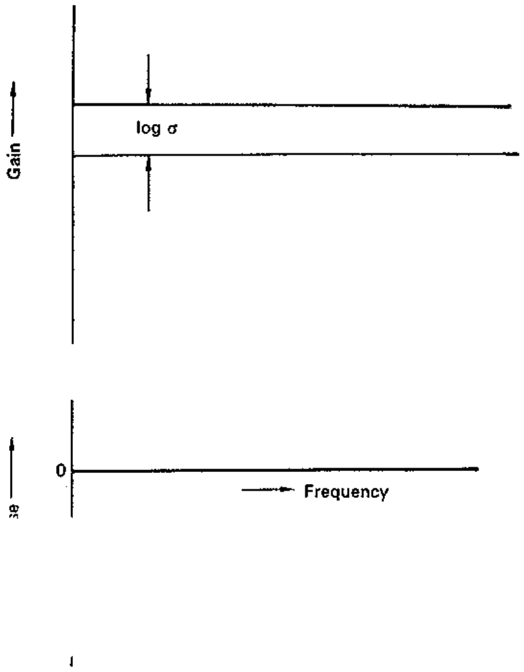


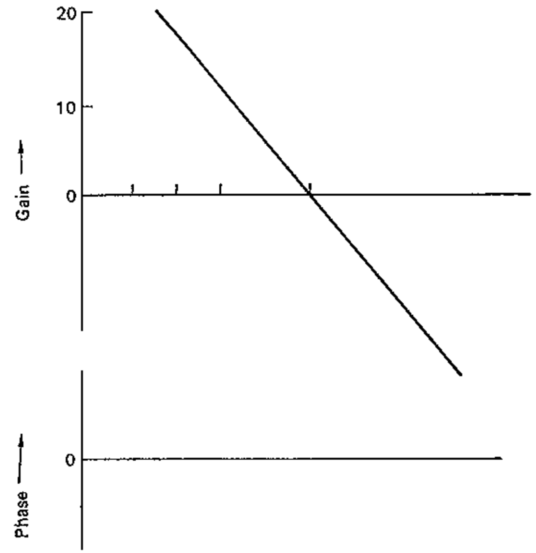
FIG. A 15 (e).

plication is replaced by graphical addition. Phase angles, of course, are added directly. Consider equation 2.7

$$\ddot{x}_f(t+T) = \sigma[\dot{x}_i(t) - \dot{x}_f(t)]$$

or T seconds earlier $$\ddot{x}_f(t) = \sigma[\dot{x}_i - \dot{x}_f](t-T).$$

This can be represented by a block diagram as in Fig. A15(c). Each element can be shown separately on a Bode diagram. Thus A15(d) shows the effect of a finite delay. Gain remains at unity, i.e. log gain $= 0$ and phase lags by ωT. This must be represented by a curve ωT plotted against log ωT. If we take T as $\pi/2$ for convenience then $-90° = -\pi/2$ occurs at pulsatance ratio unity or log pulsatance ratio $= 0$. σ is simply a constant represented by horizontal lines as in Fig. A15(e). The integration term is represented by straight lines. The modulus slopes downwards from left to right and the phase is horizontal at $-\pi/2$. The value of the modulus is $1/\omega T$ so that the slope becomes -6dB/octave. It crosses the origin when log $\omega T = 0$, i.e. when $\omega = 1/T$, Fig. A15(f). Combining Figs. A15(e) and (f) we have A15(g). This shows practically no phase or gain margin so that



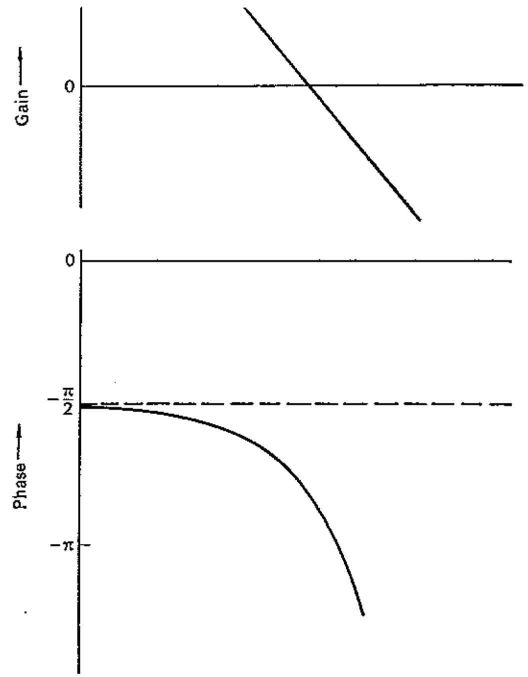


FIG. A 15 (g).

the system is on the verge of instability. If, however, we take σ as being less than unity its value is subtracted from the modulus as in Fig. A15(h) which gives phase and gain margins as indicated therein.

Referring to equation (2.9), σ may not be a constant but may take the form $\dfrac{K}{x_l(t) - x_f(t)}$ and a driver might compensate for an excessive reaction time T by

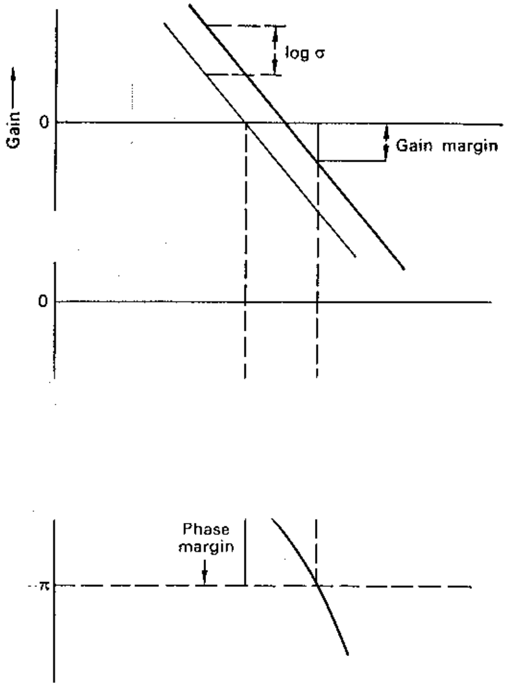


FIG. A 15 (h).

reducing K or increasing $x_l(t) - x_f(t)$. It is likely that learner drivers rapidly adopt a method of driving which avoids instability while allowing for their personal characteristics as well as those of their vehicles. Nevertheless, the car-following laws indicate that there are limits to route capacity consistent with stability. In abnormal conditions, low visibility for example, drivers may attempt to exceed these limits and so introduce the risk of multiple collisions.

