Founded 1925
Incorporated
by Royal Charter 1961
To promote the advancement of radio, electronics and kindred subjects by the exchange of information in these branches of engineering

Volume 51 Number 7/8 July/August 1981

The Radio and Electronic Engineer

The Journal of the Institution of Electronic and Radio Engineers

Optical Fibre Communications

OPTICAL fibre technology has come a long way since the mid-1960s when work on it first started in earnest. In those early days optical fibres having losses of 1000 dB/km or more were the norm. The announcement in 1970 of the achievement of fibre having losses of 20 dB/km was a major breakthrough since, with that level of attenuation, optical fibre systems could be designed which would have a similar performance to metallic cable systems. This announcement resulted in an enormous upsurge of interest. Research laboratories throughout the world took up the challenge to demonstrate that optical fibre systems were completely practicable and interest in areas such as the fibre itself and semiconductor opto-electronic devices, as well as many others, became intense.

This activity culminated during the second half of the 1970s with the installation of numerous field trials, experimental and demonstration systems in many countries. The practical experience gained from these systems was so encouraging that optical fibre systems are now being employed as a standard communications medium and a large range of systems are now becoming commercially available.

One might wonder why there should have been so much interest in this particular new transmission medium. The answer, not surprisingly, is that there are many advantages to be had by its use. For instance, one of the major attractions is that extremely large repeater spacings are attainable: the 20 dB/km fibre losses of 1970 have now been reduced to 2 dB/km or less and some laboratories have reported fibre losses of as low as 0·2 dB/km at certain wavelengths. With this level of fibre loss, repeater spacings of tens of kilometres can be achieved, as compared with just a few kilometres for systems using metallic conductors, and having the same traffic carrying capability.

Other major advantages which are particularly worthy of note are: potentially low material cost of the optical fibre, extremely large potential bandwidth capability, negligible crosstalk, immunity to electromagnetic interference, complete electrical isolation, and small cable size. With all these advantages it is perhaps not surprising that there should be so much interest in optical fibre technology.

The technology is currently in a short period of consolidation prior to a further quantum leap forward with the practical implementation of monomode optical fibre technology. It is thus a very appropriate time for *The Radio and Electronic Engineer* to publish this special issue in which the papers included have all been invited and the authors are, without exception, experts in their fields. Thus within these covers we have one of the most up-to-date collections of papers describing the technology as it currently exists. Moreover many of the papers give some details of state-of-the-art research together with some useful indications of the trends which are to be expected over the next decade.

The papers cover the whole spectrum of optical fibre technology and provide the reader with a comprehensive coverage of the subject. Each of the papers in this issue has been written so that there is something for both the specialist and the non-specialist reader. It is the hope that this approach will find widespread appeal amongst the readership.

The first paper 'Optical fibre transmission lines' by W. A. Gambling *et al.*, discusses optical

fibres as a transmission medium and explains the propagation theory and common fibre fabrication techniques. The way in which fibres are incorporated into cables and the problems involved are discussed in the second paper entitled 'Optical fibre cables' by M. H. Reeve. Fibres usually have to be terminated in suitable connectors, and the third paper by P. Mossman discusses the various connector techniques which are being employed and others which are still being developed. To form a system it is necessary to have opto-electronic sources and detectors at each end of the fibre. These items are covered in papers by P. A. Kirkby on semiconductor lasers and by A. C. Carter on light emitting diodes, and a paper by I. Garrett covers photodiodes and receivers. All the topics covered in these papers are brought together in a paper by P. E. Radley, entitled 'System applications of optical fibre transmission', which describes the theoretical and practical realization of optical fibre systems.

The subject of 'Characterization of single-mode fibres' is then discussed in a paper by K. I. White *et al.* Optical safety as applied to optical fibre systems is a topic about which very little has been published and the paper by J. D. Topping and the late J. C. North discusses some of the problems in formulating safety criteria. Finally a paper by P. A. Laybourn and J. Lamb provides an insight into the theory and practical realization of integrated optics, a subject about which we are hearing more and more with every year that passes.

Owing to space limitations it has not been possible for the papers in this special issue to cover all aspects of the technology in detail. Nevertheless the references at the end of each paper will enable the interested reader to pursue particular aspects further. Indeed it is an indicator of the tremendous activity in the field of optical fibre technology that the ten papers published in this issue should together make reference to over 200 other papers, the majority of which have been published in the last few years. These references are of course just a small fraction of the total number of papers published each year on this technology. One might conclude that, as a technology which is likely to be employed to deal with the information explosion, it is itself contributing to it!

C. J. LILLY

Christopher Lilly, guest editor for this special issue, is a Manager in the Transmission Development Division of British Telecom and has been responsible for the development of various telecommunications transmission systems for over 14 years. For well over half this time he has been involved with the development and implementation of optical fibre technology and was instrumental in the installation of many of the first field trial optical fibre systems in the UK. He was subsequently technical manager responsible for the implementation of the first optical fibre systems installed in the British Telecom network for operational use. Now that these first-generation systems are well underway, he is concerned with the development of standard optical fibre systems of the future.

Mr Lilly sits on various national and international committees concerned with optical fibres including committees of the BSI and IEC. He is the UK representative on optical fibre technology at the CCITT and is chairman of the IEC/CCITT joint working party on optical fibre terminology. He has published and presented a variety of papers on the subject of optical fibres. Elected a Member of the Institution in 1973, Chris Lilly has been a member of the Council of the IERE since 1978 and a member of the Papers Committee since 1979.

Institution Premiums for 1980

The Council of the Institution announces that authors of the following papers are to receive Premiums for outstanding contributions published in the Journal during 1980.

MAIN PREMIUMS

LORD MOUNTBATTEN PREMIUM *Value* £100

For the outstanding paper on the engineering aspects of electronics or radio

'NICAM 3: near-instantaneously companded digital transmission system for high-quality sound programmes'
C. R. Caine, A. R. English and J. W. H. O'Clarey (*British Broadcasting Corporation*)
(Published in the October 1980 issue of the Journal)

CLERK MAXWELL PREMIUM *Value* £100

Science of electronics or radio

'Data compression techniques and applications'
Dr G. Benelli, Prof. V. Cappellini (*Institute of Electronics, Florence*) and Dr F. Lotti (*Institute for Research on Electromagnetic Waves, Florence*)
(January/February)

MARCONI PREMIUM *Value* £75

Engineering of an electronic system, circuit or device

'An integrated circuit v.h.f. radio receiver'
I. A. W. Vance (*Standard Telecommunication Laboratories*)
(April)

SPECIALIZED TECHNICAL PREMIUMS

LORD BRABAZON PREMIUM *Value* £50

Aerospace, maritime or military systems

'British telecommunications trans-horizon radio links serving off-shore oil/gas production platforms'
S. J. Hill (*British Telecom*)
(August)

A. F. BULGIN PREMIUM *Value* £50

Electronic components or circuits

'Development of c.c.d. area image sensors for 625 line television pictures'
D. J. Burt (*GEC Hirst Research Centre*)
(May)

DR NORMAN PARTRIDGE PREMIUM *Value* £50

Audio frequency engineering

'Transmission of speech by adaptive sampling'
Dr J. Dunlop and Dr N. C. Changkakati
(*University of Strathclyde*)
(March)

PAUL ADORIAN PREMIUM *Value* £50

Communications or broadcasting engineering

'The measurement of Teletext performance over the United Kingdom television network'
L. A. Sherry and R. C. Hills (*Independent Broadcasting Authority*)
(October)

J. LANGHAM THOMPSON PREMIUM *Value* £50

Theory or practice of systems or control engineering

'Systems engineering: an approach to whole system design'
Prof. P. K. M'Pherson (*The City University*)
(November/December)

P. PERRING THOMS PREMIUM *Value* £50

Radio or television receiver theory or practice

'Wide range frequency synthesizers with improved dynamic performance'
Dr M. J. Underhill (*Philips Research Laboratories*)
(June)

SIR CHARLES WHEATSTONE PREMIUM *Value* £50

Electronic instrumentation or measurement

'A review of distortion and its measurement in p.c.m. telephony systems'
R. G. Rolls (*Marconi Instruments*)
(July)

CHARLES BABBAGE PREMIUM *Value* £50

Design or electronic engineering application of computers

'The choice of a recording code' (April)
and
'A superposition-based analysis of pulse-slimming techniques for digital recording' (June)
Dr N. D. Mackintosh (*Formerly with Racal Recorders; now with Burroughs Corporation*)

GENERAL PREMIUMS

LESLIE MCMICHAEL PREMIUM *Value* £50

Management techniques associated with electronic engineering

'Management in a competitive environment–the experience of Tandbergs Radiofabrikk A/S'
R. McLellan (*Anglian Regional Management Centre*)
(January/February)

Papers of sufficiently high standard were not published within the terms of the following Premiums and they are withheld:

Heinrich Hertz Premiums (*Value* £75)—Physical or mathematical aspects of electronics or radio

Lord Rutherford Premium (*Value* £50)—Electronics associated with nuclear physics or nuclear engineering

Arthur Gay Premium (*Value* £50)—Production techniques in the electronics industry

Sir Henry Jackson Premium (*Value* £50)—History of radio or electronics

Eric Zepler Premium (*Value* £50)—Education of electronic and radio engineers

Dr V. K. Zworykin Premium (*Value* £50)—Medical or biological electronics

Hugh Brennan Premium (*Value* £50)—Outstanding paper first read before any Local Section of the Institution and subsequently published in the Journal

Sir J. C. Bose Premium (*Value* £50)—Outstanding paper by an Indian scientist or engineer or any subject

INSTITUTION OF ELECTRONIC AND RADIO ENGINEERS

Notice of Annual General Meeting

NOTICE IS HEREBY GIVEN that the twentieth ANNUAL GENERAL MEETING of the Institution since Incorporation by Royal Charter will be held on THURSDAY, 29th OCTOBER 1981, at 6.00 p.m. in the Goldsmith's Hall, London School of Hygiene and Tropical Medicine, Keppel Street, Gower Street, London W.C.1.

AGENDA

1 **To receive the Minutes of the nineteenth Annual General Meeting of the Institution since Incorporation by Royal Charter, held on 23rd October 1980.**
Reported on pages 14–17 of the January 1981 issue of *The Radio and Electronic Engineer*.

2 **To receive the Annual Report of the Council for the year ended 31st March 1981.**
To be published in the October 1981 issue of *The Radio and Electronic Engineer*.

3 **To receive the Auditor's Report, Accounts and Balance Sheet for the year ended 31st March 1981.**
To be published in the October 1981 issue of *The Radio and Electronic Engineer*.

4 **To confirm election of the Council for 1981–82.**
In accordance with Bye-Law 49 the Council's nominations were sent to Corporate members by a Notice dated 5th June 1981 in the June 1981 issue of *The Radio and Electronic Engineer*. As no other nominations have been received under Bye-Law 50 for the following offices, a ballot will not be necessary and the following members will be elected.

The President

H. E. Drew, C.B.

The Vice Presidents

Under Bye-Law 46, all Vice-Presidents retire each year but may be re-elected provided they do not thereby serve for more than three years in succession.

For Re-election: Professor J. R. James, B.Sc., Ph.D., D.Sc.,; P. K. Patwardhan, M.Sc., Ph.D.

For election: Colonel W. Barker; L. A. Bonvini; Major-General H. E. Roper, C. B., B.Sc.(Eng.); D. L. A. Smith, B.Sc. (Eng.); Group Captain J. M. Walker, R.A.F.

The Honorary Treasurer

For Re-election: S. R. Wilkins

Ordinary Members of Council

Under Bye-Law 48, Ordinary Members of Council are elected for three years and may not hold that office for more than three years in succession.

FELLOWS

The following must retire: Sir Robert Clayton, C.B.E., M.A., E.Eng.; C. S. den Brinker, M.Sc.; Major General H. E. Roper, C.B., B.Sc.(Eng.); D. L. A. Smith, B.Sc.(Eng.);
For election: L. W. Barclay, B.Sc.; G. A. McKenzie; V. Maller, M.A.; Professor K. G. Nichols, B.Sc., M.Sc.

MEMBERS

The following must retire: C. J. Lilly
For election: P. Atkinson, B.Sc.(Eng.)

HONORARY FELLOW OR COMPANION

The following must retire: H. J. Kroch, O.B.E.
For election: R. B. Michaelson

The remaining members of Council will continue to serve in accordance with periods of office laid down in Bye-Law 48.

5 **To consider and, if thought fit, approve the Resolutions 1 to 5 as set out in the Schedule following this Notice.**

6 **To appoint Auditors and to determine their remuneration.**
Council recommends the re-appointment of Gladstone, Jenkins & Co., 50 Bloomsbury Street, London W.C.1.

7 To **appoint Solicitors.** Council recommends the re-appointment of Bax, Gibb and Gellatlys (incorporating Braund and Hill), 14 Gray's Inn Square, London W.C.1.

8 **Awards to Premium Winners.**

9 **Any other business.**

Notice of any other business must have reached the Secretary not less than forty-two days prior to the meeting.

6th July 1981

By Order of the Council,
S. M. DAVIDSON, *Secretary*

Corporate Members may vote at this meeting by Proxy and the President and Secretary of the Institution have agreed to act in this capacity if so desired. Members wishing to vote in this way should apply, either in writing or telephone, to the Secretary for a proxy form in accordance with Bye-law 80.

After the conclusion of the formal business of the Annual General Meeting, the Annual General Meeting of Subscribers to the IERE Benevolent Fund will take place, following which Mr Harry Drew will give his Presidential Address.

Schedule of Proposed Changes in the Royal Charter and Bye-Laws of the Institution

As amended by Resolutions dated 27th November 1963, 18th October 1967, 15th August 1973 and 15th June 1978

RESOLUTION No. 1

ITEM	EXISTING ARTICLE OF ROYAL CHARTER *Portion marked [ ] is to be deleted*	PROPOSED REVISED ARTICLE OF ROYAL CHARTER *Portion in* **bold type** *is to be inserted*
A	13. A Fellow (including an Honorary Fellow who, when elected an Honorary Fellow, was already a Fellow) shall be entitled to the use after his name of the designation [F.I.E.R.E. and a Member the designation M.I.E.R.E.]	13. A Fellow (including an Honorary Fellow who, when elected an Honorary Fellow, was already a Fellow) shall be entitled to the use after his name of the designation **F.I.E.R.E., a Member the designation M.I.E.R.E. and an Associate Member the designation A.M.I.E.R.E.**

RESOLUTION No. 2

ITEM	EXISTING BYE-LAW	PROPOSED REVISED BYE-LAW *Portion in* **bold type** *to be inserted*
B	9. (first paragraph) Each Fellow and each Honorary Fellow who when elected an Honorary Fellow was already a Fellow is entitled to use after his name the designation F.I.E.R.E. and to describe himself as a Chartered Electronic and Radio Engineer. Each Member is entitled to use after his name the designation M.I.E.R.E. and to describe himself as a Chartered Electronic and Radio Engineer.	After second sentence insert a third sentence to read: **Each Associate Member is entitled to use after his name the designation A.M.I.E.R.E.**

Explanatory Note

To identify by the use of designatory letters the specialization of qualified and experienced Technician Engineers in membership of the IERE in like manner to that now employed by similarly qualified engineers of other disciplines in membership of other Chartered and non-Chartered bodies.

RESOLUTION No. 3

ITEM	EXISTING BYE-LAW	PROPOSED REVISED BYE-LAW
C	9. (second paragraph)	To be numbered Bye-Law **10**
D	10. As a result of the Institution's membership of the Technician Engineers Section of the Engineers' Registration Board ..... and use after his name the designatory letters 'T.Eng. (CEI)'.	To be deleted

Explanatory Note

This entitlement is now adequately provided in the Bye-laws and Regulations of the Council of Engineering Institutions, from whose Royal Charter it derives its authority.

RESOLUTION No. 4

ITEM	EXISTING BYE-LAW *Portion marked [ ] to be deleted*	PROPOSED REVISED BYE-LAW *Portion in* **bold type** *to be inserted*
E	32. Any member of any class whose annual subscription remains unpaid after [30th September] shall not, until such subscription is paid in full, be entitled to attend or take part in the meetings of the Institution, or to receive the notices and publications of the Institution, or to exercise any of the rights or privileges of membership of the Institution or to vote at any meeting.	32. Any member of any class whose annual subscription remains unpaid after **30th June** shall not, until such subscription is paid in full, be entitled to attend or take part in the meetings of the Institution, or to receive the notices and publications of the Institution, or to exercise any of the rights or privileges of membership of the Institution or to vote at any meeting.

RESOLUTION No. 5

That Resolutions Numbers 1, 2, 3 and 4 be conveyed for allowance by the Lords of Her Majesty's Privy Council, with a request that, if such allowance be granted, and subject to such alterations as the Lords aforesaid may require and the Council of the Institution may agree, the amendments to which the said Resolutions refer may come into effect as soon as the allowances are granted.

Conference Report

International Conference on Antennas and Propagation

University of York, 13th to 16th April 1981

After the first International Conference on Antennas and Propagation—in 1978,—some delegates considered that little more remained to be said on the subjects, but this second Conference was remarkable for the considerable addition to our knowledge that has taken place in two years. An undoubted stimulus was clearly shown by the large number of papers dealing with space communications, and many of the contributions to new knowledge have been derived from experimental work on the many space satellites.

Over 200 papers were given, divided into 37 Sessions. A Poster Session was also used to display about 18 papers. The Sessions provided convenient groupings for papers dealing with a common topic and, in most cases, the Sessions were preceeded by an invited review paper. In order to present the large number of papers in a reasonable time, triple parallel Sessions were operated, using three of the University's lecture halls, with subjects chosen to be as diverse as possible.

Certain expected subjects generated many papers, such as Adaptive Antennas and Antenna Measurements for example, which each required two Sessions. Propagation is now the subject of intensive study so far as Earth/satellite paths are concerned, and this interest was reflected in four Sessions.

In this review it is not possible to deal with each paper or, indeed, each Session in detail. Suffice it to state that all papers were of high quality. The fact that triple Sessions were in operation meant that, inevitably, no delegate could have attended more than one third of the total number of Sessions. This has made the Conference Publication of particular interest, because it provides the opportunity for closer study of a paper of interest from a Session not attended.

The Conference Publication is in two volumes, Part 1 Antennas (538 pp.) and Part 2 Propagation (301 pp.), and is available from the IEE as Conference Publication No 195, price £48. The volumes will prove to be valuable reference works for some time to come, and they represent a useful record of the 'state-of-the-art' for 1981.

R. S. ROBERTS

Members' Appointments

BIRTHDAY HONOURS

The Council has sent its congratulations to the following members whose names appear in Her Majesty's Birthday Honours List:

KNIGHT BACHELOR

Ernest Thomas Harrison, O.B.E., D.Sc., F.C.A. (Companion 1975) Chairman and Chief Executive of Racal Electronics. (Sir Ernest was recently awarded an Honorary Doctorate of Science by Cranfield Institute of Technology.)

MOST EXCELLENT ORDER OF THE BRITISH EMPIRE

To be an Ordinary Officer of the Military Division (O.B.E.)

Lieutenant Colonel Reginald Peter Myhill, T.D., Royal Corps of Signals, Territorial Army (Member 1967, Graduate 1963) Col. Myhill is Head of Section, Crossbar Switching System Maintenance Organization, in the Network Executive at British Telecom Headquarters.

To be an Ordinary Member of the Military Division (M.B.E.)

Squadron Leader Peter Arthur George Leach, RAF (Member 1971) Officer Commanding, No. 12 Signals Unit, RAF Episkopi, Cyprus, from 1978 to 1981, now on the Staff of the Director of Signals (Air), Ministry of Defence.

CORPORATE MEMBERS

F. K. El-Jadiry, M.Sc. (Fellow 1974, Member 1960, Graduate 1965) who joined the Emirates Telecommunications Corporation Abu Dhabi in 1978 as Contracts Manager, is now Engineering Manager.

P. M. Holker (Fellow 1981, Member 1969) has been appointed to the newly created position of Marketing Director of CSA (C&S Antennas) Rochester, Kent. Before joining CSA, Mr Holker was Sales Manager with Hatfield Instruments, following ten years service with the Marconi Company in senior sales positions both in the UK and abroad.

D. Obersby, B.Sc., (Fellow 1981, Member 1977, Associate Member 1974) has taken up an appointment as General Manager of Hall Automation in Watford. He was formerly Divisional Production Engineering Manager of the UK Grinding Wheel Division of Unicorn Industries, Stafford.

D. J. Corbett (Member 1965, Graduate 1963) has been appointed Head of Programme Services and Engineering, BBC Wales. Mr Corbett joined the BBC as a Technical Assistant in Lines Department in 1958 and from 1962 to 1966, he was with the Engineering Training Department, he then went to the Film Department as Assistant Film Training Manager, becoming Head of Film Operations in 1970. Six years later Mr Corbett moved to Programme Services where he was Chief Assistant, Operations, before becoming Special Assistant to Controller, Planning and Resource Management, in 1979.

A. V. T. Dike (Member 1958, Associate 1965) has returned to Marconi Space and Defence Systems after a break of some years, mainly spent in medical electronics, and has joined the Underwater Weapons Division as Engineering Manager, Stingray Project.

D. G. Ford, Dip. El., M.Sc. (Member 1970) was recently promoted to Technical Director of Hepworth Engineering, a subsidiary of Hepworth Ceramics. Mr Ford was previously Chief Development Engineer with the company.

J. I. Grace (Member 1972, Graduate 1969) of Cable & Wireless has taken up the position of Transmission Manager in Hong Kong. Mr Grace was previously in the engineering department of the company's London Head Office as transmission system performance engineer.

S. G. Hurst (Member 1962) is now an Engineer Class 3 in the Australian Defence Science and Technology Organization, Department of Defence, Canberra. He was formerly with EMI Electronics (Aust.), Elizabeth, South Australia.

V. E. Nwebube (Member 1973, Graduate 1966) has been promoted to Principal Engineer in charge of studios and outside broadcasting with the Anambra Broadcasting Corporation in Enugu, Nigeria.

Sqn Ldr T. S. Page, RAF (Member 1973, Graduate 1967) has moved from CSDE, RAF Swanton Morley, where he was a Project Officer, and is now Officer Commanding, Engineering Squadron at RAF Gibraltar.

R. Robinson (Member 1973, Graduate 1972) has recently taken up a post as Engineering Training Manager with Systime in Leeds; since 1977 he has been with Crosfield Electronics.

D. R. Sloggett, M.Sc. (Member 1981) has been promoted from Senior Engineer to Principal Engineer at Racal-Milgo. He was previously with Software Sciences, latterly as Consultant Engineer.

NON-CORPORATE MEMBERS

R. W. Evans (Graduate 1967) is now a Senior Systems Analyst with Computer Systems Engineering in Rickmansworth Herts.

S. Ganeswaran (Graduate 1971) who was Chief Technician (Radio) with EMIRTEL in Sharjah, has taken up an appointment with the Emirates Telecommunications Corporation as Engineer (Long Distance Services) in Ras Al Khaimah, UAE

Laing Yew Jen (Associate Member 1977) has taken up an appointment as a Senior Technical Assistant in an oil company in Sarawak, Malaysia, following some ten years with the Director of Telecommunications' Department, of the Government of Malaysia.

V. Kumar (Graduate 1972) who has been working in Zambia as a sectional engineer with a copper mining company, has moved to Australia and is now an Electrical Project Engineer with Raymond Engineers Australia of Perth.

Sqn Ldr D. M. McKeown, M.B.E. (Associate Member 1975) is now Staff Officer, MOD(Air) Training Staff at RAF Abingdon following nearly two years as Officer Commanding No. 81 Signals Unit, RAF Brize Norton.

Flt. Lt. I. D. Matthews, RAF (Associate Member 1974), formerly Officer Commanding Engineering Flight at RAF Rudloe Manor, has been appointed to the Staff of RAF Support Command Signals Headquarters.

S. U. Nwoye (Associate Member 1978) who was with the Zambian Posts & Telecommunications Corporation, has returned to Nigeria where he is now working with GTE(Nigeria) as a Telecommunications Sales Engineer.

A. A. A. Olaoya (Associate Member 1974) who was formerly Assistant Chief Technologist at the MM Airport in Ikeja, Nigeria, is now head of Electronics and Telecommunications Maintenance of all Southern Airports of Nigeria.

S. Sitsabesan (Associate Member 1976), formerly with Ceylon Business Appliances as an electronics engineer, has taken up an appointment as technician with the National Printing Centre in Abu Dhabi.

A. Sorrell (Associate Member 1974) was recently promoted to Managing Director of Australasian and Asian Operations at Medtronic International in Hong Kong.

Sub Lt K. P. White, B.Eng., RN (Graduate 1979) has been posted to HMS *Collingwood* as a Weapon Engineer Officer.

Obituary

The Institution has learned with regret of the deaths of the following members.

Keith Burrows (Fellow 1977, Member 1958, Graduate 1954) of High Wycombe, died suddenly while on holiday in France on 20th March 1981. He was 50 years of age.

After graduating from Manchester University in 1951, Keith Burrows did postgraduate research in radio astronomy at Jodrell Bank Radio Experimental Station for a year to gain his Master's degree. He then served for two years National Service in the Royal Air Force, first as a Pilot, then as a Technical Signals Officer. In 1954 he went to Hunting Surveys as a geophysicist planning and conducting airborne magnetic and electromagnetic surveys worldwide. In 1957 he obtained a research bursary in the Geophysics Department at Imperial College, where he remained until 1964 and during this period he carried out geomagnetic measurements by means of rocket vehicles which gained him his doctorate in 1962. Papers on the measurement, telemetry and data processing equipment involved in this project, carried out at Woomera, were published in the Institution's Journal, for one of which he was awarded the A.F.Bulgin Premium. Three years at the NASA Goddard Space Flight Center followed during which time he carried out further rocket-borne ionospheric measurements. On his return to the United Kingdom in 1967, Dr Burrows joined the then Radio and Space Research Station at Slough (now Appleton Laboratory) as a Principal Scientific Officer and for the next few years continued his pioneering work on *in-situ* magnetic field and other measurements in the ionosphere. His most notable contributions included confirmation of the wind-shear theory of sporadic-E and measurement of mid-latitude ionospheric currents during a magnetic storm. With the change in emphasis of the Laboratory's work towards satellite communications and technology, Dr Burrows' interests changed towards laser experiments and infra-red radiometry and it was in these areas that he was working at the time of his death.

Apart from his contributions to the Institution's Journal, referred to above, Keith Burrows had over thirty publications to his credit. Between 1967 and 1969 he served on the Institution's Programme and Papers Committee and he was an Honorary Lecturer on geomagnetics at University College London. He also served on working groups of the UK Space Policy and Grants Committee and of the Commonwealth Consultative Committee on space research. Since 1973 he had been a part-time tutor/counsellor for the Open University.

Peter Richard Deacon (Member 1973, Graduate 1965) of Camberley, Surrey, died last year, aged 39. After an electrical engineering apprenticeship with the British Aircraft Corporation, Mr Deacon worked as an electronics engineer in the Electronics and Technical Office. After two years in 1964, he moved to the Research and Development Department of Louis Newmark as Flight Development Engineer working on helicopter autopilots. From 1965–67 he was Technical Liaison Officer with Tectonic (Electronics), Wokingham and in 1967 he joined Mullard as a Product Engineer. He was subsequently appointed Deputy Sales Manager of the Computer Electronics Division.

Harry John Figgest (Associate 1947, Student 1946) of West Drayton, Middlesex, died on 29th March 1981, aged 60. Following service with the Royal Air Force during the war as a Radar Mechanic, Mr Figgest joined the Ministry of Civil Aviation in 1947 and he remained with the Civil Aviation Authority until his retirement last year.

Peter Hackwood (Member 1974) of Erith, Kent, died recently, aged 46. Following five years with the Royal Air Force, latterly as an Air Fitter, he worked initially for the Plessey Company and A. H. Hunt Capacitors before joining the London Electricity Board in 1964 as an engineer in charge of development projects involving the application of electronic equipment in power systems.

Victor Royce Hunt (Member 1973, Graduate 1963) of Winnersh, Berkshire, died in July 1980 aged 59. During the war Mr Hunt served in the Royal Air Force as a Wireless Mechanic subsequently becoming, prior to his retirement, a radio instructor and a senior member of the technical literature section of the No. 1 Radio School. From 1960–62 he was a senior technical author at the Royal Radar Establishment, Malvern and he then received an appointment as Lecturer at the RRE College of Aeronautics. When this was closed in 1968 he transferred to the School of Electronic Engineering, REME, Arborfield where he continued lecturing on a variety of courses in electronics and radar.

Daniel Iu Leung (Graduate 1962) died on 23rd December 1980, aged 46. Before moving to Canada early in 1980, Mr Leung held appointments with radio manufacturing companies in Hong Kong. He obtained his technical education in London between 1955 and 1962 at Battersea College of Advanced Technology and Wandsworth College of Technology, gaining Higher National Certificates.

Robert Basil Light (Associate Member 1973, Associate 1958) of Frimley, Surrey, died recently, aged 59. Entering the Royal Air Force as a Wireless Operator Mechanic in 1938, Mr Light remained with the service until 1949, much of his time having been spent with Transport Command, latterly in charge of all ground radar at RAF Lyneham. In 1950, he joined International Aeradio, working for the company at various overseas stations until 1966 when he returned to the United Kingdom. Up to the time of his death he was Technical Services Manager at the Company's head office in Southall, Middlesex.

Jack Morton (Associate Member 1973, Associate 1953) died recently at Maidstone, Kent, aged 63. Mr Morton served with the Royal Signals during the war and on demobilization joined the War Office as a Civilian Technician. Up to 1967 he was employed at War Office Wireless Stations and on promotion subsequently became Projects Manager for various army systems, being stationed at the School of Signals, Blandford Camp, for a number of years.

Denis Manktelow Neale (Fellow 1971, Member 1964) of Brentwood, Essex, died on 15th December 1980 at the age of 59. A graduate of the City and Guilds Engineering College where he obtained his B.Sc. in electrical engineering in 1941, Mr Neale was throughout his professional career concerned with the applications of electronics in industries producing materials. From 1941–45 he was with the Morgan Crucible Company, latterly as a research engineer responsible for measurements and research on carbons for arc lights and he then moved to Ilford to design and develop electronic devices for the control of processes in photography. In 1967 he was appointed Manager of the Applied Physics Section in the engineering research division where as a Project Manager he had overall responsibility for a wide range of processes and techniques. Throughout his career Mr Neale was a prolific inventor and he had approaching fifty patents and applications to his credit. He was the author of numerous papers, two of them to the Institution on cold cathode circuit design and applications and he also was the author of a book on this latter subject.

Edward James Palmer (Member 1947) of Tauranga South, was killed in an air crash in New Zealand in December 1979; he was 63 years of age. Born and educated in Auckland, New Zealand, Mr Palmer held appointments with Australian and New Zealand radio manufacturing companies prior to the war, when he served with the Royal Air Force, commanding radar and 'Oboe' stations. On demobilization he returned to hold appointments with New Zealand manufacturing companies.

Frederick George Smith (Fellow 1978, Member 1957, Graduate 1950, Student 1942) of Bristol died on 16th January aged 59. Mr Smith was for the first few years of his professional career concerned with public address installations on land and in troop ships. From 1944–1947 he served with the Royal Electrical and Mechanical Engineers and in 1949 he joined the BBC as a shift engineer in Cardiff. He subsequently became a sound supervisor of the television outside broadcasting unit in Wales and in 1969 moved to the Midlands as Engineer in Chief, Operations for Birmingham. In 1970 he was appointed Head of Programme Services and Engineering for the network production centre in Bristol, a position he held up to his death.

Contributors to this issue

Arthur Hartog took a B.Sc. in electronics at Southampton University in 1976. He then joined the Optical Fibre Group of the Electronics Department at the University as a Research Student, where he is now working as a Research Fellow. Mr Hartog has studied propagation and measurement techniques in optical fibres, including material and intermodal dispersion, and the backscattering technique for fibre evaluation. He was awarded the 1979 John Logie Baird Travelling Scholarship by the Royal Television Society.

Catherine Ragdale received a B.A. degree in engineering and electrical sciences from the University of Cambridge in 1977. She subsequently joined the Southampton University Optical Fibre Group to study for the degree of Ph.D. She is at present holding a Research Fellowship from the Royal Commission for the Exhibition of 1851. Miss Ragdale has studied such aspects of propagation in single-mode fibres as dispersion, birefringence and bending loss. She was awarded the 1980 John Logie Baird Travelling Scholarship by the Royal Television Society.

Professor Alec Gambling (Fellow 1964, Member 1958) was appointed to a Chair in Electronics at the University of Southampton in 1964 and from 1974–80 he was Head of the Department; he is now British Telecom Professor of Optical Communications in the Department.

Professor Gambling graduated in electrical engineering from the University of Bristol, his Ph.D. was awarded by the University of Liverpool and his D.Sc. by the University of Bristol. Before going to Southampton, Professor Gambling was a Lecturer at the University of Liverpool and he also spent two years as a Postdoctoral Research Fellow at the University of British Columbia. He has held appointments as Visiting Professor at the Bhabha Atomic Research Centre, Bombay, and at the University of Colorado. He is a member of the Optics and Infra-Red Committee of the Electronics Research Council and of the Technology Subcommittee of the UGC.

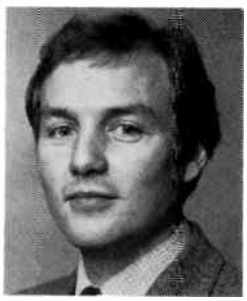
In 1966 Professor Gambling initiated Southampton University's very successful research project on optical fibre communications which has received worldwide recognition. He has contributed many original papers and reviews to the technical literature on a wide range of subjects; his earlier research interests were on microwaves and lasers, and in recent years he has been the author of numerous papers on optical fibre technology and communications.

Actively involved in Institution affairs for many years, Professor Gambling has served as Chairman of the Southern Section Committee and of conference organizing committees, on the Education Committee, and as an Ordinary Member of the Council; he was a Vice President from 1970–73 and from 1974–77, and President in 1977–78. He has been the IERE member on the Board of CEI and he was in 1979 elected to the Fellowship of Engineering.

Michael Reeve graduated from Durham University in 1973 with a degree in applied physics and joined the British Telecommunications (then Post Office) Research Department at Martlesham in Suffolk. He has worked on a number of aspects of optical fibre communications including fibre scattering and mode cut-offs, fibre strength and fibre cabling. He is currently head of a group concerned with optical cable research.

Andrew Carter received an honours degree in physics from the University of Oxford in 1973, where he continued as a postgraduate student, researching the infra-red properties of impurities in semiconductors, and received the degree of D.Phil. in 1977. Since then he has been employed at the Allen Clark Research Centre of the Plessey Company, where he has worked on III-V semiconductor devices for fibre-optic systems.

Ian Garrett graduated from Trinity College, Cambridge in 1965, and completed a Ph.D. on radiation damage in metals at Cambridge in 1969. He then joined the Post Office Research Department (now British Telecom Research Laboratories) as a Research Fellow working on the theory of chemical transport reactions. In 1971 he became a group leader, responsible for preparation of compound semiconducting films and crystals. Since 1976 Dr Garrett has led a section responsible for optical transmitters and receivers and integrated optic devices.



After receiving a B.Sc. in electronics from Southampton University in 1971, **Paul Kirkby** joined Standard Telecommunication Laboratories, Harlow. During the next three years he worked with G. H. B. Thompson on the liquid-phase epitaxy of multilayer GaAlAs laser structures with extremely thin active layers. He then joined the device technology group developing high reliability oxide stripe lasers. During this period he studied optical waveguiding and reliability for an external Ph.D. with Southampton University and was awarded his doctorate in 1978. Dr Kirkby has recently been appointed manager of the long wavelength laser group at STL and is currently involved with the development of GaInAsP lasers for submarine and land line applications.

Currently head of the Transmission Division at Standard Telecommunication Laboratories, Harlow, **Peter Radley** graduated in electrical engineering at Cambridge University. After some industrial experience he gained a doctor's degree in computer aided design for electronics; his early work also covered the growing technologies of hybrid microcircuits and large-scale integration. Responsibilities on submerged cable system developments and analogue coaxial cable systems have led to Dr Radley's position as head of a line transmission division.

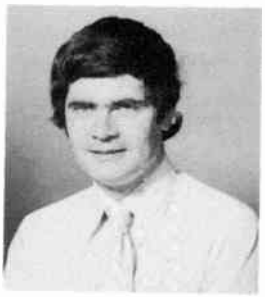
Kenneth White was educated at Edinburgh University, receiving a B.Sc. in physics in 1968 and a Ph.D. in 1972 for a light scattering study of some ferroelectric crystals. Since then he has worked at British Telecom Research Laboratories on optical communications research, developing several methods of fibre characterization including a calorimetric measurement of absorption loss and fibre refractive index profile. He now leads a group concerned with optical fibre propagation studies.

Stephen Hornung graduated from the University of Exeter in 1975. In 1980 he received the D.Phil. from the University of Oxford for work in low temperature nuclear orientation of some neutron-light isotopes of iridium. Since 1979 he has worked for British Telecom Research Laboratories on optical fibre communications research, specializing in measurements on optical fibres with particular interest in cabling aspects.

Biographies of other authors are given on pages 340 and 413.

Brian Nelson received a CNAA degree in electronic engineering following study at Liverpool Polytechnic. He worked for a Ph.D. degree in antenna theory at Leeds University which was awarded in 1974. Since then he has worked at British Telecom Research Laboratories on the propagation characteristics of optical fibres.

Michael Brierley received his degree in electronic engineering from Liverpool University in 1971. Later that year he joined British Telecom Research Laboratories where he has since worked on optical glass and fibre characterization measurements including loss mechanisms and refractive index profiles.



John Wright read natural sciences at the University of Cambridge, receiving his degree in 1970. Subsequently he worked at both Cambridge and Imperial College, London on infra-red emission studies in the atmosphere. Since joining BTRL in 1974 he has been concerned with optical communications with particular emphasis on the bandwidth of multimode fibre and fibre links.

James Topping graduated from Aberdeen University in 1972 with an honours degree in electrical engineering. Later that year he joined the Post Office Telecommunications Headquarters in London and, until 1977, worked on the development of frequency division multiplex transmission equipment. From 1977 to 1980 he was involved in the development of optical fibre transmission systems for use in both the trunk and junction networks. Now an Executive Engineer, Mr Topping is currently completing a 12 month M.Sc. course in computing science at Imperial College.

The late **James North** graduated in 1940 with a B.Sc. in physics from London University and an M.Sc. in optics from Imperial College in 1941. He joined the Post Office Research Department at Dollis Hill in 1947 as a Scientific Officer and was involved in the application of optical techniques for the latter part of his career. He was with the Optical Communications Division of the Research Department from its formation and was Head of a Group responsible for microoptic components at the British Telecom Research Laboratories at Martlesham, Suffolk. He became much involved in various aspects of optical safety over the years, and was a member of the BSI committee studying laser safety. Mr North died in February 1981, shortly after the completion of his joint paper with Mr J. D. Topping.

UDC 621.372.8:621.391.63/4
Indexing Terms: Optical communications, Fibre optics, Single-mode transmission, Multimode transmission

Optical fibre transmission lines

Professor W. A. GAMBLING, D.Sc., F.Eng., F.I.E.R.E.*

A. H. HARTOG, B.Sc.*

and

C. M. RAGDALE, M.A.*

SUMMARY

Optical fibre transmission lines have many advantages over coaxial cables. The most widely used fabrication techniques involve chemical vapour deposition resulting in losses of well below 1 dB/km. Analysis of propagation follows well-defined principles and the results obtained are broadly comparable with those at microwave frequencies, although the wavelength of operation, ~1 μm, is much shorter. By suitable design single-mode fibres can be made to give bandwidth × length products approaching 100 GHz · km. Multimode fibres are now well understood, pose fewer problems and can be excited with light-emitting diodes.

* *Department of Electronics, University of Southampton, Southampton, Hampshire SO9 5NH*

1 Introduction

The structure of optical fibre transmission lines takes the very simple form of a cylindrical glass core of refractive index n_1 surrounded by a cladding glass of refractive index n_2 where $n_2 < n_1$. Normally most of the propagating energy is contained in the core but there is always a radially-decaying evanescent field in the cladding, which may extend over several wavelengths in the case of single-mode fibres. Both core and cladding materials must therefore have very low absorption and scattering losses.

As with waveguides, when the transverse dimensions of the guiding structure, in this case the core, are comparable with a wavelength then only a single mode can be supported, whereas for larger core diameters multimode operation prevails. Section 2 indicates that the wavelength of operation is in the region of 1 μm, corresponding to a frequency of 300,000 GHz, so that *single-mode fibres* have a core diameter of 1 to 10 μm while *multimode fibres* have standardized core diameters of between 50 and 60 μm. For practical convenience the outer diameter is made 125 μm in both cases. In *step-index fibres* the refractive indices are constant in both core and cladding, whereas in (ideal) *graded-index fibres* the refractive index is a maximum n_1 at the core centre but falls monotonically to that of the cladding n_2 at the core boundary. With fibres designed for long-distance transmission $(n_1 - n_2) \ll n_1 \simeq 1{\cdot}5$ and the *relative refractive-index difference* $\Delta \simeq (n_1 - n_2)/n_1$ is about 1%.

At this point it is convenient to provide definitions of a few other basic quantities which are used in later Sections. Firstly, the maximum angle θ_m to the axis that light can enter a fibre at the input end from a medium of refractive index n_0 is defined in terms of the *numerical aperture* NA as

$$n_0 \sin \theta_m = NA = (n_1^2 - n_2^2)^{\frac{1}{2}} \simeq n_1(2\Delta)^{\frac{1}{2}} \qquad (1)$$

If the medium is air, for which $n_0 = 1$, and $\Delta = 0{\cdot}01$ then $\theta_m \simeq 12°$. The core radius a is usually normalized to the free-space wavelength of operation λ through

$$V = (2\pi a/\lambda)(n_1^2 - n_2^2)^{\frac{1}{2}} \simeq (2\pi a n_1/\lambda)(2\Delta)^{\frac{1}{2}} \qquad (2)$$

V is called the *normalized frequency* although it could equally well be referred to as the normalized core diameter or normalized wavelength.

Why, one might ask, and under what circumstances, are optical fibres preferred to other forms of transmission line? Some of their merits and drawbacks are discussed in the following Sections, and other papers contained in this issue, but may be summarized as follows:

Advantages

1. Extremely low transmission loss (down to 0·2 dB/km) giving distances between repeaters in a trunk network, or in underwater cables, of 100 km and more, compared with 2 km for coaxial cables.

0033–7722/81/070313+13 $1.50/0

2. Extremely large bandwidths of up to 1 GHz·km for graded-index multimode fibres and 100 GHz·km for single-mode fibres, compared with ~20 MHz·km for coaxial cable.
3. Small size, low weight and high degree of flexibility.
4. Freedom from electromagnetic interference and earth-loop problems.
5. Fabricated from relatively abundant materials (silica, phosphorus, germanium, boron).
6. Zero cross-talk between closely-spaced lines.
7. Larger Young's modulus and resistance to crushing than copper.

Disadvantages

1. Glass is brittle and therefore breaks when the elastic limit is exceeded.
2. Long-term (20 years) mechanical stability under strain is unknown.
3. Demountable connectors and other similar components are expensive.

The properties of the principal types of optical fibre transmission line are summarized in Table 1.

The present paper attempts to do two things. Firstly, it outlines briefly the principal methods of fabrication, and the related attenuation characteristics, so that the properties of fibres can be more easily understood. Secondly, it describes the latest work, some of which has not been published elsewhere, on propagation in single-mode and multimode fibres. The details of the analyses are omitted but the principal steps and results are described in terms acceptable to the systems designer and practising engineer.

2 Fabrication of Optical Fibres

The glassy form of pure SiO_2 has an extremely low transmission loss in the optical and near-infra-red regions of the spectrum and can be drawn into long lengths with a high degree of precision. Silica, unfortunately, has the low refractive index of 1·45 and, when used as the core of a fibre, there are comparatively few compatible materials which have a sufficiently low refractive index to act as cladding. Possibilities are silica admixed with a small proportion of boron or fluorine and certain plastics such as silicone rubber. Conversely the index of silica can be raised by adding such oxides as P_2O_5, GeO_2 or TiO_2.

2.1 Chemical Vapour Deposition Techniques

Synthetic silica can be made by vapour deposition techniques involving oxidation or hydrolysis of volatile silicon compounds such as silicon tetrachloride or silane, which are liquids at room temperature. Methods of this kind are also used, but at very low deposition rates, in the semiconductor industry. Transition metal impurities, which cause high absorption, are largely excluded because the vapour pressures of their chlorides and hydrides are very much lower than those of silicon. Thus,

Table 1

Properties of optical fibres

Type of fibre		Core diameter (μm)	Materials	Transmission loss (dB/km)				$B \times L$ product
				0·85 μm	1·05 μm	1·3 μm	1·5 μm	GHz·km
Single-mode		1–10	Core: silica based glass Cladding: silica based glass	2	1	0·38	0·2	50–100
Multimode	Step-index	50–200	Core: silica based glass Cladding: silica based glass	2	1	0·5	0·2	0·005 to 0·02
			Core: silica based glass Cladding: plastic	2·5	1·5	high		
			Core: multi-component glass Cladding: multi-component glass	3·4	6	high		
	Graded-index	50–60	Silica glass	2	1	0·5	0·2	1
			Multicomponent glass	3·5	10	high		0·4

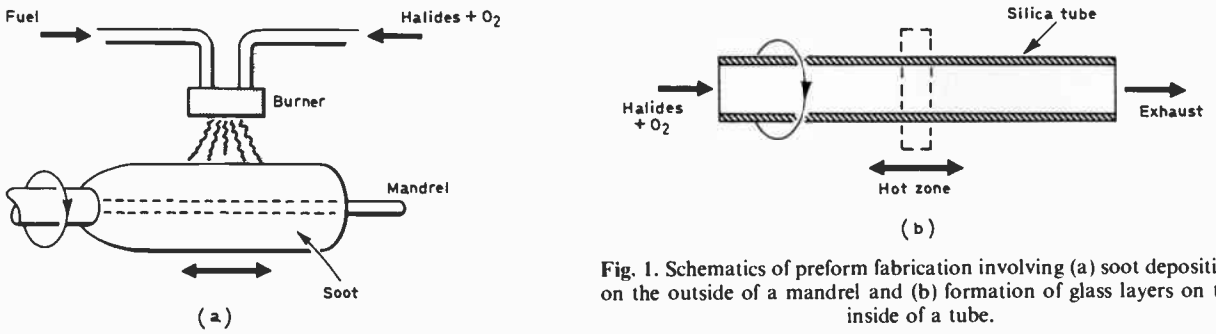


Fig. 1. Schematics of preform fabrication involving (a) soot deposition on the outside of a mandrel and (b) formation of glass layers on the inside of a tube.

when the appropriate silicon compound is vaporized by passing oxygen through it the impurities are left behind. If the mixture, now pure, is blown through an oxy-hydrogen flame then oxidation occurs and silica is generated as a fine soot which may be deposited on a cool mandrel. The oxidation reaction is the simple one: $SiCl_4 + O_2 \rightarrow SiO_2 + 2Cl_2$. When a sufficient thickness has been deposited the mandrel is removed and the porous mass of silica soot is then sintered and collapsed into solid glass.

The process can be adapted to the fabrication of fibres by starting with a mixture of two vapours such as $SiCl_4$ and $GeCl_4$, together with oxygen, as in Fig. 1(a). The mandrel is rotated and traversed in the longitudinal direction under the flame from the burner, so that the corresponding oxides deposit as a fine, mixed, porous soot on the mandrel.[1] When a sufficient thickness has been achieved the gas composition is changed so that a soot of the mixed cladding oxides forms a second layer. The mandrel is then removed and the soot preform is carefully sintered into a solid composite preform which is drawn into fibre. This technique is often referred to as Outside Vapour Phase Oxidation (OVPO) and is one of the general class of Chemical Vapour Deposition (CVD) methods.