I. 219–48; II. 142–60, 317–43; III. 255–88, 357–408; IV. 201–59; VI. 133–44; VII. 184–99; VIII. 249–74; IX. 70–84; X. 66–95; XII. 202–68; XIII. 149–209; XIV. 373–5; XV. 110–39; XVI. 309–16; XVII. 203–26; XVIII. 147–64, 200–24; XIX. 16–18, 45; XX. 91–9.

A8. Common forms of non-linearity

The formal definition of a non-linear effect is one wherein the principle of superposition does not apply. In appropriate circumstances almost any element of a system may exhibit

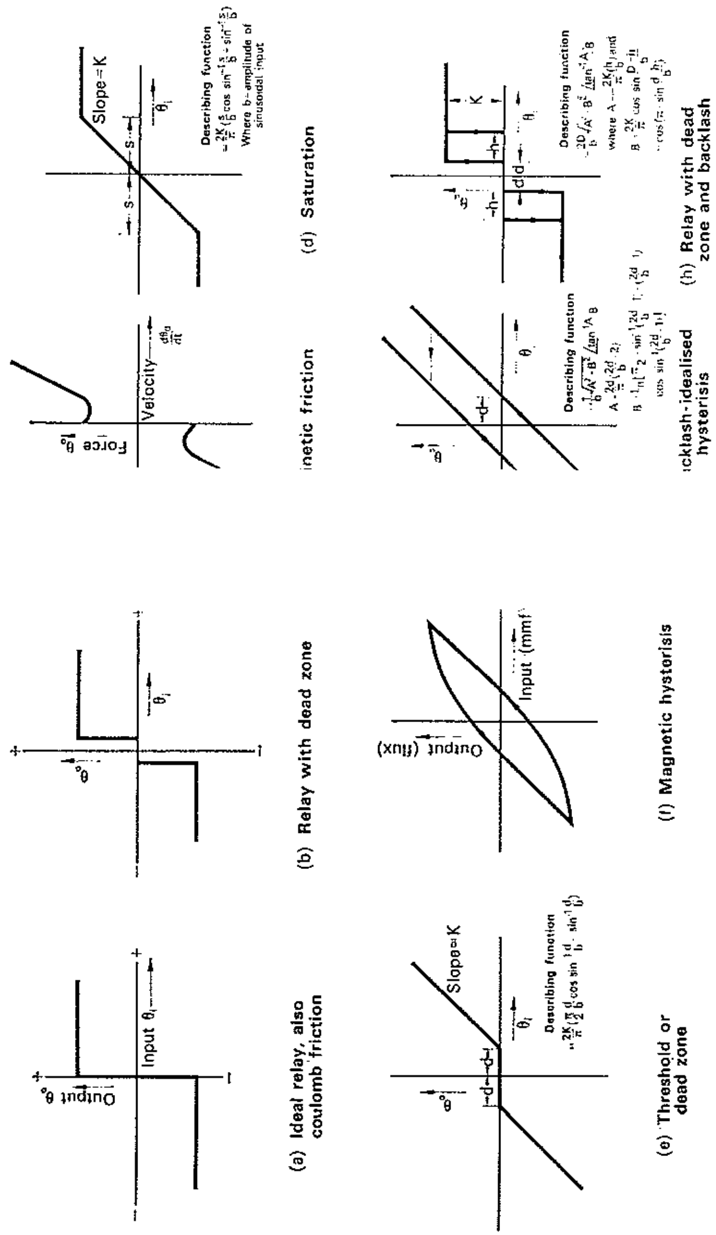


FIG. A16. Non-linear element… g functions.

non-linear behaviour. Even a hydraulic valve must eventually become fully open after which further manipulation of the control will not increase through-put.

A relay may have only two positions depending on whether the actuating signal is positive or negative in sign as shown in Fig. A16(a). It is usually impossible to achieve this in practice. For example, in an electrical circuit required, perhaps, to reverse the direction of rotation of a motor, the circuit has to be opened to interrupt current in one direction before it is closed to allow current to flow in the other direction. Otherwise there would be a dead short across the mains. An interval, however short, must occur therefore between the opening of one circuit and closing of another. One way of achieving this is to require the input signal to attain a small but definite value of either polarity before the switch or valve is closed. This is known as "dead zone" as shown in Fig. A16(b).

Figure A16(a) also represents the effect of coulomb friction in a mechanical element. The term "coulomb friction" implies that the coefficient of friction remains constant irrespective of the extent or speed of motion. The diagram illustrates the fact that friction always acts in the direction to resist relative motion of the two surfaces. A more realistic representation of friction appears in Fig. A16(c).

Elements embodying magnetic cores will become saturated when flux approaches a certain intensity and, as mentioned earlier, almost any element will possess only limited capacity. This feature is represented in idealised form by Fig. A16(d) wherein output is assumed to follow input linearly until a certain definite value is reached, after which it

proportional to input as shown in Fig. A16(f). The area within the curve represents energy which constitutes the iron losses of a machine. In mechanical systems, particularly when wear takes place, lost motion may occur so that when the direction of input changes, a delay must take place whilst "slack" is taken up. The position of the dotted line indicates that the point at which the transition takes place depends on the maximum value of the input. Figure A16(h) illustrates the case of a relay possessing both dead zone and backlash.

II. 576; IV. 321-31; V. 1-63; VIII. 337-8; XI. distributed; XIII. 352-6; XIV. 117, 213; XV. 215-22; XVI. 218-23; XVII. 479; XVIII. 255-7; XX. 223-30.

A9. The describing function

When the magnitude of the non-linear feature of a function is not large compared with its maximum value and where the system contains filtering or integrating elements subsequent to the non-linear input or element, an equivalent sinusoidal function (perhaps having amplitude and/or phase difference from the sinusoidal input) may be employed in the harmonic analysis of the system.

A Fourier analysis is made of the feature and all harmonics in excess of the first are ignored. Thus a sinusoidal input applied to an ideal relay would produce a square wave output as in Fig. A17. Input may be taken as $b \sin \omega t$. The sign of the output will change from positive to negative at $t = \pi/\omega$ and back again when $t = 2\pi/\omega$.

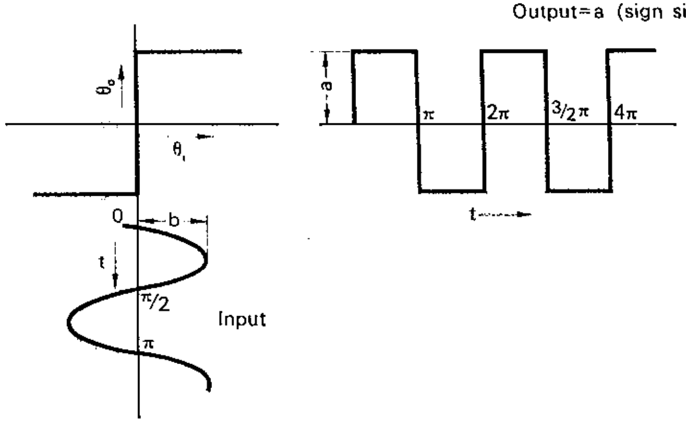


FIG. A17. Output from ideal relay.

The Fourier series describing the output will take the form expressed in equation (A31

$$f(t) = a_1 \sin \omega t + a_2 \sin 2\omega t + a_3 \sin 3\omega t, \ldots, a_n \sin n\omega t$$

where the a terms are

$$\frac{\omega}{2\pi}\int_0^{2\pi} f(t) \sin n\omega t dt \qquad \text{(A32}$$

$$= \frac{\omega b}{\pi}\int_0^{\pi/\omega} \sin n\omega t dt - \frac{\omega b}{\pi}\int_{\pi/\omega}^{2\pi/\omega} \sin n\omega t dt,$$

$$= \frac{4b}{n\pi},$$

$$\therefore f(t) = \frac{4b}{\pi}\left(\sin \omega t + \frac{1}{3}\sin 3\omega t + \frac{1}{5}\sin 5\omega t, \ldots, \text{etc.}\right).$$

Supposing this represents a torque applied to a rotating mass of moment of inertia

then $I \dfrac{d^2\theta}{dt^2} = f(t)$

or velocity $= \dfrac{d\theta}{dt} = \dfrac{1}{I}\int f(t)dt$

and displacement $= \theta = \frac{1}{I}\iint f(t)dtdt,$ (A33)

$$\therefore \text{angular velocity} = \frac{4b}{I\pi}\left(\frac{1}{\omega}\cos\omega t + \frac{1}{9\omega}\cos 3\omega t + \frac{1}{25\omega}\cos 5\omega t, \ldots, \text{etc.}\right)$$

$$\text{and displacement} = -\frac{4b}{I\omega^2\pi}\left(\sin\omega t + \frac{1}{27}\sin 3\omega t + \frac{1}{125}\sin 5\omega t, \ldots, \text{etc.}\right).$$

It will be observed that the coefficients of all terms but the first diminish rapidly with integration. In practice, therefore, these terms may be neglected. The describing function in the above case therefore becomes $4b/\pi a$.

The general expression for the describing function becomes therefore

$$f(\omega t) = \frac{a}{2} + \frac{2}{\pi}\int_{-T/2}^{T/2}\theta_0(\omega t)\cos\omega t\, d(\omega t)\cos\omega t$$
$$+ \frac{2}{\pi}\int_{-T/2}^{T/2}\theta_0(\omega t)\sin\omega t\, d(\omega t)\sin\omega t. \quad \text{(A34)}$$

Below $-S$ and above $+S$ output becomes

$$\mp K\sin\omega t.$$

The output will be in phase with the input and will be conveniently described by the approximate expression $B\sin\omega t$ where B/b is the describing function.

When b is less than S the system will respond linearly and the describing function may be written as unity.

When b exceeds S the response will be divided into two zones, behaving linearly when $\omega < \sin^{-1} b/s$ and having a constant output when $\omega > \sin^{-1} b/s$.

Let $\sin b/s = \beta$,

then
$$f(\omega t) = \frac{4Kb}{\pi}\left[\int_0^{\beta}\sin^2\omega t\, d(\omega t) + \int_{\beta}^{\pi/2}\frac{s}{b}\sin\omega t\, d(\omega t)\right]$$
$$= \frac{4Kb}{\pi}\left[-\tfrac{1}{2}\cos\omega t\sin\omega t + \tfrac{1}{2}\omega t\right]_0^{\beta} - \frac{s}{b}\left[\cos\omega t\right]_{\beta}^{\pi/2}$$
$$= \frac{2Kb}{\pi}\left(\frac{s}{b}\cos\sin^{-1}\frac{s}{b} + \sin^{-1}\frac{s}{b}\right)$$

or as a describing function

$$2\frac{K}{\pi}\left(\frac{s}{b}\cos\sin^{-1}\frac{s}{b} + \sin^{-1}\frac{s}{b}\right). \quad \text{(A35)}$$

The describing functions for some other non-linear features are shown in Fig. A16 where appropriate.

IV. 331–7; V. 63–96; VIII. 338–88; XIII. 356–66; XV. 223–67; XVII. 479–503; XVIII. 257–81.

A10. Inverse Nyquist or Whitely diagram

By using the describing function, systems embodying non-linear features may be treated as linear when investigating stability using Nyquist or Bode diagrams. However, amplitude must now be taken into account so that it is convenient to separate the system transfer function into its linear and non-linear elements for separate treatment.

Thus in the closed-loop transfer function

$$\frac{KG(s)}{1 + KG(s)H(s)} \quad \text{either } G(s) \text{ or } H(s)$$

can embody some non-linear element whose transfer function may be written $N(s)$.

Assuming unity feedback, and restricting the symbol $G(s)$ to representing the linear portion of the forward path

of amplitude and can be represented by a separate curve for each value of amplitude applied (Fig. A18(a)).

The transfer function can be written

$$\frac{1}{KG(s)N(s) + 1}$$

and its reciprocal becomes $KG(s)N(s) + 1.$

For stability $KG(s)N(s) > -1$

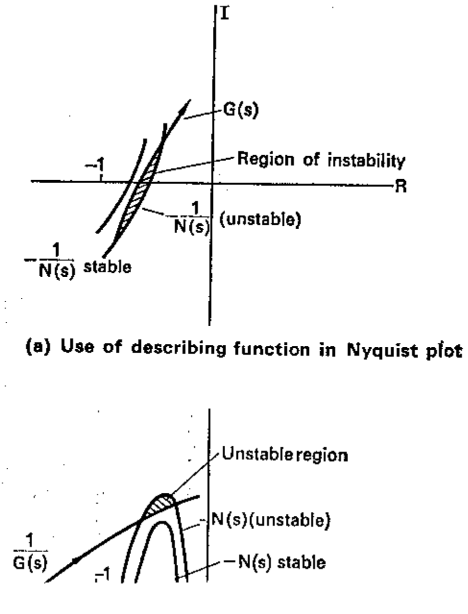
or $\frac{1}{KG(s)} > -N(s).$

If therefore one curve for $1/G(s)$ is plotted, curves for $-N(s)$ may be drawn from the describing function for a number of amplitudes. Intersection of the $1/G(s)$ and $-N(s)$ curves will indicate instability (Fig. A18(b)).