The method most widely used involves deposition on the inside of a tube in such a way that soot formation and sintering are simultaneous. It was developed independently at Southampton University[2] and Bell Telephone Laboratories[3] and is known as the modified, or homogeneous, CVD process (sometimes as the Inside Vapour Phase Oxidation or IVPO process). The tube is usually of high-quality silica, having uniform thickness and good circularity. This tube can form the cladding material but more usually acts simply as a supporting structure. In that case a low-loss silica cladding can be produced by picking up the vapour of silicon tetrachloride by bubbling oxygen through the liquid and passing the mixture along the tube, as in Fig. 1(b). By heating the tube with an outside heat source, usually a flame, the oxidation reaction again produces a fine soot of silica but it now, in the homogeneous or modified CVD method, deposits and sinters into a clear glass layer on the wall of the tube. The thickness may be up to 10 μm. In order to ensure uniformity the heating torch is traversed in the longitudinal direction and the tube is rotated. When a sufficient total thickness has been formed by successive traverses of the torch the vapours of the chlorides of phosphorus ($POCl_3$) or germanium ($GeCl_4$) are added to the flow and layers of phosphosilicate or germanosilicate glass are deposited to form the core glass. The deposition temperature is in the region of 1500 to 1600°C. In practice boric oxide (from BCl_3) or fluorine is added to the silica cladding to produce a lower refractive index and a lower deposition temperature which reduces tube distortion.

The second stage of the process is to raise the temperature of the torch so that the composite tube softens and surface tension forces cause it to collapse into a solid rod.[4] Usually two or three passes of the flame are sufficient. Finally the preform is mounted vertically and is drawn into a fibre in a fibre-drawing machine, where a short section at one end of the rod is heated to softening temperature. The lower end is pulled downwards at a constant velocity of the order 1 m/s and the resulting fibre diameter is determined by the ratio of the fibre drawing speed to the rate at which the preform is fed into the furnace. A great advantage of the CVD methods, especially the homogeneous (inside) process is that the composition of successive layers can be changed thus producing a graded refractive-index distribution.

The methods described so far are essentially batch processes in that a preform is made and then drawn into a finite length, up to 10 km, of fibre. A more recent development is that, instead of depositing the glass soot on the outside cylindrical surface of the mandrel as in the OVPO method, it is deposited on the end of the rotating boule. By slowly withdrawing the boule in the longitudinal direction at the same rate as the deposition builds up a continuous process results. This is the basis of the Vapour Phase Axial Deposition (VAD) method.[5]

2.2 Multicomponent Glass Fibres

An alternative method of making fibres involves multicomponent glasses, such as sodium borosilicate, which can be worked at the much lower temperatures of 900–1300°C. A double-crucible arrangement[1], as in Fig. 2, is normally used. Core glass is melted in the inner crucible and cladding glass in the outer, and they flow through the nozzles to form a fibre. This method may be

suitable for mass production but care has to be taken to prevent impurities entering from the crucibles and it is difficult to obtain starting materials of the required quality. Great care must therefore be taken during preparation.

As shown in Table 1, multimode fibres with both a step-index and a graded-index, including parabolic, distribution in the core have been produced from multicomponent glasses. The loss has recently been reduced to less than 3·5 dB/km at a wavelength of 0·85 μm but it rises rapidly beyond 1 μm. A graded distribution can be obtained with the double-crucible method by allowing diffusion of ions, and therefore an ion exchange, after the core glass has emerged from the inner nozzle.

The Rayleigh scattering loss may be smaller in certain compound glasses, but unfortunately multicomponent scattering is increased.

2.3 Other Fabrication Methods

There are a number of other fabrication processes which cannot all be mentioned here. Generally the losses in the resulting fibres are higher but this may not be important in short-distance applications.

A simple process is to insert a rod of core glass into a tube of cladding glass and to draw the combination into fibre. In another variation a silica rod is drawn into fibre and coated with a suitable plastic, of lower refractive index, which forms the cladding.[4]

2.4 Transmission Loss of Optical Fibres

The factors contributing to loss in an optical fibre transmission line include absorption, scattering due to inhomogeneities in the core refractive index (Rayleigh scattering), scattering due to irregularities at the boundary between core and cladding, bending loss, loss at joints and connectors and the coupling losses at the input and output.

Remarkable progress has been made in reducing the transmission loss, which has fallen by about three orders of magnitude in the past ten years and is now much lower than that of coaxial cables having a similar transmission bandwidth. The absorption loss at some wavelengths is almost negligible and below about 0·8 μm scattering is the dominant factor.

The main cause of the absorption loss is the presence of transition metals such as Fe, Cu (especially in multicomponent glasses), water in the form of OH^- ions and the intrinsic absorption of the pure glass itself. In order to reduce the absorption to an acceptable level, it is necessary to prevent a metal concentration of more than 1 in 10^9, and an OH radical concentration of more than 1 in 10^7, from occurring.

Another purely material effect in addition to absorption is the scattering due to inhomogeneities of the refractive index. These fluctuations are on a scale

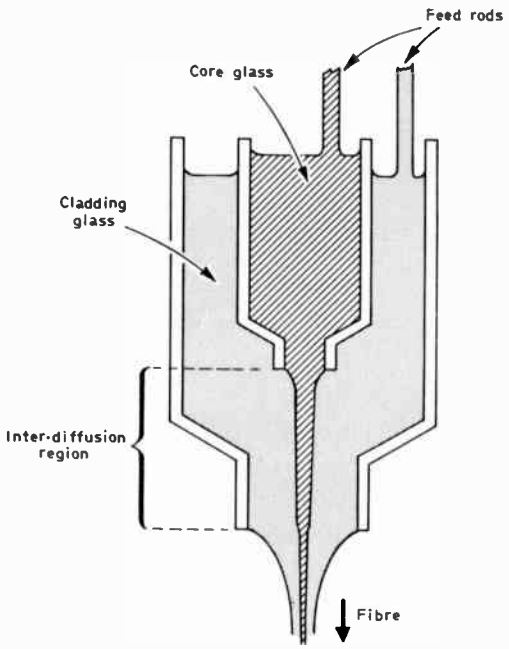


Fig. 2. Fibre drawing with double crucible showing diffusion zone for producing a graded refractive index.

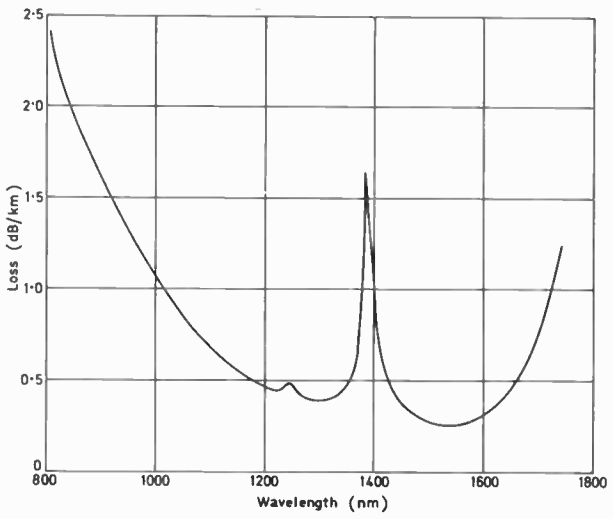


Fig. 3. Transmission loss of a single-mode fibre with a germanosilicate core of 10·5 μm diameter and phosphofluorosilicate cladding. Peaks due to OH^- are present at 1·24 and 1·39 μm.

which is smaller than the wavelength and the resulting Rayleigh scattering is inversely proportional to the fourth power of the wavelength (λ^{-4}), so that it becomes rapidly smaller at longer wavelengths.

The transmission of a single-mode fibre† with a germanosilicate core and phosphofluorosilicate cladding is shown in Fig. 3 and exhibits the characteristic features of fibres made by the modified CVD process. In this particular case the core diameter is 10·5 μm, the relative

† K. J. Beales (Private Communication).

refractive-index difference is $\Delta = 0{\cdot}17\%$ and the cut-off wavelength of the second set of modes (TM_{01}, TE_{01} and HE_{21}) is 1·2 μm. At short wavelengths the attenuation is inversely proportional to the fourth power of the wavelength and is therefore typical of Rayleigh scattering.

The rise in attenuation beyond 1·7 μm is attributed to the intrinsic infra-red absorption of the glass. The effect of OH^- impurities can be clearly seen but are at a much lower level than is normally observed. The fundamental vibration is at $\lambda = 2{\cdot}8$ μm in silica and there are overtones at 1·39 μm and 1·24 μm.

The transmission loss at 1·3 μm of 0·38 dB/km is the lowest yet reported at that wavelength and is 0·28 dB/km at 1·55 μm. The fibre illustrated in Fig. 3 was designed for operation at 1·3 μm but it has been shown that in other fibres[6] the loss at 1·55 μm can be as low as 0·2 dB/km. Large transmission distances are therefore possible (see Table 1), and British Telecom are already designing single-mode fibre systems operating at 1·3 μm over repeater spacings of 30 km at 140 Mbit/s.

A mode conversion loss, and a loss due to radiation, occur if the fibre has small irregularities at the boundary between the core and cladding. However, this interface scattering, which is referred to as 'microbending', can be reduced by increasing both the core radius and the index gradient in the core in order to minimize the light intensity at the core/cladding boundary.

Bends can also cause mode conversion to occur in addition to the energy loss due to radiation.

3 Propagation in Single-mode Fibres

The analysis of optical fibres follows the same procedure as that for any other transmission line which guides electromagentic waves. Thus solutions of Maxwell's equations and the corresponding wave equation are sought in terms of the appropriate boundary equations, using well-established techniques. For each of the propagating modes it is possible to deduce the spatial distribution of the electric and magnetic fields, the propagation constant, phase and group velocities, and so on, in the normal way. Optical fibres differ in degree only and not in principle from, say, hollow metal waveguides, in that they are designed for operation at frequencies higher by a factor of 10^4 and the guiding structure is fabricated entirely from dielectric materials since metals are very lossy at optical frequencies.

3.1 Basic Concepts

A dielectric waveguide supports a finite number of guided modes and an infinite number of radiation modes which together form a complete orthogonal set. Only guided modes are considered here.

To simplify the analysis of fibre waveguides the cladding may be assumed to be of infinite extent. In practice this simply means that the cladding diameter must be large enough for the field to decay to a negligible level at its outer edge. In single-mode fibres a significant proportion of the power is carried in the cladding which must have a diameter roughly seven times that of the core, say 30 μm.

An exact description of the modal fields is complicated, but the analysis can be simplified by making use of the fact that, in practice, $(n_1 - n_2) \ll n_1$, the well-known 'weakly-guiding approximation'. The approximate mode solutions derived in this way are very nearly linearly-polarized[7,8] and are denoted by $LP_{\nu\mu}$ where ν and μ denote the zeros of the field in the azimuthal and radial directions respectively. These linearly-polarized modes correspond to a superposition of the two modes $HE_{\nu+1,\mu}$ and $EH_{\nu-1,\mu}$ of the exact solution of Maxwell's equations. The exact modes are nearly degenerate and as $n_2 \to n_1$ their propagation constants become identical.

Maxwell's equations with the weakly-guiding approximation give the scalar wave equation as:

$$\frac{d^2\psi}{dr^2} + \frac{1}{r}\frac{d\psi}{dr} + \frac{1}{r^2}\frac{d^2\psi}{d\phi^2} + [n^2(r)k^2 - \beta^2]\psi = 0 \qquad (3)$$

where ψ is the field (E or H), $k = 2\pi/\lambda$ is the free-space wave number, $n(r)$ is the radial variation of the refractive index and r, ϕ are the cylindrical co-ordinates. The propagation constant β of a guided mode obviously lies between the limits $n_2 k < \beta < n_1 k$. The fibre is circular in cross-section and the solutions of the wave equation are separable, having the form:

$$\psi = E(r)\cos\nu\phi \exp[j(\omega t - \beta z)] \qquad (4)$$

For simplicity the factor $\exp[j(\omega t - \beta z)]$ will be omitted from later equations.

In single-mode fibres only the fundamental LP_{01} mode propagates and it has no azimuthal dependence, i.e. $\nu = 0$. It corresponds to the HE_{11} mode derived from the exact analysis. For this fundamental mode equation (3) reduces to

$$\frac{d^2E}{dr^2} + \frac{1}{r}\frac{dE}{dr} + [n^2(r)k - \beta^2]E = 0 \qquad (5)$$

In a step-index fibre, i.e. one with a constant refractive index n_1 in the core, equation (5) is Bessel's differential equation and the solutions are cylinder functions. The field must be finite at $r = 0$ and therefore in the core region the solution is a Bessel function J_ν. Similarly the field must vanish as $r \to \infty$ so that the solution in the cladding is a modified Bessel function K_ν. For the fundamental LP_{01} mode polarized in either the x or y direction the field is therefore[7]

$$\begin{aligned} E(r) &= AJ_0(UR) \quad R < 1 \quad \text{(core)} \\ &= AJ_0(U)\frac{K_0(WR)}{K_0(W)} \quad R > 1 \ \text{(cladding)} \end{aligned} \qquad (6)$$

where $R = r/a$ is the normalized radial coordinate and A

is the amplitude coefficient. U and W are the eigenvalues in the core, and cladding, respectively and are defined by

$$U^2 = a^2(n_1^2k^2 - \beta^2) \qquad (7)$$
$$W^2 = a^2(\beta^2 - n_2^2k^2)$$

therefore

$$V^2 = a^2k^2(n_1^2 - n_2^2) = U^2 + W^2$$

Related to these parameters is the normalized propagation constant b, defined as[8]

$$b = [(\beta/k)^2 - n_2^2]/2n_1^2\Delta = 1 - \frac{U^2}{V^2} \qquad (8)$$

where

$$\Delta = \frac{n_1^2 - n_2^2}{2n_1^2} \simeq \frac{n_1 - n_2}{n_1} \ll 1$$

Since, for a guided mode, the limits of β are n_2k and n_1k then b must lie between 0 and 1.

The field expressions in equation (6) are normalized so as to have the same value at $r = a$. In addition the tangential electric field components must be continuous at this point, leading to the following eigenvalue equation for the LP_{01} mode:

$$\frac{UJ_1(U)}{J_0(U)} = \frac{WK_1(W)}{K_0(W)} \qquad (9)$$

It should be noted that it is only because of the weak-guidance approximation that the boundary conditions of the magnetic field components are also satisfied by this condition.

By solving equations (7) and (9) the eigenvalue U, and hence β, can be calculated as a function of the normalized frequency V. Therefore the dependence of the propagation characteristics of the mode on the wavelength and fibre parameters can be determined.

At the lower limit of $\beta = n_2k$ the mode phase velocity equals the velocity of light in the cladding and is no longer guided, the mode is cut off and the eigenvalue $W = 0$ (equation (7)). As β increases, less power is carried in the cladding and at $\beta = n_1k$ all the power is confined to the core.

The limit of single-mode operation is determined by the wavelength at which the propagation constant of the second, LP_{11}, mode equals n_2k. For a step-index fibre this cut-off condition is given by

$$J_0(V_c) = 0$$

where V_c denotes the cut-off value of V which, for the LP_{11} mode, is equal to 2·405. The fundamental mode has no cut-off, hence single-mode operation is possible for $0 \leqslant V \leqslant 2{\cdot}4$.†

† The reader is invited to deduce, from simple physical principles, the consequences of operating in practice at too low a value of V.

3.2 Dispersion in Single-mode Fibres

The bandwidth of optical fibres is limited by broadening of the propagating light pulse which has a finite spectral width due to (i) the spectral width of the source, and (ii) the modulation sidebands of the signal. If, therefore, the fibre waveguide is dispersive the different frequency components will travel at different velocities resulting in pulse distortion.

The transit time for a light pulse propagating along a fibre of length L is

$$\tau = \frac{L}{c}\frac{d\beta}{dk} \qquad (10)$$

where c is the velocity of light.

If β varies non-linearly with wavelength the fibre will be dispersive. From equation (8) we have

$$\beta^2 = k^2n_1^2[1 - 2\Delta(1 - b)] \qquad (11)$$

Thus β is a function of the refractive indices of the core and cladding materials and of b. Equation (8) shows that b is a function of V so that pulse dispersion arises from the variation of b with the ratio a/λ. In addition, the refractive index of the fibre material varies non-linearly with wavelength and this also gives rise to pulse dispersion.

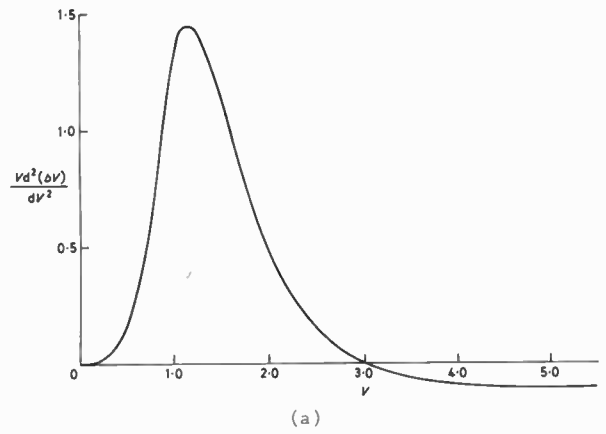
The pulse spreading caused by dispersion is given by the derivative of the group delay with respect to wavelength[9]

$$\text{pulse spread} = \left|\delta\lambda\frac{d\tau}{d\lambda}\right|L = \frac{L}{c}\frac{2\pi}{\lambda^2}\frac{d^2\beta}{dk^2}\delta\lambda \qquad (12)$$

where $\delta\lambda$ is the spectral width of the source. Substituting equation (11) into equation (12), and differentiating with respect to k, gives the dependence of the pulse spreading on the material properties and the mode parameter b. The dependence on the refractive index is given in terms of the material dispersion[9] parameter $-(\lambda/c)(d^2n/d\lambda^2)$ where $n = n_1$ or n_2 and the dependence on b is given by the mode dispersion parameter defined as $V(d^2(bV)/dV^2)$. In addition a third term, which is proportional to $d\Delta/d\lambda$, arises from the differentiation in equation (12).

The preceding three effects are inter-related in a complicated manner, but Reference 10 shows that the expression for pulse spreading can be separated into three composite dispersion components in such a way that one of the effects dominates each term. For example, a composite material dispersion term can be defined which has a dependence on both b and $d^2n/d\lambda^2$, however it becomes zero when $d^2n/d\lambda^2$ is zero.

In multimode fibres the majority of the modes are far from cut off and most of the power is carried in the core. In this case the composite dispersion components simplify to terms which depend on either material or mode dispersion, and the two effects can be separated. In addition, in step-index multimode fibres the effect of $d\Delta/d\lambda$ can be neglected.



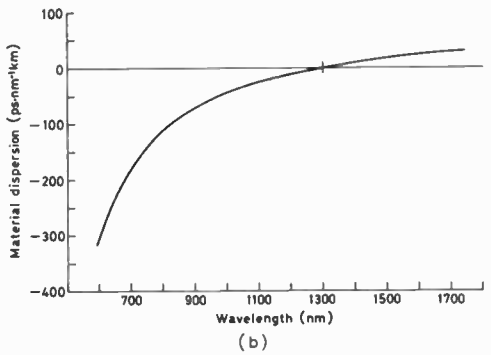


Fig. 4.(a) The mode dispersion parameter, $V(d^2(bV))/dV^2$ as a function of normalized frequency V for the fundamental LP_{01} mode.
(b) The material dispersion parameter, $-(\lambda/c)(d^2n/d\lambda^2)$ as a function of wavelength for a germanosilicate glass fibre with $NA = 0{\cdot}2$.

Material and mode dispersion also have a dominant effect in single-mode fibres but the effect of $d\Delta/d\lambda$ can no longer be neglected[10] as has been assumed by some authors.

In the absence of material dispersion the pulse spreading is controlled by the mode parameter $Vd^2(bV)/dV^2$ which is shown in Fig. 4(a) as a function of V for the LP_{01} mode. In the single-mode regime, i.e. $V < 2{\cdot}4$, the mode dispersion is always positive and reaches a maximum at $V = 1{\cdot}15$. It is seen that a change in any of the waveguide parameters, e.g. core radius or wavelength, changes V and hence the mode dispersion.

The material dispersion parameter, $(\lambda/c)(d^2n/d\lambda^2)$ is plotted as a function of wavelength in Fig. 4(b) for a germanophosphosilicate glass fibre with $NA = 0{\cdot}2$. At most wavelengths the material dispersion exceeds mode dispersion, but at 1·29 μm the material dispersion is zero[9] (i.e. $d\tau/d\lambda = 0$). Thus at wavelengths near this value the bandwidth is limited by mode dispersion.

The total dispersion of a single-mode fibre arises from the combined effects of material dispersion, mode dispersion and $d\Delta/d\lambda$ terms. As shown in Fig. 4(b) the material dispersion function changes sign at a wavelength of approximately 1·29 μm whereas mode dispersion always has the same sign in the single-mode regime. Therefore the effects of material dispersion, $d\Delta/d\lambda$ and mode dispersion can be balanced to give zero first-order dispersion at a given wavelength.[10] Hence extremely large bandwidths can, in theory and practice, be achieved in single-mode fibres.

Since the dispersive properties of the fibre depend on both the fibre core dimensions and the fibre materials the total dispersion can be altered by changing either of these parameters. The wavelength λ_0 at which the first-order dispersion is zero can therefore be tuned by appropriate choice of the core diameter or of NA. The total dispersion is plotted in Fig. 5 as a function of wavelength for different core diameters and a fixed NA of 0·23.

The range of wavelengths over which λ_0 can be tuned is limited. The maximum value depends on the usable value of NA, while the minimum value is approximately the wavelength at which material dispersion is zero ($\sim 1{\cdot}3$ μm).

If a fibre is designed to operate with zero first-order dispersion the limitations imposed on the bandwidth by secondary effects must be considered. For example, birefringence arising from ellipticity or stress in the core causes the two orthogonally-polarized modes of the 'single-mode fibres' to become distinguishable, i.e. they are no longer degenerate as in the scalar approximation.[11,12] The modes have different propagation constants which results in pulse dispersion. The dispersion caused by a difference between the major and minor axes of about 5% is less than 2 ps/km and can therefore be neglected.[11] On the other hand the pulse dispersion arising from stress birefringence may be as high as 40 ps/km if the expansion coefficients between the fibre core and cladding materials are not matched.[13]

Figure 6 shows the bandwidth of a single-mode fibre designed for $\lambda_0 = 1{\cdot}3$ μm. In the absence of second-order

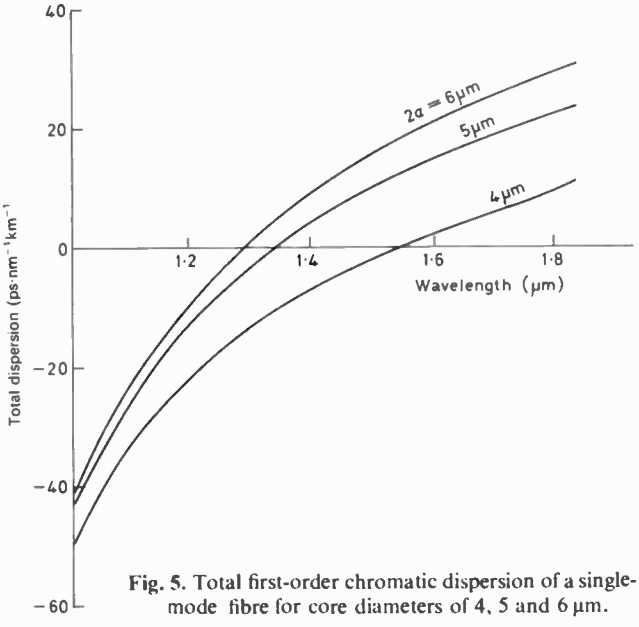


Fig. 5. Total first-order chromatic dispersion of a single-mode fibre for core diameters of 4, 5 and 6 μm.

effects the bandwidth is limited only by the spectral width of the source. Thus the solid line shows the available bandwidth with a source of linewidth 1 nm (e.g. an oxide stripe laser). On the other hand if stress birefringence introduces a pulse dispersion of 10 ps/km the bandwidth in the vicinity of λ_0 is considerably reduced (dotted curve). In the absence of polarization dispersion the bandwidth near λ_0 would be determined by higher-order effects.

Further work on this topic is required but preliminary measurements with a narrow-linewidth laser source over a 20 km length have revealed† a pulse dispersion of less than 4 ps/km.

3.3 Spot Size

The spot size of the fundamental mode is one of the most important parameters in single-mode fibre design since it largely determines the launching efficiency, jointing loss and bending loss. Usually the spot size w_0 is defined as the width to 1/e intensity of the LP_{01} mode or, alternatively, in terms of the spot size of an incident Gaussian beam which gives maximum launching efficiency.[14] The latter definition arises from the fact that the LP_{01} mode has almost a Gaussian distribution.

The spot size is a function of both V and NA, although the dependence on V is only slight[14] (w_0 changes by only 2% over the range $V = 1{\cdot}8$ to $2{\cdot}4$). The numerical aperture, on the other hand, has a strong effect since a large NA increases the guidance effect and more of the power in the LP_{01} mode is confined to the core, so that the spot size decreases.

3.4 Launching Efficiency

The ratio of power accepted by the fibre to the power in an incident beam is defined as the launching efficiency and can be calculated by integrating the product of the incident and propagating modes over the fibre cross-section.[15]

Thus launching efficiency

$$= \frac{1}{n_1}\left|\int^A E_{\text{inc}} \cdot E \, dA\right|^2 / \int^A E_{\text{inc}}^2 \, dA \int^A E^2 \, dA \qquad (13)$$

where E_{inc} is the electric field distribution of the incident mode.

Maximum power is launched into the fibre when the spot size of the LP_{01} mode is matched to the waist of the incident Gaussian beam. In practice, however, the launching efficiency of the LP_{01} mode decreases if the input beam is offset or tilted.

3.5 Joint Loss

The efficiency with which power can be coupled between two fibres is determined by the extent to which the mode patterns of the incoming and outgoing fibres can be matched. Therefore angular or lateral misalignment can considerably increase the loss at a joint. While longitudinal separation between the ends of the fibres can also occur its effect on loss in practical joints is small enough to be neglected.

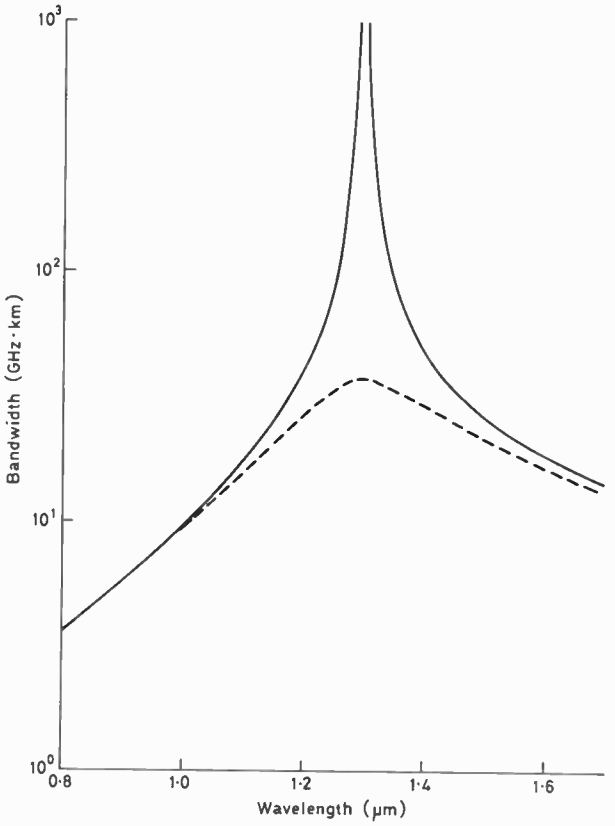


Fig. 6. Bandwidth of a single-mode fibre designed for $\lambda_0 = 1{\cdot}3$ μm, assuming a source linewidth of 1 nm and that second-order effects are small. The dotted curve shows the effect of a polarization dispersion of 10 ps/km.

If it is assumed that the spot sizes of the modes of the two fibres are the same then the joint loss can be derived simply in terms of spot size. In the absence of angular misalignment the loss caused by lateral offset[16] is

$$T_l = 2{\cdot}17\,(D/w_0)^2 \text{ dB} \qquad (14)$$

where D is the offset. The offset loss is thus inversely proportional to the square of the spot size, w_0. On the other hand the loss caused by an angular misalignment α is

$$T_a = 2{\cdot}17\,(\alpha w_0 nV/aNA)^2 \, dB \qquad (15)$$

and hence angular misalignment loss is directly proportional to the square of the spot size. Thus for a given loss there is a trade-off between the spot size, and hence NA, required for low offset loss and that required for low angular misalignment loss.

When angular and lateral misalignments occur together the combined effect is complicated,[16] but if the total loss is small it can be approximated by the sum of equations (14) and (15).

† K. J. Beales (Private communication).

3.6 Bending Loss

Radiation at bends in single-mode fibres can significantly increase the transmission loss.[17–20] The bending loss can arise either from curvature of the fibre axis or microbending, i.e. small inhomogeneities in the fibre such as diameter variations, which can arise during coating and cabling.

There are two different mechanisms giving rise to bend loss in single-mode fibres, namely transition loss and pure-bend loss.[19,20] The transition loss is oscillatory and arises because power is lost by coupling between the fundamental mode and the radiation modes. In other words the power distribution in the HE_{11} mode of the straight fibre is different from that of the corresponding mode in the curved fibre and power is lost at the interface between the two due to this mismatch.

The second mechanism, pure-bend loss, represents a loss of energy from the pure mode of the curved fibre and can be explained as follows. At a bend the phase fronts are no longer parallel and at a sufficiently large distance from the centre of curvature the increased distance between the phase fronts corresponds to a phase velocity greater than c/n_2. This part of the wave is no longer guided and radiates away from the fibre. As the curvature is increased the radius at which the phase velocity equals the velocity of light decreases, hence more energy is lost. The amount of power radiated depends on the spot size. If the spot size is reduced the power is more tightly guided in the core and the pure-bend loss decreases.

Both the transition loss and pure bend loss are strongly dependent on the NA (and hence the spot size). It is therefore possible to reduce bending loss to a negligible level by increasing the numerical aperture of the fibre.

3.7 Arbitrary Profiles

In the discussions above only a stepped refractive-index profile has been considered. In practice, however, the real profiles of single-mode fibres have a dip in the centre and some grading of the core/cladding boundary caused by diffusion of dopants during the fabrication process. In addition the refractive index in the cladding is not usually constant. The field distribution and propagation characteristics of the LP_{01} mode are thus different quantitatively, although not qualitatively, from those in the step-index fibre. Hence all of the properties discussed above will be different for fibres with different profiles.

There are two ways of dealing with this problem. The first is to use the measured refractive-index profile to produce a numerical solution[21] of the scalar wave equation (equation (3)). The second method involves matching the mode field of the real fibre to that of a fibre with equivalent step-index distribution.[22–24] This simplifies the problem since the analytical expressions for a step-index fibre can be used; however, it is not very accurate for fibre profiles which depart too far from a stepped distribution.

4 Propagation in Multimode Fibres

4.1 Basic Concepts

Although single-mode fibres have intrinsically higher bandwidths than those offered by multimode fibres, much of the research and development in optical communications has, until recently, concentrated on the latter type of waveguide. Multimode fibres have larger core diameters and numerical apertures than single-mode fibres and, as a result, can be coupled more easily to optical sources. In particular, light-emitting diodes, which are cheaper and more reliable than lasers, can be used to drive multimode fibre links. Moreover jointing and splicing losses are much lower than with single-mode fibres since the dimensions are larger and hence the alignment tolerances are much less stringent. Finally, multimode fibres are less susceptible to microbending losses.

Typical multimode fibres have a core diameter of 50 μm, a numerical aperture of 0·2 (i.e. a relative index difference Δ of slightly less than 1%) and an outer diameter of 125 μm. At a wavelength of 0·85 μm (the emission wavelength of GaAs devices) the corresponding normalized frequency is $V = 37$ and the number of guided modes (approximately $V^2/4$ for graded-core fibres) is ~ 340.

In general, therefore, power is launched into a large number of modes having different spatial field distributions, propagation constants, chromatic dispersion and so on. In an ideal fibre, having properties (e.g. core size, index difference, refractive-index profile) which are independent of distance, then the power launched into a given mode remains in that mode and travels independently of the power launched into other modes. In addition most of the modes are operated far from cut-off and their properties are, therefore, relatively independent of wavelength. This behaviour contrasts with single-mode operation where the mode parameters, such as normalized propagation constant or power confinement factor, vary rapidly with wavelength.

Since the majority of modes operate far from cut-off, and are thus well confined, most of the power carried by multimode fibres travels in the core region. The properties of the cladding therefore only significantly affect those modes which are near cut-off and whose electromagnetic fields extend appreciably beyond the core.

4.2 Intermodal Dispersion

The existence of several hundred modes, each having its own propagation constant, causes a form of pulse distortion which does not exist in single-mode fibres, namely intermodal dispersion. The energy of an impulse launched into a multimode fibre is therefore spread over

a time interval corresponding to the range of propagation delays of the modes. The number of signal pulses which may be transmitted in a given period, and hence the information-carrying capacity of the fibre, is therefore reduced. Since, in the absence of mode filtering or mode conversion, the pulse spreading increases linearly with fibre length, the bandwidth is inversely proportional to distance. The product of bandwidth B and distance L is therefore a figure of merit for the information capacity of an optical fibre. For example, Table 1 shows that the $B \times L$ product for a step-index fibre is typically 20 MHz·km. As indicated in the Introduction, a careful choice of the radial variation of the refractive index enables the transit-times of the modes to be almost equalized so that $B \times L$ products of 10—20 GHz·km have been predicted although not yet achieved in practice. The power-law, or α, class of refractive-index profiles, given by

$$n^2(r) = n_1^2\left[1 - 2\Delta\left(\frac{r}{a}\right)^{\alpha}\right] \qquad r < a$$
$$n^2(r) = n_1^2[1 - 2\Delta] = n_2^2 \qquad r > a \tag{16}$$

has been used extensively to model the grading function of multimode fibres. The profile is optimized by a suitable choice of α. It may be shown [25,26] that, neglecting the dispersive properties of the glasses forming the waveguide, the value of α which minimizes the r.m.s. pulse broadening is given by

$$\alpha_{opt} = 2 - \frac{12\Delta}{5} \tag{17}$$

and the r.m.s. output pulse width produced by a unit impulse at the input is then:[26]

$$\sigma_{opt} = \frac{L}{c}\frac{n_1}{20\sqrt{3}}\Delta^2 \tag{18}$$

The intermodal dispersion is, however, an extremely sensitive function of the index-profile. Minute departures of refractive index from the power law, or an incorrect design of the profile, lead to a much lower bandwidth than is theoretically achievable. For example, an error in α of $\sim 1\%$ degrades the bandwidth by a factor of two. The central dip caused by dopant evaporation in the high-temperature collapse stage of the CVD process, and the step-like structure caused by the deposition of individual glass layers, have been shown to contribute significantly to the pulse broadening. Nevertheless extremely high $B \times L$ products have been achieved using several variants of the CVD process and the best results reported in the literature are summarized in Table 2.

It may be seen that the best value is within a factor of 2 of the maximum predicted for α profiles, but whether fibres of this quality can be prepared under routine production conditions remains to be seen.

4.3 Effect of the Wavelength Dependence of Refractive Index

The variation of refractive index with wavelength also causes the transmitted pulses to broaden, as we have seen in the case of single-mode fibres. With multimode fibres an additional, more subtle, effect exists since the index dispersion $dn/d\lambda$ also alters the relative transit-times of the modes and hence, the intermodal dispersion.[26] This is normally referred to as 'profile dispersion' and is a result of the difference which exists between the group index $N = n - \lambda(dn/d\lambda)$ (which determines the pulse transit time) and the refractive index n.

Since $dn/d\lambda$ is, in general, a function of glass composition it varies across the core of a graded-index fibre. Hence each mode is affected differently by dispersion since the spatial distribution of power is not the same for all modes. For example, low-order modes travel, on average, in a medium of higher dopant concentration than do higher-order modes. Thus a correction to the optimum profile parameter is required to account for the difference between the functions $N_1(r)$ and $n_1(r)$. Olshansky and Keck[26] extended the analysis of power-law profiles to include the effect of glass dispersion and showed that in the presence of profile dispersion the optimum value of α is given by:

Table 2

Pulse dispersion and bandwidth of graded-index multimode fibres

			σ(ps·km^{-1})			$B \times L$ (GHz·km)	
Process	Date	Δ/%	Measured	Theory	$\sigma_{meas}/\sigma_{th}$	Measured	Theory
VAD	9/80	1	28	14	2	6·7	13·4
MCVD	3/80	1	50	14	3·6	3·7	13·4
PCVD†	1977	0·94	70	12·4	5·6	2·7	15
OVPO	1/78	0·76	60	8·1	7·5	3·1	23

† An inside vapour-phase oxidation process in which a microwave plasma is generated in the reacting gases.

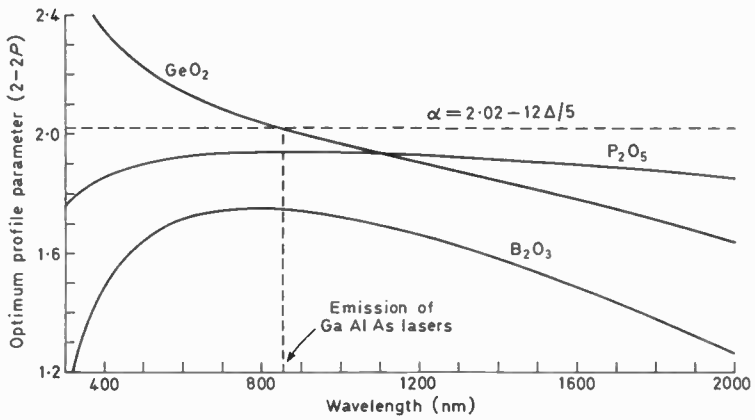


Fig. 7. Variation of the optimum profile parameter for germanosilicate, phosphosilicate and borosilicate glasses. The broken horizontal line represents the profile grading function $\alpha = 2{\cdot}02 - 12\Delta/5$ which, for germanosilicate glasses, is optimum for operation at 850 nm.

$$\alpha_{opt} = 2 - 2P - 12\Delta/5 \qquad (19)$$

where the profile dispersion parameter P is defined as

$$P(\lambda) = \frac{n_1}{N_1}\frac{\lambda}{\Delta}\frac{d\Delta}{d\lambda} \qquad (20)$$

Equation (19) is only valid if the grading function is constant, i.e. if α is independent of wavelength. P is then purely a property of the additive (e.g. germania) employed to modify the index of the base glass (e.g. silica) and does not depend on the dopant concentration. Unless this is the case, P varies across the core of a graded-index fibre and an α profile is not the optimum for equalizing the mode transit times.[27] P would thus be a meaningless parameter.

The most reliable measurements[28] of P for the germanosilicate system show no dependence on concentration. The results are shown in Fig. 7 in the form of curves of $(2-2P) = (\alpha_{opt} + 12\Delta/5)$ versus wavelength. The figure also includes similar results for the other two most important dopants in silicate fibres, namely phosphorus pentoxide and boric oxide.

It may be seen from Fig. 7 that α_{opt} is a strong function of wavelength, particularly for germania and boric oxide dopants. It follows that a profile optimized for one wavelength is, in general, inappropriate for operation at another source wavelength. Intermodal dispersion is therefore a function of wavelength. For example, a germania-doped fibre having a profile parameter of $(2{\cdot}02 - 12\Delta/5)$.(represented as a horizontal line in Fig. 7) would have least intermodal dispersion at the wavelength, 0·85 μm, where the curve of $2-2P$ falls to $(2{\cdot}02 - 12\Delta/5)$. With departure from $\lambda = 0{\cdot}85$ μm the profile is no longer optimal and the bandwidth thus decreases rapidly. Close control of the fabrication process is therefore required in order to ensure that the highest bandwidth occurs at the intended operating wavelength.

It is desirable to produce fibres having large bandwidths over a broad spectral range. This would extend the capacity of the fibre by allowing operation at several wavelengths simultaneously (wavelength-division multiplexing) and give the user the flexibility of altering the source wavelength if improved devices become available after the installation of a cable. Fibre designs employing more than one index-modifying additive have been proposed which either reduce the wavelength-dependence of bandwidth over a particular spectral range, or optimize the intermodal dispersion for several different operating wavelengths.

4.4 Material Dispersion in Multimode Fibres

The power carried by multimode fibres travels almost entirely in the core region. Because, in addition, most modes are operated far from cut-off they are almost free of waveguide dispersion. The pulse delay in multimode fibres is thus given, to first order, by[8]

$$\tau = \frac{L}{c}N_1 = \frac{L}{c}\left(n_1 - \lambda\frac{dn_1}{d\lambda}\right) \qquad (21)$$

where N_1 is the group index of the core material. (For graded-index fibres, N_1 represents a value of group index averaged over the core area.)

Semiconductor sources used in optical communications systems radiate over a finite range of wavelengths and, from equation (21), each spectral component travels at a different group velocity. The resulting pulse broadening σ_m is known as material dispersion.

For a source of r.m.s. spectral width σ_s and mean wavelength λ_s, σ_m may be evaluated by expanding equation (21) in a Taylor series about λ_s:

$$\sigma_m = \sigma_s\frac{d\tau}{d\lambda} + \frac{\sigma_s^2}{2!}\frac{d^2\tau}{d\lambda^2} + \ldots \qquad (22)$$

The first term normally dominates, particularly for sources operating in the 0·8 to 0·9 μm wavelength region. Thus for the material illustrated in Fig. 4(b) the material dispersion parameter $M = \frac{1}{L}\frac{d\tau}{d\lambda}$ is $\sim 100\text{ps nm}^{-1}\text{ km}^{-1}$ at 0·85 μm. For a typical light-emitting diode having

$\sigma_s = 18$ nm and $\lambda_s = 850$ nm the resulting pulse broadening is 1·8 ns·km^{-1} which limits the $B \times L$ product to 100 MHz·km. This level of dispersion is an order of magnitude greater than the intermodal dispersion in the best available fibres. Even with semiconductor lasers (having spectral widths of, typically, 1 nm r.m.s.), material dispersion sets an ultimate limit on the capacity of optical fibre systems.

Figure 4(b) shows that a wavelength region exists where the material dispersion parameter is negligible. The wavelength λ_m of zero material dispersion is found to vary according to the glass composition[29] but, for silica-based fibres, is always in the vicinity of 1·3 μm. Operation in this wavelength region substantially reduces the bandwidth limitations arising from material dispersion and very high $B \times L$ products are available, even with light-emitting diodes. It may be seen from Fig. 3 that, for silica-based glasses, the fibre attenuation is also extremely low (0·38 dB·km^{-1}).

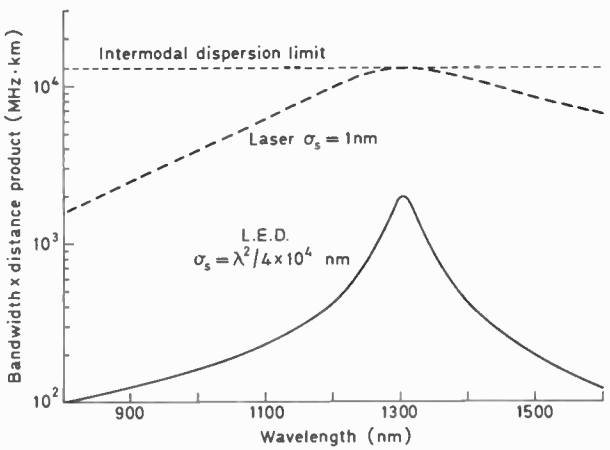


Fig. 8. Bandwidth limitation resulting from material dispersion in a typical $GeO_2-P_2O_5-SiO_2$ multimode fibre (see Fig. 4(b) for dispersion data). The dashed curve applies to a semiconductor laser source having a spectral width of 1 nm r.m.s. The solid line is for a typical l.e.d. having a spectral width given by $\lambda^2/4\times10^4$ nm. For a numerical aperture of 0·2 ($\Delta \simeq 1\%$), residual intermodal dispersion alone limits the $B \times L$ product to the value shown by the dotted line.

As shown in Fig. 8, the bandwidth available from multimode fibres increases rapidly as the wavelength of operation is increased from 0·85 μm (emission of GaAlAs devices) to the region of negligible material dispersion ($\sim$ 1·3 μm). Thus for a laser having a spectral width of 1 nm r.m.s. (dashed curve), the information-carrying capacity is limited in the region of 1·3 μm, by residual intermodal dispersion. For a numerical aperture of 0·2 the maximum available bandwidth is $\sim$ 13 GHz·km, providing the refractive-index profile is correctly designed. A large proportion of future multimode systems will therefore be operated at 1·3 μm.

Material dispersion does not completely vanish, however, even at the wavelength where the first derivative of group delay $d\tau/d\lambda$ is zero. According to equation (22) higher-order terms must then be taken into account.[30] The pulse broadening resulting from second-order material dispersion is proportional to the square of the source linewidth, and for silica-based fibres is of the order of 0·1 ps·nm^{-2} km^{-1}. With laser sources second-order material dispersion is not a serious limitation.

The linewidth of light-emitting diodes increases as the square of the operation wavelength[31] and, at 1·3 μm, values of 100 nm f.w.h.m. (42 nm r.m.s. spectral width) have been reported. With such broad spectral widths, second-order material dispersion limits the bandwidth[32] to about 2 GHz·km^{-1}, see Fig. 8, solid curve. Although the bandwidth of the best multimode fibres can, in principle, be limited in this way, in practice a counteractive effect, namely wavelength filtering,[31,33] takes place. The latter effect results from even minor variations in the loss/wavelength characteristic of the fibre, which are enhanced by transmission over long distances (e.g. 20–30 km for typical 1·3 μm systems). Thus, the effective wavelength spread is dictated[31,33] more by the attenuation characteristics of the fibre than by the spectral width of the source. In addition the wavelength at which the received power is a maximum may not coincide with the peak source wavelength.

4.5 Sources for Multimode Fibre Systems

Semiconductor sources developed for optical communications are of two types, namely injection lasers and light-emitting diodes. Semiconductor lasers allow 1mW or more to be launched into multimode fibres. With surface-emitting l.e.d.s launched power levels of a few hundred microwatts have been achieved. With edge-emitting, high-radiance, l.e.d.s, up to 1mW of power can be launched into typical multimode fibres. Unfortunately l.e.d.s having large output power usually have a reduced modulation bandwidth. Semiconductor lasers offer better performance than l.e.d.s; their cost is, however, much greater than for l.e.d.s and their reliability is still not sufficient for some systems applications. Both types of device are therefore used in the systems already installed, or planned, for the near future.

5 Conclusions

In the short space of 15 years optical fibre transmission lines have been transformed from a tentative research idea to practical realization. Optical fibre cables are already in operation in the trunk telephone network and a massive extension of their use is planned which includes submarine cable communication. The attenuation attainable is nearly two orders of magnitude lower than in coaxial cable and the bandwidth of single-mode fibres is four orders of magnitude greater.

A variety of fibre types are available commercially, ranging from single-mode fibres having a core diameter of a few micrometres to step-index, plastic-cladded,

multimode fibres approaching 1 mm diameter. Applications to short-distance links such as in buildings, aircraft and between computers are being studied, as well as in instrumentation for current measurement at high voltages, in fibre gyroscopes and in hazardous situations such as fuel tanks and other explosive atmospheres where metal conductors cannot be used. The extension of 'electronic' techniques to optical wavelengths is well under way.