I. 215–17; II. 261–9; IV. 331–6; V. 81–97; VI. 186–8; VIII. 197–8; IX. 80–97; X. 130–6; XV. 124–8; XVIII. 162–4, 280–1.

A11. Liapunov's second method

Liapunov (1892), by employing energy considerations, developed two methods for determining the steady state stability of a system. His first method is of no special interest in the light of the foregoing material but his second method is of importance.



(a) Use of describing function in Nyquist plot

(b) Whitely or inverse Nyquist diagram

FIG. A18. Stability analysis with non-linear elements.

Consider a point on a trajectory of a phase-plane diagram passing through the origin. If this is a singular point and the system is stable, a function $V(x, z)$ can be found which is always finite except when x and $z = 0$ and whose differential coefficient is never positive. V is analogous to the energy stored in the system but is not numerically equal to it.

III. 358–63; VII. 124–32; IX. 297–305.

A12. Random inputs

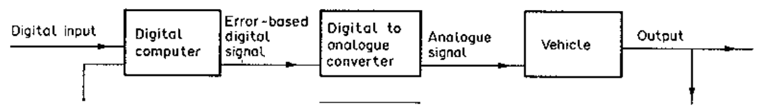
In many systems involved in transport, the input takes the form of a random signal. For example, the imperfections in a road surface as seen by the suspension of an automobile will lack regularity, but taken over a sufficiently long distance the particular construction will possess a definable characteristic. Statistical methods can be applied to such situations and treatments are available based on auto-correlation and power density spectra (see IX. 308–49).

A13. On-line control—sampled data

As control theory is applied to the automation of increasingly complex industrial systems, digital computers become necessary to manipulate the large amount of data involved. It is frequently necessary to employ discontinuous or pulsed data. For methods of treatment of such data, see IV. 379–405 and IX. 350–417.

As control theory is applied to the automation of increasingly complex industrial systems, digital computers become necessary to manipulate the large amount of data involved. It is frequently necessary to employ discontinuous or pulsed data. Moreover, the price of digital devices, notably microprocessors, has declined and their reliability has improved so dramatically in recent years that they have acquired dominant importance.

As regards systems embodying digital hardware, some form of analogue-to-digital conversion is necessary both on the forward path and the feedback loop. The basic circuit of a control system based on sampling the output data is shown in Fig. A19. The



input is a set of signals based on a sampling period T. The output is sampled by a digital-to-analogue computer and an error function in digital form is derived from the interaction of the input and output values fed to the computer. If the sampling period is short compared with the time constants of the system, then the general methods referred to in Appendix I (1 to 7) can be applied but if this is not the case the effect of sampling must be allowed for.

The term "zero order hold" is used to express the fact that whilst the primary signal which is being sampled may be varying continuously, the sampled output remains constant for the sampling period T. Thus a ramp input will produce a stepped output, as sketched in Fig. A20.

The Z Transform

The Z transform has been introduced to assist in the analysis of sampled systems involving sampling periods which are significantly long when compared with the time constants of the system. Its use can be explained as follows: The output of a sampler at time t is the sum of a number of impulses each of which can be written as $r(KT)$

$$= \Sigma\, r(KT)\delta\,(t - KT)$$

where δ is the impulse function.

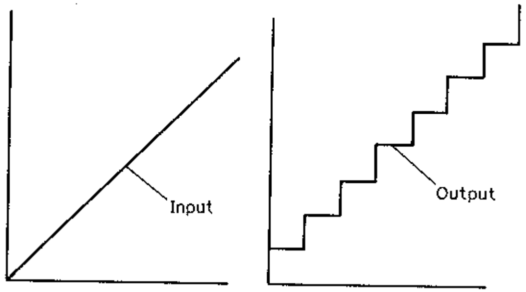


FIG. A20. Sample data, zero order hold.

When $t > 0$ the Laplace Transform becomes

...mapping from the s plane.

The Z transform is written

$$Z\{r(t)\} = \sum_{K=0}^{\infty} r(KT)\, Z^{-K}.$$

A table of Laplace and Z transforms appears as Table A1. As stated in A6, a linear feedback system is stable if all the poles of the closed-loop transfer function lie in the left half of the s-plane. From the definition $Z = e^{sT}$ we may write $|Z| = e^{\sigma T}$ or $\angle Z = \omega T$. Therefore, in a sampled system, stability is assured if all the poles of the closed-loop transfer function lie within the unit circle.

Recent treatments are as follows:

JACQUOT, R. G., *Modern Digital Control Systems*, Marcel Dekker, Inc., New York and Basel.
KATZ, PAUL, *Digital Control Using Micro-processors*, Prentice Hall, New Jersey, U.S.A.
KUO, B. C., *Digital Control System*, Holt, Reinhart and Winston, New York.
MARSHALL, J. E., *Control of Time-delay Systems*, Peter Peregrinus, London.

APPENDIX II

The S.I. Units (Système International d'Unités)

A SELF-CONSISTENT system of units is always convenient, particularly for dealing with conversion of energy between its electrical, mechanical and thermal manifestations. T international nature of the S.I. system together with its inclusion of familiar electri quantities provides justification for its use in preference to any new consistent syst based on the foot and the pound

whether within the earth's gravitational field or on the surface of the moon. Wh directly concerned with weight, i.e. the earth's gravitational pull on an object, the fo in newtons is simply the mass multiplied by the gravitational field appropriate to t particular location. Thus the force required to be exerted by a crane or the force exert by the wheel of a vehicle on the road can easily be calculated. When it is desired estimate the maximum acceleration or deceleration which can be attained with a coe cient of friction μ the force is μMg acting on mass M so that acceleration is simply It must be expected that for many years to come, engineers, particularly in the U.S. will be required to manipulate data provided in obsolescent systems of units and to qu results in these units. Nevertheless, the saving in labour resulting from use of consiste units is sufficient to justify their use even in these circumstances.

When starting a calculation, all necessary data should be assembled in whatever un it is provided. Ratios should then be extracted and recorded for further use and t remaining data converted into S.I. units. If required, the solution can be converted ba to the original system.

This last operation is not as time consuming as it may appear because most engineeri calculations proceed from a great deal of data to the determination of very few quantiti perhaps a single dimension.

Table AII provides conversion factors for most of the quantities encountered transport studies.

Certain units of measurement in this system are inconveniently small for engineeri purposes. Thus pressure in N/m^2 represents only about one hundred thousandth of atmosphere. In the c.g.s. system 1 bar is 10^6 dynes/cm^2 ∴ 1 millibar = 1 N/m^2.

Table AII. Use of S.I. System of Units

Quantity	S.I. unit and symbol	Conventional unit	Multiplying factor for converting quantity expressed in conventional units to S.I. units
Time	second, s		
Length	metre, m	inch	0·0254
		foot	0·3048
		mile	1609·343
		nautical mile	1852
Velocity	metres/second, m/s	mile per hour	0·4470
		foot per second	0·3048
		kilometre per hour	0·2779
		foot per minute	0·00508
		knot (international)	0·5144
Angular velocity	radian per second, rad/s	revolutions per second	2π
		revolutions per minute	0·1047
Acceleration	metre per second squared, m/s^2	standard gravitational acceleration, g = 32·17 ft/sec^2	9·807
		mile per hour per second	0·4470
		foot per second per second	0·3048

Quantity	S.I. unit and symbol	Conventional unit	Multiplying factor
		U.K. gallon	$4·546 \times 10^{-3}$
Mass	kilogramme, kg	ton	$1·016 \times 10^{3}$
		pound	0·4536
		slug	14·594
Density	kilogramme per cubic metre, kg/m^3	pound per cubic ft	16·02
		specific gravity (= gm/cc)	1×10^{3}
Volume flow rate	cubic metre/second, m^3/s	U.K. gallon/hour	$1·263 \times 10^{-6}$
Force	Newton, N (kg m/s^2)	pound force	4·448
		dyne	$1·00 \times 10^{-5}$
Pressure, stress	Newton per square metre, N/m^2 Pascal, Pa.	pounds per square inch	$6·894 \times 10^{3}$
		atmospheres	$1·01325 \times 10^{5}$
Viscosity (absolute or dynamic)	Poiseuille = Newton sec per metre square, Ns/m^2 = Pascal second, Pa.s	centipoise	1×10^{-3}
Viscosity (kinematic)	metre squared per second, m^2/s	centistokes	1×10^{-6}
Work, energy, heat	joule, J (Nm)	foot pound	1·356
		therm	$1·055 \times 10^{8}$
		British thermal unit	$1·055 \times 10^{3}$
		calorie	4·186
Power	watt, W (J/s)	horse power	745·7
Temperature	degree Kelvin, °K	degree Fahrenheit	5/9 (°F + 459·7)
		degree Celsius (Centigrade)	°C + 273·2
Heat flow rate	joule/sec *as power*	—	—
Thermal capacity per unit mass	joule per kilogramme degree Kelvin, J/(kg K)	calorie per gramme degree Celsius (specific heat)	$4·187 \times 10^{3}$

Table AII (*cont.*)

Quantity	S.I. unit and symbol	Conventional unit	Multiplying factor for converting quantity expressed in conventional units to S.I. units
Thermal conductivity	joule metre per metre squared second degree Kelvin, W/(m.K)	British Thermal Unit foot per foot squared hour degree Fahrenheit	1·731
Electric current	ampere, A		
Frequency	hertz, Hz		
Quantity of electricity	coulomb, C		
Electric tension or potential	volt, V		
Electric resistance	ohm, Ω		
Magnetic flux	weber, Wb		
Magnetic flux density	tesla, T Wb/m^2		

by 10^n where appropriate, but in conversation and perhaps in descriptive texts the following multiples are useful:

$\times 10^{6}$	mega	M
$\times 10^{3}$	kilo	k
$\times 10^{-3}$	milli	m
$\times 10^{-6}$	micro	μ
$\times 10^{-9}$	nano	n
$\times 10^{-12}$	pico	p

Name Index

Abele 97
Adams, C. 349
Adams, W. 55
Albert 231
Allen 231
Allinson 74
Allsop 51, 55
Alston 119
Anderson, E. 231
Anderson, J. H. 68, 74
Andrews 87, 145, 146, 188, 216, 231
Aster 91
Athans 74, 84, 86
Audinot 252

Barwell 30, 118, 188, 231, 232, 251, 349
Baseler 311
Baverez 53, 54, 56
Beadle 349
Bennett, 57, 74
Berndt 207
Berry 97
Bisbec 55
Bishop 231
Blake 293, 349
Bocock 349
Bode 365
Bollinger 355
Bolton 349
Boltzmann 51
Boon 349
Borgeaud 189
Bouladon 290, 294, 348
Boura 74
Bowler 189
Bradley 213, 231
Brentnall 44
Bretherton 74
Broadbent 205, 207
Brown, J. 95
Brown, R. G. 356
Brunel 32
Buchanan 285, 348
Buhler 207
Bunting 349
Burrell 74

Caiati 283
Carling 87
Carlson 142
Carter 215, 231
Cham 252
Chestnut 355
Chirgwin 349
Cholley 252
Chu 97
Clarke viii, 232
Cree 96, 97
Cribbens 74

Dalby 148, 188
Dalifol 293
Davies, P. D. A. 355
Davies, P. B. 188
Davies, T. 127
Davis 188, 189
Dell 252
Doppler 109
Dorf 355
Dover 145, 188
Dowty 129
Dumas 205

Eastham, P. 349
Eastham, T. 349
Eder 17
Edgeley 252
Edie 53, 54, 56
Ellis 188
English 349
Ericsson 110
Evans 361

Faraday 317
Fellows 232, 349
Ferbech 97, 283
Ferranti Ltd. 92
Finelli 283
Fitz Gerrel 97
Fitzpatrick 349
Fossard 355
Fourier 371
Freeman 251
Freiberg 56
Fricke 252
Furniss 74

Galawa 97
Garwood 55
Giles 44
Ginns 340
Goodstein 127
Gore 123, 127
Gosling 17
Goubou 92
Gough 231
Guignard 231

Hammet 188
Hanes 232
Harrison 355
Hauslen 56
Hawkes 60, 74
Hazlerigg 349
Heald 123, 124
Helly 55
Herd 188
Herman 21, 23, 30
Hertz 214
Heumann 220, 231
Hewton 65, 74
Hill 283
Hix 74, 103
Hobbs 232
Hodkinson 349
Hoel 283
Holmes 58, 74
Hosada 349
Howard 283
Humpheys 232
Hunt 74
Hurwitz 361
Hutchings 96, 97

Irving 231
Ishi 349

Jacquot 376
Jayawant 349
Jenkins 219, 231
Johnson 253

Kallberg 252
Kangas 283
Kashida 349
Katz, P. 376
Katz, R. 349
Kekonius 252
Keller 103
Kennedy 253
Kilb 119, 252
King 349
Kinoshita 349
Kirchhoff 78
Kitani 97
Koffman 231
Kolbuszewski 284, 348
Konsaka 349
Kroger 207
Kurachi 97
Kurodo 97
Kuo 376

Lagerhausen 252
Laithwaite x, 349
Lake 349
Lam 252
Langhurst 283
Laplace 356
Layton 353
Leech viii, 74, 142
Levi 349
Levine 74, 84, 86
Lewis 207
Liapunov 374
Liénard 12
Lighthill 53, 56
Littler 188
Loosli 189
Lucas 188
Lutz 283