6 References

1 Midwinter, J. E., 'Optical Fibres for Transmission' (Wiley, New York, 1979).
2 Payne, D. N. and Gambling, W. A., 'New silica-based low-loss optical fibres', *Electronics Letters*, **10**, p. 289, 1974.
3 French, W. G., MacChesney, J. B., O'Connor, P. B. and Tasker, G. W., 'Optical waveguides with very low losses', *Bell Syst. Tech. J.*, **53**, pp. 951–4, 1974.
4 Sandbank, C. P. (Ed.) 'Optical Fibre Communication Systems' (Wiley, Chichester, 1980).
5 Sudo, S., Kawachi, M., Edahiro, T., Izawa, T., Shioda, T. and Gotuh, H., 'Low-OH-content optical fibre fabricated by vapour-phase axial-deposition method', *Electronics Letters*, **14**, pp. 534–5, 1978.
6 Miya, T., Terunuma, Y., Hosaka, T. and Miyashita, T., 'Ultimate low-loss single-mode fibre at 1·55 μm', *Electronics Letters*, **15**, pp. 106–8, 1979.
7 Snyder, A. W., 'Asymptotic expressions for eigenfunctions and eigenvalues of a dielectric optical waveguide', *IEEE Trans. on Microwave Theory & Techniques*, **MTT-17**, pp. 1130–8, 1969.
8 Gloge, D. 'Weakly guiding fibres', *Applied Optics*, **10**, pp. 2252–8, 1971.
9 Payne, D. N. and Gambling, W. A., 'Zero material dispersion in optical fibres', *Electronics Letters*, **11**, pp. 176–8, 1975.
10 Gambling, W. A., Matsumura, H. and Ragdale, C. M., 'Mode dispersion, material dispersion and profile dispersion in graded-index single-mode fibres', *IEE J. Microwaves, Optics & Acoustics*, **3**, pp. 239–46, 1979.
11 Adams, M. J., Payne, D. N. and Ragdale, C. M., 'Birefringence in optical fibres with elliptical cross section', *Electronics Letters*, **15**, pp. 298–9, 1979.
12 Love, J. D., Sammut, R. A. and Snyder, A. W., 'Birefringence in elliptically deformed optical fibres', *Electronics Letters*, **15**, pp. 615–6, 1979.
13 Norman, S. R., Payne, D. N., Adams, M. J. and Smith, A. M., 'Fabrication of single-mode fibres exhibiting extremely low polarization birefringence', *Electronics Letters*, **15**, pp. 309–11, 1979.
14 Gambling, W. A. and Matsumura, H., 'Simple characterisation factor for practical single-mode fibres', *Electronics Letters*, **13**, pp. 691–3, 1977.
15 Snyder, A. W., 'Excitation and scattering of modes on a dielectric or optical fiber', *IEEE Trans. on Microwave Theory & Techniques*, **MTT-17**, pp. 1138–44, 1969.
16 Gambling, W. A., Matsumura, H. and Ragdale, C. M., 'Joint loss in single-mode fibres', *Electronics Letters*, **14**, pp. 491–3, 1978.
17 Petermann, K. 'Fundamental mode microbending loss in graded-index and W fibres', *Optical & Quantum Electronics*, **9**, pp. 167–79, 1977.
18 Miyagi, M. and Yip, G. L., 'Mode conversion and radiation losses in a step-index optical fibre due to bending', *Optical & Quantum Electronics*, **9**, pp. 51–60, 1977.
19 Gambling, W. A., Matsumura, H., Ragdale, C. M. and Sammut, R. A., 'Measurement of radiation loss in curved single-mode fibres', *IEE J. Microwaves, Optics & Acoustics*, **2**, pp. 134–40, 1978.
20 Gambling, W. A., Matsumura, H. and Ragdale, C. M., 'Curvature and microbending losses in single-mode optical fibres', *Optical & Quantum Electronics*, **11**, pp. 43–59, 1979.
21 Sasaki, I., Payne, D. N., Mansfield, R. J. and Adams, M. J., 'Variation of refractive index profiles in single-mode fibre preforms measured using an improved high-resolution spatial-filtering technique', 6th European Conference on Optical Communication, York 1980.
22 Snyder, A. W. and Sammut, R. A., 'Fundamental (HE_{11}) modes of graded optical fibres', *J. Opt. Soc. Am.*, **69**, pp. 1663–71, December 1979.
23 Stewart, W. J., 'Simplified parameter based analysis of single-mode optical fibres', *Electronics Letters*, **16**, pp. 380–2, 1980.
24 Matsumura, H., Suganuma, T. and Katsuyama, T., 'Simple normalization of single-mode fibres with arbitrary index profile', 6th European Conference on Optical Communication, York 1980.
25 Gloge, D. and Marcatili, E. A. J., 'Multimode theory of graded-core fibres', *Bell Syst. Tech. J.*, **52**, pp. 1563–78, 1973.
26 Olshansky, R. and Keck, D. B., 'Pulse broadening in graded-index optical fibres', *Applied Optics*, **15**, pp. 483–91, 1976.
27 Arnold, J. A., 'Optimum profiles for dispersive multimode fibres' *Optical & Quantum Electronics*, **9**, pp. 111–9, 1977.
28 Sladen, F. M. E., Payne, D. N. and Adams, M. J., 'Profile dispersion measurements for optical fibres over the wavelength range 350 nm to 1900 nm'. Proceedings of 4th European Conference on Optical Communication, pp. 48–57, Genoa 1978.
29 Payne, D. N. and Hartog, A. H., 'Determination of the wavelength of zero material dispersion in optical fibres by pulse-delay measurements', *Electronics Letters*, **13**, pp. 627–9, 1977.
30 Kapron, F. P., 'Maximum information capacity of fibre-optic waveguides', *Electronics Letters*, **13**, pp. 96–7, 1977.
31 Gloge, D., Ogawa, K. and Cohen, L. G., 'Baseband characteristics of long-wavelength led systems', *Electronics Letters*, **16**, pp. 366–7, 1980.
32 Adams, M. J., Payne, D. N., Sladen, F. M. E. and Hartog, A. H., 'Optimum operating wavelength for chromatic equalisation in multimode optical fibres', *Electronics Letters*, **14**, pp. 64–6, 1978.
33 Stewart, W. J., 'Wavelength filtering effects in multimode fibres', Proceedings of Optical Communication Conference, Paper 12.3, Amsterdam 1979.

Manuscript received by the Institution on 20th February 1981.
(Paper No. 1994/Comm 221)

Some Recent Optical Fibre Developments—1

Cable Television and Radio over Optical Fibres

To demonstrate the suitability and advantages of fibre optics over conventional wire systems for carrying cable colour television and stereo radio signals, twenty optical fibre systems have been supplied by Standard Telecommunication Laboratories (STL) to British Telecom's Martlesham Heath laboratories.

These broadband optical fibre transmission systems comprise two sections: the trunk transmission system between the receiving aerial and a distribution switch, and a number of links from the distribution switch to subscribers. The trunk equipment consists of five similar single fibre systems. Using pulse frequency modulation four of these each carry one colour television channel, including television sound, while the fifth carries three v.h.f. stereo radio channels.

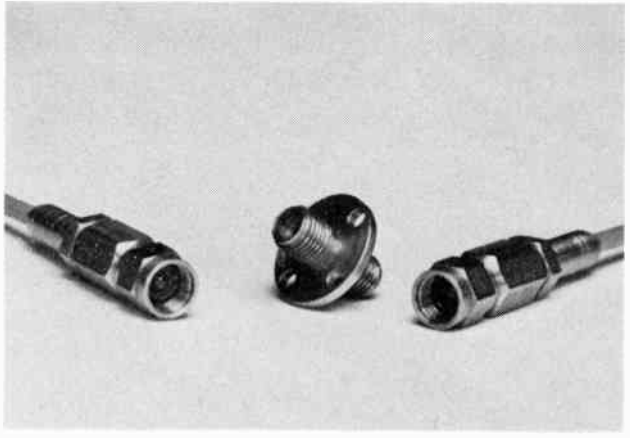
For the purposes of the demonstration the number of subscribers served has been limited to seven. Each subscriber link consists of two fibres and can simultaneously carry two television channels or six radio channels. A pair of ordinary metallic wires allows each subscriber to select the desired programme by controlling the distribution switch.

This type of transmission equipment is very similar to that which can be used in an interactive broadband network. British Telecom is planning such a system for a trial two-way visual service which will link 50 subscribers in three UK cities in 1982.

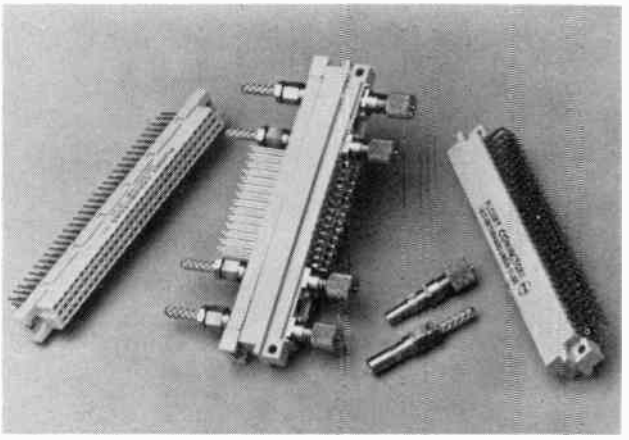
Optical fibre for the Martlesham demonstration system was supplied by STL's parent company, Standard Telephones and Cables. British Telecom provided the video distribution switch. Transmit/receive modules are now being manufactured by ITT Components — part of STC — in Leeds. Systems will be marketed by STC transmission products division from Basildon, Essex.

Connectors for Fibre Optics

Plessey Connectors of Northampton have recently launched a lens type fibre optic connector and a prototype fibre optic DIN 41612 insert. The advantages of the lens type connector are most apparent when very reliable connectors are required to work in extreme environmental conditions. In the lens type connector all the precision parts are contained behind the hard sapphire sphere lens within the connector body and are not exposed to the environment. The lens type connector is robust, easy to clean, repeatable and reliable, and for systems requiring frequent use in adverse conditions is more cost effective.



The lens type fibre optic connector developed by Plessey Connectors



Centre: A mated new version of the Series 16 with fibre optic DIN 41612 insert
Left and right: Plessey Connectors new version of the Series 16

Based on the experience of the Plessey ferrule connector type FC701, ferrule inserts are being developed to fit existing connector bodies. These inserts will enable low loss connection of optical fibres in multiway or mixed optical/copper combinations. This type of connector insert is being used with the Series 16 range of connectors which was introduced to meet the market requirement for reliable cost effective indirect interconnections between printed circuit boards or between associated printed circuit boards and backplanes.

Optical Fibre Fault Locator and Attenuation Measuring Set

Cossor Electronics has recently introduced a new optical fibre fault locator. Using time domain reflectometry techniques, the instrument can be used for the detection of faults, joints and other imperfections in optical fibres and cables. It complements Cossor's range of conventional cable fault locators and, like them, it is compact (measuring $140 \times 260 \times 290$ mm), lightweight (7·5 kg) and fully self-contained, with an internal battery pack which gives in excess of eight hours continuous use.

The instrument has a screen displaying a complete view of the cable under test, which gives an indication of the type of fault. Areas of particular interest on the display can be expanded for greater detail. Attenuation assessment is facilitated by etched reference lines on the screen. In addition, a direct digital readout, in metres, of distance to a fault is provided for accurate measurement. The maximum measurement range is 10 km and fibre core diameters from 50 μm upwards can be accommodated.

In conjunction with the optical fibre fault locater, Cossor has also introduced an optical fibre attenuation set. It is a two-part, field-portable, battery powered, instrument, comprising a transmitter and receiver for measuring optical path losses up to 60 dB. The transmitter and receiver can be carried in detachable halves of a briefcase-style container, which also houses the accessories.

UDC 621.315.2: 621.391.63

Indexing Terms: Cables, Optical fibres

Optical fibre cables

M. H. REEVE*

SUMMARY

Optical fibre cables serve to protect the fibre sufficiently to allow installation of the cable without impairment of the transmission characteristics and to ensure fibre survival over the life of the cable. The degree of protection required is examined and some possible ways it may be achieved in practice discussed.

* *British Telecommunications Research Laboratories, Martlesham Heath, Ipswich IP5 7RE*

1 Introduction

The low loss and high bandwidth of current optical fibres has led to an attractive transmission medium for the telecommunication network, as an alternative to coaxial cable or wire pairs. The reinforcement and protection of these fibres to the point where they can safely be drawn into underground ducts (or other environments), without impairment to their transmission characteristics, is the job of the optical-fibre cable maker. This paper attempts to give some idea of the problems that can arise in this process and some idea of the variety of ways these problems have been tackled. It is proposed to concentrate mainly on underground duct cable although many of the general principles will apply to the design of aerial and internal cables.

2 Functions of the Optical Cable

The various functions of the cable may be detailed as follows.

2.1 Fibre Survival

One of the main jobs of the cable is to protect against fibre breakage under the various handling and installation conditions it meets, and, after installation, over the life of the cable. To do this we will need to know something about the strength characteristics of the fibre, in particular the sort of strain levels a fibre could be expected to withstand without breakage.

2.2 Maintenance of Transmission Characteristics

A fibre's optical loss can be drastically increased by poor coating and cabling techniques. To design a good cable we need to know something of the effects involved and how these can be minimized. The effect on the bandwidth of the cabling technique, whilst usually less significant, may also need to be considered.

2.3 Identification and Jointing Aspects

Cables containing more than one fibre will need some means of separately identifying each fibre. A cable containing a large number of fibres may need to separate groups of fibres into a geometrical arrangement suitable for multiple jointing to avoid the tedium of individually jointing each fibre between cable sections. A cable with the fibres arranged in ribbons is an example.[18]

3 Fibre Strength Considerations

We will now consider the fibre strength constraints mentioned in Section 2.1.

3.1 Griffith's Theory

Telecommunications optical fibres are invariably silica-based or of a compound glass.[1,2] These materials are brittle, being almost perfectly elastic right up to the breaking point. The bulk material strength is intrinsically very high but can be drastically lowered by the presence of surface flaws. This process is usually treated by the Griffith theory,[3] which postulates that

0033–7722/81/070327+06 $1.50/0

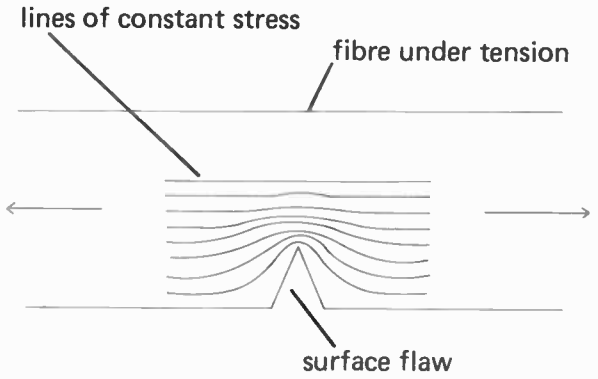


Fig. 1. Lines of constant stress in a fibre under tension 'bunch' at a surface flaw, showing concentration of stress at this point.

stress concentration occurs at the tips of surface cracks, leading to catastrophic crack growth and failure. This process is illustrated in Fig. 1, which shows stress lines concentrating at the crack tip. The deeper the crack the higher is the stress at its tip.

Griffith defined a stress intensity factor K_I, as

$$K_I = \sigma\, YC^{\frac{1}{2}} \tag{1}$$

where σ is the macroscopic stress on the fibre, Y is a constant due to the shape of the crack ($=\pi^{\frac{1}{2}}$ for elliptical cracks), and C is the crack depth. At some critical stress intensity factor, K_{IC}, fracture occurs. Griffith derived an expression for this:

$$K_{IC} = (2E\gamma_p)^{\frac{1}{2}} \tag{2}$$

where E is Young's modulus of the material and γ_p is the energy required to create two new surfaces. Combining equations (1) and (2) leads to a critical crack depth for failure at any particular stress σ on the fibre. This is shown in equation (3):

$$C = \frac{2E\gamma_p}{Y^2\sigma^2} \tag{3}$$

Using values for silica this equation predicts failure at 1% strain for a flaw in the region of only 0·5 μm deep.

We can now see that the fibre surface must be protected from abrasion to a very high degree to ensure high strength in the fibre. For this reason a primary protective coating is usually applied to the fibre on-line, as soon after the initial production of the fibre as possible and before contact with any pulleys or guides.[4]

3.2 Stress Corrosion

Before the Griffith criterion for failure given in equation (3) is reached, an additional effect due to the slow growth of flaws under the action of stress and water vapour must usually be considered—this is called stress corrosion. Stress corrosion is a mechanism whereby molecular bonds at a crack tip are preferentially attacked by water because they are under stress. This leads to a slow growth of the flaw under the imposed stress and eventual failure. Thus a fibre that survives the short-term stresses of cable making and installation may not survive lower residual stresses that are present for the life of the system. When dynamic stresses are applied to the fibre, this same process leads to a dependence of the strength of the fibre on the strain rate imposed. Flaws in a fibre that is strained slowly will have more time to grow to the Griffith limit than those that are strained more quickly, giving lower apparent strengths for the slower strain rates.[5]

Stress corrosion is commonly modelled using an empirical relationship between crack velocity and applied stress intensity factor as follows:

$$V = AK_I^n \tag{4}$$

where V is the crack velocity and A and n are constants for the particular fibre material. n is called the stress corrosion susceptibility.

Mathematical manipulation of equation (4) yields an expression for the time-to-failure of a fibre flaw under stress. For the particular case of a constant (static) stress this becomes:

$$\log t_f = n \log \sigma + \log B \tag{5}$$

where B is a constant and t_f is the time to failure.

This is a straight line on a log/log plot of time-to-failure against stress or strain. An example is shown in Fig. 2. If the value of n for the particular fibre type is known (values are usually in the range 15–25), such a plot can form the basis for the prediction of fibre lifetime under service strains.[5]

3.3 Flaw Statistics

All of the above analysis, and in particular the use of Fig. 2 to predict the life of a fibre under a particular strain, of course refers to only one flaw. In practice fibres will have many flaws, the largest being the most significant and causing fibre fracture. The strength of any length of fibre now becomes an unknown dependent on the size of the largest flaw in its surface. This also leads to a length dependency of the strength, a longer length of fibre being more likely to contain a large flaw than a shorter length.

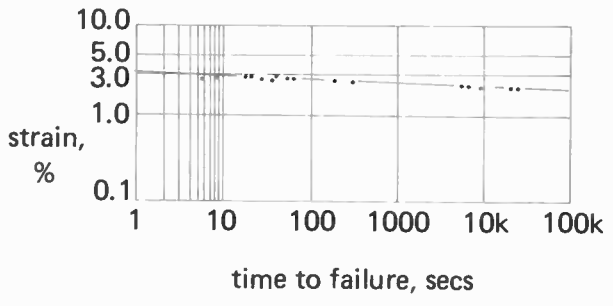


Fig. 2. Fibre lifetime under static stress. The points show the time to failure of 50 cm lengths of fibre under a range of static strains. Straight line behaviour on the log/log plot is obtained.

It is apparent that we must resort to statistical methods to proceed further. The approach usually used is that due to Weibull,[6] who evolved a theory to describe the strength behaviour of any system controlled by its weakest link. The basic relationship established empirically by Weibull is of the form of equation (6):

$$F = 1 - \exp\left[-\left(\frac{\sigma}{\sigma_0}\right)^m \left(\frac{L}{L_0}\right)\right] \qquad (6)$$

F is the probability of failure at stress, σ_m is the Weibull distribution shape parameter, σ_0 is a scale parameter and L_0 is a constant with the dimensions of length.

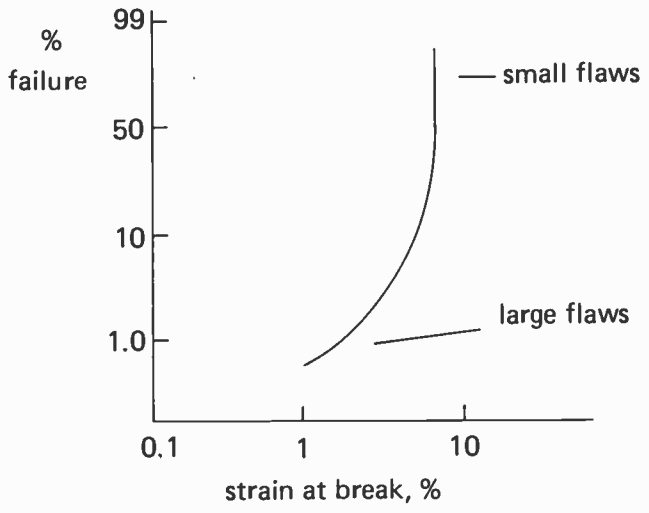


Fig. 3. Schematic Weibull plot. Fibre breaking strains are plotted against the cumulative probability of their occurrence. The majority of samples break at high strain due to numerous small surface flaws. A small number of samples contain larger flaws, leading to breakage at lower strain and a consequent 'tail' on the plot.

Data are conveniently represented in the Weibull form using plots of the type shown in Fig. 3. In this case the shape parameter becomes the slope of the curve and the scale parameter the intercept on the strain axis. The plot is typically established by breaking a large number of 10–20 metre lengths of fibre to measure the strain at break, plotting the resulting strains against the cumulative probability of their occurrence. Plots of this type usually show two distinct regions, a fairly steep region at high strains corresponding to the very numerous small flaws and a much shallower region at low strains corresponding to the much smaller number of larger flaws (probably due to external damage to the fibre). Obviously the low strength region is of most interest in determining the life of the fibre under stress.[7]

The Weibull plot can only apply to one particular breakage condition and to this must be added the additional problem of stress corrosion.

3.4 Proof Testing and Lifetime Prediction

Having emerged with this rather complex picture of a fibre under stress with many different size flaws, all subject to growth through stress corrosion, we can now begin to bring this together into something of use to the cable maker. Some confidence is required that the fibre will survive the stress conditions it meets during installation and over life. This can be approached through a combination of proof testing and use of stress corrosion data. In simple terms proof testing involves straining all the fibre used up to a certain screening level (for instance a 1% strain may be chosen).[8] Fibre which survives this screening process is accepted, that which fails is rejected. Curves such as Fig. 2, obtained for the same fibre type, can then be used to derate this proof-test level to a safe strain for longer times to take account of stress corrosion.

Unfortunately, we cannot forget that crack growth also occurs during the proof test itself, giving a finite probability that a crack could grow sufficiently for the fibre to fail the proof test if subjected to it a second time. A full treatment of this is given in Ref. 9.

The final outcome is that even with a proof test we can only give a *probability* that the fibre will survive in any situation. To add to this we have the doubt that different flaw sizes may grow at different rates[10] and that stress corrosion data obtained in short term tests may not extrapolate well over the years of life required.[11] It can be seen that the fibre strength area is one well worthy of further study.

3.5 Fibre Strength—Conclusion

As a conclusion to this Section on fibre strength it may be worth putting some figures to the sort of strength levels one can expect from silica fibre and how these must be derated to ensure survival in the long term.

In the absence of the stress corrosion mechanism (liquid nitrogen temperatures slow the reaction down), short length measurements have shown in the region of 20% extensions at break,[12] giving at least some guide as to the intrinsic strength of the material. In normal laboratory humidities the stress corrosion process reduces this figure to around 8% extension at break. Unfortunately these very high strengths are difficult to maintain over the kilometre lengths required in the cables. In fact a 1% proof test (100,000 p.s.i. – 7030 kgf/cm^2) at reasonable yield is fairly difficult to achieve in a production process but is now being attained by some manufacturers on kilometre lengths. The proof test theory and stress corrosion data presently available show that this level will need to be derated to about 0·3% for a reasonable probability of surviving a strain existing for 20 years. The cable design must be capable of limiting the fibre strain to these levels under all conditions of use.

4 Maintenance of Transmission Characteristics

In parallel with the need to ensure fibre survival the cable must not increase the optical loss or reduce the bandwidth of the fibre. The transmission characteristics of the fibre can be altered by bending of the fibre axis causing coupling of power between the modes of a multimode fibre and radiation loss in both multimode and monomode fibres. Random bending of the fibre axis on a microscopic scale is called *microbending* and can easily result in unacceptable loss being added to the fibre (Fig. 4). Such bends can result from fibre coating variations or other constraints of the cable and must be minimized by the cable design.[13]

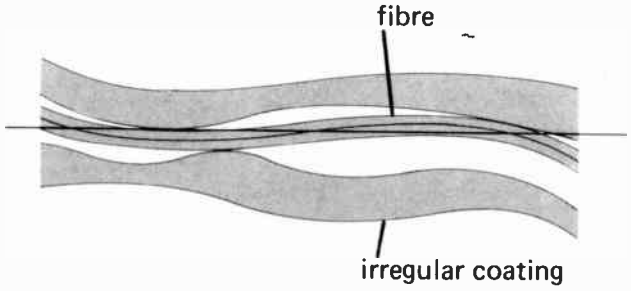


Fig. 4. Fibre microbending caused by an irregular coating. These minute bends couple light out of the fibre, leading to large optical losses over long distances.

It is found that the mode coupling caused by microbending deviations is particularly sensitive to certain spatial frequencies in the disturbance, leading to a range of microbending pitch lengths which must be avoided. The analysis of Ref. 14 shows that these spatial frequencies are in the range of 1–2 mm or so.

Much longer pitch bends or macrobends can also cause transmission loss but in general need to be radii of a few centimetres or less before the effect is significant. This may need to be borne in mind when choosing the fibre stranding lay in the cable.

In addition to adding minimal loss to the fibre as the cable is made, the various components of the cable must maintain this situation under all the temperature and strain conditions the cable is likely to meet.

Two general approaches have been taken in optical cable design to solve the problem of microbending loss. Accepting that some on-line primary coating is necessary to prevent fibre abrasion during handling, these two approaches revolve around the next coating applied to the fibre—the secondary coating. They may be roughly described as the 'loose' or 'tight' packaging methods. The first relies on mechanically isolating the fibre from its surroundings by encasing it in a loose tube or cavity in the cable. The second relies on 'stiffening' the fibre against outside microbending with a thick buffer coating applied directly on to the primary coating (up to 1 mm of nylon applied by extrusion has been used). In this case the secondary coating must itself be applied so that it does not cause microbending loss.

The main differences between the two approaches emerge on consideration of the large thermal expansion differences between the fibre and the polymeric materials used for the secondary coating and the effects this difference may have.

Considering first the loose tube method, we can see that if the tube is cooled, and contracts more than the fibre, the fibre 'buckles' inside the tube.[15] If this is not to lead to microbending loss the tube must be made large enough to absorb the differential contraction and only give a mild buckling effect. In Ref. 16 this problem is tackled by orienting the molecules of the tube to reduce its thermal expansion coefficient.

In the case of the tightly extruded coating the 'fit' of the coating and its inner surface must be good enough that this buckling process is not allowed to happen, the fibre being held straight at all times.

One of the above philosophies has been used in most cables to date, although in some cases the fibres are not individually coated, the secondary coating forming part of the main cable structure.

5 Cable Design

Having considered the fibre constraints of Section 3.5 and Section 4 we can now look at how these might be met in a practical cable design. In arriving at any particular design we must always consider the twin aspects of protecting the fibre from strain (both in tension and bending of the cable) and added loss.

5.1 Components

We first consider the components necessary to any cable design.

Fibre and primary coating. The primary coating merely prevents abrasion of the fibre surface during production of the cable.

Secondary coating. Protects the fibre from microbending induced loss. May be a tight coating or a loose tube configuration. Often individually applied to the fibre but may be part of the cable structure as a whole.

Strength member. Some component is usually added to the cable to take the loads involved in installation and limit the cable strain to a value low enough to protect the fibre. Stranded steel wires or Kevlar (DuPont Ltd.) yarns are commonly used.

Sheath. A fairly substantial outer sheath is usually added to the cable to provide abrasion and crush resistance.

Water barrier. Since most polymer sheaths are rather poor at preventing water ingress to the cable an additional barrier is usually added. In the British Telecommunication network this takes the form of an

axially laid aluminium foil/polythene film laminate just inside the outer sheath.[17] The foil is overlapped and the welding of the polyethylene film at the overlap by the heat of the sheath extrusion leaves only a very thin path for water ingress.

Tape wraps. Paper or mylar tapes are commonly used to keep the components of the cable in register between manufacturing stages.

5.2 Design Stages

We can now look at how the above components might be made up into a practical cable. Figure 5 shows a design for an eight-fibre cable using a loose tube secondary coating. The secondary coated fibres are helically laid around a steel wire strength member. This

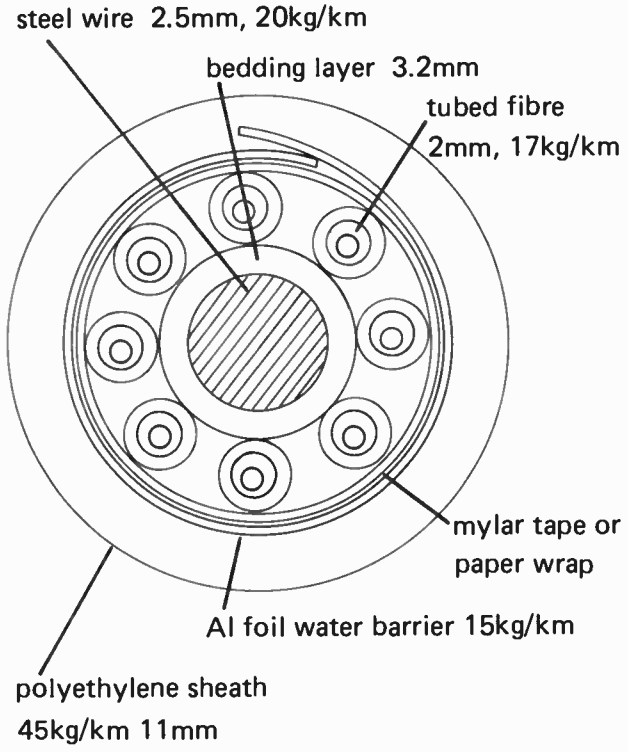


Fig. 5. A possible 8-fibre optical cable design.

cable would be suitable for a junction or trunk transmission route within the British network. We will look at the stages in the design of this cable and then compare it with other designs under production.

1. Decide on the number of fibres. Current systems use one fibre for each direction of transmission.

2. Choose the secondary coating type and size. In Fig. 5 a loose tube approach has been chosen. The internal bore of the tube will be determined by factors such as the temperature behaviour as mentioned in Section 4. It is convenient to choose 1 mm.

3. Determine the water barrier and sheath thicknesses. These are necessary for a calculation of the total weight of the cable. As these are common components in wire-pair type cables we can draw on experience with these and use a standard water barrier of 150 μm thickness, with a sheath thickness of 1·5 mm.

4. Calculate the maximum cable load. In a straight duct installation this would come from the duct coefficient of friction and the weight of the maximum installation length, the modulus of the cable and the maximum strain that could be allowed. Unfortunately very few ducts are at all straight but usually some correlation between cable weight and pulling-in load remains. To aim for an installation length of about 1 km at a load equal to the weight of 1 km of the cable seems the most appropriate at the moment in this country recognizing that the installation length must be reduced for difficult routes. From Section 3 it can be seen that the fibre strain in the cable must be kept below around 0·25% to be reasonably safe.

5. A summation of the weight/kilometre of each component of the cable now allows the total weight to be calculated. If it is assumed that the only component holding load is the strength member the size of this to keep the strain below the set limit can be calculated. A bedding layer may be needed around the steel to increase its size sufficiently to fit the eight fibres around it, as shown in the diagram.

6. It remains to set the helical pitch of the fibre packages around the strength member and the design is complete. This must be short enough to include at least one pitch length in the smallest bend likely to be encountered but not so small as to give fibre bending radius problems. Between 20 and 50 cm would be appropriate for the cable of Fig. 5.

The resulting sizes of the component parts of the cable are shown in Fig. 5. The design of a cable using tightly coated fibres would follow the same lines as that shown.

5.3 Other Cable Designs

Other cable structures are shown in Fig. 6 and whilst they may be quite different in geometrical layout, the principles to be followed will be similar to those already described.

Of particular interest because of their differences to the cable of Fig. 5 are those of Fig. 6(b) and Fig. 6(c). Figure 6(b) is a cable in which the primary coated fibres are laid in helical grooves in a solid rod of polymer, the grooves effectively forming a loose secondary coating. Figure 6(c) is a cable containing a very large number of fibres (144 max.) arranged in a stack of 12 ribbons, each of 12 fibres.[18] The aim of the cable is to allow multiple jointing of fibres at the end of a cable section, each ribbon being jointed as a single unit. The ribbons are formed as a sandwich of the 12 fibres between two tapes.

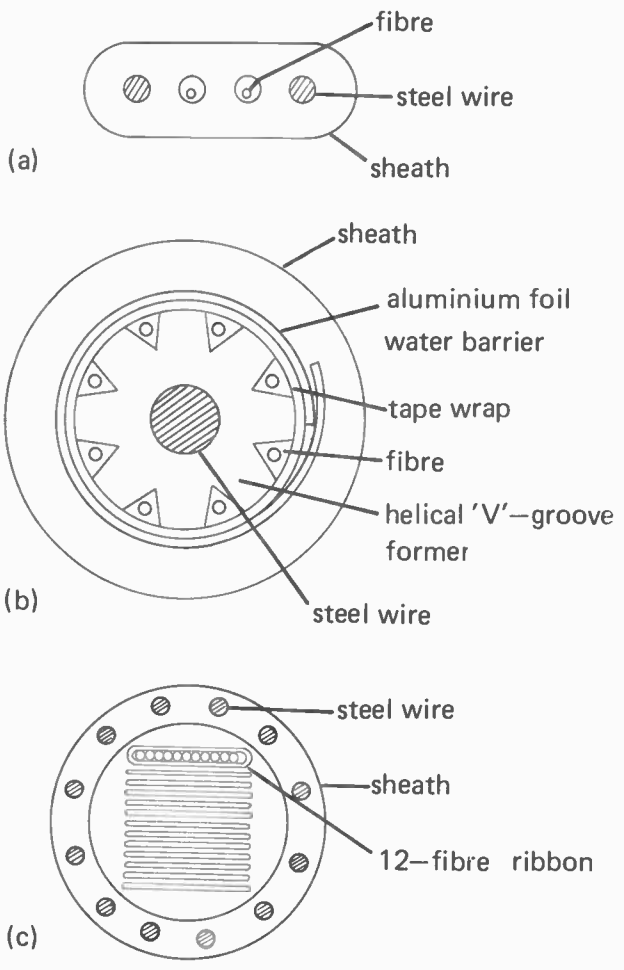


Fig. 6. Some alternative cable configurations: (a) BICC flat cable, (b) LTT helical-grooved-core cable, (c) Bell Telephone ribbon cable. Each of the stack of 12 ribbons holds a possible 12 fibres. The ribbons are twisted to aid bending properties.

6 Conclusions

The optical fibre cable designer must take as his starting point the properties of the fibre itself, in both its mechanical characteristics and its susceptibility to bending induced loss increases. Once these constraints are known the design proceeds through the buffering of the fibre against loss increase, the addition of a strength member of high tensile modulus, to the addition of a final sheath for crush and abrasion protection. A wide range of geometrically different solutions is possible in meeting these aims. A variety of trials involving the laying of optical fibre cables in telecommunications ducts have taken place, worldwide, since 1977.[19,20] Many of these trials have seen continuous operation since installation and none, to our knowledge, has failed as a result of any of the potential degradation mechanisms discussed here. In the UK a number of proprietary optical line systems are currently being installed: the first of these in the Birmingham area has been carrying a full traffic load as an integrated part of the British telecommunications network since last year. From this evidence it may be that current commercial optical fibre cables have been 'over engineered' but they are certainly working very successfully. Over the next decade an evolution and simplification of optical cable designs will occur as optical systems penetrate extensively into traditional telecommunications networks.

7 Acknowledgments

The author wishes to thank the Director of Research of British Telecommunications for permission to publish this paper. The views expressed are his own.

8 References

1 Ainslie, B. J. *et al.*, 'Preparation of long lengths of ultra-low-loss single-mode fibre', *Electronics Letters*, **15**, no. 14, pp. 411–2, 5th July 1979.
2 Beales, K. J. *et al.*, 'Low loss graded-index fibre by the double-crucible technique', 5th European Conference on Optical Communication, Amsterdam 1979, Paper 3.2–1.
3 Griffith, A. A., 'Phenomena of rupture and flow in solids', *Phil. Trans. R. Soc., Lond.*, Ser. A, **221** pp. 163–98, 1920.
4 France, P. W., Dunn, P. L. and Reeve, M. H., 'Plastic coating of glass fibres and its influence on strength', *Fiber & Integrated Optics*, **2**, no. 3–4, p 267, 1979.
5 Reeve, M. H., France, P. W. and Jackson, L. A., 'Strength and stress corrosion of sodium borosilicate optical fibres', 4th European Conference on Optical Communication, Genoa 1978, p. 107.
6 Weibull, W., 'A statistical theory of the strength of materials', *Proc. R. Swedish Inst. Res.*, No. 151, publication no. 4, 1939.
7 Olshansky, R. and Maurer, R. D., 'Tensile strength and fatigue of optical fibres', *J. Appl. Phys.*, **47**, no. 10, p. 4497, October 1976.
8 Tariyal, B. K., Kalish, D., Santana, M. R., 'Proof testing of long length optical fibres for a communication cable', *Am. Ceram. Soc. Bull.*, pp. 204–5, 1977.
9 France, P. W. and Duncan, W., 'Proof testing of optical glass fibres', Conf. on Physics of Fibre Optics, Chicago, April 1980.
10 Helfinstine, J. D. and Maurer, R. D., 'Effect of flaw distribution on fatigue characterisation in optical waveguides', 6th European Conf. on Optical Communication, York 1980, p 117.
11 Krause, J. T., 'Transitions in the static fatigue of fused silica fiber lightguides', 5th European Conf. on Optical Communication, Amsterdam, 1979, 19.1–1.
12 France, P. W. *et al.*, 'Liquid nitrogen strengths of coated optical glass fibres', *J. Mat. Sci.*, **15**, pp. 825–30, 1980.
13 Gardner, W. B., 'Microbending loss in optical fibres', *Bell Syst. Tech. J.*, **54**, no. 2, pp. 457–65, 1975.
14 Gloge, D., 'Optical-fibre packaging and its influence on fiber straightness and loss', *Bell Syst. Tech. J.*, **54**, no. 2, pp. 245–62, 1975.
15 Reeve, M. H., 'Investigation of optical fibre buckling in loose tube packaging', *Electronics Letters*, **14**, no. 3, pp. 47–8, 2nd February 1978.
16 Jackson, L. A., Reeve, M. H. and Dunn, A. G., 'Optical fibre packaging in loose fitting tubes of oriented polymer', *Optical & Quantum Electronics*, **9**, pp. 493–8, 1977.
17 Harrison, J. C., 'The metal foil/polyethylene cable sheath and its use in the Post Office', Institution of Post Office Electrical Engineers Paper No. 229, November 1968.
18 Saunders, M. J. and Parham, W. L., 'Adhesive sandwich optical fibre ribbons', *Bell Syst. Tech. J.*, **56**, no. 6, pp. 1013–4, 1977.
19 Eve, M. *et al.*, 'Transmission studies on three graded index fibre cable links installed in operational ducts', *Optical & Quantum Electronics*, **10**, p. 253–6, 1978.
20 Sandbank, C. P., (Ed.) 'Optical Fibre Communication Systems', ch. 11 (Wiley, New York, 1980).

Paper first received by the Institution on 20th February 1981 and in final form on 13th April 1981.

(*Paper No. 1995/CC 334*)

UDC 621.315.68: 621.372.83: 621.391.63/4
Indexing Terms: Fibre optics, Optical communication, Connectors

Connectors for optical fibre systems

P. MOSSMAN, M. Inst. P.*

SUMMARY

This paper outlines the requirements of demountable connectors for optical fibres and cables used in data links and telecommunications applications. Problems associated with alignment and termination of the mating fibres are discussed. The various methods for overcoming or minimizing these problems are reviewed and specific examples of connector technology are highlighted. Some techniques for measuring the optical properties of connectors are given and their value and limitations in the assessment of system performance are detailed. Finally future requirements are discussed.

* *The General Electric Co. Ltd., Hirst Research Centre, Wembley, Middlesex HA9 7PP*

1 Introduction

Demountable connectors for electrical transmission line systems have always presented a design problem. This problem is based on the need to preserve continuity of the transmission line parameters at the connector interface whilst maintaining the mechanical requirements such as adequate retention force and durability. As is well known, the problem of transmission line continuity becomes more acute at the higher frequencies. It is perhaps then not too surprising to find, with the advent of optical fibre transmission line systems, that the design and production of optical connectors of low insertion loss has yielded very severe difficulties.

One obvious criterion in deciding what is an acceptable level of insertion loss for a connector is the attenuation rate of the transmission medium. State-of-the-art monomode and multimode optical fibres[1,2] have very low losses with attenuation rates in the range 0·2 dB km^{-1} to 0·5 dB km^{-1} at wavelengths of 1·55 μm and 1·3 μm respectively. Clearly, the designer of a long distance telecommunication system (e.g. in the trunk or junction telephone networks) wishes to use, to the full, the potential of these modern fibres by increasing the spacing between repeaters. Connectors having insertion losses in excess of say, 1·0 dB will not generally be acceptable. Furthermore, in the interests of optical safety and source reliability power levels should be minimized. Good connector performance helps with the achievement of this requirement.

For short length, low bandwidth data links the problems are not quite so acute. Some of the more severe mechanical problems can be alleviated by using large core fibres and power considerations are relatively unimportant. Nevertheless variations in the insertion loss with repeated matings of a given connector set, and variations between groups of similar connector sets, must be kept to narrow limits.

In recent years several connector designs have been produced employing a variety of techniques to secure good optical and mechanical performance. It is the main purpose of this paper to review the more promising of these techniques.

2 Alignment Problems

Deviations in the geometrical and optical parameters of optical fibres at their interface within a mated connector set will affect the insertion loss of the connectors. It will not be possible to allow for all of these variations in any particular termination technique. The variations to be considered are in: core and reference surface diameters, their circularity errors, core and concentricity errors, and the numerical apertures of the two fibres. These terms are defined in CCITT Publication G651.

Such deviations must be considered in addition to the quality of the fibre alignment provided by the connector mechanism. Connector alignment errors, which must be

0033–7722/81/070333+08 $1.50/0

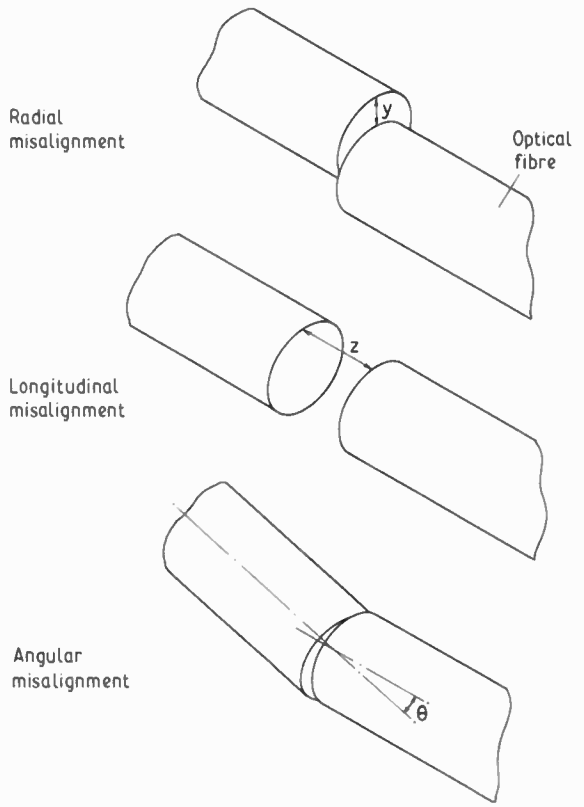


Fig. 1. Three types of misalignment between coupled fibres.

kept to a minimum, are illustrated in Fig. 1. The loss arising from any of these three errors is dependent on the fibre type, core size and the distribution of power between the various propagating modes. The effect of radial and longitudinal displacements can be assessed relatively simply and have been studied by several authors.[3-5] However, assessment of angular errors requires a great deal of experimental care in the avoidance of simultaneous radial displacements. Some degree of longitudinal displacement is, of course, inherently part of the angular misalignment. The effects of radial and longitudinal displacements are given for a 50 μm core graded index fibre in Fig. 2, from which it will be seen that radial alignment presents by far the more important mechanical constraints, a displacement of 10 μm leading to an insertion loss of 1 dB. The implications then for a connector manufacturer is that, if less than 1·0 dB is a satisfactory level of insertion loss, each connector of a set must have radial alignment controlled to 5 μm. It should also be noted that for radial deviations in excess of 10 μm, the loss increases rapidly. This then is the key problem in connector technology for small-core optical fibres. Since the tolerances required are outside normal machining limits, the problem is indeed severe. Figure 3 gives calculated values of insertion loss as a function of core size for various values of radial displacement. The curves indicated that only for fibres with core sizes of 150 μm or greater can standard machining practices be used in the manufacture of low loss connectors.

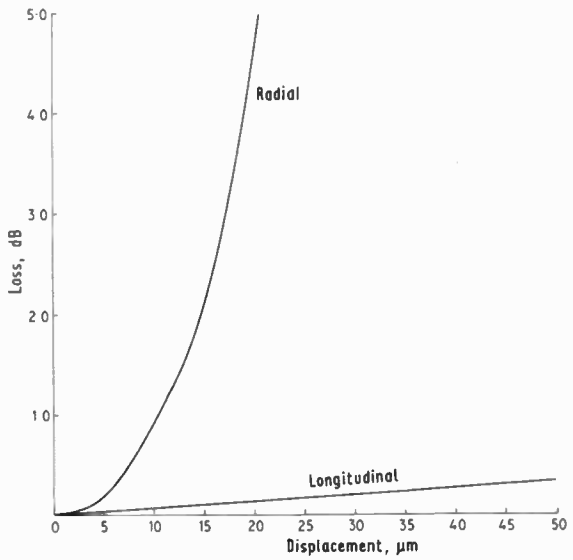


Fig. 2. Insertion loss contribution due to radial or longitudinal misalignment.

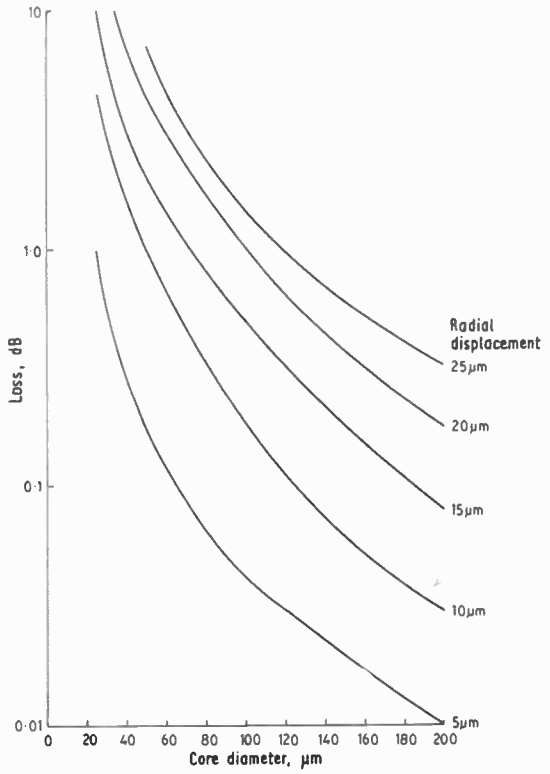


Fig. 3. Variation in insertion loss with fibre core diameter for given values of radial displacement.

In addition to the very severe problems of loss due to misalignment with small-core fibres some mention must be made of other possible effects on system performance. Since misaligned connectors will act as a mode scrambler the bandwidth of a system with connectors remote from the transmitter or receiver will

be modified. Also a misaligned connector or joint will act as a modal filter and may give rise to modal noise in systems using coherent sources.[6] These are highly specialized problems and cannot be dealt with more extensively here.

3 Termination Problems

Termination techniques vary between the different connector technologies, which will be discussed in more detail in Section 6. However, to provide an understanding of the basic termination problems we shall consider a very simple connector system comprising two ferrules supporting the mating fibres and a centre section with an alignment sleeve in which the ferrules are inserted. Figure 4 illustrates the arrangement of these principal components. Assuming that adequate radial and angular alignment is achieved with the connector mechanism, there remains the problem of ensuring:

(a) that the end-faces of the mating optical fibres are prepared so that they are smooth, flat and perpendicular to the axis of the fibre,
(b) that, when mated, they butt in a manner that provides a maximum of coupled power without damage to their end faces.
(c) that the fibres are not unduly stressed in the mated or unmated condition.

Three techniques for preparation of the fibre end-faces are currently in use. They comprise: cleaving the fibre[7] prior to insertion into the ferrule; inserting and bonding the fibre into the ferrule, and then cleaving the fibre close to the ferrule end-face;[8] thirdly, using either of the above procedures and then polishing the end-face flush with that of the ferrule. None of these methods is entirely satisfactory. It is not difficult to cleave a fibre in such a way that the end face is smooth and flat. However, a perpendicular end-face is not easy to secure.[8,9] There is also a risk of damage to the fibre end-face if it is inserted into the ferrule after cleaving. Polishing after insertion and bonding undoubtedly can provide a good end-face.[10,11] However, polishing techniques are more expensive, time consuming, and are much less convenient for adverse in-field conditions than are simple cleaving techniques. One solution to the problem of field termination is to terminate the connectors with pig-tails in the factory. For the pig-tail type of connector the assembly is spliced into the system. The additional costs and degradation is performance due to these splices must of course be taken into account in the overall assessment of the technique.

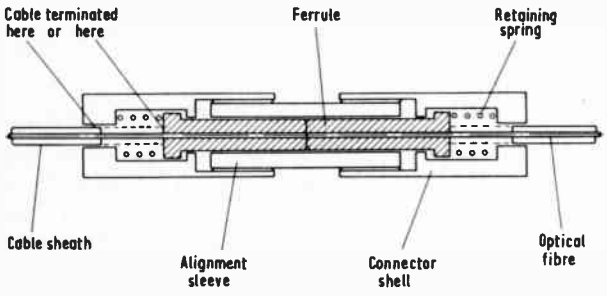


Fig. 4. Basic structure of a ferrule-type optical fibre connector.

A further problem arises from the need to secure close abutment of the fibre end-faces. The coupling must be secured without damage to the fibres or ferrules and it is normally necessary to allow the ferrules to float so that when mated they are held in contact by means of a spring or other resilient member. Under these circumstances one must decide whether the fibre cable sheath and tensile strength members are to be terminated within the fixed part of the connector shell or within the floating ferrule. In the former case the fibre has a tendency to buckle as the resilient member is compressed, particularly when tight-jacketted cables are employed: in the latter the cable termination is much more complex and the mated ferrules may be parted by a tensile load applied to the cable. On balance it would appear that termination of the cable within the fixed part of the connector provides the better solution, but the design and precision of the connector parts must be such that the floating action is limited to avoid bending loss in the fibre.[12]

Before moving on to discuss reliability, mention should be made of the use and limitations of refractive index matching media at the fibre interfaces. From the point of view of optical performance use of a low absorption medium of refractive index close to that of the core material at the fibre interface is entirely desirable. Not only are reflection losses reduced to negligible proportions, but the condition of the fibre end faces becomes less critical. However, none of the available media, whether liquid, gel or resilient solid appears to be satisfactory in practice. All are subject to contamination and require frequent replacement if a connector set is to undergo several matings during its useful life.

4 Reliability

Damage, wear and contamination are important factors affecting the reliability of optical fibre connectors. The materials for the component parts must be selected with great care to ensure sufficient manufacturing ease and precision combined with robustness. Even with fine machined parts some degree of wear occurs in the initial stages of use, due to bedding-in of the inevitable machining marks. Allowance must be made for this process in the connector design. Similarly any polymer based components must be free from distortion and relaxation during prolonged mating. Contamination of the fibre end-face is an inevitable problem which is aggravated with decrease in core size. Ideally the connector design should incorporate some self-cleaning mechanism. However, most of the available connectors require some manual cleaning procedures as part of the handling rules for the installed connectors.

5 Safety

Accompanying the use of high-radiance light-emitting diodes and semiconductor laser diodes as sources for optical fibre systems, there is a conceivable risk of eye damage due to radiation. The risk is increased by the fact that the wavelength range is normally outside the visible region. The most likely point at which eye damage could occur is at the connector interface, where fibre end-faces will frequently be closely examined to check for contamination etc. Ideally the connector should provide some safety shutter mechanism or electrical interlock to switch off or turn down the source power level when the connectors are unmated. Standards for safety procedures are currently being considered at National and International level. Until such standards are published, one can only underline the need for reasonable care on the part of designers and users of optical fibre equipments. Some of the criteria involved are discussed elsewhere in this issue.[25]

6 Available Technology

The very severe problems involved in providing mechanisms of sufficient accuracy for the achievement of low-loss connectors for small-core optical fibres has led to a wide variety of approaches to the solution. Individual designs, shall be discussed, but first the broad categories will be examined without reference to specific implementations.

Almost all optical fibre connector sets comprise three main parts: a central alignment section and two nominally identical insertion members attached to the two mating fibres. This construction stems mainly from the need for access to both optical surfaces for cleaning purposes. Perhaps the most common type of connector is the ferrule type, the basic design of which has already been described in Section 3. A very interesting class of connectors is comprised of those using pseudo-kinematic design principles, namely those in which the fibre is located in the interstices between accurately machined spheres or rods or in which the fibres themselves are positioned in V-grooves. Yet another technique employs lenses to collimate the light from one fibre and to refocus it on to the core of the other fibre. Among these various broad categories there are several variants. Many have appeared in the patent literature. Unless otherwise stated, the results given in the following examples are for connectors using 50 μm core graded-index fibres.

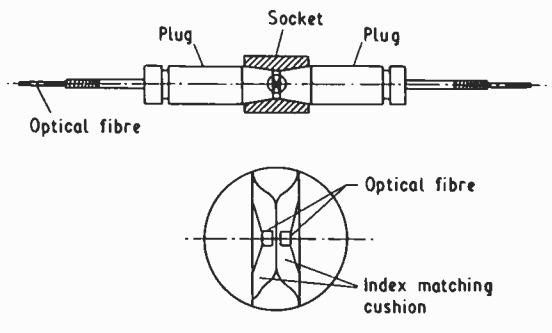


Fig. 5. Cross-section of the single fibre connector detail shows elastomeric index matching cushions (After Runge and Chang.[14])

6.1 Ferrule Connectors

A ferrule-type connector which has received considerable attention is the moulded biconical connector developed at Bell Laboratories.[11,13,14] The basic elements of this connector are illustrated in Fig. 5. The biconical ferrules (or plugs) are cast around the fibre using a silica-loaded epoxy resin. Alignment of the fibre is controlled by the moulding die and radial and angular alignment is held to better than 5 μm and 1·0 degree respectively. The reported insertion loss levels are considerably better than 1·0 dB under equilibrium mode distribution conditions. The significance of modal distribution will be discussed in Section 7. By loading the epoxy material with silica particles good dimensional stability and wear resistance is achieved. Initially the biconical socket was of machined brass: more recently a moulded socket has been developed. A special feature of the original design was the use of transparent silicone resin pads over the fibre end-faces. These pads provided a refractive index match medium at the mating interface. Currently the fibres are being polished flush with the ferrule end-faces and the fibres are in direct contact. The gap and parallelism of the end-faces can be controlled to a degree that allows loss levels better than the level due solely to Fresnel reflections.[15] This connector has been featured in field trial and other installations in the USA.

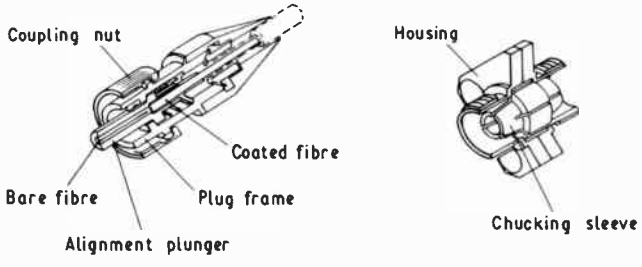


Fig. 6. Connector structure (After Kurochi.[10])

A novel manufacturing technique for metallic ferrule-type connectors has been developed by workers at Nippon Electric Co Ltd.[10] The special features of this connector are that the fibre is mounted within a glass capillary which is set in the tip of the ferrule and allows the fibre end-face to take a good polish. The outer surface of the ferrule is then trimmed by a cylinder grinding technique. Optical centring of the fibre core with respect to the axis of rotation of the grinding machine ensures that the fibre core is concentric with the mating surface of the ferrule after the grinding process. Figure 6 shows the connector structure with its centre section comprising a sprung chucking sleeve. Insertion losses of better than 0·5 dB have been reported for this type of connector. Further development of this type of

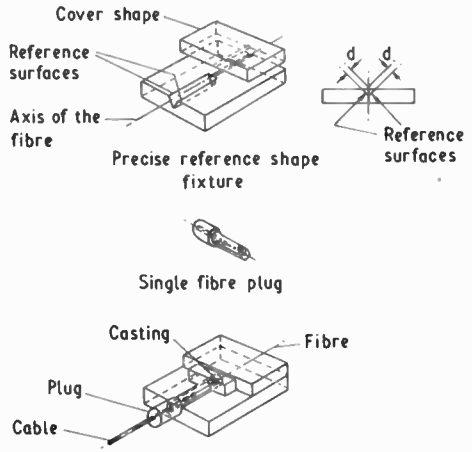


Fig. 7. Principle for preparing accurate plugs (After Le Noane.[17])

connector has been carried out by workers at NTT,[15] who have advocated a ceramic capillary rather than a glass one. Also a C-shaped alignment sleeve has been developed. Results for a monomode version of this connector have also been published.[16]

A third ferrule-type connector employing a novel alignment method has been described by Le Noane[17] of Centre National d'Etudes des Telecommunications, France. The alignment principle is illustrated in Fig. 7. The termination procedure requires a special jig comprising two V-grooves, one of which holds the ferrule by means of two mutually perpendicular reference surfaces. The second V-groove holds and positions the fibre with respect to the reference surfaces. The fibre is then cemented into the ferrule using a fast setting epoxy. Finally the fibre is cleaved and the end-face polished flush with the ferrule. The centre section of the connector set contains a V-groove in which two ferrules are aligned and mated. Insertion losses between 0·1 and 0·5 dB have been reported.

Another ferrule-type connector worthy of note is the double eccentric type[18,19] illustrated in Fig. 8. In this

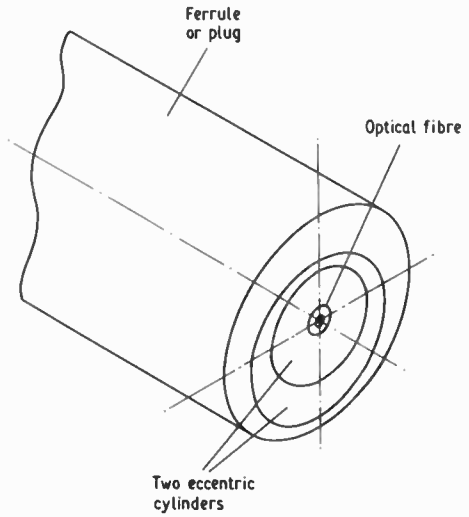


Fig. 8. Double eccentric adjustable connector

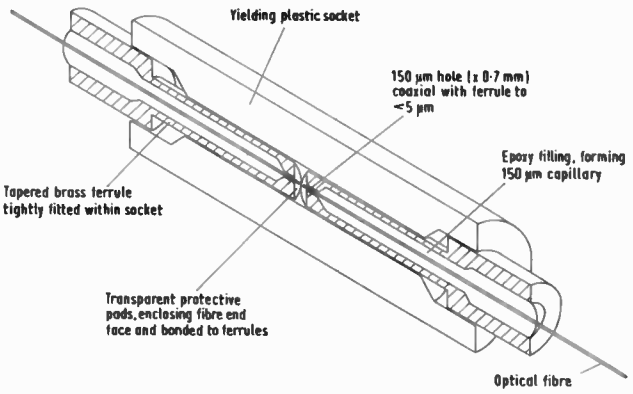


Fig. 9. Demountable connector alignment mechanism.

type of connector the fibre core can be centred on the reference surface of the ferrule by rotation of the mechanism. The parts are then locked to give permanent alignment. This type of design has also been used to produce a connector suitable for use with monomode fibres.

Ferrule-type connectors have been developed[20] in which slightly tapered ferrules are mated within a yielding plastic sleeve as shown in Fig. 9. The advantage of this approach is that mechanical free play between the ferrules and the sleeve is eliminated and wear of the parts is negligible. A machining technique was used to achieve optimum concentricity. The procedure involved drilling and tapering the ferrules sequentially without disturbing the chuck settings. The insertion losses achieved were around 1·0 dB under fully-filled modal power conditions.

6.2 Pseudo-kinematic Designs

Of the connectors using kinematic design principles a noteworthy example is the triple-ball connector described by Hensel[21] and illustrated in Fig. 10. The fibres are positioned by clusters of three accurately ground spheres making point contact. In the mating of two connectors the fibres are aligned by interpenetration of the two clusters until they contact at six points. The reported average insertion loss was 0·49 dB without index matching and 0·18 dB with matching.

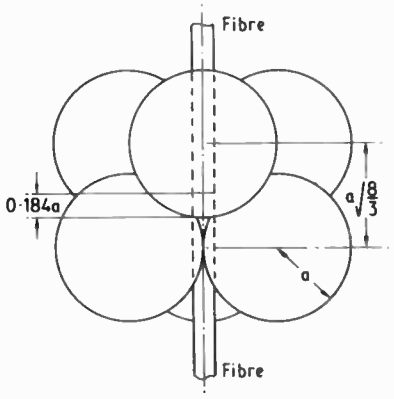


Fig. 10. Two interlocking sets of three balls showing position of fibre ends. (After Hensel.[21])

6.3 Expanded-beam Connectors

A schematic arrangement of the so-called expanded-beam connector[22] is shown in Fig. 11. The connector comprises two lenses for collimating and refocusing the light from one fibre to the other. With this arrangement radial alignment of the mating parts is much less critical than for a connector in which the fibres are in butting

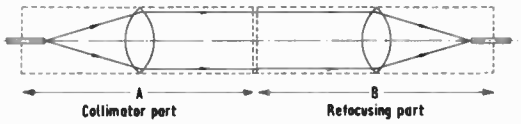


Fig. 11. Schematic diagram of an expanded-beam connector

contact. Furthermore, separation between the mating parts A and B is not at all critical. However, angular alignment of the parts is very critical. For example, with a lens of, say, focal length 10 mm, a misalignment of 6 minutes of arc would produce an excess loss of 3 dB with 50 μm core fibres. A further disadvantage is that even with perfect alignment there is an intrinsic loss due to spherical aberration when simple lenses are used. Nevertheless, in applications such as multiway connectors and edge connectors for printed circuit boards where good radial and longitudinal alignment might be difficult to achieve, the expanded-beam technique could offer clear advantages. In addition, the problems of contamination are much less severe with expanded-beam connectors and they are therefore preferred for military and industrial environments where the presence of dirt, dust or mud would preclude the use of butting connectors.

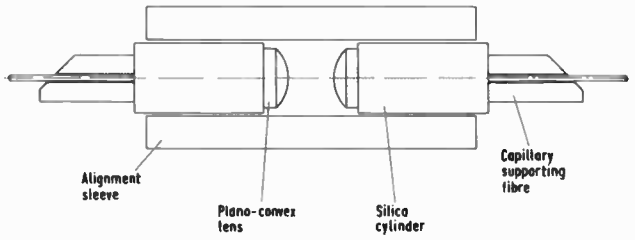


Fig. 12. Practical expanded-beam connector

One example of an expanded beam connector is that developed by the author and is illustrated in Fig. 12. A plano-convex lens is cemented at one end of a silica cylinder with its optical axis accurately aligned with the geometrical axis of the cylinder. The fibre is butted at the focal point of the lens at the other end of the cylinder and is supported by a glass capillary. The insertion losses achieved with this type of connector were 2·0 dB to 2·5 dB.

7 Assessment Techniques

At the time of writing there were no agreed standards for the assessment of optical fibre connectors, although National and International Standards Committees were then currently engaged in the drafting of specifications for optical test methods. In the case of insertion loss measurement it has been recognized that more than one technique may be required depending on whether the connector is the factory terminated pig-tail type, in-field terminated, or a connector/cable assembly in which no free end to the optical fibres is available.

We shall describe one technique suitable for in-field terminated connectors and another for pig-tail connectors as examples. However, before doing so it is necessary to consider some of the problems associated with insertion loss measurements. There are two important questions to be answered. They are: (a) bearing in mind that defects or variations in the geometrical and optical characteristics of fibres will contribute to the measured insertion loss of a connector, how do we determine the intrinsic quality of the connector parts; (b) how do we provide practical information on the optical properties of connectors for the purposes of system design. A possible solution to the first is by the use of selected optical fibres with geometrical and optical parameters controlled to within agreed limits, in other words, virtually perfect fibres. There is, in fact, a precedent to this approach in the testing of radio frequency connectors using selected coaxial cable.

The second question poses some very difficult problems. Because of the possible deviations in the multiplicity of fibre parameters and the present lack of statistical information on production spreads, a reliable estimation of the effects on overall connector performance cannot yet be made. A further difficulty arises from the fact that connector insertion loss is significantly dependent on the distribution of power between the various propagating modes in a multimode optical fibre. Selective attenuation of power within the various modes leads to a variation in attenuation rate with propagation distance in the fibre until an equilibrium distribution is attained. A connector addressed by power from a fully-filled fibre (i.e. equal power in all modes) will show an insertion loss of perhaps three times that obtained under equilibrium conditions. The possible values of insertion loss for a given connector are therefore infinite. A further consequence of this situation is that a connector inserted at some point in a fibre may lead to redistribution of the optical power between modes such that the attenuation rate in the transmission path following the connector may be anomalously increased. Under these circumstances the locally-measured insertion loss would not provide the true insertion loss of the connector. Many optical fibre systems will have connectors only at points close to the source and the detector. However, it can be envisaged that with field deployable systems connectors may be used at any of a number of intermediate points.

It does not therefore appear practicable or even

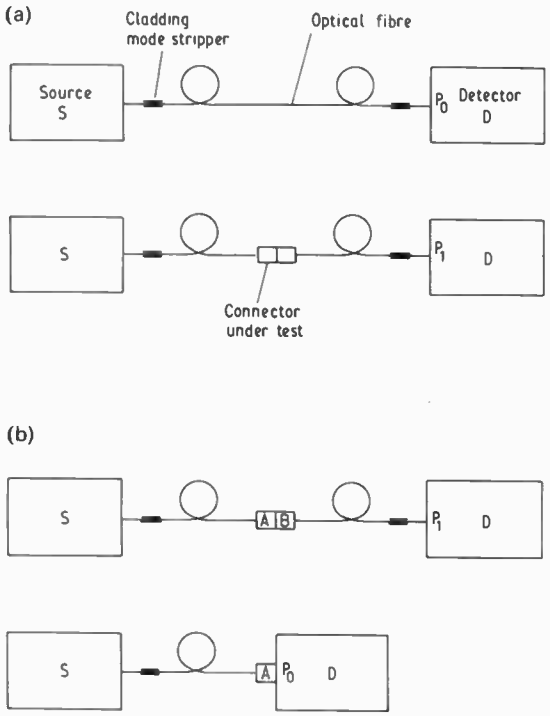


Fig. 13. Test arrangement for two possible measurement techniques for connector insertion loss.

possible to provide information on connector insertion loss for all conceivable conditions of use. However, standardized test conditions can be prescribed which should serve as a valuable basis for calculating system performance. It seems likely that standard test conditions will call for 'fully-filled' and 'equilibrium-mode' conditions both of which can, in principle, be provided[11,24] in the testing laboratory. However, the methods for achieving these conditions requires final agreement for standards purposes.

Figure 13 illustrates the arrangement for measurement of connector insertion loss. Techniques usable for in-field terminated connectors (a), and for pig-tail connectors (b), are given. If both fully-filled and equilibrium-mode distributions are to be used then the source and launch optics must be capable of producing a spot of light covering the fibre core with a cone half-angle greater than that given by $\theta = \text{arc sin } (NA)$. With this arrangement the fully-filled condition is automatically satisfied. An equilibrium-mode distribution can then be achieved either by appropriately restricting the launched light beam spot size and cone half-angle to less than the core size and θ respectively. Alternatively a mode filter comprising several turns of fibre wound onto a mandrel can be used to obtain equilibrium conditions.[11] Throughout the measuring procedure the launched power must be kept at a constant level. Care must be taken to prevent light propagating in the cladding from affecting the measurement and a cladding mode stripper is placed before the connector and before the detector. Light launched into cladding modes cannot be regarded as useful power and must be removed by locally destroying the waveguide properties of the cladding. A cladding mode stripper can be simply a bath of liquid of refractive index matching that of the cladding material.

The detector would comprise a photodiode of suitable spectral response. The response to incident light should either be linear or should be calibrated. The cone of light should have its axis perpendicular to the active area of the photodiode, within which all the light should fall.

For method (a) a suitable length of fibre is placed, between source and detector as shown in Fig. 13. The power level P_0 (generally in arbitrary units) is measured. Without disturbing the fibre at the source and detector the fibre is cut at midpoint and terminated with the connector set under test. The new power level P_1 is measured and the insertion loss is given by:

$$-10 \log P_1/P_0 \text{ dB}$$

For method (b) where the connector set is already terminated a reference power level P_0 is obtained by presenting the light from connector A directly to the detector. The clear disadvantage of this method is that any losses in the connector A alone would not appear in the calculation. Also allowance must be made for the attenuation of the fibre attached to connector B, which may or may not be negligible compared with the insertion loss of the connector set.

8 Future Requirements

Although very important advances have been made in the technology for multimode optical fibre connectors it does not seem that the definitive answer to the question of a simplified in-field termination procedure is yet available. Very little attention has been given to safety features for systems carrying high optical power levels. Similarly, methods for preventing contamination in the unmated state do not appear to have been studied nor do techniques for self cleaning.

Several monomode connector designs have been reported. However, more work is needed in dealing with the problems of deviations in fibre geometry. Practicable techniques and termination procedures have yet to be firmly established. Since monomode fibres are likely to be employed on a large scale in the coming decade there is an urgent need for a satisfactory solution.

Until agreed standards are established we shall undoubtedly continue to see a wide variety of technologies being pursued in the future development of optical fibre connectors.

9 References

1 Hanawa, F. *et al.*, 'Fabrication of OH-free VAD fibre', *Electronics Letters*, **16**, no. 18, pp. 699–700, August 1980.
2 Ainslie, B. J. *et al.*, 'Optimized structure for preparing long ultra-low-loss single-mode fibres', *Electronics Letters*, **16**, no. 18, pp. 692–3, August 1980.
3 Miller, C. M. and Mettler, S. C., 'A loss model for parabolic-

profile fibre splices,' *Bell Syst. Tech. J.*, **57**, no. 9, pp. 3167–80, November 1978.
4 Di Vita, P. and Rossi, N., 'Theory of power coupling between multimode optical fibre', *Optical & Quantum Electronics*, **10**, pp. 107–17, 1978.
5 Chu, T. C. and McCormick, G. R., 'Measurement of loss due to offset, end separation and angular misalignment in graded index fibres excited by an incoherent source', *Bell Syst. Tech. J.*, **57**, no. 3, pp. 595–602, March 1978.
6 Epworth, R. E., 'The phenomenon of modal noise in analogue and digital optical fibre systems', Proceedings of the Fourth European Conference on Optical Communication, pp. 492–501, 1978.
7 Gloge, D. *et al.*, 'Optical fibre end preparation for low-loss splices', *Bell Syst. Tech. J.*, **52**, no. 9, pp. 1579–88, November 1973.
8 Cundy, S. L. *et al.*, 'An insertion loss meter and reflectometry of prepared fibre ends', IEE Colloquium on Measurement Techniques for Optical Fibre Systems, IEE Digest no. 1979/33 14/1 May 1979.
9 Millar, C. A., 'A measurement technique for optical fibre break angles', *Optical & Quantum Electronics*, **13**, pp. 125–31, 1981.
10 Kurochi, N. *et al.*, 'A development study on design and fabrication of an optical fibre connector', Proceedings of Third European Conference on Optical Communication, pp. 97–9, 1977.
11 Young, W. C. *et al.*, 'A transfer-moulded biconic connector with insertion losses below 0·3 dB without index match', Proceedings of Sixth European Conference on Optical Communication, pp. 310–13, 1980.
12 Gloge, D., 'Bending loss in multimode fibres with graded and ungraded core index', *Applied Optics, (USA)* **11**, no. 11, pp. 2506–13, November 1972.
13 Cook, J. S. and Runge, P. K., 'Exploratory fibreguide interconnection system', Proceedings of Second European Conference on Optical Communication pp. 253–6, 1976.
14 Runge, P. K. and Cheng, S. S., 'Demountable single-fibre optic connectors and their measurement on location', *Bell Syst. Tech. J.*, **57**, no. 6, pp. 1771–91, August 1978.
15 Suzuki, N. and Nawata, K., 'Demountable connectors for optical fibre transmission equipment', *Rev. Elect. Commun. Lab.* **27**, no. 11–12, pp. 999–1009, November–December 1979.
16 Suzuki, N. *et al.*, 'Ceramic capillary connector for 1·3 μm single-mode fibres', *Electronics Letters*, **15**, no. 25, pp. 809–10, December 1979.
17 Le Noane, G., 'Low-loss optical-fibre connection systems', *Electronics Letters*, **15**, no. 1, pp. 12–13, January 1979.
18 Zemon, S. *et al.*, 'Eccentric coupler for optical fibres: a simplified version', *Applied Optics, (USA)* **14**, pp. 815–6, 1975.
19 Tsuchiya, H. *et al.*, 'Double eccentric connectors for optical fibres', *Applied Optics, (USA)* **16**, no. 5, pp. 1323–31, May 1977.
20 Cundy, S. L. *et al.*, '8·448 Mbit/s optical system for long range transmission', Proceedings of the Fourth European Conference on Optical Communication, pp. 616–25, 1978.
21 Hensel, P., 'Triple-ball connector for optical fibres', *Electronics Letters*, **13**, no. 24, pp. 734–5, November 1977.
22 Nicia, A., 'Practical low-loss lens connector for optical fibres', *Electronics Letters*, **14**, no. 16, pp. 511–2, August 1978.
23 North, J. C. and Stewart, J. H., 'A rod lens connector for optical fibres', Proceedings of the Fifth European Conference on Optical Communication, p. 9.4–1, 1979.
24 Holmes, G. T., 'Launch dependent attenuation measurements on a 10 kilometre concatenation experiment', Proceedings of the Sixth European Conference on Optical Communication, pp. 144–7, 1980
25 Topping, J. D. and North, J. C., 'Optical fibre transmission systems—the problems of formulating optical radiation safety criteria,' *The Radio and Electronic Engineer*, **51**, no. 78, pp. 393–6, July/August 1981.

Manuscript first received by the Institution on 14th April 1981 and in final form on 11th May 1981.
(Paper No. 1446/CC 340)

The Author

Peter Mossman joined the General Electric Company in 1953 and from 1960 to 1964 worked on magnetic film storage cells; for the next eleven years he carried out development work on ferrite materials for microwave and low-frequency applications. Since 1976 he has been Group Leader of the Optical Fibres and Components Group at Hirst Research Centre. He participates in BSI, IEC, and CCITT standards activities and is the current Chairman of IEC, SC46E Working Group 2 on optical fibre components and accessories.

UDC 621.383.2: 621.391.3: 621.391.63/4
Indexing Terms: Semiconductor devices, light-emitting diodes, Fibre optics

Light-emitting diodes for optical fibre systems

A. C. CARTER, M.A., D.Phil.*

SUMMARY

Light-emitting diodes are important fibre optic sources not only for short haul, low cost data links, but also in the long haul, high data rate telecommunications field. This paper discusses the range of applicability of l.e.d. sources, and the performance, design and optimization of such devices. Results are given for devices operating both in the low-loss 1·3 μm window and in the 0·8–0·9 μm window.

* *Plessey Research (Caswell) Limited, Allen Clark Research Centre, Caswell, Towcester, Northants NN12 8EQ.*

1 Introduction

A wide variety of sources is now available to the fibre optic system designer. These devices range between simple spontaneous emitters (l.e.d.s), through superluminescent devices to sophisticated single-mode lasers. The choice of a source for a given system depends on a number of parameters, including performance, temperature dependence and reliability. In general a l.e.d. will be used where an adequate performance can be achieved, since these devices have simpler circuit requirements, immunity to self-pulsation and modal noise, and proved reliability and performance at elevated temperatures. Lasers, on the other hand, will always be used in systems which demand very high launch powers, wide modulation bandwidths and narrow spectral linewidths.

With the development of low-loss multimode fibres in the 1·3 μm band, where fibre material dispersion is a minimum, the l.e.d. has become a viable source for long haul, high data rate telecommunication links, an area where the laser has been traditionally regarded as the essential source element.

In this paper the range of applicability of the l.e.d. in the fibre optics field is discussed and the structure, design and optimisation of suitable l.e.d.s described. L.e.d. results will be stressed for 1·3 μm devices, but results obtained for 0·8–0·9 μm devices, which have important applications in short haul and military systems, will be given for comparison.

2 Performance Limits of L.E.D.-sourced Systems

A wide range of l.e.d. sources suitable for fibre optic communications has now been developed, including the surface emitting l.e.d. first adopted by Burrus,[1] spontaneous edge emitters[2,3] and high power superluminescent devices.[4]

Suitable devices for most multimode fibre systems can be chosen from the range of l.e.d.s which have now been developed. GaAlAs/GaAs devices emitting in the 0·8 to 0·9 μm range are suitable for moderate length (up to ~5 km) analogue and digital transmission links, and are finding wide application in both civil and military systems. Such devices show reliable operation at high temperatures (> 125°C), impose only modest demands on supply current and drive circuitry and their immunity to modal noise and self-pulsation gives excellent performance predictability.

The high-power superluminescent devices are suitable for moderate length, high-loss systems, where the use of high-attenuation, low-cost fibre will give economic advantages. Such devices have the potential advantages of high power and reduced spectral linewidth, but generally operate at higher current densities than lasers fabricated in similar material, so that further research to establish their long-term reliability is necessary.

The maximum link lengths and data rates which can

0033–7722/81/070341+08 $1.50/0

be achieved in multimode optical fibre links are determined by the tolerable link loss (limited by fibre attenuation, source-coupled power and receiver sensitivity) and fibre dispersion. The latter is composed of two parts, chromatic or 'material' dispersion, which is a property of the fibre material, and 'modal' dispersion, which is determined by the accuracy of the refractive index profile in the fibre.

The degree to which chromatic dispersion is significant depends on source spectral linewidth and centre wavelength. L.e.d.s have an inherently large spectral linewidth, typically $2–3kT$ where k is Boltzmann's constant and T is the absolute temperature, and as such can be much more susceptible to this form of dispersion than lasers, which have an inherently narrow spectral width. For instance, for l.e.d.s with emission wavelength centred near 0·9 μm, material dispersion severely limits the transmission distance at data rates above a few Mb/s. For this reason, all 'first generation' 140Mb/s links, which operate in the 0·8–0·9 μm band, necessarily use laser sources.

The material dispersion in silica-based fibres falls through zero near 1·3 μm, so that even with the wide spectral width of the l.e.d., material dispersion can impose much less of a limit on link length than modal dispersion.

Figure 1 shows estimates of the maximum link lengths and data rates possible with l.e.d. sourced systems. The three parallel curves in the centre of the figure represent the fibre modal dispersion limits. Three examples are given, 0·75, 1·5 and 3 GHz·km, which straddle the range between what is now commercially available and what is near 'state-of-the-art' for multimode fibres. Link lengths and data rates for two l.e.d. wavelengths (0·9 and 1·3 μm) are given, and are calculated assuming no modal dispersion, but allow for some narrowing of the l.e.d. linewidth by the filtering action of the fibre.[5] To the left of the Figure, where the link length falls only slowly with data rate, the system is attenuation limited. To the right, where link length falls very rapidly, the link is limited by chromatic dispersion. It can be seen from the Figure that the 0·9 μm system is limited by material dispersion, whereas the 1·3 μm system is limited by the fibre modal dispersion.

In the power-limited region of the diagram (left side) no allowance is made for splice or connector losses or for system margin. Consequently practical link lengths will lie well within the bounded region. However, it would seem that links in the range 15–30 km at 140 Mb/s are feasible, depending on the availability of wideband (> 1·5 GHz·km) multimode fibre. It is thus probable that l.e.d. sources will be used for the majority of single-channel point-to-point multimode links, as the extra power and narrower linewidth given by laser sources will be superfluous.

For all single-mode systems, it is essential to use laser sources. Only the high power and narrow linewidths of such devices can fully exploit the potential of low-loss, ultra-wide-bandwidth single-mode fibres.

A particularly important area for such systems is in undersea applications. Here several laser sources, which can be optically and electrically switched into the link, may be required to achieve the necessary ultra-high reliability. The extra cost and complexity of the laser stabilization circuitry, device redundancy and thermo-electric coolers could be easily absorbed in these systems.

3 L.E.D. Structure and Materials

Highly efficient light emitting diodes can be made in many of the direct bandgap III-V semiconductors. To cover the wavelength range 0·8–1·65 μm two material systems in particular are employed, GaAlAs/GaAs (0·8–0·9 μm and GaInAsP/InP (0·95–1·65 μm), both of which enable full lattice matching of double heterostructures resulting in efficient, highly reliable devices.

The lens-coupled surface emitting l.e.d. developed by Plessey is illustrated in section, in Fig. 2. Two different active layer configurations are shown for both of the

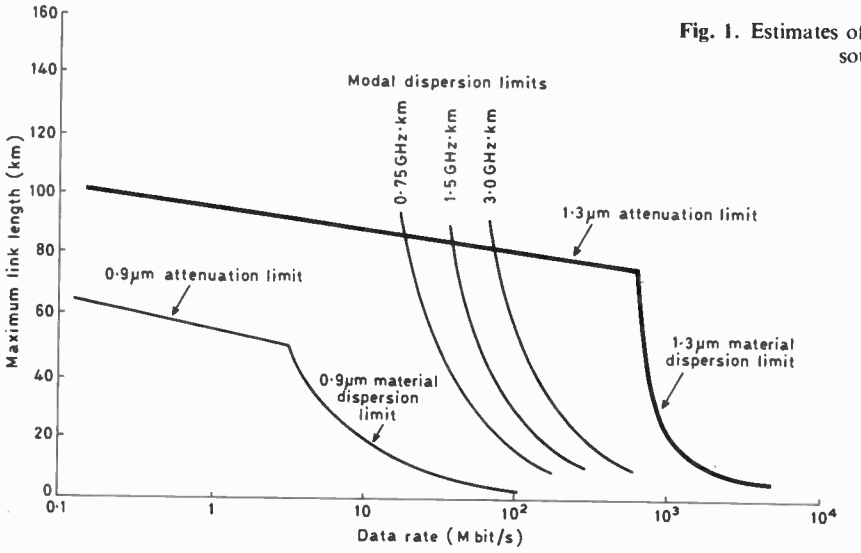


Fig. 1. Estimates of maximum link lengths and data rates for l.e.d.-sourced systems at 0·9 and 1·3 μm.

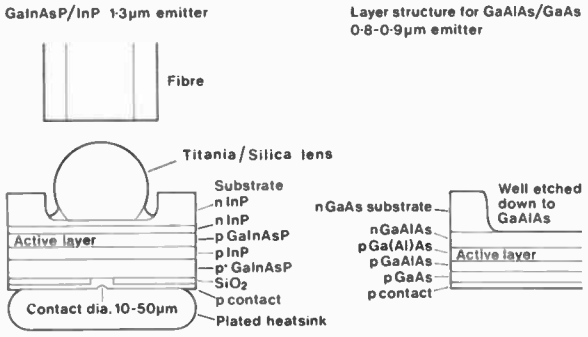


Fig. 2. The lens-coupled high radiance l.e.d.

double heterojunction l.e.d.s mentioned above. A third alternative is the simple diffused device, where the junction is made by zinc diffusion into an n-type GaAs substrate. The basic unlensed device is derived from the configuration first adopted by Burrus.[1] The current is confined to flow through a small central area (10–50 μm in diameter) by either dielectric isolation by a film of SiO_2 or by proton implantation of the passive area of the device, which renders the material semi-insulating. Efficient radiative recombination occurs between electrons and holes in the relatively narrow bandgap active layers of the double heterojunction devices and between injected electrons and majority holes in the p side of the junction in the case of the Zn-diffused device. Light, which is radiated isotropically, exits from the top face of the device (into which is etched a well to reduce absorption), is collected by the lens and coupled into the fibre. Heat is effectively removed from the device by the integral plated gold heatsink, which is 5–10 μm thick. Overall chip dimensions are typically 200 × 300 μm, the etched well being offset to enable a thermocompression gold wire bond to be made to the n-side of the device. A bonded and lensed device is shown in Fig. 3. Here the lens diameter is ~110 μm and the device is of the GaAlAs/GaAs type.

The fraction of internally generated light which can be coupled into a fibre is severely limited by the high refractive index of the semiconductor material. In a GaAs device, with a refractive index of 3·5, only ~2% of the light escapes through a planar GaAs/air interface, and only ~4% of that 2% would be coupled into a 0·2*NA* optical fibre.

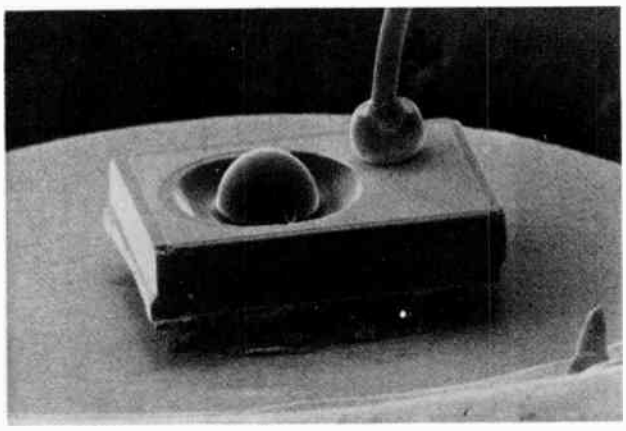


Fig. 3. Scanning electron micrograph of lensed, surface-emitting GaAlAs/GaAs l.e.d.

This situation can be greatly improved by the use of microlens coupling, which can increase the power launched into a fibre by up to ~ 30 times,[6] depending on fibre parameters. For efficient coupling, the unlensed emitter diameter must be considerably smaller than that of the fibre, typical area ratios being 10–20.

Figure 4 shows the theoretical maximum powers which can be coupled into step index fibres for optimally lensed, 100% efficient l.e.d.s with reflecting rear contacts.

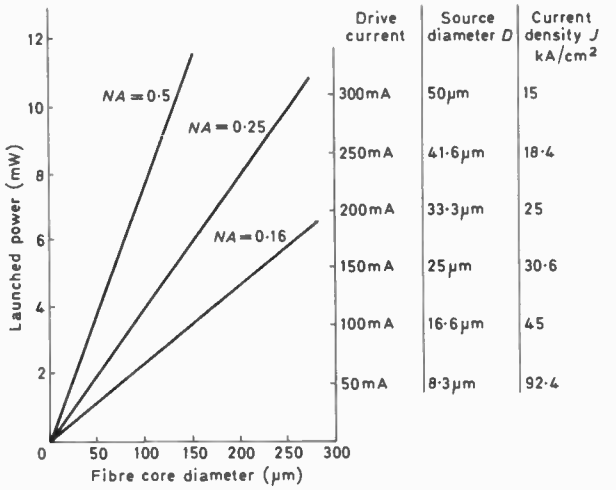


Fig. 4. 1·3 μm power coupled vs fibre size for optimum diode and lens geometry.

The figure must be read by selecting fibre parameters, then reading unlensed source diameter, operating current and coupled power. The lens parameters, diameter and truncation, are selected for given fibre parameters.[6] In practice, power levels up to ~$\frac{1}{3}$ of these predicted values have been obtained from GaAlAs/GaAs and GaInAsP/InP l.e.d.s into certain fibres. Difficulties arise, reducing the practical launched powers, particularly with the small-core telecommunications type fibres, as the optimum source diameter is ~10 μm. Saturation of the l.e.d. output can occur with such small devices, due to both thermal and non-thermal effects.[7] In addition, it becomes difficult to control current spreading in the device, which gives rise to wasted light generation outside the desired emission area. Consequently some design compromises have to be made. At present best performance with 50 μm core fibre is obtained for l.e.d.s with emitting areas of 20 μm diameter.

4 L.E.D. Performance

4.1 Bandwidth and Efficiency

The bandwidths of the surface emitting l.e.d.s are in general determined by three mechanisms, these being the

doping level in the active layer, device parasitic capacitance and the reduction in radiative lifetime due to the injected plasma. The first mechanism results in a radiative lifetime equal to $1/Bp$, where B is the radiative recombination coefficient, typically $1\text{–}2\times10^{-10}\,\text{cm}^3\cdot\text{s}^{-1}$, and p is the doping level in the active region. Consequently to achieve a switching speed of 5 ns, $\sim 10^{18}$ acceptors cm^{-3} are required.

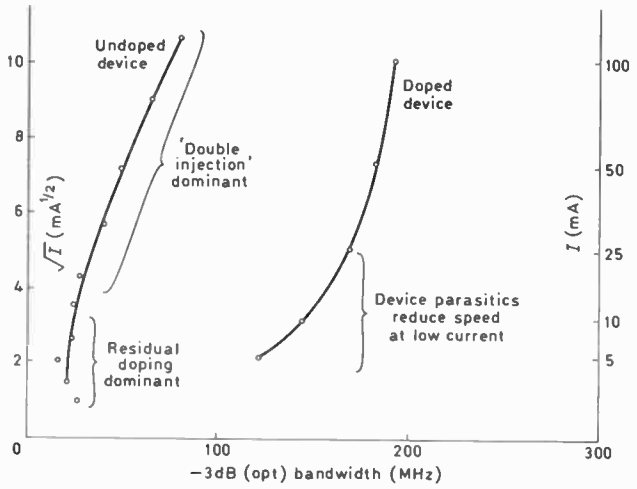


Fig. 5. Small signal frequency responses of oxide isolated 1·3 μm l.e.d.s.

If a device is made with a thin, lightly or undoped active layer, the injected hole density can exceed that from the acceptor doping. As the probability for recombination is proportional to the product of the electron and hole populations, the lifetime varies with current drive, giving a frequency response proportional to $\sqrt{(J/d)}$, where J is the operating current density and d is the active layer thickness.[8] Figure 5 shows the small signal frequency responses of two quaternary 1·3 μm l.e.d.s with different active layer doping levels. Here the frequency response is plotted against $\sqrt{I}$, where I is the drive current, so that the different speed determining mechanisms are apparent. The reduction in bandwidth of the heavily doped device at low current due to device parasitics in the oxide isolated device shown can be improved by the use of proton implant isolation, which reduces the effects of passive junction capacitance.

The technique of using thin active layers, giving a high recombination plasma density, can result in devices with frequency responses in excess of 300 MHz,[3,9] but such devices can have relatively long fall-times for pulse modulation or suffer more from saturation than doped devices with thicker active layers.[10]

The sub-linear light current characteristic generally observed for small area surface emitting l.e.d.s is due to the combined effects of junction temperature rise, increasing non-radiative recombination and in-plane superluminescence.[7] The non-thermal saturation mechanisms can be controlled by increasing the active layer doping to decrease the radiative lifetime and hence increase the current for the onset of plasma-density-dependent non-radiative effects. Figure 6 shows the light current characteristics, recorded under pulsed conditions to eliminate heating effects, for two quaternary 1·3 μm l.e.d.s with different active layer doping levels. It can be seen that more linear but reduced low-current efficiency is seen for the heavily doped devices. For small area devices operating at high current density, there can be an efficiency advantage in using heavy doping, as undoped or lightly doped devices could be heavily saturated.

Internal efficiencies fall from around 60% for lightly doped ($5\times10^{17}\,\text{cm}^{-3}$) 1·3 μm devices (assuming 100% reflecting back contact) to $\sim$ 25% for heavily doped devices ($\sim 10^{19}\,\text{cm}^{-3}$) which exhibit rise and fall times of 1·5 ns. Consequently improvements in power-bandwidth products of up to ten times can be achieved by fabricating devices with heavily doped active layers.

Rather higher efficiencies are obtained for fast GaAlAs/GaAs 0·85 μm double heterostructure devices. Here internal efficiencies above 50% have been obtained for devices with 2 ns rise and fall times.

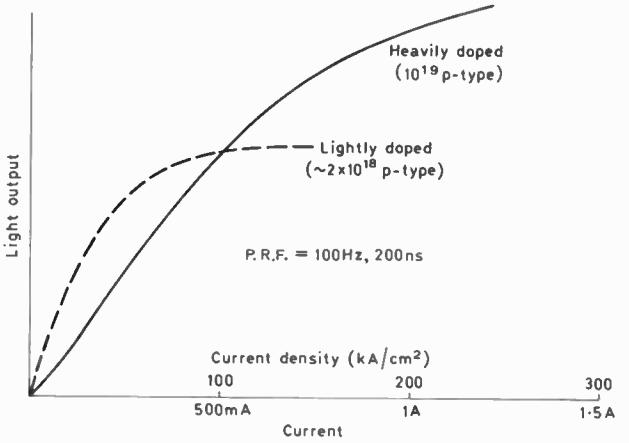


Fig. 6. Pulsed light current characteristics for GaInAsP/InP l.e.d.s with lightly and heavily doped active layers.

Figure 7 (a) shows the typical pulse response (0–150 mA) for a heavily doped 1·3 μm l.e.d. Figure 7 (b) shows the response to a 160 M baud NRZ pseudo-random binary sequence for lightly and heavily doped devices showing a clean, open 'eye' for the heavily doped device.

4.2 Launch Powers

Table 1 shows the launch powers which have been obtained into various fibres from GaAs, GaAlAs and GaInAsP surface emitting l.e.d.s. In all cases the cladding modes were stripped and the fibre apertured into a detector at the stated NA of the fibre.

For telecommunications applications, the 50 μm, 0·2 NA graded index fibre is most important. Here power levels in excess of 250 μW (d.c.) have been obtained with GaAlAs/GaAs devices and 70 μW (d.c.) has been obtained with 1·3 μm emitters. The lower power from the 1·3 μm device is due to saturation effects and should be

improved by further device optimization. Nevertheless, this power level is sufficient for many long-haul, high data rate applications.

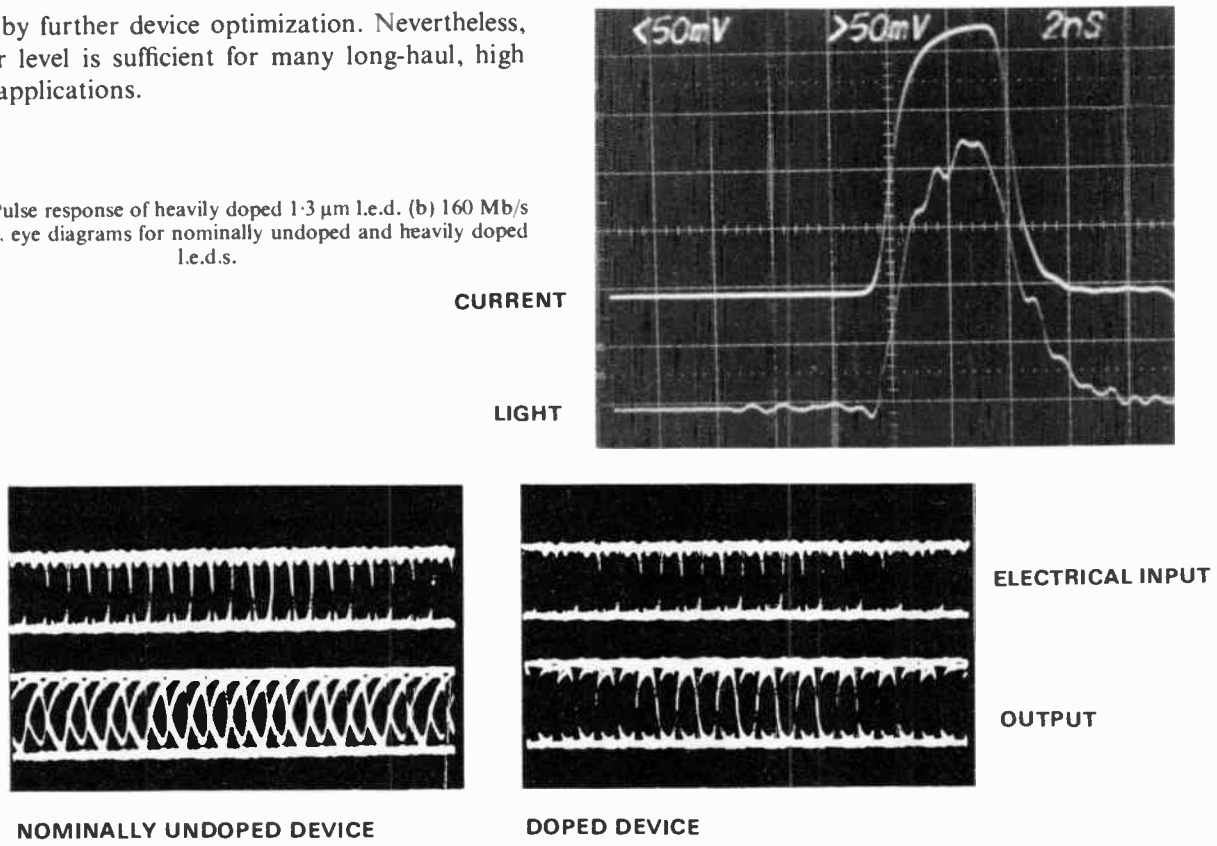


Fig. 7. (a) Pulse response of heavily doped 1·3 μm l.e.d. (b) 160 Mb/s NRZ p.r.b.s. eye diagrams for nominally undoped and heavily doped l.e.d.s.