Macdermot 44
Mach 146
Macmillan 188, 355
Maddison 207
Madigan 283
Magnien 252
Majou 252
Manser 252
Marshall 376
Martin 355
Masakaya 349
Matsui 349

Maxwell 151, 188
Medechi 97
Meinecke 231
Mew 251
Meyers 74
Michaels, D. 348
Michaels, R. M. 24, 30
Migamoto 97
Millard 74, 251
Murphy 355

Nakaham 349
Nakahara 97
Naslin 17, 355
Newcomb 355
Nouvion 189
Nyquist 83, 87, 364, 373

O'Flaherty 349
Oborne viii, 232
Ogilvy 91, 118, 252
Oliver 55

Parkinson, T. 349
Parkinson 251
Partek 97
Paul 80, 87
Peacock 44, 207
Pernot 283
Perrot 334
Pick 103
Place 87
Poisson 45
Porter, A. 355
Porter, S. R. M. 231
Potts 21, 23, 30
Poupé 112, 118
Powell 74
Powner 67, 74
Prigogine 51, 56
Prince 349

Qualtrough 74
Quinn 348

Rach 65, 74
Readman 102
Rees 189
Reynolds 146, 213
Richardson 355
Robertson 74
Rocard 213, 231, 355
Roscher 74, 103
Rose 87
Rosenkranks 283
Routh 361
Ryland 74

S.N.C.F. 77, 86, 305
Sadler 141
Sage 355
Saise 74
Sales 189
Saumebber 207
Savage 74
Savant 355
Scales 44, 207
Schmitz ii, 119, 252
Segal 213, 231
Seifert 349
Sell 349
Shearer 355
Shelley 58, 74
Shinners 355
Short 74, 127
Sinha 349
Smith 252
Soberman 283
Sommerfeld 92
Steeds 145
Stenow 142
Stevens 348

TSO 252
Takaoha 349
Takemura 97
Taketsusi 349
Tanner 55
Taylor 355
Thaler 356
Touch 349
Trabold 348
Trapeznikov 233, 251
Tresidder 74
Trollope 213, 231
Twine 349
Tyler 89, 119
Tyson 283

Umemori 349

Wagner 74
Waldmann 283
Walker 231
Wallace 213, 231
Warder vi
Wardrop 30, 51, 55
Ware 252

Watt 151, 200
Webb 356
Webster 55, 65, 67, 74
Weiss 55
Welbourne 154, 188, 356
West 17, 355
Westinghouse 200
Wheler 349
Whiteley 87, 373
Whorlow 349
Wickens 222
Willison 74
Wilson 195, 207
Winton 74
Wise 207
Woolacott 189

Yasatrsugu 349
Yoshida 97
Young 45

Subject Index

ALVEG System 261
A.P.T. vii, 301–9
A.W.R.S. 101–3
A.W.S. 106, 118
Acceleration 4, 143–51, 185, 290
Action of signalman 34, 61, 70, 236
Active control systems 350–5
Adaptive control 182, 184, 238
Adhesion 169–84
  control 158, 169
  effect on capacity 19
  limit 11, 143
  O.R.E. tests 172–7
  remedical measures 177–82
  test methods 172
Aerodynamic drag coefficient 145–7, 234
Aerotrain 316
Air bearing 315
Air brake 200–4
Air cushion support 314–36
Airtrans system 260, 264–8
Amplifiers, operational 78–80
Analogue computing 75–85
Analogy, electrical 78, 350–5
"And" function 123
Anti-slip brake 182
Application of digital computers 57–74
"Aster" jointless track circuit 90–1
Attenuation characteristics of railway track 113
Audio-frequency spots 243, 244
Automatic railways 233–51
Automatic retarders 129
Automatic train control 104–7
Automatic transmissions 154–7
Automatic vehicle systems 26–30
Automatic wagon recording system 101–3
Automatic warning system 104–7
Automatic wheel slip protection 169–82, 195–8
Autonomics 184
Axle counting 36

BART 247
BESI 98–101
Back e.m.f. 158
Backlash 370
Balancing speed 14, 144, 150
"Ball" 32
Ballast conductance 113
Bang-bang control 9
Beacon devices 107–11
Bell code 34–5
Binary decimal system 101–2
Binary numbers, transmission of, 115
Blake system 293–4
Block diagram 2
Block sheet 36
Block system 33–44, 47–50, 250
"Board" 32
Bode plot 365–9
Bogie oscillation 220–2
Bond, impedance 90
Boolean algebra 122–4
Booster retarder 130–41
Borough Market Junction 60–1
Bouladon system, 290–1, 294
Brake
  air 200
  E.P. 203
  vacuum 204
Braking distance 19, 26, 34, 48, 191
Braking, control of 190–206
British Rail 31, 37, 43, 57–61, 92, 106, 114–8, 162, 178–80, 197, 222, 235, 301–5, 305–8, 323, 341–3
Bus
  capacity 294
  electronic scanning indicator 98–101

CORECT system 281
Cab signalling 104–18
California code 23
Cant deficiency 223
Capacitance 78, 350
Capacity of route 18–30, 41–2, 55, 284, 289–92
Car following theory 19–26, 365–8
Carol mine 239, 240, 241
Cascaded vehicles 67–8, 83–5
Characteristic speed 21, 53
Chassis dynamometer 75
Chinese railways 34
Circle diagram 330–2
"Citizens Band" radio 287
Claw motor 337
Closed loop system 6–9, 353–5
Coasting 150–1, 237, 243
Coded track circuit 111, 237, 240, 242

Coefficient
of adhesion 19, 169–82, 190, 195–8, 213, 216, 222
of friction 18, 169, 171, 198, 213
Collective flow 52
Collisons 29, 54
Combination of magnetic and inductive loop systems 117
Comfort of passengers 184–8, 195, 225–31
Commutating thyristor 165
Complement 122
Complementary function 4, 7, 356
Complex roots 357–8
Compliance 78
Compliant air bearing 315
Compressed air brake 200–3, 244
Computer control 55, 238, 251, 281–2, 293
Congestion effect on multiple aspect signalling 47–50
Coning of railway wheels 208, 217
Constraints 3, 239
Consumption of fuel 155
Continuous conductors 113–18, 237, 247
Continuous platforms 289–93
Continuous progress control 57–60
Continuous train 292, 294
Control
of braking 86, 190–206, 244, 331
of cascaded vehicles 67–8, 83–5
of diesel power 182–4
of electric motive power 157–69, 293
of engine speed 151–4
of intersections 125
of linear motors 331–3
of slip (slow-up) 169–82
function 358
proportional 6, 354
table 121
Coulomb friction 370
Crash test of automobiles 328–9
Creep 214–5, 222
Current collection 305–8
Curved track 218–20, 223