4.3 Spectral Linewidths

Although the disadvantage of large spectral linewidth inherent in the l.e.d. largely disappears at 1·3 μm because of the low material dispersion of the fibre, control and reduction of linewidth is still highly desirable to improve reproducibility and to reduce the need for receiver equalization. The linewidth of 1·3 μm l.e.d.s increases with increasing doping level due to the formation of band tail states. Figure 8 shows the linewidths of two devices, with Zn doping levels of 5×10^{17} cm^{-3} and 10^{19} cm^{-3} in the active layer, and such devices exhibit rise times of 15 ns and 1·5 ns respectively. The linewidths (full width at half maximum) increase from 75 to 95 nm with increasing doping level. A shift to lower energy is also apparent so that the active layer composition must be adjusted to maintain the same centre wavelength.

4.4 Thermal Behaviour

The optical power from a l.e.d. changes only relatively slowly with temperature so that optical feedback to stabilize the working point is unnecessary. Figure 9 shows the relative optical output as a function of temperature for a 1·3 μm GaInAsP/InP l.e.d. In the range 20–100°C the output falls by $\sim 0{\cdot}7\%$ deg C^{-1}, reaching half power at ~ 130°C. GaAlAs/GaAs devices tend to be even less temperature sensitive, falling by $\sim 0{\cdot}5\%$ deg. C^{-1}.

The peak spectral output shifts to lower energy with

Table 1

Device Type	Bandwidth at max. current	τ_r	Coupled power 50 μm 0·2 *NA*, graded index	Current	Coupled power 100 μm 0·3 *NA*, step index	Current
GaInAsP/InP double heterostructure	50 MHz	7 ns	70 μW (d.c.)	100 mA	400 μW (d.c.)	100 mA
GaInAsP/InP double heterostructure	300 MHz	1·5 ns	38 μW (d.c.)	100 mA	—	
GaAlAs/GaAs double heterostructure	150 MHz	3 ns–5 ns	250 μW (d.c.)	150 mA	2·05 mW (d.c.)	150 mA
GaAlAs/GaAs double heterostructure	300 MHz	1·5–2 ns	—	—	1·2 mW (d.c.)	150 mA
GaAs zinc diffused	30 MHz	10 ns	150 μW (d.c.)	150 mA	800 μW (d.c.)	150 mA

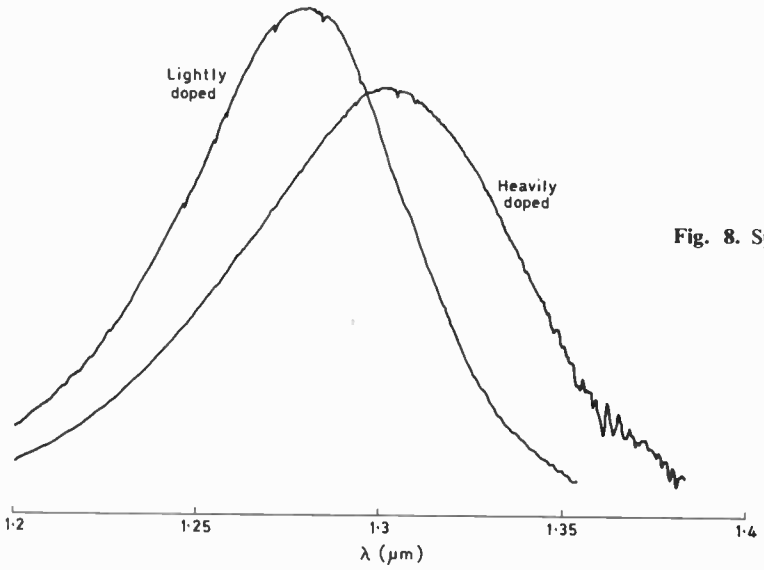


Fig. 8. Spectral linewidths of heavily ($\sim 10^{19}$ cm^{-3}) and lightly ($\sim 5\times10^{17}$ cm^{-3}) doped 1·3 μm l.e.d.s.

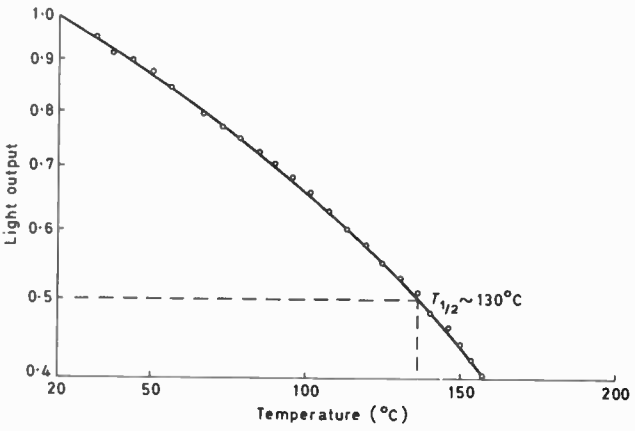


Fig. 9. Optical output vs temperature for 1·3 μm l.e.d.

increasing temperature, by 0·4 nm deg. C^{-1} for 1·3 μm quaternary devices. This shift is normally insignificant compared to the linewidth.

5 Advanced Devices

A number of devices have been investigated which should show an improvement in efficiency over the lensed planar surface emitting l.e.d. Two devices in particular are the 'hybrid' edge-surface emitting device[11] and the integral lens l.e.d.[12] The former uses a technique of selective etching to produce a faceted structure on the rear surface of the l.e.d. and has a potential gain of three times over a conventional device. Light travelling in the plane of the device (Fig. 10) is reflected by the 90° facets towards the 45° diagonal facet and consequently exits from the top face of the l.e.d. Figure 11 shows a s.e.m. of an etched mesa and the front-face near-field electroluminescence from a complete device. The increased light output along the diagonal facet can clearly be seen. In present devices inaccuracy in the facet etching limits the increase in front-face efficiency to a factor of two over planar devices.

The integral lens l.e.d. has been investigated by several authors[12,13,14] and has great potential both as a low-current, high-power source for small core fibres and as a very-high-power device for large core fibres. In such a device, a lens is formed directly in the semiconductor material rather than being fabricated in a glass and then fixed to the l.e.d. chip with adhesive. This leads to the elimination of the planar semiconductor/adhesive/lens interfaces, which limit the maximum lens gains of the device discussed in Section 3. For optimized devices, efficiencies for coupling into fibres should exceed 15%.[15] Figure 12 shows a s.e.m. of an integral lens fabricated in GaAlAs/GaAs. Here the lens has been formed by infilling hemispherical depressions in a GaAs substrate by liquid-phase epitaxy with GaAlAs. After growth of the rest of the diode structure, the lens is revealed by a preferential

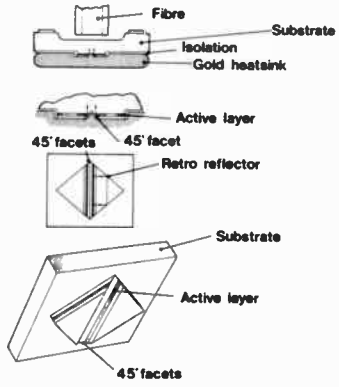


Fig. 10. Hybrid double heterostructure surface emitter.

etch. Such devices have given increases in coupled power of five times compared to planar l.e.d.s, close to the theoretical for the geometries employed. However, to exploit fully the potential of such a device, the unlensed emitting area must be less than 5 μm in diameter whereas present devices are in the range 20–25 μm.

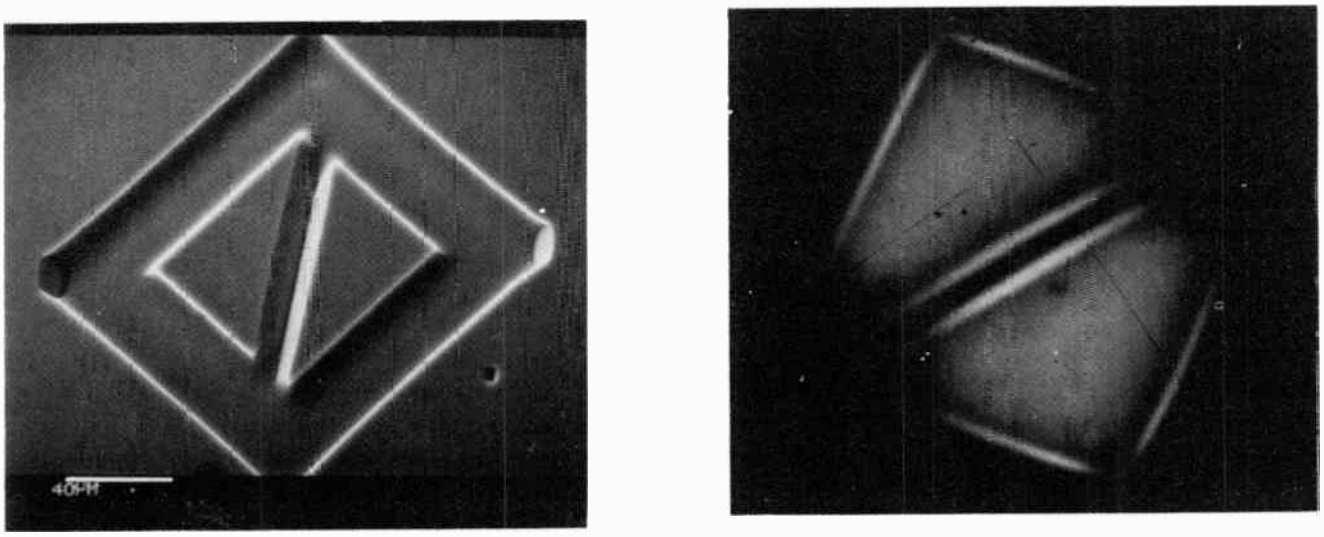


Fig. 11. Scanning electron micrograph of hybrid mesa and front-face electroluminescence of hybrid device.

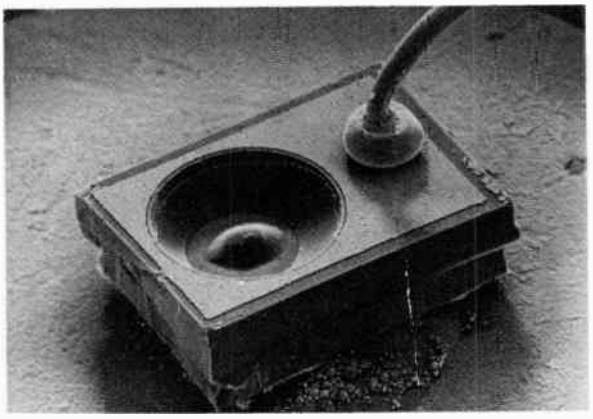


Fig. 12. Scanning electron micrograph of integral lens l.e.d.

6 Reliability

The surface emitting high radiance l.e.d. is a particularly long lived, reliable device. GaAs zinc-diffused devices have been operating at Caswell since 1974 and results of high temperature accelerated ageing indicate lives at room temperature around $\sim 5 \times 10^5$ hours, the ultimate failure mechanism being gold diffusion from the contact. Similarly good lives have been found by ourselves and others[16,17] for GaAlAs/GaAs double heterostructure l.e.d.s.

Although the quaternary GaInAsP/InP l.e.d.s are less well developed than the corresponding GaAlAs devices, they seem particularly resistant to degradation. Devices at 45°C have operated for over 15 000 hours with little change in light output. Devices operated at 170°C, in an attempt to increase degradation rates, have operated for over 5000 hours, with only two failures in a batch of 16. At this temperature, GaAs devices would fall to half output within 1000 hours. Further lifetest data are needed at intermediate and higher temperatures to enable reliable predictions to be made of device life at room temperature.

7 Conclusions

High-power, fast, reliable l.e.d. sources have been developed for all wavelengths of current interest in fibre optics. The range of devices includes 1·3 μm l.e.d.s suitable for long-haul, high data rate telecommunication links. Such devices will play an important role in 'second generation' 140 Mb/s land based systems.

GaAlAs/GaAs double heterostructure l.e.d.s capable of launching more than 2 mW into 100 μm core 0·3 *NA* fibre are eminently suitable devices for short-haul, high-loss analogue and digital links. The long life and high power of these devices makes them particularly attractive for avionic and military as well as civil applications.

8 Acknowledgments

The data given in this paper are the results of the effort of many people in the infra-red source research and development groups of the Allen Clark Research Centre. In particular the author would like to thank R. C. Goodfellow, R. Davis, R. Plastow, M. Harding, I. Griffith, J. Humpage, J. Ure, R. Bradley and M. Faultless. This work was sponsored in part by the Procurement Executive, Ministry of Defence, and by British Telecom Research Laboratories, Martlesham Heath, Ipswich.

9 References

1 Burrus, C. A. and Miller, B. I. 'Small area, double heterostructure aluminium-gallium arsenide electroluminescent diode sources for optical fibre transmission lines', *Opt. Commun.*, 4, no. 4, pp. 397–9, 1971.

2 Zargar'yants, M. N., Mezin, Yu. S. and Kolanenkova, S. I. 'Electroluminescent diode with a flat surface emitting continuously 25 W/cm²-steradian at 300 K', *Sov. Phys. Semicond.*, 4, pp. 1371–2, February 1972.

3 Ettenberg, M., Hudson, K. and Lockwood, H. 'High-radiance light-emitting diodes', *IEEE J. Quantum Electronics*, **QE-9**, no. 10, pp. 987–91, October 1973.

4 Davies, I. G. A., Bricheno, T., Goodwin, A. R. and Plumb, R. G. 'A GaAlAs edge emitting l.e.d. with a launched power in excess of 1 mW', Proc. 6th European Conference on Optical Communication York. September 1980, pp. 199–202.

5 Stewart, W. J. 'Wavelength filtering effects in multimode fibres', Proc. 5th European Conf. on Optical Communication Amsterdam, September 1979, Paper 12.3.

6 Abram, R. A., Allen, R. W. and Goodfellow, R. C. 'The coupling of light-emitting diodes to optical fibres using sphere lenses', *J. Appl. Phys.*, **46**, p. 3468, 1975.

7 Goodfellow, R. C., Carter, A. C., Rees, G. J. and Davis, R. 'Radiance saturation in small area GaInAsP/InP and GaAlAs/GaAs l.e.d.s', *IEEE Trans. on. Electron Devices*, **ED–28**, no. 4, pp. 365–71, April 1981.

8 Namizaki, H., Kan, H., Ishii, M. and Ito, A. 'Current dependence of spontaneous carrier lifetime in GaAlAs/GaAs double heterojunction lasers', *Appl. Phys. Letters*, **24**, pp. 486–7, May 1974.

9 Goodfellow, R. C., Carter, A. C., Griffith, I. and Bradley, R. R. 'GaInAsP/InP fast, high radiance, 1·05–1·3 μm wavelength l.e.d.s with efficient lens coupling to small numerical aperture silica optical fibres', *IEEE Trans. on Electron Devices*, **ED—26**, no. 8, pp. 1215–20, August 1979.

10 Carter, A. C., Goodfellow, R. C. and Davis, R. '1·3 μm GaInAsP/InP l.e.d.s for 140 and 280 Mb/s fibre optic systems', Proc. 6th European Conference on Optical Communication, York, September 1980, pp. 211 4.

11 Carter, A. C., Goodfellow, R. C. and Griffith, I. 'A new surface emitting GaInAsP 1·3 μm l.e.d. with up to threefold enhancement in external quantum efficiency', Proc. IEDM, Washington, December 1979, pp. 118–21.

12 Ure, J., Humpage, J., Carter, A. C. and Goodfellow, R. C. 'The integral lens light emitting diode', Paper presented at the 4th IEEE Specialist Conference on the Technology of Electroluminescent Diodes, Brighton, September 1980.

13 King, F. D. and Springthorpe, A. J. 'The integral lens coupled l.e.d.', *J. Electronic Materials*, **4**, no. 2, pp. 243—52, 1975.

14 Alferov, Z. I., Andreev, W. M., Egoros, B. V. and Syrbu, A. V. 'Heterojunction light emitting AlGaAs diodes formed by negative profiling of the substrate', *Sov. Phys. Semicond.*, **7**, no. 10, pp. 1123–27, October 1977.

15 Abram, R. A. and Goodfellow, R. C. 'Coupling efficiency calculations on an integrated l.e.d. sphere lens source for optical fibres', *Electronics Letters*, **16**, no. 1, pp. 14–16, 3rd January 1980.

16 Abe, M. *et al.*,'High efficiency long lived GaAlAs l.e.d.s for fibre optical communications', *IEEE Trans. on. Electron Devices*, **ED 24**, no. 7, pp. 990–4, July 1977.

17 Dawson, L. R., Keramidas, V. G. and Ziptel, C. L. 'Reliable, high speed l.e.d.s for short haul optical data links', *Bell Syst. Tech. J.*, **59**, no. 2, pp. 161–8, February 1980.

Manuscript first received by the Institution on 16th February 1981 and in final form on 13th April 1981
(Paper No 1997/CC 341)

Standard Frequency and Time Service

Communication from the National Physical Laboratory

Relative Phase Readings in Microseconds NPL—Station

(Readings at 1500 UTC)

APRIL 1981	MSF 60 kHz	GBR 16 kHz	Droitwich* 200 kHz
1	2·4	12·2	42·0
2	2·5	10·2	42·5
3	2·5	10·3	43·0
4	2·6	9·8	43·1
5	2·5	10·3	43·5
6	2·5	10·3	43·9
7	2·6	10·3	44·3
8	2·3	10·1	44·7
9	2·2	12·2	45·1
10	2·4	10·9	45·4
11	2·3	10·8	45·7
12	2·3	10·3	46·0
13	2·3	10·2	46·3
14	2·2	9·8	46·6
15	2·1	9·9	46·9
16	2·2	10·9	47·2
17	2·2	10·9	47·5
18	2·1	10·8	47·8
19	2·1	10·6	48·2
20	2·1	10·6	48·6
21	2·2	10·9	48·9
22	2·2	11·2	49·3
23	2·2	11·0	49·7
24	2·1	11·8	50·1
25	2·1	11·0	50·5
26	2·2	11·4	50·7
27	2·3	11·8	51·1
28	2·2	11·9	51·5
29	2·1	9·7	51·9
30	2·1	10·3	52·3

MAY 1981	MSF 60 kHz	GBR 16 kHz	Droitwich* 200 kHz
1	2·2	11·7	52·6
2	2·2	10·8	52·9
3	2·1	10·6	53·1
4	2·2	10·8	53·6
5	2·2	10·6	54.0
6	2·1	10·5	54·4
7	2·2	10·5	54·7
8	2·0	10·8	55·0
9	2·3	9·8	55·3
10	—	—	55·5
11	2·2	11·1	55·8
12	2·1	10·5	56·2
13	2·2	18·8	56·4
14	2·3	19·1	56·7
15	2·2	19·0	57·0
16	2·2	18·9	57·3
17	2·2	18·8	57·6
18	2·0	19·2	58·0
19	2·0	18·6	58·3
20	1·9	18·4	58·6
21	2·0	18·6	58·8
22	2·0	18·6	59·0
23	1·8	18·8	59·2
24	1·9	18·7	59·5
25	1·8	17·8	59·9
26	1·8	17·1	60·1
27	1·7	16·8	60·3
28	1·6	18·3	60·6
29	1·6	—	60·9
30	1·5	—	61·2
31	1·5	—	61·4

Notes: (a) Relative to UTC scale (UTC_{NPL}–Station) = +10 at 1500 UT, 1st January 1977.
(b) The convention followed is that a decrease in phase reading represents an increase in frequency.
(c) 1 μs represents a frequency change of 1 part in 10^{11} per day.
*It may be assumed that the satellite stations on 200 kHz at Westerglen and Burghead will follow the day to day changes in these phase values.

UDC 621.383.52: 621.391.63/4
Indexing Terms: Semiconductor diodes, Photodiodes, Fibre optics, Optical communication

Receivers for optical fibre communications

I. GARRETT, M.A., Ph.D., C.Eng., M.I.E.E.*

SUMMARY

This paper is a review of today's state-of-the-art in receivers for optical fibre communications, both in practice and in theory. The functions of an optical receiver are discussed, and a detailed description of unity-gain photodiodes is given. The problem of detection in the presence of noise is treated next, and the difference between optical and radio-frequency systems is highlighted. The theory of avalanche photodiodes is summarized, and state-of-the-art devices and their performance described. The final section is a review of the theory of optical receivers, with comparisons of theoretical predictions with practical achievements, especially for the sensitive p-i-n f.e.t. receivers.

* *British Telecom Research Laboratories, Martlesham Heath, Ipswich IP5 7RE.*

1 Introduction: Functions of an Optical Receiver

At the receiving end of an optical fibre communication system, we have a very weak optical signal carrying information which we want to convert into something more manageable, usually a strong electrical signal carrying a good estimate of the signal originally transmitted. The receiver thus has three functions:

1. Conversion from optical to electrical signal
2. Amplification
3. Estimation of the message originally transmitted.

The first function is performed by a photodetector, and that is the subject of the first part of this paper. The second function, amplification, is not specific to optical systems, and nothing will be said about it here. Estimation of the message involves dealing with noise and various system impairments, and is a very large subject: only the more basic ideas will be covered here. In these three functions, an optical receiver seems similar to a radio receiver; however, it is quite different in the way it performs these functions. Typically a radio receiver uses heterodyne detection. The arriving signal is mixed with a local oscillator to produce a modulated intermediate frequency carrier which is then amplified and filtered as required. Heterodyne detection is very sensitive and provides excellent rejection of adjacent channels. It also allows one to make use of the phase and frequency information in the modulated carrier. But it requires that the arriving signal is coherent, that the local oscillator can be made to track the carrier frequency accurately, and that the signal and local oscillator are matched in polarization and phase-front geometry. These conditions are not met in optical fibre communication links using present fibres and semiconductor lasers. Other types of laser, being more coherent, have successfully been used for coherent homodyne detection in atmospheric systems, but these lasers have engineering disadvantages. The most recent semiconductor lasers have much greater coherence length than the earlier ones. And there is the possibility of using temperature control to alter the wavelength of the local oscillator and so track the carrier frequency. So coherent homodyne detection is not ruled out by any means, but is very difficult to implement at present. All practical fibre optic communication systems use incoherent (direct) detection, in which only optical power variations are detected, there being no information in the phase and frequency variations.

2 Unity Gain Detectors

2.1 Requirements for an Optical Detector

The device used for converting the optical signal to an electrical form must be efficient at the operating wavelength and must respond at a speed appropriate to the message data rate or frequency band. We may also require a linear response, operation at ambient temperature from a convenient voltage supply, and we

0033–7722/81/070349+13 $1.50/0

would have preference for a small, light, cheap and reliable device. Semiconductor photodiodes fit all these requirements remarkably well, and there is little interest in other types of detector for optical communications, at least in normal terrestrial environments. Photoconductive detectors have inferior noise performance except when the incident optical power level is high; pyro-electric detectors can only be made fast at the expense of sensitivity, and photomultipliers offer no advantage in sensitivity when, as is normally the case, the optical power level on zero bits is not zero.

2.2 Principle of Operation

A photodiode is a reverse-biased p-n junction formed in a semiconductor material. Photons are absorbed in the semiconductor and create electron-hole pairs. These carriers can be separated by an electric field, such as exists in the depletion region of a p-n junction, and so give rise to a photocurrent in the external circuit. To convert light efficiently into electrical current, the semiconductor material must have a high absorption coefficient at the wavelength of the light, so that different materials are appropriate for different wavelength ranges. The speed of response is governed by:

(i) the time taken for the photogenerated electrons and holes to reach the terminals of the device,
(ii) the RC time-constant of the measuring circuit, which may be affected or even dominated by the junction capacitance of the photodiode.

Photogenerated carriers travel across the device to the terminals from the points at which they are generated by diffusion and by drift in any internal field. The rate of diffusion is generally so slow that except in very thin layers most carriers are lost by recombination and do not contribute to the photocurrent. The device is made fast and efficient by ensuring that the incident photons are absorbed in the high-field depletion region of the junction.

2.3 Device Structures and Materials

Figures 1 and 2 illustrate two photodiode structures which are used in practice. The first is a silicon device designed for the wavelength range 0·8 to 0·9 μm, and has a thick depletion region, 30 to 100 μm thick, formed in low-doped material. The absorption coefficient of silicon in this wavelength range is 950 to 350 cm^{-1}, so that several tens of microns of material are needed for almost complete absorption. Very little of the incident radiation is absorbed in the undepleted n^+ layer at the surface, which is only about 1 μm thick. The device is designed so that the field required to deplete it fully is well below the breakdown field strength, but sufficiently high to accelerate the carriers to their scattering-limited velocity, around 10^7 cm s^{-1} in many semiconductors at room temperature, resulting in a response time of about 10 ps per micron of depletion region (see Sect. 2.6).

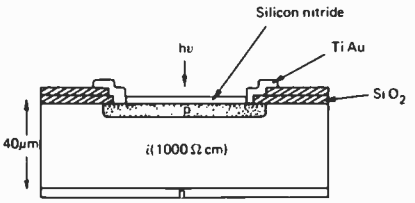


Fig. 1. Schematic cross-section of a silicon p-i-n photodiode for wavelengths from 0·8 to 1 μm.

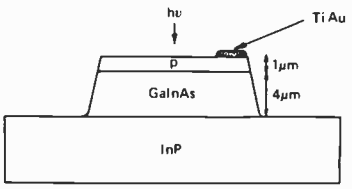


Fig. 2. Schematic cross-section of a GaInAs/InP p-i-n photodiode for wavelengths from 1 to 1·6 μm.

For wavelengths in the near-infra-red up to about 1 μm, good photodiodes may be made in silicon. The absorption coefficient is greater than 600 cm^{-1} at a wavelength of 0·85 μm,[1] typical of gallium-aluminium arsenide lasers, which means that 95% of photons entering the material are absorbed in a 50 μm thick layer. At longer wavelengths the absorption coefficient is smaller, and a thicker layer is needed to absorb the light efficiently. The surface of the device is usually coated with a dielectric layer to reduce reflection. High-speed silicon photodiodes, such as are required for fibre optic communications, are designed to have a wide depletion region, typically 50 μm, with a very low doping level, so that fast response is obtained with a moderate applied voltage. Such a device is known as a p-i-n photodiode, the i-region being nearly intrinsic. The wide depletion layer reduces the junction capacitance, too. The device shown in Fig. 1 is 100 μm in diameter, and has a capacitance of less than 0·1 pF.

At wavelengths beyond 1 μm, silicon becomes increasingly transparent. Although silicon photodiodes have been made which respond at 1·1 μm, the sensitivity is poor and the response is slow because of the thickness of the absorbing region. For wavelengths beyond 1 μm, a different material is required for photodiodes intended for communication systems. An obvious choice is germanium, which has a bandgap of 0·66 eV and so should be sensitive out to 1·8 μm or so, well beyond the optimum transmission wavelengths of 1·3 and 1·55 μm. The small bandgap of germanium is something of a disadvantage in fact, because, coupled with the high density of states in the conduction band, it means that the reverse bias dark current is large,[1] which degrades the performance of an optical receiver. The other possible materials are the so-called III–V compounds; binary compounds of elements from groups IIIb and Vb, such as gallium arsenide and indium phosphide. To detect light at 1·55 μm, a material with a bandgap no greater than 0·8 eV is needed, and in

fact none of the binary III–V compounds has such a bandgap. But many of the III–V compounds form extensive solid solutions with each other, and the mixed compounds have properties intermediate between those of the binaries. So it looks as if there ought to be a wide choice of materials, and in principle that is so. In practice the choice is limited because of the techniques available for preparing these materials in sufficiently pure and perfect form. The most usual materials for detectors in this range are the ternary compound InGaAs and the quaternary InGaAsP. In either material, the bandgap can be adjusted over a wide range by selecting a suitable composition. The reverse bias dark current is smaller than in germanium by one or two orders of magnitude typically, because of the much smaller density of states in the conduction band. The device illustrated in Fig. 2 has an absorbing layer of InGaAsP deposited on an InP substrate, with the p-n junction formed by diffusing a dopant such as zinc into the absorbing layer. This device is designed for use in the wavelength range 1 to 1·6 μm.[2,3] In this range of wavelengths, the InGaAs layer has a high absorption coefficient, around $10^4\,\mathrm{cm}^{-1}$, so that a large fraction of the incident radiation is absorbed in the undepleted p^+ region at the surface even if it is only 1 μm thick. Many of the carrier pairs formed in this region are lost by surface recombination or by recombination within this layer, so that the efficiency is reduced considerably. This problem can be surmounted by arranging for the light to be incident through the back of the device, i.e. through the InP substrate, which is transparent in the wavelength range of interest. Because the absorption coefficient is high, only a thin layer is needed, 3 to 10 μm, so that the response is fast too.

2.4 Quantum Efficiency

With absorption coefficient α and depletion region thickness d, assuming that no photons are absorbed outside the depletion region, the internal quantum efficiency is given by:

$$\eta_{\mathrm{int}} = 1 - \mathrm{e}^{-\alpha d}. \tag{1}$$

If some photons are absorbed near the surface in an undepleted layer of thickness l, the internal quantum efficiency is reduced:

$$\eta_{\mathrm{int}} = \mathrm{e}^{-\alpha l} - \mathrm{e}^{-\alpha d}. \tag{2}$$

Generally one would aim to make the fraction of carriers absorbed in the undepleted layer as small as possible.

The external quantum efficiency is less than the internal quantum efficiency because not all photons incident on the device enter it; a fraction are reflected. If the semiconductor has refractive index n_1 and the surrounding medium has index n_0, the fraction of the optical power reflected is:

$$R = \frac{(n_1 - n_0)^2}{(n_1 + n_0)^2}. \tag{3}$$

Thus a silicon–air interface ($n_1 = 3{\cdot}4$) has 30% reflection, GaAs ($n_1 = 3{\cdot}3$) has 29%, InAs ($n_1 = 3{\cdot}8$) has 34%. By coating the surface with a half-wavelength thickness of a material of intermediate refractive index, the reflection is much reduced. For normal incidence, zero reflected power is obtained if the refractive index n_f of the film is given by:

$$n_f^2 = n_0 n_1 \tag{4}$$

and the thickness d_f is an odd number of quarter wavelengths:

$$n_f d_f = (2k+1)\lambda/4. \tag{5}$$

The optimum film thickness is also dependent on the wavelength and the angle of incidence, but not critically. Although it is not possible to produce completely non-reflecting single-layer coatings for practical photodiodes, the reflected power can be reduced to a few percent.

2.5 Depletion Voltage

The voltage required to fully deplete a layer of thickness d and dopant concentration N_b in a material of relative dielectric constant ε is given by:

$$V + V_{bi} = \frac{qN_b d^2}{2\varepsilon\varepsilon_0} \tag{6}$$

where V_{bi} is the built-in potential at the p-n junction.

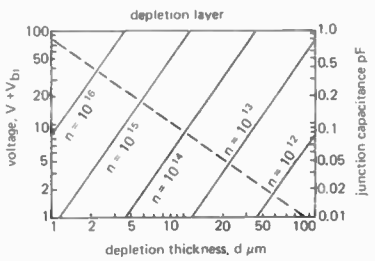


Fig. 3. Depletion voltage and junction capacitance as functions of depletion layer thickness for a 100 μm diameter diode.

Figure 3 plots this relationship for a material with relative dielectric constant 12, which is approximately correct for typical photodiode structures. Doping levels of 10^{12} to $10^{13}\,\mathrm{cm}^{-3}$ are available in silicon, so that a few tens of microns can be depleted at 5 to 10 volts, as required. In the III–V mixed compounds, the lowest doping levels achieved to date are around $10^{15}\,\mathrm{cm}^{-3}$, so that 15 to 20 volts are required to deplete a few microns.

2.6 Junction Capacitance

The capacitance of the photodiode is an important parameter in determining the receiver noise properties as well as the speed of response. The main contribution is generally the depletion layer capacitance which is given by:

$$C_d = A\varepsilon\varepsilon_0/d \tag{7}$$

where A is the junction area. This expression is plotted

on Fig. 3 as the broken line for a 100 μm diameter device. The contribution from the metallization in practical devices can be kept to below 0·1 pF on devices with an active region diameter of 100 μm or so. One can see from the figure that silicon devices with depletion layers a few tens of microns thick have a depletion layer capacitance of below 0·1 pF on a 100 μm diameter device. Devices fabricated in III–V materials with depletion layers a few microns thick have junction capacitance around 0·2 to 0·5 pF.

2.7 Impulse and Step Response

We consider here the speed of response as determined by the transit time for carriers across the depletion region. Generally, holes have the lower mobility and lower scattering-limited velocity, so that the response time is longer if the light is incident on the n^+ side of the device. For this analysis we assume that the carrier velocity is constant across the depletion region, so that either carriers are travelling with their scattering-limited velocity or else the electric field is essentially constant.

Considering first light incident on the p side, let the rate of generation of carrier pairs as a function of depth be given by:

$$g(x) = g_0 \, e^{-\alpha x} \, cm^{-3} \, s^{-1}. \tag{8}$$

The hole density $N_h(x)$ produced by an impulse of light of duration dt is then:

$$N_h(x) = g_0 \, e^{-\alpha x} \, dt. \tag{9}$$

The hole contribution to the photocurrent is then:

$$i_h(t) = v_h g_0 q \, dt \int_{v_h t}^{d} e^{-\alpha x} \, dx \tag{10}$$

$$= v_h g_0 q \, dt \, (e^{-\alpha v_h t} - e^{-\alpha d})/\alpha \qquad 0 < t < d/v_h \tag{11}$$

$$= 0, \qquad t > d/v_h$$

Similarly, the electron contribution to the impulse response is given by:

$$i_e(t) = v_e g_0 q \, dt \, (1 - e^{-\alpha(d - v_e t)})/\alpha \qquad 0 < t < d/v_e \tag{12}$$

$$= 0, \qquad t > d/v_e$$

The step response is found by integrating the impulse response with respect to time. We normalize so that the step response tends to unity current as t becomes large. The total response is then:

$$I(t) = [\alpha v_e t + (1-\eta)(1 - e^{\alpha v_e t}) + 1 - e^{\alpha v_h t} - \alpha v_h t(1-\eta)]/\eta\alpha d \qquad 0 < t < d/v_e \tag{13}$$

$$= [\alpha d - \eta + 1 - e^{-\alpha v_h t} - \alpha v_h t(1-\eta)]/\eta\alpha d \qquad d/v_e < t < d/v_h \tag{14}$$

$$= 1, \qquad t > d/v_h$$

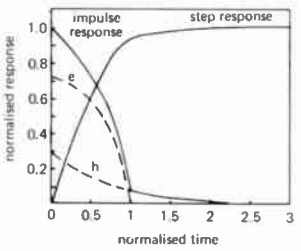


Fig. 4. Impulse and step responses of a p-i-n diode illuminated from the p-side.

This result is plotted in Fig. 4 as the normalized response as a function of normalized time. The time axis has been normalized by putting $t' = tv_e/d$. The absorption coefficient is such that $d = 3$, and we have taken the hole velocity to be 0·4 times the electron velocity, typical of many semiconductors. It can be seen that there is a fast rise component followed by a slow component, but the fast component accounts for 90% of the total response, and is, of course, mainly due to the electron contribution. The impulse response and the separate components are also shown in the Figure.

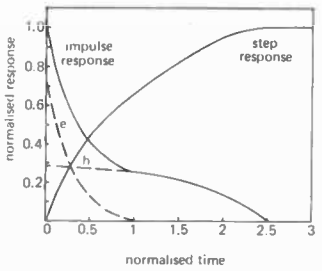


Fig. 5. Impulse and step responses of a p-i-n diode illuminated from the n-side.

For light incident on the n^+ side, we can perform the corresponding analysis and obtain the result for the step response:

$$I(t) = [\alpha v_h t + (1-\eta)(1 - e^{\alpha v_h t}) + 1 - e^{-\alpha v_e t} - vt(1-\eta)]/\eta\alpha d \qquad 0 < t < d/v_e \tag{15}$$

$$= [\alpha v_h t + \eta + (1-\eta)(1 - \alpha d - e^{\alpha v_h t})]/\eta\alpha d \qquad d/v_e < t < d/v_h \tag{16}$$

$$= 1, \qquad t > d/v_h$$

This expression is plotted in Fig. 5, along with the impulse response and the separate contributions, with the same device parameters as before. Again there is a fast component and a slow component of the step response, but this time the fast component is only 30% of the total response, and the 10 to 90% response time is about twice that for the diode with light incident on the p^+ side. One concludes that it is better to design photodiodes so that the light is incident from the p^+ side and for unity gain devices this is true unless material or fabrication technology rule otherwise. In Section 4 we investigate photodiodes with internal gain using the

avalanche multiplication process (avalanche photodiodes or a.p.d.s) where considerations of excess noise may override speed of response.

2.8 Leakage Current

The reverse-bias leakage current ('dark current') of a photodiode is an important parameter since the shot noise on this current can be the dominant noise source in the optical receiver in some situations. The dark current arises from current leakage over the surface of the device as well as through the depletion layer (bulk leakage). Surface leakage is minimized by careful device processing and by 'passivating' the surface by suitable coatings or chemical treatment: methods used vary between different materials. Bulk leakage is due to diffusion of minority carriers from the undepleted material and to thermal generation and recombination of carrier pairs in the depletion region. For an abrupt, single-sided junction with $p \gg n$, the bulk leakage current density $J_{r,b}$ is given by:[1]

$$J_{r,b} = q\left(\frac{D_p}{\tau_p}\right)^{\frac{1}{2}} \frac{n_i^2}{N_D} + \frac{qn_i w}{\tau_e} \qquad (17)$$

in which D_p is the hole diffusion coefficient, n_i is the intrinsic carrier concentration, N_D is the donor concentration in the n-region, τ_p is the hole life-time in the n-region, w is the thickness of the depletion region, and τ_e is the effective carrier life-time in the depletion region.[1] The first term is the diffusion contribution and the second is the generation-recombination term. The first term tends to dominate in photodiodes made from material with a large intrinsic carrier concentration, such as germanium. The second term is generally most important in silicon photodiodes since the band-gap is large, and in III–V photodiodes since the density of states in the conduction band is low compared with germanium.

The generation and recombination responsible for the dark current may be directly from band to band in some instances, but more generally states within the band gap play an important part. These states are due to material imperfections such as vacancies and to defects and to complexes of defects and impurities. The generation and recombination rates via these interband states are functions of the density of the states and their capture cross-section which is partly determined by their position within the band. Low generation–recombination current therefore usually implies low concentration of deep-state impurities and good crystalline quality.

3 Detection in the Presence of Noise

3.1 Poisson Statistics and Quantum Noise

The most important parameter of any receiver is its sensitivity, and there are several factors which prevent us from being able to handle arbitrarily weak signals. The signal will have suffered various impairments during transmission, because of the dispersion and attenuation of the fibre. In addition to being distorted, the signal leaving the optical receiver has wideband random fluctuations added to it, i.e. noise, produced by the components of the amplifier, etc. Lastly, even with an infinite fibre bandwidth and a noiseless amplifier, the optical signal itself is still statistical because of the quantum nature of light. Radio waves are also quantized, of course, but the quantum energy hv is much less than the thermal energy kT of electrons in the amplifier components, so that quantum effects do not show up at radio frequencies. At room temperature, kT/h is about 6000 GHz, well above the highest frequencies used in radio transmission, which are about 60 to 100 GHz, and well below the frequency corresponding to a wavelength of 1 μm, which is 300 THz. Photons arrive at the detector with Poisson statistics, and one property of the Poisson probability distribution is that the variance is equal to the mean. If the expected number of photons in some time interval is m, then the probability that n will be detected is given by:

$$p(n) = \text{Pos}\,[n, m] = m^n\, e^{-m}/n! \qquad (18)$$

Consider a binary digital system in which we need to decide whether or not a pulse was received during each bit period. The number of detected photons n is counted for each bit period, and if that number exceeds some threshold number d a 'one' pulse is recorded, otherwise a 'zero' is recorded. We make errors when $n < d$ when a 'one' pulse was transmitted, and this happens with probability:

$$P_e = \sum_{i=0}^{d'} \text{Pos}\,[i, m] \qquad (19)$$

where d' is the largest integer smaller than or equal to d, and m is the expectation number of photons per pulse. The best situation is obviously $d' = 0$, i.e. the threshold d is set between 0 and 1 photons. Then:

$$P_e = e^{-m} \qquad (20)$$

and we see that we cannot have zero error probability for finite m. For $P_e = 10^{-5}$, $m = 11{\cdot}5$, and for $P_e = 10^{-9}$, $m = 20{\cdot}7$.

In an analogue system, we are interested in the signal-to-noise ratio (SNR) at the receiver output with a post-detection bandwidth B, which smooths fluctuations over an integration time $t = 1/2B$. If the mean photon arrival rate is r, then the number m which arrive, on average, during the time t is:

$$m = r/2B. \qquad (21)$$

At the output of the receiver, the signal power is proportional to m^2, while the noise power is proportional to the variance of m, which is just m. Thus:

$$\text{SNR}_{out} = m^2/m = r/2B. \qquad (22)$$

For example, for a 50 dB SNR and a 1 MHz bandwidth, we require, on average, 2×10^{11} photons/s, or 40 nW at a wavelength of 1 μm.

3.2 Thermal Noise

The sensitivites above are the best performance we could expect, even with a perfect detector and a noiseless amplifier, limited only by the quantum fluctuations in the incoming optical signal. In real life, amplifiers are not noiseless because electrons in the conductors move with randomized velocities with energy kT, and the amplifier has to have non-zero input conductance. Using conventional components, an amplifier with input capacitance of 10 pF and a bandwidth of 10 MHz would need to have an input resistance of about 10 kΩ or less loading the photodiode. The mean square thermal noise voltage in a bandwidth B due to a resistance R is:

$$\langle v_T^2 \rangle = 4kTRB = 8{\cdot}3 \times 10^{-10}\ V^2 \qquad (23)$$

at room temperature for $R = 5$ kΩ and $B = 10$ MHz. The signal voltage generated across R due to m photons detected in time t is:

$$V_s = mqR/t = 1{\cdot}6 \times 10^{-8}\ m\ \text{V}. \qquad (24)$$

The signal-to-noise ratio is:

$$\text{SNR} = (1{\cdot}6 \times 10^{-8}\, m)^2/(8{\cdot}3 \times 10^{-10})$$

so that in a digital system with SNR = 22 dB, m is about 20 000 photons in a bit period t (taken as $1/2B$ here). This is 1000 times or 30 dB greater than the quantum noise limit, which justifies us in ignoring quantum noise in this calculation. Obviously we would like to improve on this situation, as 30 dB can be translated into perhaps 100 km of extra fibre at 1·55 μm—by no means a small benefit. There are two ways of increasing the receiver sensitivity which we need consider. Reducing the amplifier noise is one way, and we will return to that later. The other way is discussed in the next Section.

4 Avalanche Photodiodes

4.1 Avalanche Gain

An electron or hole accelerated by an electric field may gain sufficient energy so that, when it is scattered by the lattice, a lattice atom is ionized, creating an electron-hole pair. This is the process of impact ionization. The newly created carriers can, in turn, cause impact ionization and so lead to an avalanche process with current gain.

Figure 6 is a hypothetical reverse-biased p-n junction with a field E across the depletion region of width w. First we will assume that only one type of carrier, say electrons, can cause impact ionization. Usually the ionization rate α is represented as an exponential function of field:

$$\alpha = A \exp(-B/E) \qquad (25)$$

in units of ionizing collisions per unit length. In silicon, $A = 3{\cdot}8 \times 10^6\ \text{cm}^{-1}$, $B = 1{\cdot}75 \times 10^6$ V/cm, so that $\alpha = 1\ \mu\text{m}^{-1}$ for a field of 30 V/μm. A readable theoretical treatment of gain in avalanche photodiodes may be found in McIntyre *et al.*,[4] who show that if:

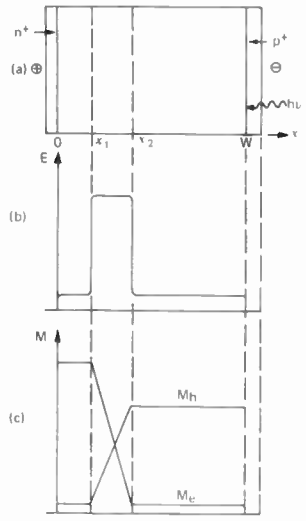


Fig. 6. Schematic avalanche photodiode: (a) the structure with high field region from x_1 to x_2; (b) the field as a function of distance; (c) the multiplication factors for holes and electrons.

$$G(x) = \exp - \int_x^{x_2} (\alpha - \beta)\, dx' \qquad (26)$$

then the solution for the gain for carrier pairs generated at x is:

$$M(x) = \frac{G(x)}{1 - \int_{x_1}^{x_2} \alpha G(x)\, dx} \qquad (27)$$

and the average multiplication is:

$$\bar{M} = \frac{\int_0^w g(x) m(x)\, dx}{\int_0^w g(x)\, dx}. \qquad (28)$$

where β is the ionization rate for holes and $g(x)$ is the generation rate of carrier pacers by other processes, e.g. photogeneration.

The equations for $M(x)$ cannot generally be solved. However, it turns out that the ratio $k = \alpha/\beta$ is often insensitive to M and can be considered constant. Then $M(x)$ can be expressed as:

$$M(x) = \frac{G(x)(1-k)}{G(x_1) - k} \qquad (29)$$

with

$$G(x_1) = \exp[-(1-k)\delta] \qquad (30)$$

$$\delta = \int_{x_1}^{x_2} \alpha\, dx. \qquad (31)$$

These equations can be solved quite easily for a diode with a uniform field between x_1 and x_2, and typical results are shown in Fig. 7 for the electron gain M as a function of electric field (values of A and B appropriate to silicon have been used) and various values of the ratio k. Inspecting equation (29), we see that $M(x)$ becomes infinite if $G(x_1) = k$, and that this breakdown condition is

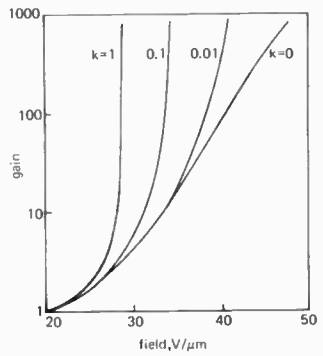


Fig. 7. Avalanche gain as a function of field (the breakdown characteristic) for different values of the ratio k of ionization rates for electrons and holes.

reached at a finite field unless $k = 0$. So the gradient of all the curves in Fig. 7 becomes infinite for some finite field, except for $k = 0$. The implication of this is as follows: to get useful current gain from the a.p.d. it must be biased fairly close to breakdown—very close if k is near to unity. But any variation in field, due to the diode not being perfectly uniform, or the supply voltage being imperfectly regulated, causes a change in the current gain, and this change can be large if k is near unity. Operation of the device then becomes unpredictable and also noisy. In silicon k can be as low as 0·01, and silicon a.p.d.s can be operated at gains of a few hundred or even thousands in some cases. In germanium and many III–V compounds, k is 0·3–1 and it is hard to fabricate and control a device for a gain above 10 to 15. There are also noise problems associated with a value of k close to unity.

4.2 Benefits of Avalanche Gain

The current gain arising from avalanche gain increases the signal voltage across the amplifier input and so improves the signal-to-noise ratio, since the amplifier noise is unaffected. However, the current gain also increases the quantum noise by the same amount as the signal, so that one cannot get beyond the quantum noise limit. In practice one cannot even get very near to it, because of extra noise introduced by the impact ionization process. To see how avalanche gain affects the sensitivity of an optical receiver, consider a sinusoidally varying incident optical power:

$$p(t) = \bar{P}(1 + \sin \omega t) \tag{32}$$

where $\bar{P}$ is the mean, and $2\bar{P}$ the peak optical power. The resultant multiplied photocurrent (peak-to-peak) is:

$$\langle i_s \rangle_{p-p} = 2\bar{P}\eta qM/h\nu. \tag{33}$$

The mean electrical signal power is proportional to $\langle i_s \rangle^2_{p-p}$. The shot noise power is proportional to $2q\langle i_s \rangle BM^x$, where $\langle i_s \rangle = \langle i_s \rangle_{p-p}/2$ is the mean multiplied photocurrent, B the receiver bandwidth and M^x the excess noise factor (see Sect. 4.3). The thermal noise power, which we shall assume to be dominated by the noise of a bias resistor R, is proportional to $4kTB/R$. In each case the constant of proportionality is the same, and depends on the receiver amplifier. So the output signal-to-noise ratio is:

$$\text{SNR} = \frac{(2\bar{P}\eta qM/h\nu)^2}{2\bar{P}\eta q^2 BM^{2+x}/h\nu + 4kTB/R}. \tag{34}$$

We have already seen that, with $M = 1$, the thermal noise term dominates. As M is increased from unity, the signal power increases as M^2, but, so long as the thermal noise term dominates, the total noise power is little affected and SNR increases. When M is large, thermal noise is insignificant and SNR decreases with increasing M as M^x. We see that there is an optimum avalanche gain given by:

$$M^{2+x} = (4kT/R)(h\nu/x\bar{P}\eta q^2) \tag{35}$$

so that

$$\frac{\text{shot noise power}}{\text{thermal noise power}} = \frac{2}{x}.$$

Equation (34) is plotted in Fig. 8 with different values of x. We see that if x is small, as with a silicon a.p.d., the optimum gain is large and the maximum in SNR is broad. The a.p.d. can, in fact, be used to vary the gain of the receiver and provide a.g.c. When x is near unity, less improvement in SNR is possible, the optimum gain is lower and the maximum is much sharper. Such an a.p.d. is difficult to control to get maximum performance, because with x (hence k) close to unity, the avalanche gain is a very critical function of applied voltage.

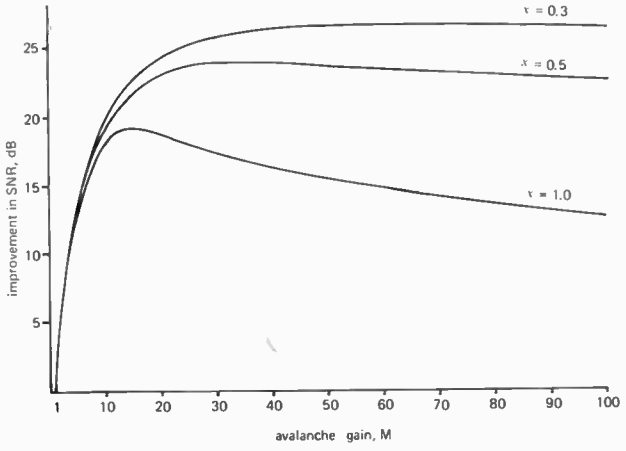


Fig. 8. Improvement in SNR as a function of avalanche gain M.

4.3 Excess Avalanche Noise

The avalanche process is statistical and therefore noisy. The number of carriers produced by avalanche gain is not deterministic, but has a mean value and some variation about it which is equivalent to an excess noise factor. The statistics of the avalanche gain and the resulting noise spectral density[4] are of central importance to the theory of optical receivers, particularly when the expected signal is very weak, but a detailed treatment is outside the scope of this paper. We will assume that the effect of the avalanche noise is to

increase the multiplied mean-square shot noise current by a factor $F = M^x$, where M is the mean current gain and x ($0 < x < 1$) is a parameter of the a.p.d., its design, material and so on. We also assume that the resulting noise is white and Gaussian. McIntyre[4] gives a more accurate expression for the excess avalanche noise factor F:

$$F = M\left[1-(1-k)\left(\frac{M-1}{M}\right)^2\right] \qquad (36)$$

if the only carriers injected are electrons, and

$$F = M\left[1+\left(\frac{1-k}{k}\right)\left(\frac{M-1}{M}\right)^2\right] \qquad (37)$$

if the only carriers injected are holes. Here k is again the ratio α/β of hole to electron ionization rates. Notice that as k tends to zero, F in equation (36) tends to 1 for M tending to 1, while for k close to unity, F tends to M. The approximate representation M^x thus covers the same range for F if x is chosen appropriately. Typical values for x are 0·3 for a good silicon a.p.d. ($k = 0{\cdot}02$) and 0·7 to 1 for III–V or germanium a.p.d.s ($k = 0{\cdot}3$ to 1).

4.4 Silicon Reach-through A.P.D.s

To make an a.p.d. in silicon with a fast response, a simple p-n junction will not do, because most photons will be absorbed in undepleted material where the field is negligible. It is necessary to use the 'reach-through' structure shown in Fig. 9, in which the depletion region consists of a high-doped, high field gain region followed by a lower field, low-doped absorbing region. The problem is to ensure that the absorbing region is fully depleted well before the gain region breaks down, and this demands great control over the fabrication of the device. Nevertheless, good commercial silicon reach-through a.p.d.s (r.a.p.d.s) have been on the market for several years.

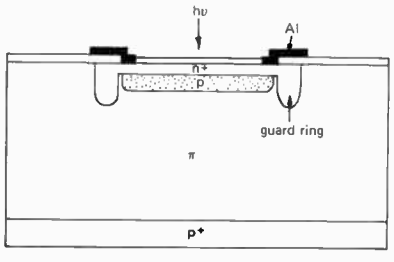


Fig. 9. Schematic cross-section of a silicon reach-through a.p.d.

4.5 Germanium A.P.D.s

Apart from the recent results using the p-i-n f.e.t. receivers, all system work at longer wavelengths has been carried out using Ge a.p.d.s. Germanium seems an obvious material, since the photodiodes can be made sensitive out to 1·6 μm and beyond by reducing the thickness of undepleted material near the surface. Germanium is not ideal for a.p.d.s because the ratio of

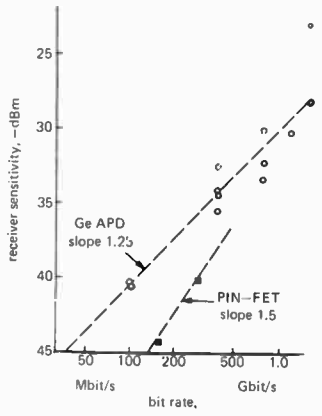


Fig. 10. Published receiver sensitivities against data-rate.

ionization coefficients k is close to unity (i.e. $x = 1$) so that the excess noise factor is high. More importantly, the reverse bias leakage current density is high, because the high intrinsic carrier concentration results in a large diffusion contribution to the leakage current. The unmultiplied leakage current density is typically 3×10^{-4} A·cm^{-2} at room temperature, sufficient to cause a system penalty of a few decibels at a data-rate of a few hundred Mbit/s. The leakage current is dependent on temperature, and at 50°C is about an order of magnitude greater than at 20°C, resulting in a large system penalty and reducing the optimum gain to about 3 to 5, since the dominant noise source is multiplied bulk leakage. At room temperature, receiver sensitivities of −34 dBm at 400 Mbaud and −30 dBm at 800 Mbaud have been reported[5] using germanium a.p.d.s. These figures would be several decibels worse at 50°C. Brain[6] has recently carried out a detailed comparison of the performance of germanium a.p.d.s with p-i-n f.e.t. receivers and concluded that the germanium a.p.d. becomes marginally superior in performance at data rates above 1·2 Gbaud for temperatures below 35°C. Published receiver sensitivities at 1·3 and 1·55 μm are plotted in Fig. 10 for the available range of bit-rates, and it can be seen that the bit-rate dependence is approximately the 1/25th power, as one would expect from an a.p.d. with an x-factor near unity. Also shown are the results for p-i-n f.e.t. receivers with a 1/5th power dependence as expected, and lending some experimental support to Brain's analysis.

4.6 III–V A.P.D.s

In pursuit of the excellent performance achieved with silicon a.p.d.s, very considerable effort has been expended in research on a.p.d.s made in III–V compounds. To date, no system results have been reported, although there is much published material on the devices themselves. As with lasers, the main work has been carried out on the GaInAsP/InP system, although several workers, notably Law *et al.*[7] have investigated other systems as well. Until recently, avalanche gains in

the region 10 to 20 were typical, limited probably by non-uniformity of the material of the high-field region leading to micro-plasma breakdown.[8] More recently, a structure with the high-field region in InP has been described,[9-11] as shown in Fig. 11, and gains of up to several thousand have been reported.

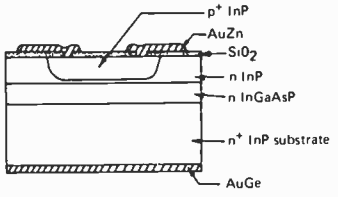


Fig. 11. Schematic cross-section of a III–V heterostructure a.p.d. with the high-field region within the InP layer.[10]

The ionization coefficients for electrons and holes have been measured in several III–V compounds as functions of field strength and orientation. Law[7] has summarized several results, and Pearsal[12-14] has reported on GaInAs and GaInAsP in particular. Generally the ratio of ionization coefficients lies between 5 and 0·2 and, as will be shown in Section 5, such values yield little improvement in receiver sensitivity over a unity ratio. The ionization coefficients are known to be sensitive to material composition.[14] Pearsal *et al.*[15] have found that the ionization coefficients in GaAs are sensitive to field strength in particular directions. There is thus the possibility of attaining a k-value usefully different from unity.

A common feature of all III–V a.p.d.s reported for the 1·3 and 1·55 μm wavelength range is the soft breakdown characteristic. The reverse bias dark current does not saturate as one would expect for a generation-recombination current, but continues to rise approximately exponentially with voltage. The multiplied dark current increases faster with applied voltage than one would expect from a normal avalanche multiplication. Thus, although devices may have quite low reverse bias dark currents at 0·5 and 0·9 of the breakdown voltage, the dark current becomes very large in the region of useful avalanche gain, which in these materials with k-values near to unity is close to breakdown. Forrest *et al.*[16] and Takanashi *et al.*[17] have explained this behaviour in terms of tunnelling, investigating several materials over considerable ranges of temperature to demonstrate that the current density varies as $\exp(-AE_g^{3/2})$. The implication is that the dark current can be reduced to an acceptable level only by keeping the high field region to low-doped, large band-gap material such as InP. The excess avalanche noise properties of the device then depend on this material.

Meaningful measurements of quantum efficiency, avalanche gain and excess noise are not straightforward, and reported results must be interpreted with care. Noise measurements in particular are complicated by the leakage current which increases rapidly with bias voltage. Even quantum efficiency can be voltage-dependent as the depletion layer varies in thickness, and it is necessary to know the quantum efficiency to measure the avalanche gain. The final arbiter must be system measurements, which to date have been eloquent in their absence.

5 The Digital Optical Receiver

5.1 Error Rate and Gaussian Statistics

In a receiver for a binary digital system, the aim is to process the signal in such a way as to be able to distinguish between two hypotheses, which we label 'zero' and 'one', with the minimum possible error. In this way we seek the best estimate of the original message from the attenuated, distorted and noisy signal in the receiver. Commonly the signal is detected, amplified and filtered and then presented to a decision gate which is 'opened' for a short interval at the centre of each bit period by a pulse from a clock circuit. This interval is called the 'decision time'. Let us assume that, for a received 'zero' bit, the receiver output voltage $v(0)$ at the decision time has a mean value m_0 and variance s_0, while for a received 'one', the mean is m_1 and the variance s_1. Because the quantum noise is signal-dependent, s_0 and s_1 are different, in contrast to transmission systems operating at radio or microwave frequencies. We will also assume for simplicity that $v(0)$ has a Gaussian distribution, although we know that the multiplied quantum noise has, in fact, a compound Poisson intensity distribution. The probability of an error in decision for a received 'zero' is:

$$\begin{aligned} P_{e,\text{zero}} &= Pr[v > v_D|\text{zero transmitted}] \\ &= \frac{1}{s_0\sqrt{2\pi}} \int_{v_D}^{\infty} \exp\left[-(v-m_0)^2/2s_0^2\right] dv \end{aligned} \tag{38}$$

where v_D is the decision threshold voltage. By substituting $x = (v-m_0)/s_0$ and $Q_0 = (v_D-m_0)/s_0$, we obtain:

$$P_{e,\text{zero}} = \tfrac{1}{2}\text{erfc}\,(Q_0/\sqrt{2}). \tag{39}$$

Similarly, with $Q_1 = (m_1-v_D)/s_1$ we find:

$$P_{e,\text{one}} = \tfrac{1}{2}\text{erfc}\,(Q_1/\sqrt{2}). \tag{40}$$

If our message is such that, *a priori*, 'zeros' and 'ones' are equally likely, it is reasonable to set v_D so that the error probabilities $P_{e,\text{zero}}$ and $P_{e,\text{one}}$ are equal, i.e., $Q_0 = Q_1$. Then:

$$m_1 - m_0 = Q(s_1 + s_0) \tag{41}$$

$$P_e = \tfrac{1}{2}\text{erfc}\,(Q/\sqrt{2}). \tag{42}$$

Equation (41) tells us what difference we must have in optical power between the 'zero' and 'one' bits in terms of the variance (noise) and Q, which is related to the signal-to-noise ratio (SNR = $4Q^2$). Equation (42) tells us what value of Q is needed for a given acceptable error rate. For example, $Q = 6{\cdot}00$ for $P_e = 10^{-9}$; small

changes in Q produce large changes in error rate. Note that for design error rates of this magnitude, errors arise from the far tails of the noise distribution—six standard deviations away from the mean. That is why accurate models of noise statistics are important in optical systems. In fact, the Gaussian approximation used here is successful at predicting error rate as a function of mean signal power, but is poor at giving the correct signal threshold level and the optimum avalanche gain, for this reason.

5.2 Theoretical Model

We consider (Fig. 12) a photodiode with capacitance C_d and bias resistor R_b working into an amplifier with voltage gain A, and input resistance and capacitance R_a and C_a. The amplifier is followed by filters $U(\omega)$ and $H(\omega)$, where

$$U(\omega) = G + j\omega C \tag{43}$$

with

$$G = 1/R_b + 1/R_a \tag{44}$$

$$C = C_d + C_a \tag{45}$$

and $H(\omega)$ filters the pulses to the shape required for the decision gate. The first filter $U(\omega)$ compensates for the frequency response of the bias and input circuit. It is useful conceptually to separate it from $H(\omega)$ although it is not separately realizable.

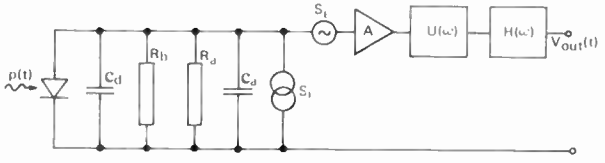


Fig. 12. Model for a digital optical receiver, with principal noise sources.

Let the received pulse shape be $h_p(t)$, normalized so that:

$$\int_{-\infty}^{\infty} h_p(t)\,dt = 1. \tag{46}$$

If b_k is the energy in the kth pulse (for a binary system b_k has two possible values which we designate b_{zero} and b_{one}), the received optical signal as a function of time is:

$$p(t) = \sum_k b_k h_p(t - kT) \tag{47}$$

where T is the bit period (not to be confused with absolute temperature, for which θ is used in the rest of this paper). This optical power gives rise to a photocurrent:

$$E[i_p(t)] = \frac{\eta q M}{h\nu} \sum_k b_k h_p(t - kT) \tag{48}$$

where M is the mean of the avalanche gain distribution, and $E[\ ]$ denotes expectation value. The output voltage is:

$$E[v_{out}(t)] = \frac{A\eta q M}{h\nu} \sum_k b_k h_p * h(t - kT) \tag{49}$$

where $h(t)$ is the Fourier transform of $H(\omega)$. We will put:

$$h_p * h(t) = h_{out}(t) \tag{50}$$

and we will normalize so that $h_{out}(0) = 1$ and $A\eta qM/h\nu = 1$. Then:

$$E[v_{out}(0)] = \sum_k b_k h_{out}(kT) \tag{51}$$

so that b_k is now the output voltage at the decision time for an isolated pulse.

The noise sources which cause the variability in the output pulse train are the quantum noise, the excess avalanche noise, and the thermal noise from the various components of the front end of the receiver. In this simplified treatment we will approximate the quantum noise by the usual expression for the shot noise on the photocurrent averaged over the bit period T:[23]

$$E[v_s^2] = 2q\langle i_0\rangle_T R^2 A^2 B_N \overline{M^2} \tag{52}$$

where

$$B_N = \frac{1}{4\pi R^2} \int_{-\infty}^{\infty} |H(\omega)|^2\,d\omega. \tag{53}$$

More exact analysis of the quantum noise is beyond the scope of this paper, and those interested are referred to the papers by Personick,[18] Andreucci,[19] and House.[20]

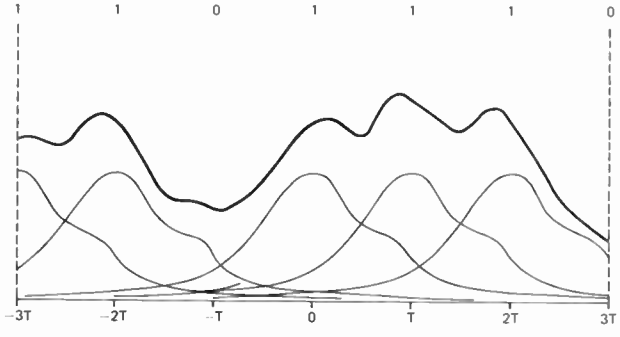


Fig. 13. Illustrating dispersion and overlapping of pulses resulting in substantial 'zero'-level noise.

Equation (52) contains the mean photocurrent over the bit period T. If the transmitted pulses are significantly spread by the fibre dispersion, the photocurrent during any given bit period will contain contributions from the tails of neighbouring pulses and will not be zero during a 'zero' bit in general. Figure 13 shows a typical section of a train of overlapping pulses. We need only consider the worst cases, as these will cause the great majority of errors. The worst case noise for either a 'zero' or a 'one' pulse occurs when all neighbouring pulses are 'ones', i.e. maximum overlap from neighbouring pulses. The energy received during a

'one' pulse is then:

$$b_{one} \sum_{k} \int_{kT-T/2}^{kT+T/2} h_p(t-kT)\,dt, \quad = b_{one} \qquad (54)$$

and during a 'zero' pulse, the received energy is:

$$b_{one}\left[1 - \int_{-T/2}^{T/2} h_p(t)\,dt\right], \quad = b_{one}(1-\gamma) \qquad (55)$$

where γ is the fraction of the pulse energy contained within the bit period. Thus:

$$\langle i_0 \rangle_T = \frac{\eta q b}{h\nu} b_{one} \quad \text{or} \quad \frac{\eta q b}{h\nu} b_{one}(1-\gamma). \qquad (56)$$

For thermal noise, we consider terms due to the bias resistor, and to the amplifier current and voltage noise sources which have spectral densities S_I and S_E.

$$E[v_R^2] = 4k\theta A^2 R^2 B_N / R_b \qquad (57)$$

$$E[v_I^2] = 2S_I A^2 R^2 B_N \qquad (58)$$

$$E[v_E^2] = S_E A^2 \frac{1}{2\pi} \int_{-\infty}^{\infty} |U(\omega)H(\omega)|^2\,d\omega. \qquad (59)$$

The amplifier voltage noise source is not filtered by the front end components.

It is useful to normalize the time and frequency dimensions with respect to the bit period, putting $\phi = \omega T/2\pi$. With the normalizations preceding equation (51), it can be shown[18] that the total mean-square noise voltage at the decision time ($t = 0$) is then:

$$E[v_n^2] = \left(\frac{h\nu}{\eta}\right)^2 \left\{\frac{M}{q}\langle i_0 \rangle_T T I_2 + \frac{Z}{M^2}\right\} \qquad (60)$$

where Z is the mean square number of thermal noise electrons per bit period (dimensionless):[18]

$$Z = \left\{\frac{T}{q^2}\left[S_I + \frac{2k\theta}{R_b} + \frac{S_E}{R^2}\right] I_2 + \frac{(2\pi C)^2}{Tq^2} S_E I_3\right\} \qquad (61)$$

and I_2 and I_3 are dimensionless normalized bandwidth integrals:

$$I_2 = \int_{-\infty}^{\infty} \left|\frac{H'_{out}(\phi)}{H'_p(\phi)}\right|^2 d\phi \qquad (62)$$

$$I_3 = \int_{-\infty}^{\infty} \left|\frac{H'_{out}(\phi)}{H'_p(\phi)}\right|^2 \phi^2\,d\phi. \qquad (63)$$

We now have all we need to calculate the receiver sensitivity, which we now formally define for digital systems as the minimum mean power required to achieve an error probability below the maximum acceptable. Equation (60) with equation (56) substituted for $\langle i_0 \rangle_T$ gives us s_0 and s_1. Then equation (41) gives us m_1 (i.e. b_{one}), since we obviously want to make m_0 (b_{zero}) as small as possible. Notice that from equation (61) we can reduce the thermal noise by making the resistances R_a and R_b as large as possible, even though this means reducing the bandwidth of the front end of the optical receiver so that it is necessary to reshape (equalize) the waveform subsequently to recover the pulses.

5.3 Unity Gain Receiver: the p-i-n F.E.T.

As a simple example, consider a p-i-n photodiode, which has unity gain only. We know that the quantum noise is insignificant, so from equation (48):

$$s_1 = s_0 = \frac{h\nu}{\eta}\sqrt{Z} \qquad (64)$$

$$b_{one} - b_{zero} = 2Q\frac{h\nu}{\eta}\sqrt{Z}. \qquad (65)$$

With typical component values, Z might be 10^6. So with $b_{zero} = 0$ and $Q = 6$, we need 12000 photogenerated electrons per 'one' pulse, in agreement with the earlier rough calculation. Using discrete components, a unity gain photodiode provides a receiver sensitivity typically 10 to 15 dB worse than an a.p.d. However, Smith *et al.*[21] have pointed out that by using hybrid integration, the input capacitance of the receiver front end can be reduced so that Z falls to below 10000 and the receiver sensitivity is then better than that attainable with a packaged a.p.d. with $x = 1$, particularly if the leakage current of the a.p.d. is appreciable. The p-i-n f.e.t. hybrid approach also offers the advantages of low-voltage operation, no need for feed-back to control the avalanche gain, simpler device technology and probably greater reliability. Typical p-i-n photodiodes for use in p-i-n f.e.t. receivers are shown in cross-section in Figs. 1 and 2.

The photodiode is hybrid integrated with a low-noise amplifier using a gallium arsenide m.e.s.f.e.t., as the front end. The receiver noise parameter Z is proportional, at high data rates, to C^2/g_m, where C is the total input capacitance (photodiode, f.e.t. gate-source and stray capacitance) and g_m is the transconductance of the f.e.t.[21] In state-of-the-art receivers, C is around 0·5 pF and g_m is 20 ms. Such receivers have achieved a sensitivity of −44·2 dBm at 160 Mbaud and −40·1 dBm at 294 Mbaud, at 1·3 μm wavelength.[2] These results are the best of any receiver at 1·3 μm. Less attention has been devoted to 1·55 μm so far, but the best results (−41·5 dBm at 160 Mbaud, −38 dBm at 294 Mbaud) are better than have been achieved with a.p.d. receivers.

The p-i-n f.e.t. receiver uses a high impedance (integrating) front-end amplifier for the best performance, although a trans-impedance amplifier could be used with a slight penalty. The integrating characteristic (time-constant typically 1000 times the bit period) has to be equalized, which can be done using a high-speed comparator in a single-tap transversal equalizer,[22] or simply by differentiating with a capacitor-resistor arrangement.

The sensitivity of the p-i-n f.e.t. type of receiver can be further improved by reducing the input capacitance and increasing the f.e.t. transconductance. Small-area

photodiodes, perhaps down to 30 μm diameter, can be used with monomode fibre without undue alignment problems. The gate length and gate width of the f.e.t. can be reduced appreciably from the values used at present, typically 0·7 by 300 μm. Anti-reflecting coatings on the photodiodes will improve the quantum efficiency by about 2 dB. An improvement in receiver sensitivity of 7 dB or more seems possible, with performance extended beyond 1 Gbaud.

5.4 Receiver with A.P.D.

It can be shown from equations (41), (60) and (61) that the sensitivity of an optical receiver is described by the equation:

$$n^2 - 2nY + Y^2\rho^2 - 4(1-\varepsilon)^2 Y n_d - 4(1-\varepsilon)Q^2 Z/M^2 = 0 \quad (66)$$

$$n = 2\bar{n}(1-\varepsilon)/(1+\varepsilon) \quad (67)$$

in which $\bar{n}$ is the mean number of photogenerated electrons per bit time, n_d is the number of dark-current electrons in the receiver per bit time, Z is the receiver's signal-independent noise parameter (Ref. 18, eqn. 61)), M is the avalanche gain, ε is the extinction ratio at the source, $Y = Q^2 I_2 M^x$ where x is the excess avalanche noise parameter, and ρ is a parameter related to ε and to the fraction γ of the pulse energy contained within the bit-time T.[24] The avalanche gain M can be optimized by setting $d\bar{n}/dM = 0$, so that:

$$\frac{Y^2\rho^2}{2(1-\varepsilon)} - \frac{(2-\rho)n}{(1+\varepsilon)} - n_d Y + \frac{2Q^2 Z}{xM^2} = 0. \quad (68)$$

The receiver sensitivity under optimum gain can be found by solving equations (66) and (68) simultaneously. Figure 14 shows some theoretical results as the mean number of photoelectrons required per bit time $\bar{n}$ and optimum avalanche gain M as a function of the excess avalanche noise parameter x, for values of n_d from zero to 10 000 electrons per bit period. The receiver sensitivity in dBm at a line rate R (Mbaud) and wavelength λ (μm) is related to $\bar{n}$ by

$$\text{dBm} = 10 \log (\bar{n}R/\lambda\eta) - 97,$$

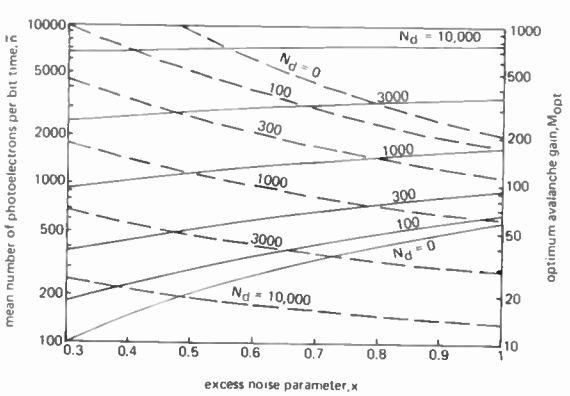


Fig. 14. Receiver sensitivity and optimum avalanche gain as functions of the excess noise parameter x, for values of the dark leakage count n_d from 0 to 10000 counts per bit-time.

where η is the quantum efficiency of the photodiode, For example at a line rate of 100 Mbaud and a wavelength of 1 μm with $\bar{n} = 1000$, the receiver sensitivity is −47 dBm. Figure 14 is calculated assuming $Z = 10^6$, typical of a receiver using discrete components at a few hundred megabaud, and with $\varepsilon = 0$ and $\gamma = 1$, i.e. zero extinction ratio and no pulse spreading. It can be seen that:

(1) The excess noise parameter x has rather little effect when $x > 0{\cdot}7$ or so, particularly if the dark current is large ($n_d > 1000$).
(2) For $n_d > 1000$ or so, $\bar{n}$ is nearly proportional to n_d as expected, as multiplied dark current is the largest receiver noise source then.
(3) The optimum gain decreases markedly for x going from 0·5 to 1, particularly if n_d is small.

Clearly it is important to minimize n_d and to a lesser extent to reduce x. Note that a leakage current of 160 nA gives n_d of 1000 at 1 Gbaud, which is large enough to affect the optimum gain and the receiver sensitivity. At lower data rates the effects would be greater still.

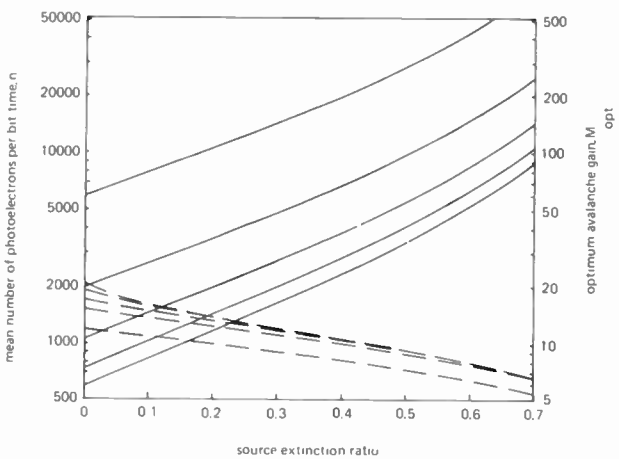


Fig. 15. Receiver sensitivity and optimum avalanche gain as functions of the source extinction ratio ε for fibre dispersion α_F from zero to 0·5, assuming an x-value of 1·0.

Figure 15 shows how $\bar{n}$ and M vary with extinction ratio ε and pulse spreading. The pulse spreading is represented by α_F, the normalized r.m.s. width of the fibre impulse response, assumed to be Gaussian. The launched pulse is taken as being half-width rectangular, and the dark current is assumed to be zero. It can be seen that:

(1) The receiver sensitivity is strongly affected by pulse spreading and by non-zero extinction.
(2) The optimum gain is reduced by 'zero'-level noise and by fibre dispersion, the effect being greatest when x is small. It must be remembered that this type of calculation, which assumed Gaussian noise statistics, tends to overestimate the optimum gain, although relative magnitudes are predicted more accurately.

Obviously combinations of appreciable pulse

spreading, non-zero extinction and considerable dark current ($n_d = 100000$) reduces the receiver sensitivity very much, and also reduce the optimum avalanche gain to near unity.

6 Conclusions

This paper has reviewed the current state of optical detectors and receivers, both in theory and in practice, with the emphasis on devices for high-performance systems. During the next few years we can confidently expect to see optical fibre systems used for long-distance telecommunications, with the emphasis on performance—greater bandwidth and greater unrepeatered system length. In such a rapidly developing subject it is essential to be aware not only of the latest published results, but also of the underlying principles. Only then can the potential of optical communication be properly appreciated. With this in mind, this paper has reviewed both the best reported performance in detectors and receivers and the areas where there is still room for better performance.

7 References

1 Sze, S. M., 'Physics of Semiconductor Devices' (Wiley, New York, 1969).

2 Hooper, R. C., Rejman, M. A. Z., Ritchie, S. T. D., Smith, D. R. and White, B. R., 'Pin-fet hybrid optical receivers for longer wavelength optical communication systems', 6th European Conference on Optical Communications, York, 1980, IEE Conference publication 190.

3 Leheny, R. F., Nahory, R. E. and Pollack, M. A., 'InGaAs p-i-n photodiodes for long-wavelength fibre-optic systems', *Electronics Letters* **15**, pp. 713–5, 1980.

4 Webb, P. P., McIntyre, R. J. and Conradi, J., 'Properties of avalanche photodiodes', *RCA Rev.*, **35**, pp. 234–78, June 1974.

5 Nakagawa, N., Hakamada, Y. and Suto, K., '800 Mbit/s fibre transmission test using low loss and low dispersion single mode cable', *Electronics Letters*, **15**, pp. 747–8, 1979.

6 Brain, M. C., *Optical & Quantum Electronics*, to be published.

7 Law, H. D., Nakano, K. and Tomasetta, L. R., 'III–V alloy hetero structure high-speed avalanche photodiodes', *IEEE J. Quantum Electronics*, **QE-15**, pp. 549–58, 1979.

8 Lee, T. P., Burrus, C. A. and Dentai, A. G., 'InGaAsP/InP microplasma limited avalanche multiplication at 1·0 to 1·3 μm wavelength', *IEEE J. Quantum Electronics*, **QE-15**, pp. 30–5, 1979.

9 Kanbe, H., Susa, N., Nakagome, H. and Ando, H., 'InGaAs avalanche photodiode with InP p-n junction', *Electronics Letters*, **16**, pp. 163–5, 1980.

10 Nishida, K., Takaguchi, K. and Matsumoto, Y., 'InGaAsP heterostructure avalanche photodiodes with high avalanche gain', *Appl. Phys. Letters*, **35**, pp. 251–3, 1979.

11 Diadiuk, V., Groves, S. H. and Hurwitz, C. E., 'Avalanche multiplication and noise characteristics of low dark-current GaInAsP/InP avalanche photodetectors', *Appl. Phys. Letters*, **37**, pp. 807–10, 1980.

12 Pearsal, T. P., Nahory, R. E. and Pollack, M. A., 'Impact ionisation coefficients for electrons and holes in $In_{0.14}Ga_{0.86}As$', *Appl. Phys. Letters*, **27**, pp. 330–2, 1975.

13 Pearsal, T. P., 'Impact ionisation rates for electrons and holes in $Ga_{0.47}In_{0.53}As$', *Appl. Phys. Letters*, **36**, pp. 218–20, 1980.

14 Pearsal, T. P., Nahory, R. E. and Pollack, M. A., 'Impact ionisation rates for electrons and holes in $GaAs_{1-x}Sb_x$ alloys', *Appl. Phys. Letters*, 28, p. 403, 1976.

15 Pearsal, T. P., Capasso, F., Nahory, R. E., Pollack, M. A. and Chelikovsky, J. R., 'The band structure dependence of impact ionisation by hot carriers in semiconductors: GaAs', *Solid-State Electronics*, **21**, pp. 297–302, 1978.

16 Forrest, S. R., DiDomenico, M., Smith, R. G. and Stocker, H. J., 'Evidence for tunnelling in reverse-biased III–V photodetector diodes', *Appl. Phys. Letters*, **36**, pp. 580–2, 1980.

17 Takanashi, Y., Kawashima, M. and Horikoshi, Y., 'Required donor concentration of epitaxial layers for efficient InGaAsP avalanche photodiodes', *Japan J. Appl. Phys.*, **19**, pp. 693–701, 1980.

18 Personick, S. D., 'Receiver design for digital fibre optic communication systems', Parts I and II, *Bell Syst. Tech. J.*, **52**, pp. 843–86, 1973.

19 Andreucci, F. and Mengali, U., 'Timing extraction in optical transmissions', *Optical & Quantum Electronics*, **10**, pp. 445–58, 1978.

20 House, K., 'Detection and equalisation in fibre-optic communication', Ph.D. Thesis, University of Essex, 1979.

21 Smith, D. R., Hooper, R. C. and Garrett, I., 'Receivers for optical communications: a comparison of avalanche photodiodes and p-i-n f.e.t. hybrids', *Optical & Quantum Electronics*, **10**, pp. 293–300, 1978.

22 Hooper, R. C., Smith, D. R. and White B. R., 'p-i-n f.e.t. hybrids for digital optical receivers', 30th IEEE Electronic Components Conference, San Francisco, April 1980.

23 Smith, D. R., Hooper, R. C. and Garrett, I., 'A simplified approach to digital optical receiver design', *Optical & Quantum Electronics*, **10**, pp. 211–21, 1978.

24 Hooper, R. C. and White, B. R., 'Digital optical receiver design for non-zero extinction ratio using a simplified approach', *Optical & Quantum Electronics*, **10**, pp. 279–82, 1978.

Manuscript first received by the Institution on 5th March 1981 and in final form on 30th April 1981
(Paper No. 1998/CC342)

UDC 621.375.826: 621.372.8: 621.391.63/4

Indexing Terms: Lasers, semiconductor, Fibre optics, Optical communication

Semiconductor laser sources for optical communication

P. A. KIRKBY, B.Sc., Ph.D*

SUMMARY

The paper is a general review of double heterostructure GaAlAs and GaInAsP injection lasers and their application in long-distance optical communication systems. After briefly describing the physics of semiconductor laser operation, the performance of oxide stripe and advanced laser structures at 0·85, 1·3 and 1·55 μm wavelength is discussed. The parameters of greatest importance in optical communication systems such as laser spectrum and modulation response are covered in some detail and estimates of device parameters and manufacturing tolerances are given.

The relative merits of the different types of laser for both multimode and single-mode optical fibre systems are then discussed. It is shown that single-mode systems operating close to the dispersion minimum are capable of data rates above 390 Mbit/s over distances of 30 km even when fairly conservative worst case assumptions are made. The key parameters that will allow single-mode systems to operate at data rates of more than 1 Gbit/s over distances of 50 to 100 km are identified.

*ITT Standard Telecommunication Laboratories Ltd., Materials and Components Laboratory, London Road, Harlow, Essex CM17 9NA.

1 Introduction

The semiconductor laser and the light emitting diode are universally used as the source in optical communication systems because no other type of optical source can be directly modulated at the high bit rates required with such low drive and high output powers. The choice between laser and l.e.d. depends on the system. For short-range, low bandwidth systems where the lower output power, lower frequency response and much broader spectrum are not limiting factors the l.e.d. is usually chosen because the drive and control circuitry is more simple. For wide-bandwidth, long-haul communication links the laser offers the best performance.

Semiconductor lasers are now widely available at 0·85 μm operating wavelength.[1,2,3] The considerably lower losses and lower dispersion of silica fibre at 1·3 and 1·55 μm wavelength have encouraged the rapid development of lasers operating at these wavelengths.[4,5,6] The most successful of these devices are based on the InP/GaInAsP quaternary III-V semiconductor alloys. The author is currently involved in developing such lasers for long-haul submarine and land line optical communication systems. For completeness this paper briefly describes the physics of semiconductor laser operation; more detailed descriptions can be found in Refs. 1–3. The performance of lasers for both multi- and single-mode optical fibre systems is then discussed in more detail.

2 Principles of Operation

Figure 1 is a diagram of an oxide-insulated stripe geometry InP/GaInAsP laser, showing the double heterostructure layers which form the active part of the device. The laser chip is typically 300 μm long and carefully cleaved facets form the end mirrors of the laser. Direct electron/hole recombination in the GaInAsP active layer is the source of the photon emission. The

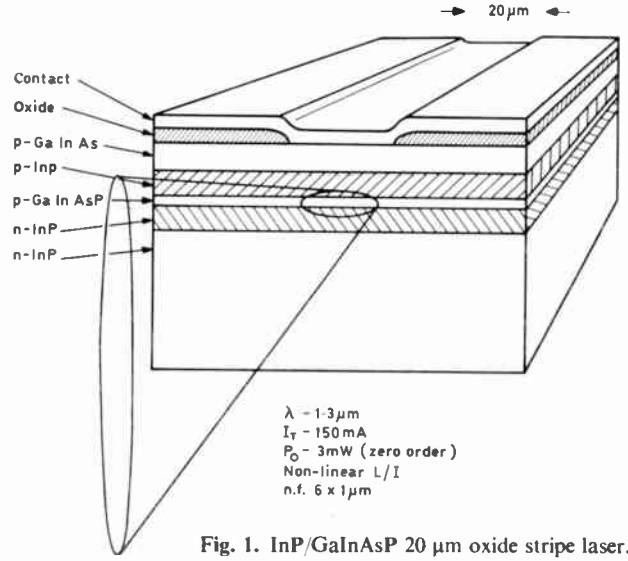


Fig. 1. InP/GaInAsP 20 μm oxide stripe laser.

The Radio and Electronic Engineer, Vol. 51, No. 7/8, pp. 362–376, July/August 1981

0033–7722/81/070362+15 $1.50/0

operating wavelength is determined by the photon energy which in turn is determined by the band gap of the active layer. When the injected carrier density in the active layer exceeds approximately 10^{18} cm^{-3}, population inversion occurs and optical gain is obtained. When the current is increased so that the optical gain becomes greater than the internal loss and the end losses, laser oscillation begins and the excess current above threshold is very efficiently converted into the intense laser output. The volume current density required to reach threshold is very high because the electron/hole recombination time is very short (2–5 ns). In order to obtain c.w. operation at the lowest possible current it is therefore necessary to minimize the active volume of the laser and confine both the carriers and the light to the narrowest possible filament between the end mirrors.

2.1 Heterostructure Confinement

In the vertical plane, carrier and optical confinement are obtained by cladding the GaInAsP active layer with n- and p-type InP passive layers. These passive layers have a wider band gap than the active layer thus forming a potential well which prevents out-diffusion of the injected carriers. The passive layers also have a lower refractive index than the active layer thus forming an optical waveguide which confines the light to the plane of the active layer as it propagates between the mirrors. The minimum thickness of active layer is between 0·1 and 0·2 μm being limited by the breakdown of this optical waveguiding. The room temperature threshold current density is then 1000–1500 A/cm^2.

For 0·8–0·9 μm wavelength semiconductor lasers the active layer is $Ga_{1-x}Al_xAs$ with $x = 0$–0·1. In this case the n- and p-type passive layers are $Ga_{1-x}Al_xAs$ with $x = 0{\cdot}3$–0·4. The passive layers perform the same functions of carrier and optical confinement. It is essential in all heterostructure lasers that there is a good crystal lattice constant size match between the active and passive layers of the double heterostructure and the substrate on which they are grown. If the lattice constant mismatch exceeds 0·3 to 0·4% then mismatch dislocations are formed at the layer interfaces causing rapid non-radiative recombination and higher thresholds.

In the $Ga_{1-x}Al_xAs$ system the aluminium atoms are almost exactly the same size as the gallium atoms they replace, so the lattice constant is virtually independent of x and good heterojunctions can be grown across the entire range of composition. This simplifying factor is the main reason why GaAlAs lasers were successfully developed many years earlier than GaInAsP lasers where the gallium and arsenic atoms are very mismatched in size compared to the indium and phosphorus atoms that they replace. This is illustrated in Fig. 2 which is a plot of lattice constant vs. bandgap for the $Ga_xIn_{1-x}As_yP_{1-y}$ system. To maintain lattice constant match to the InP

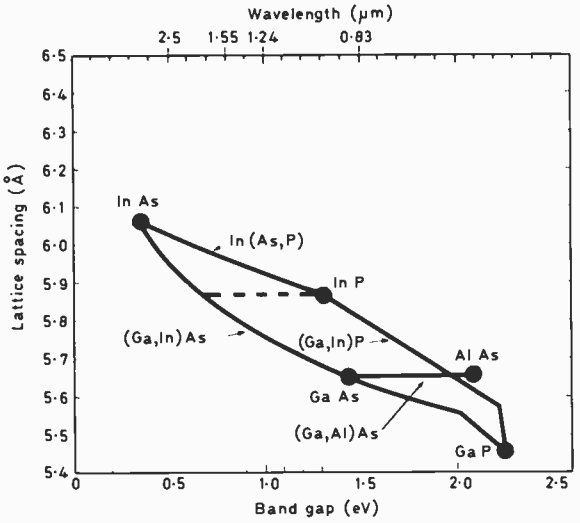


Fig. 2. Lattice spacing versus band gap for the (Ga In) (AsP) and GaAlAs systems.

substrate 2·1 arsenic atoms must be added for every gallium atom ($y/x = 2{\cdot}1$). The dashed line shows that the range of lattice matched GaInAsP compositions covers band gaps from 1·3 eV for InP down to 0·75 eV for the ternary $Ga_{0{\cdot}47}In_{0{\cdot}53}As$. Laser operation can be obtained from about 1·1 μm wavelength out to 1·67 μm for the ternary, but most effort is concentrated on the quaternary compositions with band gaps corresponding to 1·3 and 1·55 μm wavelength.

2.2 Lateral Confinement

In the plane of the active layer optical and carrier confinement can be obtained in several different ways depending on the device structure. In the oxide insulated stripe laser illustrated in Fig. 1, lateral carrier confinement is produced simply by restricting the area of the contact to typically 10–20 μm wide. Lateral current spreading occurs in the p-passive and p-capping layers and lateral diffusion of injected carriers occurs in the active layer. This causes the current distribution in the active layer beneath the stripe to be wider and more diffuse than at the contact itself. Typically an 'effective lateral diffusion length' of 2·5 to 4 μm is found depending on the resistivities and thicknesses of the p-type layers.[1] Thus, even if the stripe width approaches zero it is impossible to reduce the width of the pumped active layer below 5–8 μm.

In oxide stripe (and proton-bombarded stripe) lasers there is no 'deliberate' confinement of light in the lateral direction. However, the combined effects on refractive index of carrier concentration,[7] photoelastic[8] and thermal effects can produce either real dielectric optical waveguiding (usually only for stripes of greater than about 15 μm width) or antiwaveguiding (particularly strong in very narrow stripe lasers). In the case of antiwaveguiding the laser operates by 'gain guiding'[9,10] and behaves differently in many respects from lasers with a positive optical waveguide. The antiwaveguide causes

high internal optical losses which lower the quality factor of the laser cavity and results in a broad spectrum short coherence-length laser. This type of laser can minimize problems with 'modal noise'[11,12] in multi-mode fibre systems[13] although, as described in Section 4.2, there are other practical solutions. Typical room temperature c.w. threshold currents for oxide stripe lasers of either GaAlAs or GaInAsP are in the range 70–150 mA. They require a drive voltage about 0·4 V greater than the band gap—about 1·8 V for GaAlAs lasers and 1·3 V for GaInAsP at 1·3 μm.

2.3 Advanced Lasers

Lasers in which there is a deliberate lateral waveguide are classified as advanced laser structures. These lasers come close to the ideal of a single, very narrow lasing filament less than 3 μm wide propagating between the laser mirrors. For both GaAlAs and GaInAsP lasers the lateral waveguide can be produced by several different structures.[14–19] Two examples, the channelled substrate and the buried heterostructure laser, are shown in Figs. 3(a) and (b) respectively. The

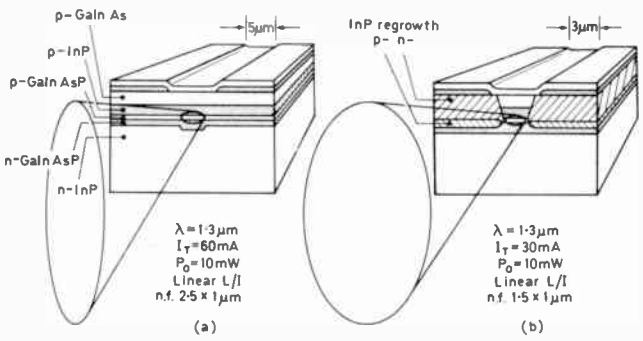


Fig. 3. Comparison of (a) channelled substrate and (b) buried heterostructure advanced lasers.

channelled substrate is the simplest to fabricate: a channel is etched in the substrate and the epitaxial layers are then grown by normal liquid phase epitaxy. The lateral waveguide is formed either by the direct optical loading effect of the substrate on the effective refractive index of the active layer[14,15] or in other types of channelled substrate laser[16,17] by the effect of the channel on the thickness (and hence effective refractive index) of the active layer. The waveguide dimensions are carefully controlled so that only the zero-order lateral mode propagates and a single elliptical spot is emitted from the output face as illustrated. Typical threshold currents for this type of laser, which usually suffers the same degree of lateral current spreading as oxide stripe lasers, are in the range 30–60 mA.

The buried heterostructure laser[18,19] is more complicated to fabricate but has the advantage of lateral current confinement as well as lateral optical confinement. The structure is made by growing a conventional double heterostructure wafer, etching to leave mesa stripes 1–2 μm wide, and then carrying out a second LPE growth to encase the remaining narrow active layer filament with the lower refractive index passive layer. The regrown layer is either high resistivity or includes a reverse biased p–n junction to minimize leakage current. The threshold current of the best of these devices is in the range 10–30 mA although good reproducibility is still difficult to achieve.

3 Laser Performance

Apart from the major differences between GaAlAs and GaInAsP lasers of operating wavelength and temperature sensitivity of threshold current, the performance of the lasers is more determined by the device structure than the material. So, for instance, the light vs. current curves, near-field and far-field patterns of oxide insulated stripe lasers of about 20 μm width are more or less indistinguishable between 0·85, 1·3 and 1·55 μm operating wavelength. This is because they have similar lateral waveguide mechanisms beneath the stripe contact. This Section therefore first illustrates the major difference between GaAlAs and GaInAsP lasers and then discusses the performance of the different types of device structure in more detail.

3.1 Light vs. Current Characteristics

The light vs. current characteristics of a 0·85 μm GaAlAs laser and a 1·3 μm GaInAsP laser are illustrated in Fig. 4. The curves are plotted at several temperatures to show how the threshold current increases with temperature. These lasers are both of the oxide insulated stripe variety with a stripe width of approximately 20 μm (subsequently referred to as 20 μm stripe). Both lasers have a room temperature threshold in the usual 100–200 mA range and require an extra 30–40 mA drive current above threshold to reach the maximum rated output power of 7 mW into a numerical aperture of 0·5 (60° cone). The most striking difference between the two sets of curves is the much more rapid increase in threshold current with temperature for the GaInAsP laser. The threshold current of semiconductor lasers

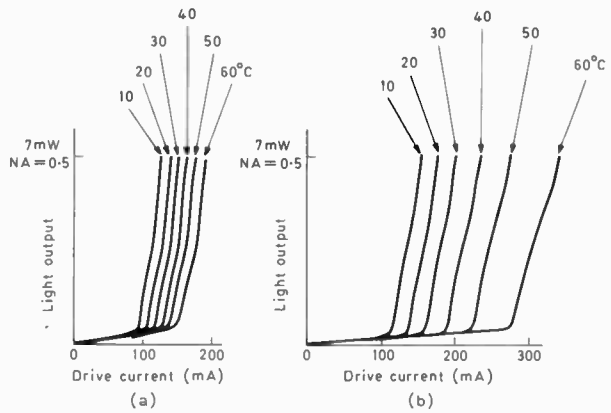


Fig. 4. Comparison of the light vs. current characteristics of (a) GaAlAs laser (b) GaInAsP laser as a function of temperature.

usually increases approximately exponentially with temperature so that threshold current sensitivity is expressed in terms of a characteristic temperature T_0 where:

$$I_t(T) = I_t(T_1) \exp \frac{T - T_1}{T_0} \qquad (1)$$

where I_t is threshold current

T is temperature.