D.A. valve 205
DAIR concept 287–9
Daf control system 155, 157
Dalby diagram 148–9
Dalifol's continuous transport system 293
Damping 82–3, 222, 350–2
Data processing 57–8
Dead zone 369–70
Degrees of freedom 2
Delay, finite (dead time, transport lag or distance velocity lag) 360
Density of traffic 18–26, 50–5, 284, 294
Derivative and integral controls 154, 354
Describing function 370–3
Detectors, presence 88–96
Diesel-electric transmission 182–4
Differential analysers 76
Differential equation of motion 3–6, 144–9
Digital computers 57–74, 238, 251
Digital slip control 181
Dingwall 39
Direct admission valve 205
Directional stability 208–32
Disc and crossbar signal 33, 286
Disc brake 199
Distributer valve 202
Distance-velocity lag 360
Distant signal 34, 104, 120
Doppler radar 109
Double rail track circuit 89
Dowty retarder 129–41
Drag 11, 145–7
Driving function 2
Droop 355
Drum test 339
Dusseldorf tunnel signals 55
Dynamic programming 57–60
Dynamic systems 1, 80–3, 350–5
Dynamometer cars 75–8

E.P. brake 203–4
Economical speed-torque curve for diesel engine 183
Effect
of oscillation on passengers 225
of station stops on route capacity 290–3
Electric motive power 157–69
Electrical analogy of mechanical system 80–3, 350–5
Electrical quantities 78, 379
Electro-lane 26–8, 293
Electromagnetic induction 103, 117–18
Electromagnetic support and guidance 337–48
Electronic token 39
Equations of motion 4–6, 11–15, 144–9
Equivalence of dynamic and active systems 350–5
Error 6, 9–10, 78, 83–5, 151–4, 353–5
Exclusion 122
Exponential lag 5
Exponential service times 47
Expressway, Transit 268–70

Fail-safe capability (Failure to safety) 44, 89 107, 113, 118, 242, 243, 250
Feedback 3, 6–9, 28–30, 154, 333, 353–5
Finland 36
Fixed progression 63
Flange forces 218–20
Flow chart 69, 70
Flow variable 78
Forward path transfer operator 7
Fourier analysis 370–3
Friction coefficient 18, 172, 198

Fuel consumption 155
Function
control 358
describing 370–3
driving 2
forcing 2
non-linear 368–74

"G" line 92–5
Gain margin 365–9
Gain 361, 362
Gauge conversion 300
General Motors Technical Centre 22, 26
General solution 4, 351
Glasgow cable subway 34, 246
Glasgow traffic control 63
Goubau line 92–5
Governor, engine 151–4
Gradients 14, 110, 116, 144
Graphical time table 62, 234
Gravity yard 130
Great Western Railway 104, 205
Guidance of vehicles 208–32, 293
Guided radar 91–5

H.S.B. 309–14
H.S.G.T. 315
Hard limit cycle 10
Harmonic motion 16, 350–3
Hawk locomotive 163–7
Headway 18–23, 40–2, 47–50
Hertzian stress 214–15
High capacity systems 290–3
High-speed ground transport 295–333
Highway code 18–19, 23–4, 26
Hill-climbing technique 185
Holland tunnel 54–5
Homopolar linear motor 317–18
Hong Kong 246
Hovercraft 314–27
Human contribution 234–7
Hump yard 128–41
Hybrid computers 85
Hydrodynamic pressure 213
Hyposynchronous braking 319
Hysteresis 369–70

Identra vehicle identification system 98, 245
Impedance bonds 90
Imprints of railway wheels 216
Improvements in control on highways 286–9
Impulse function 358
Inclusion 122
Individual flow 51–3
Induction motors 163–9, 293
Inductive loops 26–8, 114–18
Indusi train control system 107
Infra-red 103
Inscription within sharp curves 218–20
Instability 10, 22, 211–14, 220–2, 360–70
Institution of Railway Signal Engineers 44
Integrating transporter 290–1
Inter-modal transfer ix, 285–6
Interaction 51–3
Interlocking 43, 120–4
Intermediate systems 256
Intersection 122
Inverness 38
Inverse Nyquist diagram 374
Isocline 12, 17, 30
Isotropic radiation 103

JZG automatic train control 110
Jam concentration 53–4
Japanese Airlines 344
Japanese National Railways 295, 347
Jointless track circuit 90–1
Junction
Borough Market 60–1
control table 121–2
layout of signals 121
optimisation technique (JOT) 61

Kinetic control system 234
Kinetic friction 370
Kinetic theory of gases 51
Kingston, Ontario 278–81
Kirchhoff's law 78
Kyle of Lochalsh 39

LIMRV 323
Lag, exponential 5
Lane capacity 18–24, 41–2, 55, 284, 289–92
Laplace transform 356–9
Laws of vehicle following 18–26
Leningrad 254
Level crossings 125
Liapunov's second method 373
Lienard's construction 12–14
Light electric vehicles 256
Limit cycle 10, 222
Line conductors 114–18
Line rating 227
Linear induction motor 315–36
in marshalling yard 136
Linear synchronous motor 336
Liverpool traffic control 63
Lock and block 40
Logarithmic plotting 365–6
Logic elements 123–4
London Transport 98–101, 192, 239, 240, 242–5
Low level traction bars 170
Lyon metro 262

MIRA 147, 328–9
Mach number 146

Magnetic braking 193–4
Magnetic slip detection 181
Magnetic support and guidance 337–48
Manifold depression 156
Manual block system 33–40
Marginal stability 16
Marshalling 128–41
Matrix 60, 84–5, 222
Maxaret system 195
Mechanical interlocking 43, 120–3
Metcalfe-Oberlikon air/vacuum brake 206
Mexico City Metro 259
Microprocessors 69
Microwave reflection 103
Mini car 155
Mode ix, 285–6, 294
Monitor braking system 196–7
Monorail 261
Morgantown 271–5
Moscow 254
Motion on curved track 217–19, 223
Motor
  induction 163–9, 293
  linear 315–37
Motorway capacity 284, 294
Motorway signalling 286–9
Motorway space requirement 285
Motorway traffic 284
Moving block system 47–50, 250
Multiple collisiins 54
Multiple-aspect signalling 40–4, 47–50

N-X system 43, 236
Narita 344
Natural frequency 221, 351
New York Subway 239, 240, 241–2
New transport modes 284–349
Node 16, 78
Non-linearities 11, 368–74
Non-metallic brake blocks 198
Notching 158
"Not" function 123
Nyquist representation of stability 364–8