Between 0 and 70°C, T_0 has a value of 100–180°C for GaAlAs lasers and 50–80°C for GaInAsP lasers. This corresponds to an increase of threshold current of 0·6–1·0%/deg C for GaAlAs lasers and 1·2–2·0%/deg C for GaInAsP lasers. The maximum temperature of c.w. operation is determined by the threshold current, the temperature sensitivity and the thermal resistance[13] of the heat sink. For oxide stripe lasers the maximum temperature is typically 100–120°C for GaAlAs devices and 50–60°C for GaInAsP devices. The best advanced structures can operate at temperatures 30 to 40°C higher than oxide stripe devices.

The increase in threshold current with temperature for GaAlAs lasers can be accounted for with reasonable accuracy simply in terms of the increasing energy spread of the electrons and holes injected into the conduction and valence bands. For 1·3 μm GaInAsP lasers experimental evidence[20] shows that the high temperature sensitivity is caused by an additional large and rapidly rising non-radiative component of the recombination current in the active layer. Other processes such as intervalence band absorption[43] may also be important, particularly at 1·5–1·7 μm wavelength. If the non-radiative recombination could be eliminated, the threshold current of 1·3 μm wavelength GaInAsP lasers would be reduced by a factor of about 3 at room temperature and the characteristic temperature T_0 increased to 180°C. Despite considerable world-wide efforts this problem remains unsolved and it appears increasingly likely that the non-radiative recombination is a basic property of GaInAsP. It is believed to be predominantly caused by the Auger recombination process.[20] If this is correct then the 60–90°C maximum temperature of operation of GaInAsP lasers will remain a limitation to their use. However, this is not too serious in the vast majority of long-haul land-based and submarine optical communication systems when compared with the tremendous advantages of operating at these wavelengths.

3.2 Optical Output Distributions

The near-field and far-field output distributions of semiconductor lasers are well understood and are determined by the strength of the waveguide guiding the light in the horizontal and vertical directions. In the vertical direction the strong waveguide formed by the double heterostructure layers confines the light tightly.

Table 1.

Near and far-field dimensions of semiconductor lasers.

Wavelength	λ = 0·85 μm				λ = 1·3 μm			
	near-field		far-field		near-field		far-field	
Structure	(μm)		(deg)		(μm)		(deg)	
	horiz.	vert.	horiz.	vert.	horiz.	vert.	horiz.	vert.
20 μm stripe	5	0·5	6	50	7	0·8	6	40
Channelled substrate	2·0	0·5	12	50	3	0·8	15	40
Buried hetero-structure	1·0	0·5	30	50	1·5	0·8	30	40

All dimensions are full-width half-maximum

For a typical active layer thickness of 0·2 μm only the zero-order transverse mode is guided and the optical distribution is sharply peaked in the active layer with evanescent tails penetrating into the transparent passive layers. The full width to half intensity of the distribution is typically 0·5 μm for an operating wavelength of 0·85 μm and 0·8 μm for 1·3 μm lasers. The far-field pattern formed by diffraction from such a narrow near-field is very wide, typically 40–60°. It can be calculated directly in terms of the Fourier Transform of the near-field distribution.[1–3]

The distributions in the lateral direction are determined entirely by the nature of the lateral waveguide which, in turn, is determined by the device structure. Typical figures for three different types of laser are given in Table 1. For a 20 μm oxide-insulated stripe laser the near-field distribution is broad as illustrated in Fig. 5, typically 5 μm wide for 0·85 μm wavelength and 7 μm wide for 1·3 μm wavelength. The distribution is wide because it is guided only by very small changes in refractive index in the plane of the active layer. At low output power the waveguide is predominantly formed by photoelastic and thermal effects.[8] As the output power

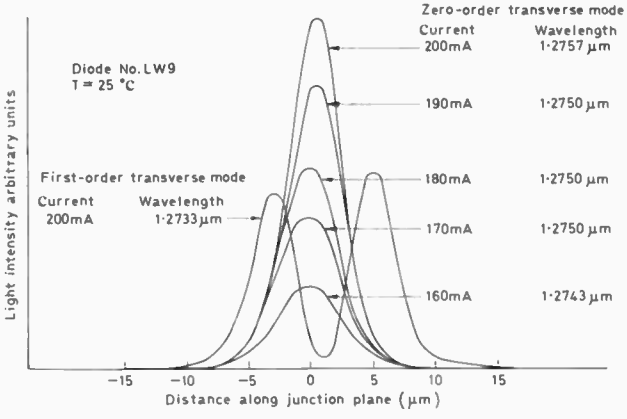


Fig. 5. Near-field of a 20 μm oxide-insulated GaInAsP stripe laser as a function of both current and wavelength.

increases there is a spatial hole burning effect which forms a dip in the carrier concentration along the lasing filament. Refractive index is inversely related to carrier concentration so the dip in carrier concentration acts as a hump in refractive index which increases the strength of the waveguide as output power increases. The effect of this is clearly seen in Fig. 5. As the output power increases, the width of the zero-order transverse mode near-field reduces slightly until the waveguide becomes sufficiently strong to guide the first-order transverse mode. The near-field then rapidly widens and becomes double peaked as the first-order mode begins to operate in addition to the zero-order mode. For 20 μm stripe lasers the maximum zero-order mode output is typically 3 mW–5 mW. In Fig. 5 the transverse modes have been separately resolved using a spectrometer. The onset of the first-order transverse mode can also be seen in the slight increase in slope efficiency of the light-current curves at 3–5 mW in Fig. 4.

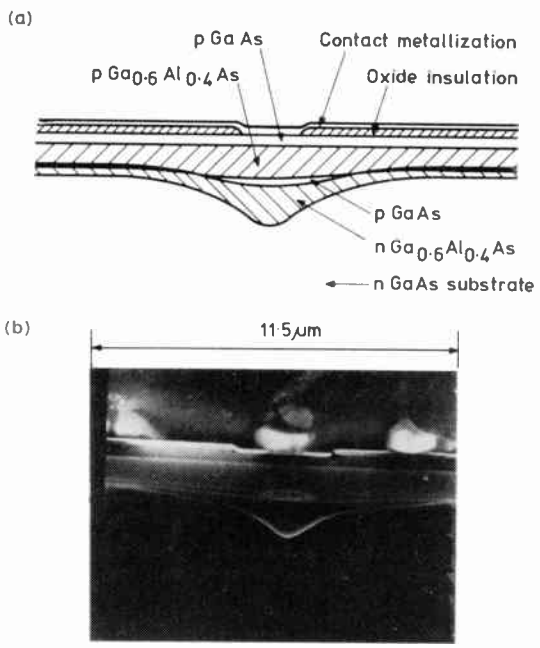


Fig. 6. Scanning electron micrograph of a channelled substrate narrow stripe (CNS) laser showing elliptical light output spot.

The lateral optical waveguides in advanced laser structures are much stronger than in oxide stripe lasers. The near-field distributions are therefore much narrower. In addition the effects of spatial hole burning are negligible compared with the deliberate waveguide so that the optical distribution is stable. Provided that the laser is correctly designed to suppress unwanted high-order transverse modes, considerably increased zero-order mode output power can be obtained. This is illustrated for a GaAlAs channelled substrate narrow stripe laser.[16] Figure 6 is a scanning electron micrograph of the device showing the crescent-shaped active layer grown above the V-shaped channel in the substrate. The elliptical spot shows the shape of the light output. Figure

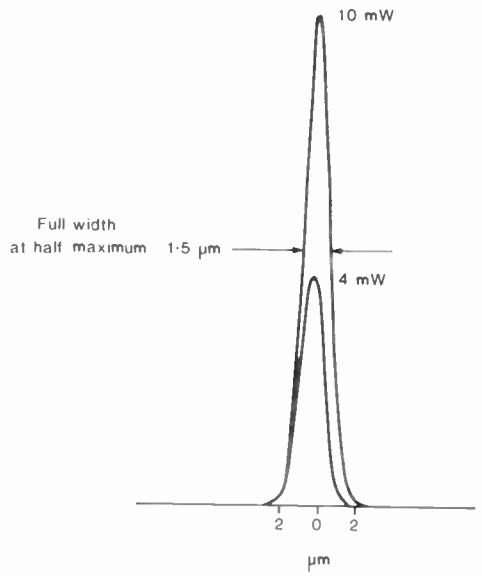


Fig. 7. Near field of a CNS laser, parallel to junction showing zero-order transverse mode operation up to 10 mW output.

7 shows that the near-field distribution is zero order mode and stable up to more than 10 mW output. Figure 8 shows the extremely linear transfer characteristics and that the device will operate at c.w. up to 160°C (the highest reported figure for any type of semiconductor laser).

3.3 Spectrum

The output spectrum of a semiconductor laser is a function of output power, temperature and modulation conditions. From the point of view of optical systems design it is important to know the peak wavelength λ_p

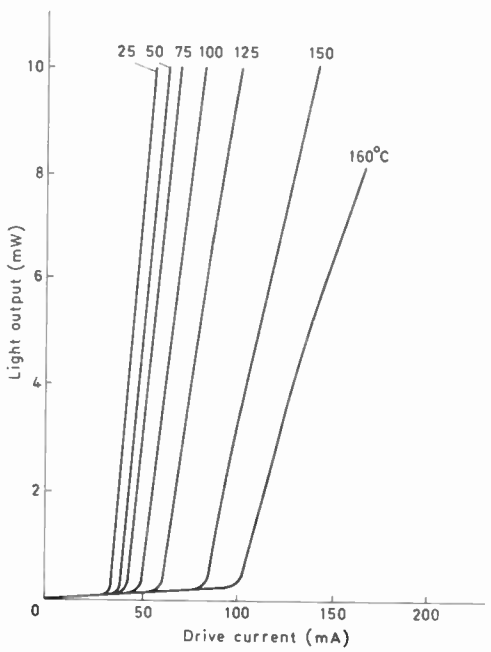


Fig. 8. C.w. light vs. current characteristics of a CNS laser showing operation up to 160°C.

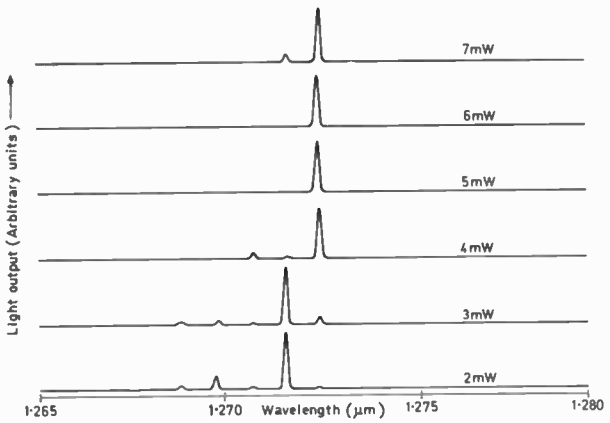


Fig. 9. D.c. spectrum of a c.w. laser at constant case temperature as output power is increased.

and spectral width $\Delta\lambda$ over the range of likely operating conditions. The spectrum of a 1·3 μm wavelength 20 μm stripe laser is plotted at increasing output powers in Fig. 9. The laser operates at one or more of its longitudinal mode wavelengths. The longitudinal modes correspond to the wavelengths at which there are an integral number of half-wavelengths along the length of the laser cavity. The spacing between the modes $\delta\lambda$ is therefore given by:

$$\delta\lambda = \frac{\lambda^2}{2L\mu} \qquad (2)$$

where L is the length of the laser

μ is the effective refractive index (about 4·5)

The effect of temperature on the laser spectrum is illustrated in Fig. 10. As temperature increases the peak of the laser gain spectrum moves towards longer wavelengths (because of the reducing band gap of the semiconductor). The effective refractive index μ also slowly increases so that the wavelength of the individual longitudinal modes increases at about one quarter of the rate of increase of the peak wavelength. The effect of this is for the peak wavelength to move from one mode to the next. The overall width of the spectrum $\Delta\lambda$ depends on the quality factor (finesse) of the laser cavity. For lasers with a low loss waveguide the $\Delta\lambda$ is often less than the longitudinal mode spacing $\delta\lambda$ so that the laser operates predominantly in a single longitudinal mode as illustrated in Fig. 9.

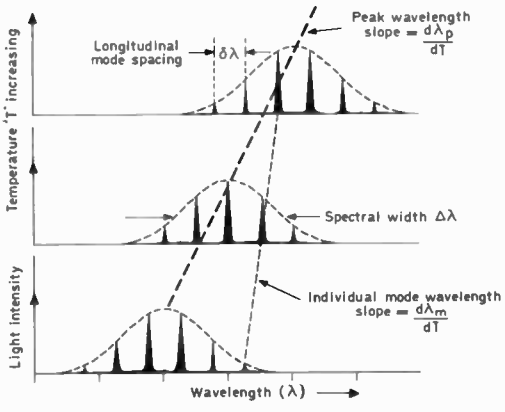


Fig. 10. Diagram to illustrate the effect of temperature on a laser spectrum.

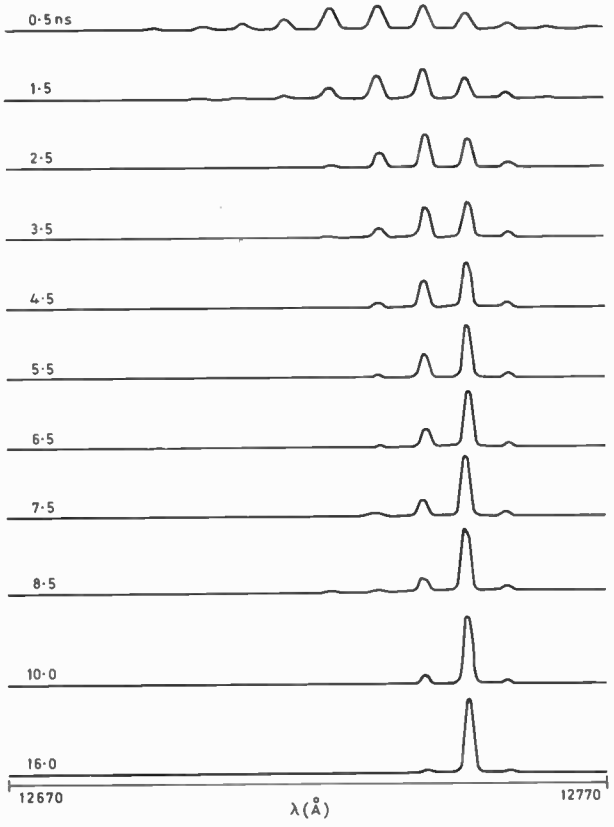


Fig. 11. Transient spectral evolution of a 20 μm stripe laser.

The line width of the individual longitudinal modes is well below the resolution of a conventional spectrometer. Interferometric measurements show a coherence length of 1–5 metres corresponding to a line width of less than 0·01 Å or 300 MHz. However, unless sophisticated wavelength-selective mirrors are used there will always be particular temperatures at which the laser jumps from one longitudinal mode to the next and two modes are momentarily operating. In practice there are other factors which broaden the spectrum. For instance if the laser is switched on from below its threshold there is a period of several nanoseconds during which the laser operates in many modes. This is illustrated in Fig. 11 for a 1·3 μm laser. The spectral width $\Delta\lambda$ decreases from 2·5 nm at 0·5 ns after switch on to single longitudinal mode at 16 ns. Thus when the laser is being modulated with pulses shorter than about 10 ns and the laser goes below threshold between pulses there can be considerable spectral broadening.

As several parameters affect the laser spectrum it is only possible to specify the spectral width under a precise set of operating conditions. The system designer is usually interested in the worst-case spectral width $\Delta\lambda_{max}$ under the specific modulation conditions used in the system. For the purposes of this paper, $\Delta\lambda_{max}$ is a conservative estimate of worst-case spectral width when the laser is modulated at several hundred Mbit/s at normal output powers (less than 7 mW). Table 2 gives typical values of this parameter together with the longitudinal mode spacing and the temperature coefficients of the peak wavelength $d\lambda_p/dT$ and individual mode wavelengths $d\lambda_M/dT$. It is assumed that the laser is 300 μm long. These figures are only approximate and should be checked in detail for specific laser types. It can be seen that the temperature coefficients, as well as the longitudinal mode spacing, scale approximately with the square of wavelength. This is because a given difference in energy corresponds to a difference in wavelength which increases as the square of the wavelength concerned.

Table 2.

Spectral properties of semiconductor lasers.

λ_p	850 nm	1300 nm	1550 nm
$\delta\lambda$ (nm) ($L = 300$ μm)	0·27	0·6	0·9
$\Delta\lambda_{max}$ (nm)	2·0	5·0	7·0
$d\lambda_p/dT$ (nm deg^{-1}C)	0·22	0·5	0·73
$d\lambda_M/dT$ (nm deg^{-1}C)	0·06	0·12	0·18

3.4 Modulation Response

The semiconductor laser is usually modulated by adding the modulating signal (usually 40–60 mA p-p) to a d.c. bias approximately equal to the laser threshold current. The d.c. bias is usually controlled by a feedback signal from a detector monitoring the output from the back of the laser. The feedback circuit is designed to keep the laser biased to the correct operating part of its transfer characteristic even as the laser threshold varies with temperature or ageing. The lasers are normally used for transmitting digital signals although pulse frequency modulation and analogue signals can be applied.

Once above threshold the laser responds extremely rapidly having a resonance in modulation response at a frequency f_R given by:

$$f_R = \frac{1}{2\pi}\left(\frac{I - I_{th}}{I_{th}(B\tau_p\tau_n)}\right)^{\frac{1}{2}} \qquad (3)$$

where I is the drive current
I_{th} is the threshold current
τ_p is the photon lifetime
τ_n is the carrier recombination time

and $B = 1 - \dfrac{I_t}{I_{th}} \simeq 0{\cdot}25$

where I_t is the current required to render the active layer transparent (gain = 0).

The resonant frequency response of an 850 nm 20 μm oxide stripe laser is illustrated in Fig. 12. The maximum rate of modulation is limited to about 300 Mbit/s by ringing effects as the data rate approaches within about a factor of 2 of the resonant frequency. The resonance is about 1 GHz when the device is 10% above threshold.

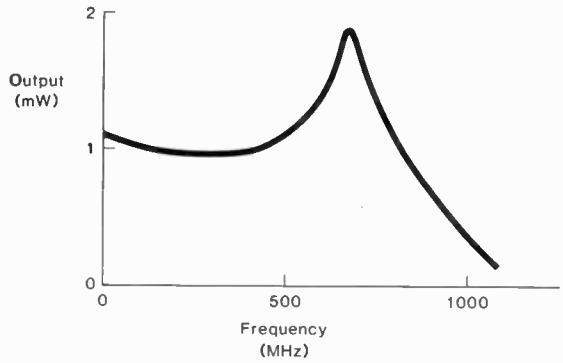


Fig. 12. Experimental frequency response of an SiO_2 stripe laser with a sinusoidal modulation.

The equation shows that the resonant frequency increases as the square root of the *proportional* drive current above threshold. This factor causes the bandwidth of advanced lasers to be nearly double that of simple oxide stripe lasers because the threshold

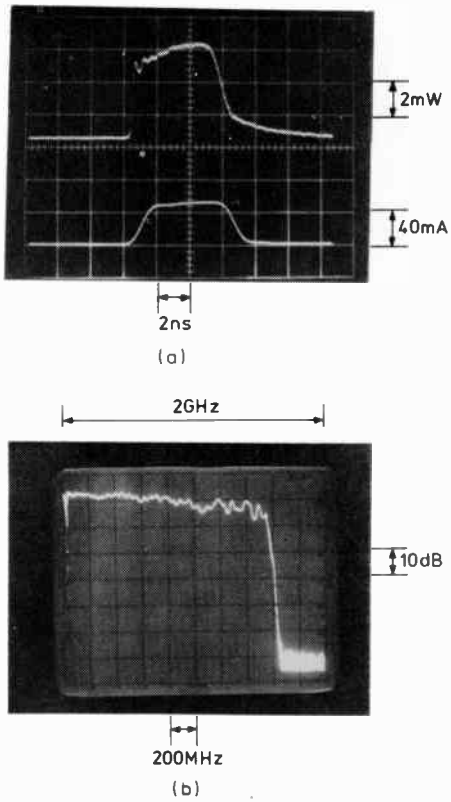


Fig. 13. Pulse response (a) and frequency response (b) of CNS laser showing flat response to beyond 1·4 GHz.

current is reduced by a factor of 3 to 4 times whereas the drive current above threshold is similar. In practice many advanced laser structures show virtually no ringing and have an almost flat frequency response out to well beyond 1 GHz. Figure 13 shows the pulse response and frequency response of a channelled substrate narrow stripe laser. The response is approximately flat out to beyond 1·4 GHz which was the limit of the measuring equipment. This device could be modulated at data rates well beyond 1 Gbit/s. The reason for this desirable damped resonant response of many advanced lasers is not understood in detail, but contributing factors are thought to be the effect of higher spontaneous emission coupled into the lasing modes, and the effects of lateral diffusion of carriers in the active layer.[1] The frequency response of GaInAsP lasers is in principle about 40% higher than the frequency response of GaAlAs devices simply because the carrier recombination time τ_n is approximately halved (2 ns for GaInAsP compared with 4 ns for GaAlAs). This is one desirable result of the otherwise undesirable rapid non-radiative recombination in GaInAsP devices. On this basis it seems likely that the best GaInAsP advanced lasers will have approximately flat response out to at least 2 GHz.

3.5 Laser Reliability

The double heterostructure GaAlAs laser was invented in 1970,[2] but it was not until 1973 that any laboratory achieved more than 1000 hours operating life from a c.w. laser.[21] During this period, semiconductor lasers gained a reputation for poor reliability. Since then there has been considerable worldwide effort aimed at improving operating life so that room temperature lives well in excess of 10^5 hours can now be routinely achieved by many manufacturers. The cause of the rapid degradation of early GaAlAs lasers was the development of 'dark line defects'[22] which are arrays of dislocations spreading from an initial threading dislocation. The dislocation motion is 'lubricated' by the electron-hole recombination in the active layer. The reproducible and long operating lives of present day GaAlAs lasers can be attributed to minimizing the dislocation density of substrate material, minimizing applied mechanical stress (by, for example, the use of indium solder for dice bonding), the use of passivating facet coatings, and developing contact metallizations that do not degrade with time. The residual ageing in 10^5 hour time scales is believed to be caused by the motion and accumulation of residual point defects.

The operating life of the GaAlAs 20 μm oxide stripe laser which is now commercially available[23] is extremely good. A total of more than 3·0 million lifetest hours have now been accumulated and the longest running devices have been operating on c.w. at room temperature for more than 44 000 hours (5 years).[24] The ultimate lives

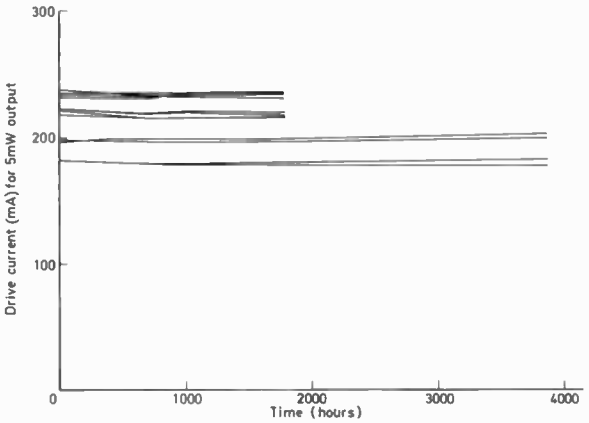


Fig. 14. Drive current as a function of time for GaAlAs lasers from four different wafers operating at 70°C 5 mW output.

can only be estimated by extrapolation and accelerated testing at high temperature. The median extrapolated life to 50% increase in threshold at 70°C and 5 mW output is at least 30 000 hours, with a standard deviation $\sigma = 0{\cdot}4$ (90% of lasers exceed 18 000 hours and 10% exceed 50 000 hours). The activation energy of the residual ageing in these devices is approximately 0·3 eV to 0·4 eV,[25] giving a predicted median life at room temperature to twice initial threshold of the order of 400 000 hours which is approaching 50 years. To illustrate the good reliability the threshold current of four recent sets of lasers operating at 5 mW output at 70°C is plotted over 4000 hours in Fig. 14. These curves indicate virtually zero change in threshold in this time showing evidence that even the excellent figures quoted here can still be improved upon.

The operating life of advanced GaAlAs laser structures is also good.[26] The 70°C extrapolated lives of the CNS devices described in Section 3.2 are greater than 10 000 hours. The life tests also show that in both the 20 μm stripe and the CNS laser there is no serious increase in laser noise with ageing. This is believed to be a direct result of the stability of the built-in waveguides of both devices.

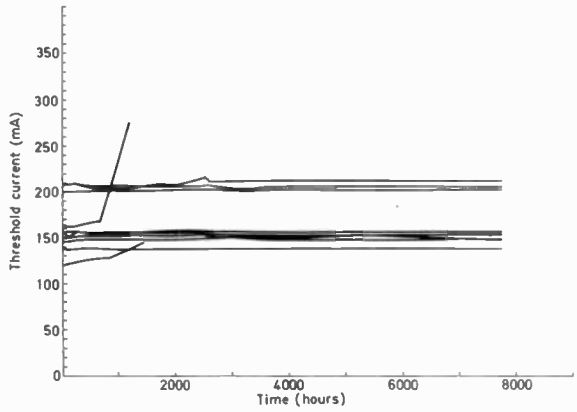


Fig. 15. Threshold current as a function of time for GaInAsP lasers operating at 20°C and 2 4 mW output power.

There is much less data available on the operating life of long wavelength GaInAsP lasers, because they are still under rapid development. Figure 15 shows the life test results on a batch of 10 early GaInAsP lasers.[27] Eight of the devices have now reached 8000 hours with virtually *no* measurable change in threshold. These results are remarkable when compared with GaAlAs devices because they are known to contain a very high density (about 10^5 cm^{-2}) of dislocations as low dislocation density InP substrates were not available when the devices were made. This indicates that at normal temperatures the movement of dislocations is strongly inhibited compared with GaAlAs devices. This may be because of the 'dislocation pinning' effect[28] of the different sized atoms of the GaInAsP compound alloy. The output facets of GaInAsP devices are also found to be much less susceptible to oxidation and erosion under operating conditions than those of GaAlAs devices.[30]

These factors support the view that it is considerably easier to obtain long operating lives with GaInAsP devices and that the ultimate operating lives will be well in excess of those of GaAlAs devices.

3.6 Packaging

The packaging of semiconductor lasers for high reliability applications is a very demanding task. In addition to the requirements of a normal high reliability electrical package it must allow ± 3 μm precision line up with a lens-coupled multimode fibre, have low inductance drive leads, contain an integral detector for feedback control and monitoring, and have hermetic seals on all leadthroughs including the optical fibre output. An example of such a package developed by ITT Paignton[20] is shown in Fig. 16. 'Stretched' versions of this package are available which include a hybrid circuit for feedback control. Versions are also being developed with the $\pm 0{\cdot}3$ μm alignment precision required for lens-coupled single mode fibres. Experimental packages have already succeeded in launching more than 1 mW at 40% efficiency from a 1·3 μm laser into a single mode fibre using a miniature cylindrical lens of 8 μm radius[31] as illustrated in Fig. 17.

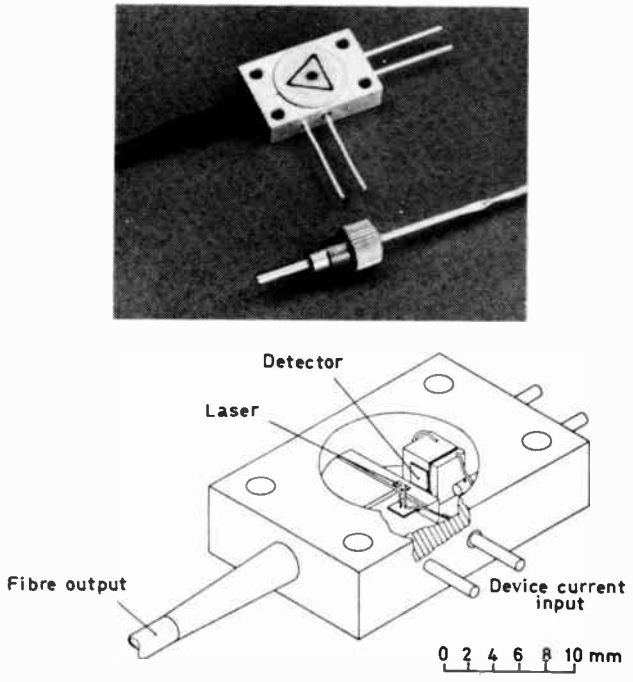


Fig. 16. Modular laser package with integral monitor detector (ITT Paignton).

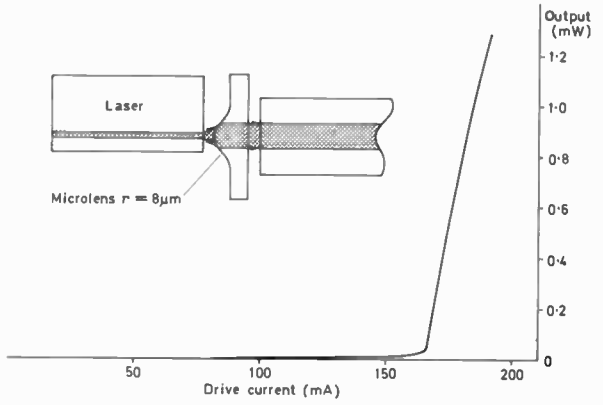


Fig. 17. Experimental launching of more than 1 mW into a single-mode fibre using cylindrical microlens.

4 Lasers for Multimode Optical Fibre Systems

4.1 Distance and Bandwidth

Until recently the majority of optical fibre transmission systems have used multimode graded index silica fibres typically of 125 μm o/d and 50 μm core diameter. The performance of these fibres is summarized in Table 3 which gives typical and best figures based on information presented at the Sixth European Conference on Optical Communication.[29] The best figures are now coming very close to the theoretical limitations so that it seems likely that commercially available fibres will only gradually improve upon the typical figure quoted. In order to illustrate the relative performance of laser/multimode fibre systems at the three important wavelengths the following simplifying assumptions can be made which are fairly typical of those made for high data rate PTT applications.

Laser mean launched power	−3 dBm
Receiver sensitivity	−43 dBm
System margin and splicing	10 dB
Allowed fibre loss	30 dB

These figures are only approximate, in real systems there are many variables. The receiver sensitivity for instance is correct for a typical 1·3 μm GaInAsP p-i-n f.e.t. receiver at 10^{-9} error rate and 160 Mbit/s.[32] The sensitivity increases slowly (about 5 dB/decade) for lower data rates and decreases at up to 15 dB/decade for

Table 3.

Multimode graded index silica fibre performance.

Wavelength		0·85 μm	1·3 μm	1·55 μm
Attenuation ($dB\ km^{-1}$)	typical	2·5	1·0	0·5
	best	2·0	0·5	0·22
Modal bandwidth ($GHz \cdot km^{-1}$)	good typical	1·5	1·5	1·5
	best	2–4	6	2–4
Wavelength dispersion ($ps \cdot nm^{-1} \cdot km^{-1}$)		−85	−5	+20

higher data rates. At 1 Gbit/s the maximum sensitivity for a very low capacitance p-i-n f.e.t. detector is about −38 dBm. The sensitivities for 0·85 μm silicon detectors are typically 5 dB better. However, in this paper, for the purposes of comparing the different wavelength laser sources for communication links, a figure of 30 dB for maximum fibre loss has been adopted.

Table 4.

Maximum distance and data rate for laser multimode fibre systems.

Wavelength	0·85 μm	1·3 μm	1·55 μm
Maximum transmission distance (km)	12	30	60
$\Delta\lambda_{max}$ (nm)	2	5	7
Wavelength dispersion (ns)	2†	0·75	8·4†
Modal bandwidth (MHz)	125	50†	25
Maximum bit rate (Mbit/s)	125	100	30

† Limiting factor on data rate

The maximum repeater spacing can then be calculated from the 'typical' attenuation figures of Table 3. The results are given in Table 4 and show that 12, 30 and 60 km can be covered at 0·85, 1·3 and 1·55 μm respectively. These figures show the tremendous advantages of moving to longer wavelengths. In order to estimate the limiting effect of wavelength dispersion on the system operating at its maximum repeater spacing it is assumed that the maximum time dispersion that can be tolerated is one quarter of a bit period. This is sufficiently conservative that possible noise effects caused by laser mode partition noise[1] and possible jitter effects caused by mode hopping between individual longitudinal modes can be neglected.

At 0·85 μm wavelength the bandwidth is limited to 125 Mbit/s by the effect of the high wavelength dispersion of silica on the 2 nm maximum line width of the laser. At 1·3 μm wavelength the wavelength dispersion of the fibre is very close to zero so that this factor is not a limitation. Instead the 1·5 GHz·km assumed fibre bandwidth which is caused by the slightly differing velocities of propagation of the many modes of the fibre, (intermodal dispersion), limits the maximum data rate to about twice the 3 dB bandwidth i.e. 100 Mbit/s over 30 km. At 1·55 μm wavelength the maximum data rate is limited to only 30 Mbit/s over 60 km by the severe effect of material dispersion on the 7 nm maximum linewidth of the laser.

4.2 Modal Noise

The phenomenon of 'modal noise' must be taken into account when designing laser multimode fibre systems. Modal noise occurs when a highly coherent laser source is used in a multimode fibre system which has mode-selective loss (e.g. misaligned connectors).[11,12] This is because coherent light forms a speckle pattern in the fibre core as the many different fibre modes interfere with one another (multi-path interference). The speckle pattern varies rapidly as the laser is modulated because of minute variations in emission wavelength; and it also varies if the fibre is flexed or vibrated. The varying speckle pattern is converted into amplitude noise at a misaligned connector because it transmits a varying proportion of the light in the core. Modal noise can severely limit the performance of multimode optical communication systems unless it is minimized by using an adequately low coherence laser source and/or sufficiently high quality connectors. The majority of optical links are digital, where a well proved technique to obtain the required signal-to-noise ratio is to use return to zero (r.z.) pulse modulation of the laser. For example in a 140 Mbit/s link the laser can be modulated using half width (3·5 ns) pulses which return the laser to below its threshold after each pulse. The laser spectrum is then consistently broadened (and the coherence length reduced) as previously illustrated in Section 3.3 and Fig. 3. This technique however, will not work if the pulse length is more than about 7 ns because the laser becomes more coherent as time into the pulse increases.

For low data rates, and for analogue modulation, recent experiments[33] indicate that the spectrum can also be broadened by applying an r.f. bias signal of 6–20 mA p-p amplitude at 700–1500 MHz frequency, which has the effect of rapidly scanning the precise laser frequency. Any modal noise effects then occur at very high frequency and are not detected in the signal bandwidth. This technique is not yet fully proved but deserves more detailed investigation.

The third alternative is to use a low coherence laser as briefly described in Section 2.2. These devices are stripe lasers with a stripe width of less than about 3 μm. They can be either oxide insulated or proton bombarded or planar zinc diffused. The advantage of this type of laser is that the spectral width never reduces below 1–2 μm (at λ_p = 850 nm) even for c.w. operation, so modal noise is

low whatever type of modulation is employed.[12] There are, however, doubts about the long-term stability of this type of laser. The lossy gain guide which operates is much less stable than the built-in waveguide of the wide oxide stripe and advanced lasers. Problems with increasing laser noise and self pulsation have been widely reported during the ageing of this type of laser.[34,35]

The laser noise is fluctuation of laser output for a steady d.c. input. It has a very wide bandwidth of greater than 1 GHz with peak amplitude at the laser resonance frequency. In good quality lasers this noise is negligible for digital communication links, being 40–60 dB down on the signal level on 100 MHz bandwidth. In poor lasers this intrinsic noise is greatly amplified by instability of the laser filament (often associated with kinks in the light versus current characteristic). Self pulsation is the extreme form of instability when the output consists entirely of narrow pulses usually less than one hundred picoseconds wide at a repetition frequency above several hundred MHz. In our experience, lasers with a built-in waveguide such as the 20 μm stripe and CNS laser show only slight changes in laser noise with ageing and have *never* been observed to develop self pulsation.[24] This is not true of gain guided narrow stripe lasers made on the same wafer material. As laser fabrication technology improves, the problems with narrow stripe lasers may be overcome.

4.3 Choice of Laser Type

Provided that the system is correctly designed to minimize modal noise then all types of semiconductor laser will perform well in multimode systems. Advanced lasers have the advantage of lower threshold current but often this is not particularly significant. The most important factor in the long-haul systems is usually the laser reliability. This factor tends to favour the first-generation lasers which have built up the most reliability data. A good example of this at 850 nm is the 20 μm oxide stripe laser which, as described in Section 3.5, has extremely good reliability and is well proved in this type of system. The fact that it has a low zero-order mode power is of no great significance as the zero-order and first-order transverse modes which are operating at the maximum rated output of 7 mW couple equally efficiently into the multimode fibre.

4.4 Comparison with L.E.D. Systems

The distance-bandwidth product of multimode fibre–l.e.d. systems at 0·85 μm and 1·55 μm is severely limited by the effect of material dispersion on the broad spectrum of the l.e.d. However, at 1·27 μm wavelength the material dispersion is zero and this problem is minimized. In this case, based on the 'typical' fibre of Table 3 an l.e.d. with −13 dBm output can transmit over 20 km at up to 150 Mbit/s. The bandwidth is limited by the 1·5 GHz·km assumed for the fibre modal bandwidth (assuming a high-speed l.e.d. is used). The complex effects of second-order dispersion on the l.e.d. spectrum, which in fact spreads from less than 1·2 μm to beyond 1·4 μm, also cause considerable pulse waveform distortion at about this data rate.

In summary the four prime sources for long-haul high-data rate multimode systems are the 1·55 μm laser (60 km at 30 Mbit/s), the 1·3 μm laser (30 km at 100 Mbit/s), the 1·3 μm l.e.d. (20 km at 150 Mbit/s), and the 0·85 μm laser (12 km at 125 Mbit/s). At 1·3 μm wavelength bandwidth is limited by the fibre assumed modal bandwidth of 1·5 GHz·km. In principle it is possible to improve considerably upon the bandwidth of 1·3 μm multimode systems by very careful control of the fibre refractive index profile but there are serious doubts about what sort of yields would be possible in production. There is now a wide body of opinion which believes that it will be simpler and more economical to eliminate inter-modal dispersion entirely by using laser single-mode fibre systems.

5 Lasers for Single-mode Fibre Systems

Single-mode fibre systems are now being rapidly developed. The elimination of the intermode dispersion of the fibre allows even higher data rates to be transmitted over long distances. The achievement of reproducible connectors, splices and laser packages with high optical coupling efficiency is of course more difficult with the very small (about 6–10 μm) core size of single-mode fibres. However, the initial work in the UK is very encouraging and much effort is now directed towards developing engineered systems for submarine and land line applications.

5.1 Distance and Bandwidth

The maximum distance of transmission is limited, as with multimode systems, by the fibre attenuation. Since the material for single-mode fibres is doped silica, as with multimode fibres, the attenuation of properly designed fibres is just as low. The material dispersion with wavelength is also similar. At 1·3 μm wavelength the same fairly conservative assumptions used for multimode systems, i.e. −3 dBm launched power, −43 dBm receiver sensitivity (at 160 Mbit/s), 10 dB for joints and margin, therefore give the same transmission distance of 30 km at 1 dB/km attenuation. However, with intermodal dispersion eliminated and material dispersion close to zero at 1·3 μm the bandwidth is increased enormously. In fact with single-mode fibres the wavelength for zero dispersion can be adjusted over a wide range by counteracting the effect of material dispersion with waveguide dispersion.[36] The waveguide dispersion is determined by such factors as the fibre core diameter and refractive index step. A plot of fibre dispersion against wavelength is given in Fig. 18, showing that the dispersion minimum can be adjusted

over the range 1·36–1·55 μm.[37] (It should be noted that at present there is still a significant attenuation penalty in fibres whose dispersion minima have been moved out as far as 1·55 μm.)

When a system is operating close to the dispersion minimum it is useful to work not in terms of the dispersion $d\tau/d\lambda$ (where τ is the delay of a signal at wavelength λ) but to use the rate of change of dispersion with wavelength $d^2\tau/d\lambda^2$. This second-order dispersion parameter has a value of approximately 0·085 ps·km^{-1}·nm^2 for the curve with a dispersion minimum at $\lambda_0 = 1{\cdot}36$ μm of Fig. 18 reducing to 0·055 ps·km^{-1}·nm^2 for the curve with $\lambda_0 = 1{\cdot}55$ μm.

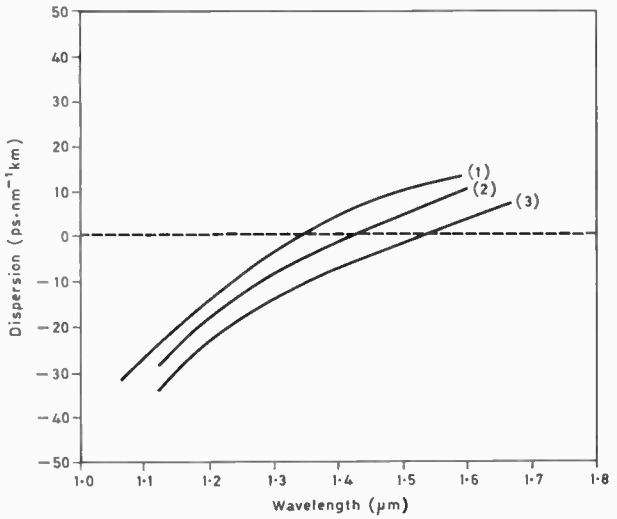


Fig. 18. Dispersion versus wavelength for three germania-doped silica single-mode fibres.

The time dispersion $\Delta\tau$ of a pulse with a wavelength spread of $\Delta\lambda_{max}$ and a peak wavelength λ_p is then given by:

$$\Delta\tau = \frac{d^2\tau}{d\lambda^2}\,\Delta\lambda_{max}\,(\lambda_p - \lambda_0) \qquad (4)$$

if

$$\frac{\Delta\lambda_{max}}{2} < |\lambda_p - \lambda_0|$$

and

$$\Delta\tau = \frac{d^2\tau}{2d\lambda^2}\left(|\lambda_p - \lambda_0| + \frac{\Delta\lambda_{max}}{2}\right)^2 \qquad (5)$$

if

$$\frac{\Delta\lambda_{max}}{2} > |\lambda_p - \lambda_0|$$

Thus for example if a laser with a waveform spread of 5 nm has its peak wavelength λ_p coinciding exactly with the fibre dispersion minimum λ_0 at 1·36 μm, then the dispersion (eqn. (5)) is only $\Delta\tau = 0{\cdot}26$ ps/km or 8 ps over 30 km. This corresponds to an immense data rate in excess of 30 Gbit/s.

In practice the system designer needs to know the data rate under worst case conditions rather than the maximum data rate calculated above. This requires that the worst case of $\lambda_p - \lambda_0$ is estimated. The following example shows that even with relatively large tolerances very wide bandwidths can still be obtained. Taking $\lambda_p = 1300 \pm 20$ nm, $\lambda_0 = 1300 \pm 20$ nm and allowing the laser temperature to vary by ± 20°C at 0·5 nm deg^{-1}C the worst case value of $\lambda_p - \lambda_0$ is 50 nm. For $\Delta\lambda_{max} = 5$ nm the dispersion is then given by equation (4) as 21 ps/km which is 640 ps over the 30 km. The maximum data rate under these conservative assumptions is therefore still greater than 390 Mbit/s. ($\Delta\tau$ less than one-quarter of a bit period.) These figures illustrate that single-mode fibre systems operating at nominal laser wavelength equal to the dispersion minimum have very high bandwidths even when worst case assumptions are made. Such systems have been demonstrated in several laboratories[44] world wide. For instance in recent experiments at STL, 320 Mbit/s were transmitted over 35 km of single-mode fibre with 6–8 dB system margin.[38] British Telecom now has plans to expand the trunk communication network in Britain using 1·3 μm single-mode fibre systems with a maximum length of 30 km at 140 Mbit/s. If approved, these systems will be installed between 1985 and 1990. There will clearly be considerable scope for increasing the data rate should this become necessary.

5.2 System Noise Effects

Unwanted noise effects in laser single-mode systems have been the subject of considerable recent speculation. At first sight, since the fibre is single-mode there should be no equivalent of modal noise. However, single-mode fibres do in fact have two orthogonal polarization modes. Because all real fibres are slightly birefringent (due to slight ellipticity for instance), these two modes propagate at slightly different velocities. In general this causes the state of polarization of a launched wave to change along the fibre. R. Epworth has demonstrated that if there is any polarization-selective loss such as might be caused by misaligned joints or sharp bends in the fibre, then polarization modal noise can occur.[39] Such noise can certainly be observed if the output from a single-mode fibre optical communication system is passed through a polarizing filter before the detector. However, provided that deliberate polarizing devices are avoided there are no reports of significant polarization modal noise in any of the experimental systems fabricated to date. The present writer's current view is therefore that this effect is not sufficiently large to be significant in single-mode systems using pulse modulation.

One form of noise that single-mode systems are particularly sensitive to is laser reflection noise.[38,40,41] A severe case of this is illustrated in Fig. 19.[38] This was caused by the effect of a reflection from the bare end of a single-mode fibre pigtail on a rather unstable 20 μm

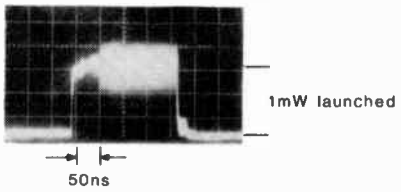


Fig. 19. A severe case of laser reflection noise in 350 MHz bandwidth (4% reflection at 2 m followed by 5 km of single mode fibre).

stripe laser ($\lambda = 1{\cdot}3$ μm). It can be seen that there is very little noise until 20 ns into the pulse when the back reflection from the fibre end first returns; the noise increases to very high levels after the third round trip at 60 ns. The noise has a very wide bandwidth in excess of 1 GHz. The cause of the noise is interference between the back reflected wave and the laser internal wave inside the laser cavity as illustrated in Fig. 20. Single-mode systems are more likely to suffer from this effect than multimode systems because back reflected light is coupled just as effectively into the laser as the laser is coupled into the fibre. In multimode systems, however, the back reflected light is spread in many modes, most of which do not couple well into the laser cavity.

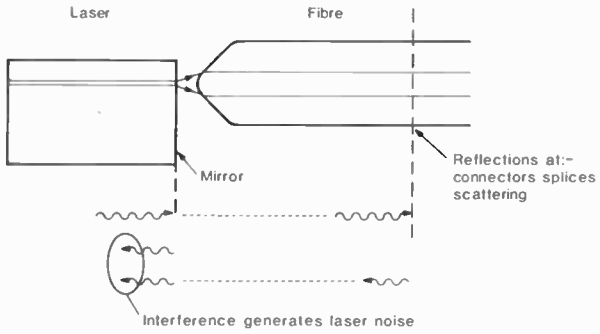


Fig. 20. Diagram illustrating the cause of reflection noise.

There are two obvious methods to eliminate or minimize reflection noise, these are:

1. Eliminate reflections by using spliced or index matched fibre joints. This technique has worked satisfactorily in all STL experimental systems to date.
2. Use an optical isolator. This uses a Faraday rotation and polarizing element to cut out back reflected waves.[38]

Perhaps less obviously however, the stability of the laser itself plays an important role. Laser instability can be caused by several factors, one of the most common of which is instability of the optical distribution in oxide stripe lasers with very weak built-in guiding effects. Such a laser was used to illustrate the severe noise effects of Fig. 19. Properly-designed oxide stripe lasers and advanced lasers with strong built-in waveguides do not suffer from this form of instability so that reflection noise effects are much less significant. An example of satisfactory modulation of a 1·3 μm 20 μm stripe laser coupled into a single mode fibre is shown in Fig. 21. This shows modulation at 320 Mbit/s (non-return to zero). The signal is detected through 6 km of single mode fibre and shows no sign of reflection noise. The ringing, clearly visible at the beginning of each pulse, is the laser resonance. This is removed by filtering in the detector circuits.