Oerlikon control valve 202
Offitread braking 198–9
On-line control 375
On/off control 9, 10
Open loop system 3
Operational amplifiers 78–80
Optimum train sequence 60–2
Optimum velocity 19–24, 52
Oscillation
  effect on passengers 225–31
  of bogies 220–2
Overlap 34, 48

PHAR system 109–10
PTACV 327
Pantograph 306
Paratransit 253
Paris Metro 260
Particular integral 4, 7, 351, 356
Passenger reaction 184–8, 194–5, 225–31
Passive system 350–2
Pendulum 16–17
People mover drive 334
Performance Index (PI) 67
Permissive working 31–2
Personal rapid transit (PRT) 271
Phase-plane diagram 9–17, 29, 352–3, 308
Phasor 357
Physiological aspects 194, 225–31
Pick-up 195–7, 199
Pitching 80
Plasma arc 178
Platoons 51
Plenum chamber 315
Plugging 319, 331, 333
Pneumatic trail 209
Poisson distribution 45–7
Poles and zeros 358
Post Office Railway 238–41
Potential and flow 78
Power control 143–89
Power law acceleration 186
Power-assisted steering 210–11
Presence detectors 88–95
Probability 45–7
Programming, dynamic 57–61
Proportional control 6–8, 353–4
Pueblo test track 309
Pulsatance, complex 357
Pulse width Modulation 168

Quebec, North Shore and Labrador Railroad
  239, 240, 241
Queues 45–7

R.A.P.T. 239, 240, 247
R.T.V.31 327
ROCOC 195, 225
Radar, guided 91–5
Radial bogie 223–4, 278
Radio link 36, 38
Railway signalling 31–4, 70, 88–91, 104–7, 120–2
Ramp function 358
Random events 45–7, 50–3, 374–5
Real time optimization programme (RTOP) 65
Reciprocal spacing law 21, 368
Regulating signalman 236
Relaxation function 52
Relay describing function 370–3
Relay, railway signal 89–90
Resistance
  aerodynamic 4, 145–7
  rolling 4, 145, 147

Retardation 192–4
Retarder tube 192
Reynolds number 146
Rheinallee Tunnel 55
Ride index 225
Road Research Laboratory 24, 195
Robertson's Law of Growth 188
Rolling and pitching 80
Root locus 345, 361–3
Rope hauled classification yards 133
Route capacity 18–30, 41–2, 55, 284, 289–92
Routh-Hurwitz criterion 360–1
Rule book 235
"Rule 55" 88–9
Running of coned wheels 217

SELTRAC system 281
SI units 54, 377–9
SNCF 77, 86, 181, 305
"S" curve 31–2
"s" plane 358, 360–3
Saddle point 16–17
Sampled data 375
Sandite 179
Sao Paulo 249, 255
Sapporo system 264
Saturation 369
Secondary resistance 331–2
Sequential operational release (SOR) 248
Series motor characteristics 157–8
Service environment simulation 86
Servo assistance 199–200
Servo-pantograph 305–8
Shinkansen railways 295
Shiohama yard 136
Shock factors 184–8, 195
Sighting distance 24–6, 31–2, 34, 41–2, 48
Signalling 31–44, 70, 88–91, 104–7, 120–2, 286–9
Signum system 108
Silicon-control rectifiers (thyristors) 161–5
Simulation 71, 138
Signle line working 37
Single-rail track circuit 89–90
Sinusoidal acceleration 187
Slip angle 209
Slip detection 180–2
Slow up 195–8
Slug (in signal relay) 90
Soft limit cycle 10
Sommerfeld wave 92–5
Sorting 128–41
Space requirements for road and railway 285
Spacing of vehicles 18–26, 50–3
Specific fuel consumption 155
Spiral-type hydraulic retainer 132
Split cycle and offset operation technique (SCOOT)
  66
Square loop characteristics 159
Stability 10, 21, 211–14, 220–2, 360–8
Stable focus 16
Stable node 16
Static deflection 351
Station stops, effect on capacity 290–3
Statistical aspects 45–7, 50–3, 172
Steerable wheels 272
Steering 208–32
Step function 358
Stimulus-response law 20, 365–8
Stockholm Sparvager 245
Strings of vehicles 67–9, 83–5
Summing element 2, 6, 82
Surface waves 91–5
Surveillance of traffic 62–4
Suspension 80–3, 220–2, 351
Switching-active and passive 334
Switching-off circuit 165
Synchronous linear motor 336
Synchronous speed 330, 332
System flow chart 70

TGV 305
TLRV 324
TOPS 60
Tachometer 244
Talgo train 300
Tappet 120
Taylor's series 360
Television 243
Thinking distance 26
Three-term controller 154, 355
Threshold values 185, 370
Thyratrons 78
Thyristors 161–7
Time constant 5
Time delay 360, 365–8
Token instruments 38
Tokyo monorail 261
Toronto traffic control 63, 88
Torque reaction 169–72
Torque-speed curves 156, 166, 330
Track circuit 88–91, 237, 240, 242
Traction bars 170
Tractive effort 11, 143–5, 331
Tractive reistance 4, 145–8
Traffic intensity 46
Traffic responsive control 63, 65
Traffic surveillance 62–6
Traffic waves 53–5
Train Consist Record 58–9
Train describer 69
Train order system 33
Train Regulation Advisory Control 61
Train starter tablet 37
Train stop 241
Trajectory 12, 30
Transductor control 159–61, 180
Transfer function 356
Transfer operator 7–8
Transistor control 168
Transit Expressway 268–70

Transmission of information in automatic railways 240
Transrapid 04 to 06 340
Trip valve 243
Triple Modular Redundancy (TMR) 71
Triple valve 201
Turbo-train 297
Two-pedal control 155
Two-stage disc brake 197
Tyne and Wear System 256

US Department of Transportation 253, 271
Ultrasonic reflection 103
Ultrasonic train detection 96
Understeer 209
Universal class 122
Unstable focus 16
Unstable node 16
Urban Mass Transportation 253, 264
Urban Transportation Development Corporation of Toronto (UTDC) 277–83

VAL system 276
Vacuum brake 204–6
Variable frequency control 163–9, 332
Variable progression traffic control 63
Vehicle detection 88–97
Vehicle identification 98–103
Vehicle ride quality 220
Vehicle suspension 80–3, 220–2, 351
Viscous friction 78
Visual estimation of relative velocity 24–6
Volume density traffic control 63

Waiting time 46
Washington Metro 249, 254
Wave guide 108
West London Traffic Control 63
Westinghouse brake 200–3
Wheel slip control 169–82, 195–8
Whiteley diagram 373
Wick 39
Wind resistance 145–8
Working time table 235

Yaw 146

Z transforms 359
Zeros 358