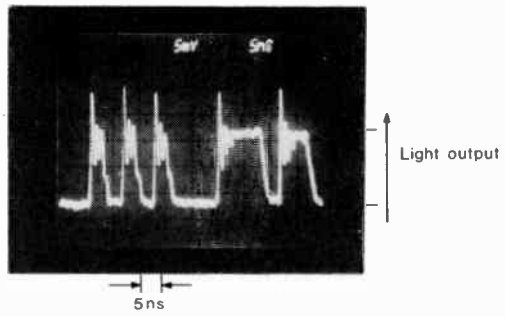


Fig. 21. Modulation of a 1·3 μm laser at 320 Mbit/s (n.r.z.) output from 6 km of single-mode fibre showing no reflection noise effects.

5.3 Choice of Laser Type for Single-mode Systems

Although single-mode systems are still in their infancy several factors support the view that advanced lasers are strongly to be preferred. The simplest factor is that a single-mode fibre requires a single-transverse-mode laser. The higher-order transverse modes of the 20 μm stripe laser simply do not couple into the fibre core, limiting the useful output power of this type of laser to about 3 mW.

Since modal noise problems appear to be insignificant in single-mode systems there are no apparent advantages in using the 'low coherence' narrow stripe type of laser. The astigmatic output beam, broad spectrum and potentially unstable optical distribution are all undesirable features. Advanced lasers have the advantages of a stable high zero-order transverse-mode output that can be effectively coupled into single-mode fibres using suitable lenses. The much wider frequency response of advanced lasers will also prove necessary as the data rates of single mode systems are pushed up above a few hundred Mbit/s.

The best choice of the many different types of advanced laser will probably depend more on factors such as reliability rather than on detailed differences in specification. At 850 nm wavelength the most reliable advanced lasers are of the channelled substrate type of structure[25] probably because of the high quality of the heterojunction interfaces. This may also apply for GaInAsP lasers but several more years of lifetest data is required before this will become certain.

6 Future Developments

Since fibre attenuation is such a key parameter it seems very likely that during the coming decade fibre attenuation will gradually fall towards the scattering loss

limits of 0·4 dB/km and 0·2 dB/km at 1·3 and 1·55 μm respectively. For a 30 dB allowed fibre loss, this would increase the maximum transmission distances towards 75 and 150 km respectively. The fact that fibre dispersion and laser spectral widths are large at 1·55 μm will make the attainment of wide bandwidths at this wavelength much more difficult than at 1·3 μm. It therefore seems likely that 1·3 μm systems will be pushed to their limits first. This will involve closer control of the laser peak wavelength and spectral width and the minimum dispersion wavelength of the fibre. The maximum direct modulation rate of an advanced GaInAsP laser is probably around 2 Gbit/s (Sect. 3.4). In order to achieve this data rate over 50 km the dispersion must be below 2·5 ps/km. Such a system also requires the use of maximum sensitivity receivers, fibre attenuation very close to the ultimate, and better launch efficiency from the laser to increase the launched power. This might in fact be achieved with the better present-day lasers. For example, if the peak laser wavelength is allowed to vary from the dispersion minimum by up to ±20 nm (±10 nm for manufacturing tolerances and ±10 nm for ±20°C), then, according to equation (4), the maximum spectral width must not exceed 1·47 nm, which would allow up to three longitudinal modes to operate. Provided that a high quality advanced laser were used, biased above threshold at all times this is certainly not out of the question. The difficulty comes in being certain that the laser will *never* exceed this spectral width. It will require much development work to determine the practical limits.

In applications such as submarine communications there is a high premium placed on maximizing repeater spacing. This can best be achieved at 1·55 μm where distances of over 100 km will soon be practical. In order to achieve, let us say, 320 Mbit/s over 100 km, equation (4) indicates that with the conservatively specified laser spectral width from Table 2 of 7 nm, the fibre dispersion minimum would need to be within 20 nm of the laser wavelength. Whether such a 1·55 μm fibre could be made with low attenuation and reproducible minimum dispersion wavelength is speculative. For the purposes of this paper, let us assume that this cannot be achieved and that the lowest loss 1·55 μm fibres have their dispersion minima close to 1·3 μm. The fibre dispersion will then be in the region of 15 $ps \cdot nm^{-1} km^{-1}$ at 1·55 μm. If the maximum allowed time dispersion is one quarter of a bit period then in order to achieve the 320 Mbit/s data rate the laser line width will have to be reduced below 0·5 nm. This is less than the longitudinal mode spacing. That is to say the laser must *always* remain truly single longitudinal mode. This cannot be achieved using conventional semiconductor lasers unless the temperature is stabilized to better than ±0·5degC and there are no ageing effects. The basic problem is not that the laser wavelength varies with temperature, but that the peak gain wavelength moves approximately four times more rapidly with temperature than the individual mode wavelength (Table 2), so that mode hopping occurs about every 1½degC. Such mode hopping produces a jitter on the received pulses. One feasible solution to this problem is therefore to completely quench the internal laser modes (for example by the application of quarter wave anti-reflection coats) and design a wavelength-selective external cavity which is temperature compensated to track correctly the wavelength of the peak gain of the laser with temperature. The cavity dimensions would have to be kept small or direct modulation of the laser at high bit rates would become difficult. Again much effort would be necessary to prove that the laser *never* jumped wavelengths. However, the linewidth of a single longitudinal mode is so small that transmission at data rates well in excess of 1 Gbit/s could be achieved over more than 100 km using such techniques.

In conclusion, using present-day 1·3 μm lasers with single-mode silica fibres and making fairly conservative assumptions about spectral width and wavelength control of the laser, worst-case data rates in excess of 390 Mbit/s over 30 km are obtained. As fibre losses and control of the laser centre wavelength are improved, data rates up to 2 Gbit/s over 50 km can be expected close to the 1·3 μm dispersion limit. If 1·55 μm lasers can be controlled to operate *always* in a true single longitudinal mode then data rates of well in excess of 1 Gbit/s become feasible over distances of more than 100 km. The almost unlimited bandwidth potential of single-mode silica fibre will provide a constant stimulus for the development of even higher data rate communication systems during the coming decades.

7 Acknowledgments

This paper is based on the work of many other people in STL: in particular, G. D. Henshall, P. D. Greene, G. H. B. Thompson, D. S. O. Renner and S. E. H. Turley, together with others in the long wavelength laser group, and A. R. Goodwin, R. G. Plumb and J. P. Curtis and others in the short wavelength laser group. The Engineering group of the Microwave and Opto Electronic Unit, ITT Components Group, Paignton, Devon, UK have contributed considerably to the work on reliability of GaAlAs lasers and carried out extensive packaging work referred to in the text. The author thanks STL for permission to publish this paper. This work has been carried out with the support of British Telecom Research Laboratories at Martlesham and the Procurement Executive, UK Ministry of Defence, sponsored by CVD.

8 References

1 Thompson, G. H. B., 'Physics of Semiconductor Laser Devices' (Wiley, Chichester, 1980).
2 Casey, H. C. and Panish, M. B., 'Heterostructure Lasers' (Academic Press, New York, 1978).

3 Kressel, H. and Butler, J. K., 'Semiconductor Lasers and Heterojunction LEDs' (Academic Press, New York, 1977).
4 Hsheih, J. J., 'Room temperature operation of GaInAsP/InP double heterostructure diode lasers emitting at 1·1 μm', *Appl. Phys. Letters*, **28**, p. 283, 1976.
5 Yamamoto, T., Sakai, S. and Suematsu, Y., 'InGaAsP/InP d.h. lasers fabricated on InP(100) substrates', *IEEE J. Quantum Electronics*, **QE-14**, p. 95, 1978.
6 Henshall, G. D. and Greene, P. D., 'Growth and characteristics of GaInAsP/InP double heterostructure lasers', Solid State Electron Devices, **3**, pp. 174–8, 1979.
7 Kirkby, P. A., Goodwin, A. R., Thompson, G. H. B. and Selway, P. R., 'Observation of self-focusing in stripe geometry semiconductor lasers and the development of a comprehensive model of their operation', *IEEE J. Quantum Electronics*, **QE 13**, no. 8, pp. 705–19, August 1977.
8 Kirkby, P. A., Selway, P. R. and Westbrook, L. D., 'Photoelastic waveguides and their effect on stripe geometry GaAs/GaAlAs lasers', *J. Appl. Phys.*, **50**, no. 7, pp. 4567–9, July 1979.
9 Cook, D. D. and Nash, F. R., 'Gain induced guiding and astigmatic output beam of GaAs lasers', *J. Appl. Phys.* **46**, pp. 1660–72, April 1975.
10 Kobayashi, T., Kawaguchi, H. and Furakawa, Y., 'Lasing characteristics of very narrow planar stripe lasers', *Japan. J. Appl. Phys.*, **16**, p. 601, 1977.
11 Epworth, R. E., 'The phenomenon of modal noise in analogue and digital optical fibre systems', Fourth European Conference on Optical Communication, Genoa 1978.
12 Goodwin, A. R., Davis, A. W., Kirkby, P. A., Epworth, R. W. and Plumb, R. G., 'Narrow stripe semiconductor laser for improved performance of optical communication systems'. Proceedings of Optical Communication Conference, Amsterdam, September 1979.
13 Steventon, A. G., Spillett, R. E., Hobbs, R. E., Burt, M. G., Fiddyment, P. J. and Collins, J. V., 'C.w. operation of GaInAsP stripe lasers', *IEEE J. Quantum Electronics*, **QE-17** May 1981. (To be published).
14 Furuse, T., Sakuma, I., Ide, Y., Nishida, K., and Saito, F., 'Transverse mode stabilised AlGaAs having a built-in plano convex waveguide'. Proceedings of Optical Communication Conference, Amsterdam, September 1979.
15 Aiki, K., Nakamura, M., Kuroda, T., Umeda, J., Ito, R., Chinone, N. and Maeda, M., 'Transverse mode stabilised AlGaAs lasers with channelled substrate planar structures', *IEEE J. Quantum Electronics*, **QE-14**, p. 89, 1978.
16 Kirkby, P. A., 'Channelled substrate narrow stripe GaAs/GaAlAs injection lasers with extremely low threshold currents', *Electronic Letters*, **15**, no. 25, pp. 824–6, 6th December 1979.
17 Oomura, E., Higuchi, H., Hirano, R., Namizaki, H., Murotani, T., Susaki, W. and Shirahata, K., 'InGaAsP/InP buried crescent laser diode ($\lambda = 1{\cdot}3$ μm) with very low threshold current and fundamental transverse mode'. Proceedings of the Third International Conference on Integrated Optics and Optical Fibre Communication, San Francisco, California, April 1981.
18 Tsukada, T., 'GaAs GaAlAs buried heterostructure injection lasers'. *J. Appl. Phys.*, **45**, p. 4899, 1974.
19 Nakano, T., Takahei, K., Naguchi, Y., Tokunaga, M., Nagai, M., '1·5 μm InGaAsP/InP BH lasers on p-type InP substrates', *Japan. J. Appl. Phys.*, **19**, p. 1612, 1980.
20 Thompson, G. H. B. and Henshall, G. D., 'Non-radiative carrier loss and temperature sensitivity of threshold in 1·27 μm GaInAsP/InP d.h. lasers', *Electronics Letters*, **16**, no. 1, pp. 42–4, 3rd January 1980.
21 Hartman, R. L., Dyment, J. C., Hwang, C. J. and Kuhn, M., 'Continuous operation of GaAs–GaAlAs d.h. lasers with 30°C half lives exceeding 1000 hours', *Appl. Phys. Letters*, **23**, no. 4, pp. 181–3, 15th August 1973.
22 Hutchinson, P. W. and Dobson, P. S., 'Defect structure of degraded GaAlAs GaAs double heterojunction lasers'. *Philos. Mag.* (*GB*), **32**, pp. 745–54, no. 4, October 1975.
23 'Summary Catalogue–Laser Diodes', ITT Components Group, Brixham Road, Paignton, Devon.
24 Goodwin, A. R. and Plumb, R. G., private communication.
25 Ritchie, S., Godfrey, R. F., Wakefield, B. and Newman, D. H., 'The temperature dependence of degradation mechanisms in long lived (GaAl)As d.h. lasers', *J. Appl. Phys.*, **49**, p. 3127, 1978. (and recent private communication).
26 Curtis, J. P., Plumb, R. G., Goodwin, A. R. and Kirkby, P. A., 'Characteristics and analysis of channelled substrate narrow stripe GaAs/GaAlAs lasers', *J. Appl. Phys.* **52**, 1981, (to be published).
27 Henshall, G. D. and Greene, P. D., 'GaInAsP/InP DH lasers as sources for optical fibre communications operating at 1·3 μm wavelength', Electro-optics/Laser International Brighton U.K., March 1980, pp. 260–7.
28 Kirkby, P. A., 'Dislocation pinning in GaAs by the deliberate introduction of impurities', *IEEE J. Quantum Electronics*, **QE-11**, no. 7, pp. 562–8, July 1975.
29 Sixth European Conference on Optical Communication, York, September 1980.
30 Morimoto, M., Imai, H., Hori, K., Takusagawa, M. and Fukuda, M., 'Facet degradation of InGaAsP/InP double heterostructure lasers', *Appl. Phys. Letters*, **37**, no. 12, pp. 1082–4, 15th December 1980.
31 Bricheno, T., Eales, B. A. and Leggett, N. D., private communication.
32 Smith, D. R., Hooper, R. C., Ahmad, K., Jenkins, D., Mabbitt, A. W. and Nicklin, R., 'p-i-n/f.e.t. hybrid optical receiver for longer wavelength optical communication systems', *Electronics Letters*, **16**, no. 2, pp. 69–71, 17th January 1980.
33 Vanderwall, J. and Blackburn, J., 'Suppression of some artifacts of modal noise in fibre optic systems', *Optics Letters*, **14**, no. 9, pp. 295–6, September 1979.
34 Hartman, R. L., Paoli, T. L. and Dixon, R. W., 'Stabilization of ageing induced self pulsations and the elimination of an initially saturable mode of degradation in AlGaAs lasers by means of facet coating', *Appl. Phys. Letters*, **35**, p. 905, 1979.
35 Nelson, J. and Dutta, N. K., 'Self sustained pulsations and negative resistance behaviour in InGaAsP ($\lambda = 1{\cdot}3$ μm) double heterostructure lasers', *Appl. Phys. Letters*, **37**, p. 709, 1980.
36 Cohen, L. G., Lin, C. and French, W. G., 'Tailoring of zero chromatic dispersion in single mode fibres into the 1·55 μm region', *Electronics Letters*, **15**, pp. 334–5, 1979.
37 Byron, K. C., 'Dispersion measurements on single mode fibres in the wavelength range 1 to 1·7 μm', European Conference on Optical Systems and Applications', September 1980, Utrecht.
38 Farrington, J. G., 'Reflection noise effects on error rate in a 35 km single mode transmission system'. Submitted to Seventh European Conference on Optical Communications, Copenhagen, September 1981.
39 Epworth, R. E. and Pettitt, M. J., 'Polarisation modal noise and fibre birefringence in single mode fibre systems', Third International Conference on Integrated Optics and Optical Fibre communication', San Francisco, April 1981.
40 Mazurczyk, V. J., 'Sensitivity of single mode buried heterostructure lasers to reflected power at 274 Mbit/s', *Electronics Letters*, **17**, no. 3, pp. 143–4, 5th January 1981.
41 Hirota, O. and Suematsu, Y., 'Noise Properties of Injection Lasers due to reflected waves', *IEEE J. Quantum Electronics*, **QE-15**, no. 3, pp. 142–9, March 1979.
42 Iwamura, H., Hayashi, S. and Iwasaki, H., 'A compact optical isolator using $Y_3Fe_5O_{12}$ crystal for near infra red radiation', *Optical & Quantum Electronics*, **10**, pp. 393–8, 1978.
43 Adams, A. R., Asada, M., Suematsu, T. and Arai, S., 'The temperature dependence of the efficiency and threshold current of $In_{1-x}Ga_xAs_yP_{1-y}$ lasers related to intervalence band absorption', *Japan J. Appl. Phys.*, **19**, pp. L621–4, 1980.
44 Hooper, R. C., Smith, D. R. and White, B. R., 'Long repeater span monomode optical fibre systems', IEE Colloquium on Optical Fibre Systems–Present and Future, London, May 1981.

Manuscript first received by the Institution on 24th March 1981 and in final form on 8th May 1981.
(*Paper No. 1999/CC 343*)

UDC 621.372.8: 621.391.63/4
Indexing Terms: Optical communications, Fibre optics

Systems applications of optical fibre transmission

P. E. RADLEY, B.A., Ph.D.*

SUMMARY

This paper describes a number of systems to demonstrate the wide application of fibre optic technology. It covers both undersea and landline trunk applications together with a discussion of power budgets and an indication of the benefits of long wavelength technology. Lower-speed digital systems for junction applications and avionics, military and industrial use are described as well as the transmission of analogue signals such as television.

* *Standard Telecommunication Laboratories, London Road, Harlow, Essex CM17 9NA*

1 Introduction

This paper is intended to be a view, as seen from an industrial laboratory, of the market for optical communications and a set of products which have been, or are being, developed in response. The products are based largely on the UK technology developments reported elsewhere in this issue.

In the subsequent Sections a number of aspects of systems for public network applications as well as products for industrial and military use are described. These are intended to demonstrate the wide area of applications already being considered for fibre optics.

2 Public Network Applications

2.1 Trunk Landline Systems

Trunk systems operating at 140 Mbit/s have figured largely in the development of optical technology for PTT systems. With STC and STL this began with a demonstration 140 Mbit/s system installed in 1977. The primary objective was to demonstrate its suitability for application in PTT networks, and that optical fibre systems could be installed and operated alongside conventional telecommunications equipment.

A 9 km long optical fibre communications system, including two intermediate repeaters, was installed in existing cable ducts between two British Post Office exchanges in Hitchin and Stevenage.[1] The route was particularly complex in the variety of tight bends and level changes, with most of the duct system being permanently flooded.

A major accomplishment was the production of large quantities of fibre with good and consistent attenuation, dispersion, geometry and strength, and the laying-up of the fibre into practical cables with minimal changes in performance.

A full hierarchy of digital multiplex equipment was included in the demonstration to provide flexibility in testing system traffic. Continuous testing and measurement of telephone traffic signals has provided vital data confirming the long-term stability and performance of this new technology. Nearly four years of field data on lasers and fibres has uncovered no unexpected problems.

The link has also been used in co-operation with the British Broadcasting Corporation to carry out a successful series of colour television test transmissions. A 140 Mbit/s signal comprising two digitally encoded 625-line colour television channels and twelve 15 kHz sound programme channels was transmitted without any discernible impairment in signal quality.

Figure 1 shows a fibre cable designed for trunk system use being pulled in during this exercise and contrasts with the large drums of thick coaxial cable we often see by the roadside. This drum is approximately 1 metre in diameter, it contains up to 2 km of optical cables and can be carried by a man; in contrast, a drum carrying up to 2 km of multi-tube coaxial cable is approximately 2·4

0033–7722/81/070377+08 $1.50/0

Fig. 1. Landline optical cable installation.

metres in diameter and has to be carried by a 5 tonne truck.

The key advantages of fibre optics in the trunk system are a smaller cable, with potentially a lower cost than conventional cable, and, most importantly, increased length of repeater section. With mature optical technology at a wavelength of 0·85 μm, section lengths of 8 km can be achieved with good margins; this gives lower costs relative to p.c.m. trunk transmission on coaxial cable.

Figure 2 illustrates the relative cost of various trunk systems with dotted trend lines indicating likely changes

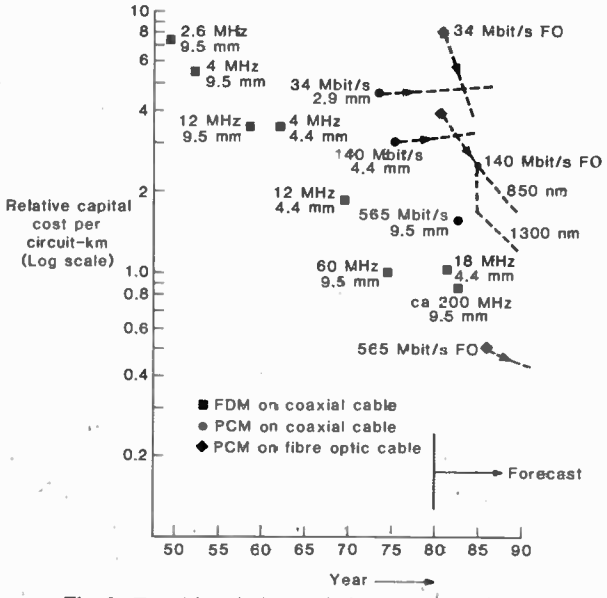


Fig. 2. Trend in relative capital cost per circuit-km.

with time. This shows clearly that the optical 140 Mbit/s system is expected to be lower in cost than its coaxial cable counterpart in the very near future.[2]

Optical technology is still however developing very rapidly and Fig. 2 shows a step decrease in cost per channel by using a longer wavelength of 1·3 μm. This has the effect of considerably increasing the section lengths achievable to the extent that 30 km of cable at 1·3 μm represents the same optical loss as an 8 km section at 0·85 μm. However, the increased lengths mean that the bandwidth limitation, caused by dispersion along the fibre, must be reduced; monomode fibres are very suitable to meet this requirement. Practical fibres are now being repeatably demonstrated which will meet both the loss and dispersion figures for 30 km sections.

This 30 km figure, in the UK network at least, is of great significance since it is the maximum distance between power feed points, and hence surface stations. This can therefore result in further cost savings because neither power feed nor supervisory equipments are needed and a completely non-metallic cable can be used. In some countries a shorter distance may be applicable, whereas in others, where trunk routes are very long, the long repeater spacings still provide the benefits of lower cost and higher reliability, despite still requiring power feed and supervision of buried repeaters.

2.2 Submerged Systems

Undersea cable systems began in the nineteenth century with telegraphy and developed for telephony in 1956. Despite the much vaunted threat of communication satellites, cable remains an integral part of the international network. Cables find their place in application to shorter routes, particularly in Europe and on the long routes, such as across the Atlantic, where they are used with satellites to provide route diversity and hence security of service.

One significant technology change has already taken place with the advent of transistors in 1964 to replace thermionic valves. The change to optical technology now proposed will probably be even more significant.

The number of routes and their capacities have increased substantially and steadily over recent years and this trend is predicted to continue. As a result pressure is being applied to evaluate methods of achieving reductions in cost per channel particularly for high capacity routes. Optical technology can potentially provide this cost reduction.[3]

Figure 3 shows the comparison of cost per channel for current f.d.m. coaxial cable systems when compared with digital optical systems; the coaxial cable curve assumes that different systems are used for each capacity whereas the optical cable curve assumes 140 Mbit/s per fibre and capacities greater than 1920 circuits by using multiple pairs of fibres. The reduction in cost per channel for the digital optical systems largely arises out of the assumed regenerator spacing of 50 km (as compared with 5 km for

today's 45 MHz analogue system).

A programme is now under way to provide such systems in service from the second half of the 1980s onwards. The first milestone of this programme has already been achieved with the laying of a trial cable.[4] Figure 4 shows the deep water cable design.

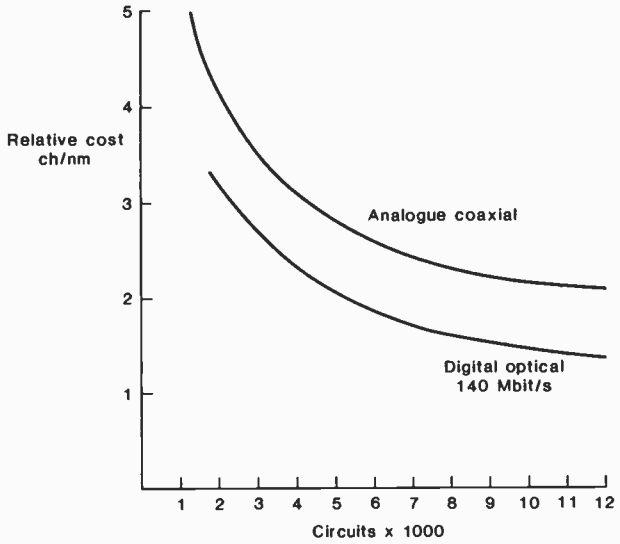
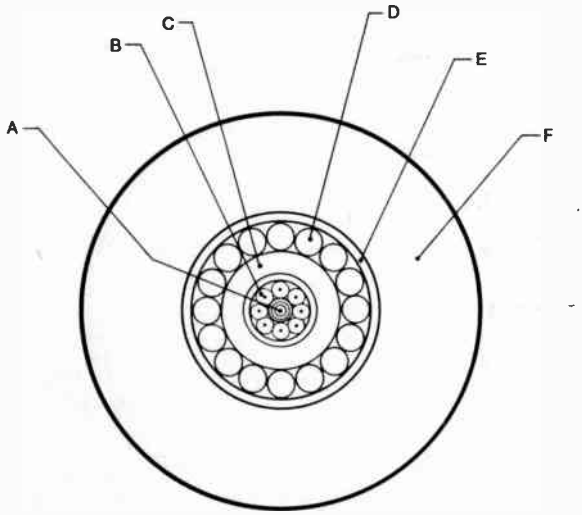


Fig. 3. Submerged system cost comparison.



A – Central strength member
B – Optical fibres
C – Aluminium tube
D – High tensile steel wires
E – Copper tube
F – Polythene insulant

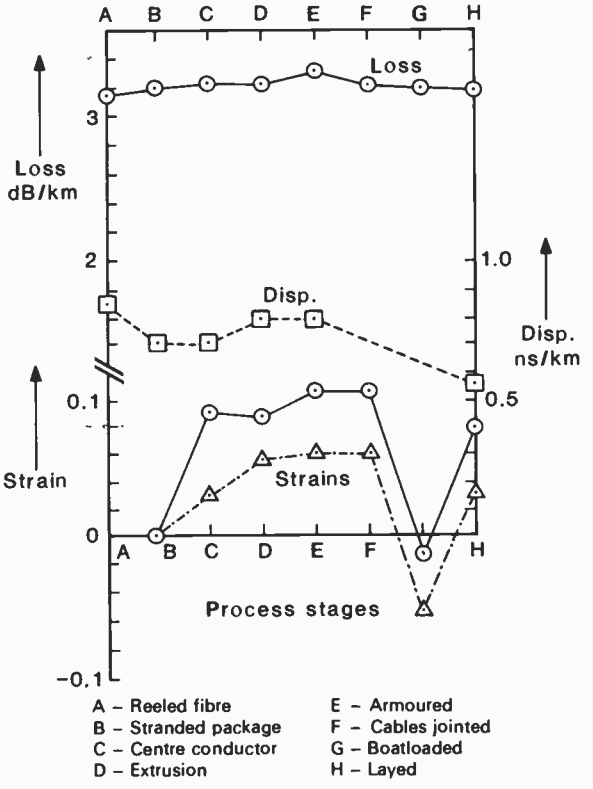
Fig. 4. Deep water cable design.

The cable made for the trial is an armoured version of the lighter weight deep water design. Although it is unconventional to armour cables having a high tensile steel central strength member, for optical cable there are good reasons for so doing. The central strength member contributes significantly to the overall cable strength, thus minimizing fibre elongation during armouring, coiling and subsequent laying. The basic cable characteristics are:

Overall diameter	26 mm
Weight in water	0·42 tonnes/km
Tensile strength	92 kN
D.c. resistance	0·49 ohms/km
No. of fibres	8

The cable for the trial system was manufactured in the latter part of 1979 and laid by the Post Office Cable Ship *Iris* in February 1980. The cables were manufactured in nominal 2 km section lengths and jointed after armouring in the factory to produce 4 km and 6 km section lengths. Each cable has six fibres, four multimode graded index for operation at 0·85 μm, two single-mode fibres for operation at longer wavelengths, plus two constantan wires for strain monitoring.

After loading onto the ship the cable sections were connected to a repeater housing and the fibres spliced inside the housing to provide continuous paths for monitoring. The cable was laid in a loop in Loch Fyne (Scotland) at an average depth of about 100 m. This is a tidal loch with similar sea bed conditions to those typical of systems laid across the North Sea. The two cable ends were brought ashore to a terminal station for transmission measurements. The multimode fibre loss and dispersion measurements during processing and laying are summarized in Fig. 5. These are the average of 20 individual fibres and show that there is negligible attenuation change at any stage. The decrease in dispersion per unit length that was observed when the 2 km section lengths were jointed up to 10 km



A – Reeled fibre
B – Stranded package
C – Centre conductor
D – Extrusion
E – Armoured
F – Cables jointed
G – Boatloaded
H – Layed

Fig. 5. 10 km trial cable fibre strains and loss.

corresponds to a (length)$^{0.8}$ law.

Also shown in Fig. 5 is the cable strain of two of the cable sections during manufacture and laying. The large decrease in strain that occurred on loading was produced by coiling the cable anticlockwise in the ships' tank. During laying, this negative strain was relieved leaving a nett strain in the fibres approximately equal to that existing prior to armouring.

The second phase of the programme to recover the repeater housing from the sea bed and relay it was successfully completed in the middle of 1980. This tested the capability of the cable and repeater housing design to withstand this operation, which provides the highest strain experienced by the cable. In addition a 140 Mbit/s regenerator was spliced into the housing, operating over the multimode fibre.

2.3 Junction Systems

Unlike trunk and submerged systems, the economics for optical fibres in the junction network are not as clear. The benefit of low losses allowing long distances to be spanned is of no great advantage since the vast majority of links are below 10 km in length. Also cost per channel savings due to high capacities are not found in many circumstances.

Transmission rates of 8 Mbit/s and 34 Mbit/s are finding favour for this type of optical system both in the UK and elsewhere. The shorter distances and smaller capacities in the junction network allow fibre of much lower bandwidth to be used and it is possible that a lower grade and less costly fibre could be adopted for these routes. In the UK the greatest need for such systems could arise in urban areas where ducts may already be crowded. Some countries have coaxial cable instead of pair cables in their junction areas, and fibre systems are already attractive there since repeaters are seldom required even with operation at wavelengths of 0·85 μm. Since there are more junction than trunk routes, we can expect this to be the larger equipment market once fibre systems prove to be economic.

The 8 Mbit/s system[5] uses a laser operating at 0·85 μm with an avalanche photodiode (a.p.d.) detector. This combination offers operation over repeater sections of up to 10–12 km with allowances for joints, connectors and operating margin. For distances greater than 12 km, regenerative repeaters are inserted at regular intervals.

Terminal equipment is shown in Fig. 6. In the transmit path the incoming 8448 kbit/s data stream is converted to a simple binary form together with extraction of the incoming 8448 kHz clock. The binary signal is then scrambled and converted to provide the line signal. The output of the laser is stabilized by automatic feedback control. A demountable connector provides the connection to the optical fibre cable.

In the receive path, the avalanche photodiode and pre-amplifier convert the optical signals into an electrical signal. Short sections can be catered for by adding an optical attenuator in the detector assembly. The binary pulses are regenerated and the receive clock in the terminal repeater is extracted. The signal is then examined for any signal errors, and after descrambling to recover the original binary signal, it is converted into the HDB3 interface code and passed to the output port.

The line terminal equipment also includes the associated alarm circuits and indicators used for fault diagnosis and the power stabilizer units for operation from a 24V or 50V battery supply.

2.4 Power Budget Planning

The performance of the systems described above can be tabulated in the form of a 'power budget' to apportion the various losses in a system, within the limits of launched and received power.

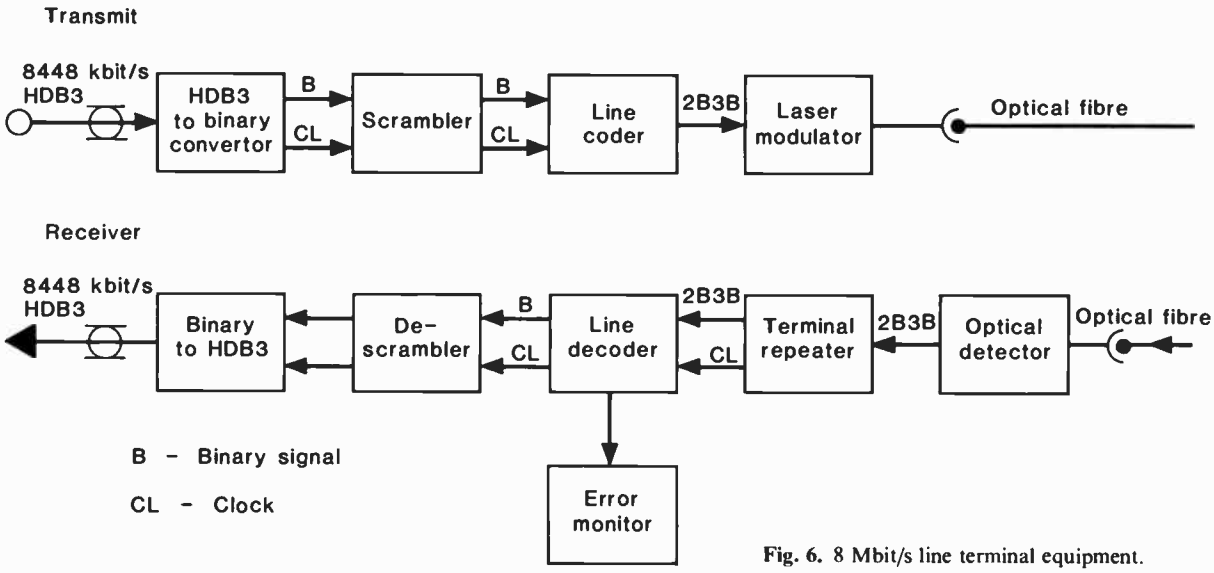


Fig. 6. 8 Mbit/s line terminal equipment.

A typical power budget for a 140 Mbit/s system operating over 8 km of multimode fibre at a wavelength of 0·85 μm is as follows:

Launched power	−4·5 dBm
Received power	−48·0 dBm
Allowable loss	43·5 dB
System margin	7·0 dB
2 connectors (1 dB each)	2·0 dB
Dispersion penalty	6·0 dB
Fibre attenuation (8 km)	24·0 dB
9 slices (0·3 dB each)	2·7 dB
Total optical loss	41·7 dB
Additional margin	1·8 dB

The launched power is for a laser with a butt launch into a fibre; launched powers over 3 dB greater can be achieved with lens coupling into the fibre. A peak power of 1·5 mW in the fibre leads to a mean power of 0·375 mW (−4·5 dBm) assuming equal marks and spaces and half-width pulses. The received power relates to a bit error rate of 1 in 10^9 with a silicon a.p.d. and is dependent upon the type of pulse launched; for example the receiver sensitivity is decreased by launching full-width pulses because fibre dispersion causes greater difficulty in differentiating seaparate pulses at the receiver.

The system margin is a relatively arbitrary figure to cover ageing and temperature effects in the optical path, electro-optic components and electronics; this is a figure based on experience on p.c.m. systems in general. The dispersion penalty is caused by the bandwidth limitation of a fibre; the relationship is shown in Fig. 7. This graph assumes that the pulse spectrum is 100% raised-cosine at the decision point and that the fibre transfer function is a Gaussian low-pass filter.

The power budget for an 8 Mbit/s system operating at 0·88 μm over 12 km is as follows:

Launched power	−6·0 dBm
Received power	−63·0 dBm
Allowable loss	57·0 dB
System margin	7·0 dB
2 connectors (1 dB each)	2·0 dB
Dispersion penalty	0 dB
Fibre attenuation (12 km)	42·0 dB
13 splices (0·3 dB each)	3·9 dB
Total optical loss	54·9 dB
Additional margin	2·1 dB

The launched power in this case is lower than for 140 Mbit/s because of the duty factor when using a 2B3B code, whilst maintaining the same peak power. The receiver power for the same type of silicon a.p.d. is lower due to the greater sensitivity which can be achieved at 8 Mbit/s; receiver sensitivity figures for various bit rates are compared in Table 1.

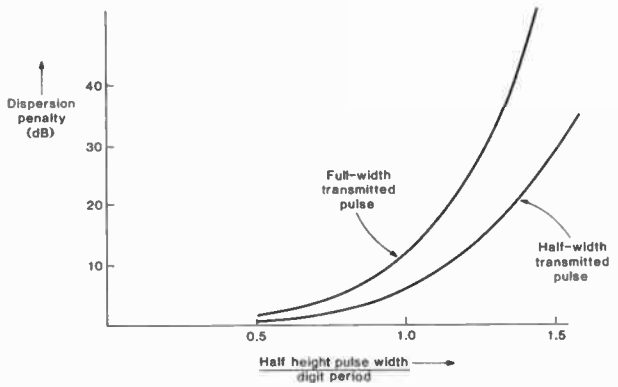


Fig. 7. Dispersion penalty for silicon a.p.d.

Table 1

Typical sensitivities for a.p.d. and p-i-n f.e.t. receivers

System Bit Rate	Possible Line Code	Line Rate	A.P.D. (at 0·85 μm)	p-i-n F.E.T. (at 0·85 μm and 1·3 μm)
Mbit/s		Mbaud	dBm	dBm
2	1B–2B	4	−67	−53
8	2B–3B	12	−63	−52
34	3B–4B	45	−54	−48
140	7B–8B	110	−48	−42
565	7B–8B	645	−41	−34
1150	7B–8B	1320	−37	−29

The other key difference is that no dispersion penalty is included because of the lower bit rate which is well within the bandwidth capability of the multimode fibre.

Long wavelength systems can be characterized in the same way; for a 140 Mbit/s system operating at a wavelength of 1·3 μm there are one or two points worth highlighting. Multimode systems give the same figures for splice and connector losses as at 0·85 μm, while the launched power is also similar for a laser. Splice and connector losses are in general greater with monomode fibres due to the much smaller core size, with lower launched power for the same reason. These effects are, however, amply compensated by lower dispersion. The p-i-n f.e.t. receiver sensitivity is 6 dB worse at 1·3 μm when compared with the 0·85 μm receiver using an a.p.d.; comparative figures are given in Table 1.

The key difference is that the fibre loss is much lower, for example around 0·7 dB/km or less. This therefore means that a splice loss of 0·3 dB represents almost one-third of the total loss allocation in a system if we assume splices every 1 km; this is a typical length of cable pulled into ducts. This demonstrates that splice technology needs careful attention at long wavelengths and points to

an examination of cable installation techniques, to allow longer cable lengths to be pulled in for reduction of the contribution due to splices.

3 Military and Industrial Applications

3.1 Multiterminal Data System for Aircraft

It is becoming widely accepted that optical fibre data transmission will find applications in avionics because of its electrical isolation, low weight, wide bandwidth, and immunity to crosstalk and radio frequency interference. The principle has already been proved and advantages demonstrated for avionics applications in point-to-point links, much of this preliminary work having been carried out with fibre bundles. As a result of a study of the potential of optical-fibre multiterminal data systems for avionics, a design approach has been investigated that is applicable to a wide range of applications.[6] This is a time-division-multiplexing system, which has features of being highly immune to problems of optical loss and multipath effects in optical highways having redundant paths, and of avoiding the need for any master terminal. This system approach is tailored to the characteristics of optical fibres, and leads to good integrity and ruggedness necessary in this type of application.

The time division multiplex system operates in such a way that each transmitter in turn has complete use of the optical highway for a burst period. It sends a fixed length message in that period, complete with an address code to specify the terminal that is to receive it. The transmitters send in strict rotation, and if a terminal is inoperative or if no information is available for sending, either that period is left empty or a dummy message is sent.

Each terminal contains its own clock oscillator and divider chain to provide all the timing information within the terminal. These clocks must be kept approximately in step so that the transmission periods do not overlap. This can be achieved if each transmitted message contains the address of the sending terminal. The address can be received by all other units and used to correct small timing errors that have developed. Special attention has been given to the problems of synchronization from switch-on and to the effects of a single faulty terminal.

The system's timing parameters allow up to 10 terminals to be accommodated. The transmission data rate is 100 kbit/s per terminal, and the baud rate is 1·5 Mbd. The optical receivers use p-i-n photodiodes, and the optical transmitters have been built both with light emitting diodes and lasers. A 30 dB optical loss can be tolerated with l.e.d.s, with approximately a 20 dB increase available using a laser. Experimental results demonstrate rapid settling down to synchronization after switch-on. The receiver dynamic range is of the order of 50 dB.

The multiple-access fibre optical data highway used operates so that an optical signal injected to any access point appears at all other access points. Therefore, no optical switching is necessary. The approach adopted also allows 'ring' or 'star' configurations, or mixtures of both to be used to realize a bus network.

A bus harness consists of single-fibre 2-way and 4-way equal-division transmission stars and a set of fibres of various lengths for connection. These components allow the interconnection of terminals with various degrees of redundancy; see Fig. 8 for two examples. Typical insertion losses for a 4-way transmission star are 10 dB between any pair of ports. Optical losses between the mixers are only due to connectors as very low loss fibres are used, and their attenuation can be neglected for the length likely to be used within an aircraft.

This type of design approach offers a viable architecture for multi-terminal data systems especially for applications where a central bus controller is undesirable.

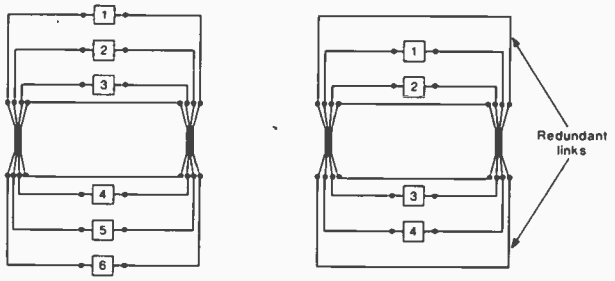


Fig. 8. Harness configuration for up to six terminals employing two 4-way transmission stars.

3.2 Military Communication System

An optical fibre system for use in a military tactical trunk communication system has been developed as a direct alternative to the high-frequency quad cable system used for the intra-nodal multichannel cable links.

The system is designed to operate over a range of up to 2 km, in four 500 m cable sections, at a traffic data rate of 256, 512 or 2048 kbit/s without the use of intermediate repeaters. The system margin achieved gives the potential for longer distances where required. An independent engineering order wire is provided at 16 kbit/s.

The optical fibre cable assemblies are about half the weight of the h.f. quad type, and are designed for quick and easy deployment in the field. The especially ruggedized hermaphroditic optical connectors are suitable for use in conditions of dust, dirt, rough handling and extreme climates. The metal-free cable design overcomes electromagnetic compatibility and electromagnetic pulse problems encountered in quad cables using metallic conductors. Figure 9 shows the layout of typical links. Figure 10 shows the special optical connector developed for this rugged application.

An 0·85 μm l.e.d. is used with a silicon avalanche photodiode detector; an alternative transmitter employing lasers can also be used. The electro-optic units are designed as direct mechanical replacements for

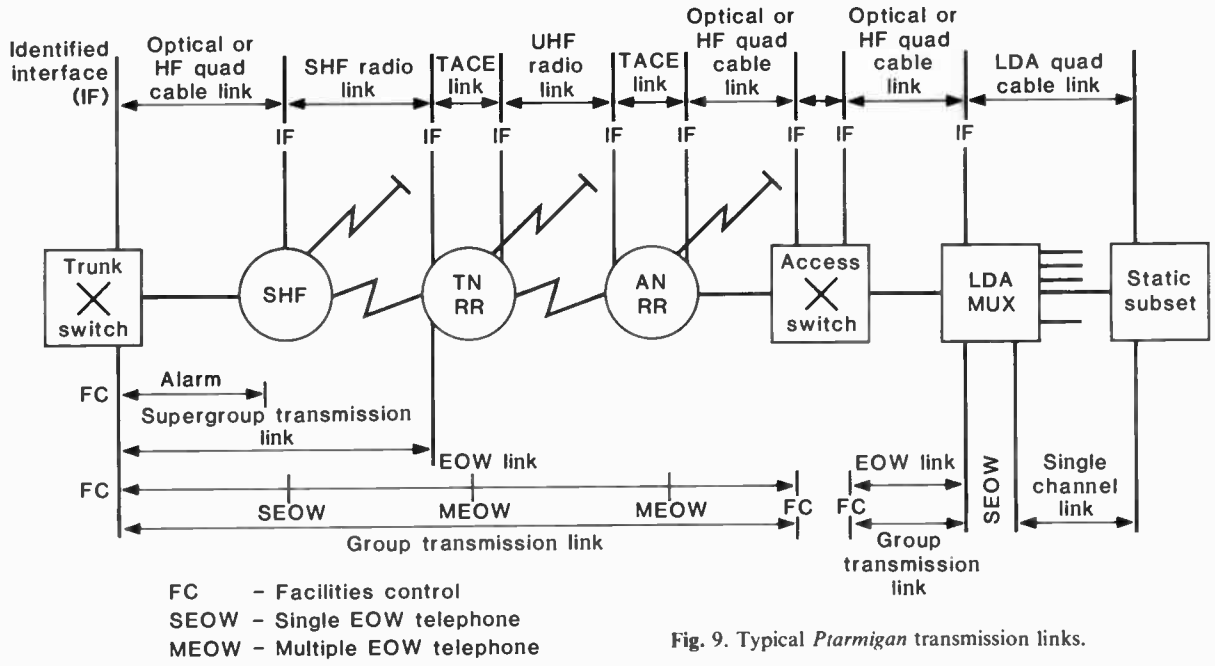


Fig. 9. Typical *Ptarmigan* transmission links.

the existing line modems used with the h.f. quad system. This system demonstrates the use of fibre optics in a military environment where ruggedness and integration into a more complex telecommunication system are both important factors.

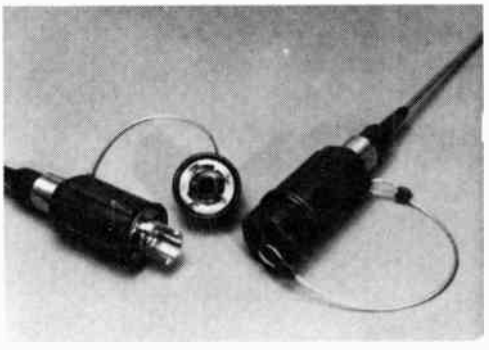


Fig. 10. *Ptarmigan* connector.

3.3 Versatile Optical System

The great flexibility of existing line plant, consisting of paired cables, overhead lines and coaxial cables, has permitted remarkable development of these media. They now carry p.c.m. and f.d.m. as well as audio signals, and in some cases routes are allocated to data or broadcast (including video).

Recent work shows that fibre optic systems of similar flexibility can be achieved.[7] Just as with existing media, it is possible to change from one application to another by connecting the appropriate traffic to the terminals. The traffic need not be encoded before transmission. This allows simple interfacing to both analogue and digital traffic of any format.

The simplicity of the line equipment used is clearly illustrated by Fig. 11. It contains neither the retiming circuitry of digital systems nor the regulating circuitry of analogue systems. The presence of a limiting amplifier defines the signal level and simplifies the equipment.

The line equipment is designed to work up to 35 MHz. This speed can be achieved with l.e.d.s as well as lasers.

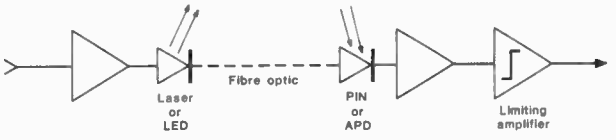


Fig. 11. Versatile line equipment.

The bandwidth of the fibre is either limited by material dispersion when using l.e.d.s, or by mode dispersion when using lasers. Figure 12 summarizes the range achievable in meeting cable performance requirements when using l.e.d.s and lasers.

The line equipment described above can be connected to p.c.m. terminal equipment to operate at any speed up to approximately 50 Mbit/s. A typical example is a 34 Mbit/s di-phase terminal; this code occupies a similar spectrum to f.m. described below. Its strong clock content would be particularly useful in asynchronous packet switched networks. This speed will be frequently encountered now that it is a preferred CCITT equipment interface speed.

The line equipment can also carry a great variety of analogue, or digital signals, or a combination of several types of traffic. In this respect an f.m. channel is as versatile as a metallic line. The traffic need not be encoded before transmission, thereby keeping terminal costs low. A typical f.m. terminal is shown in Fig. 13. If composite traffic is to be carried on subcarriers, then linearity becomes very important. The use of a beat frequency technique achieves high modulator linearity.

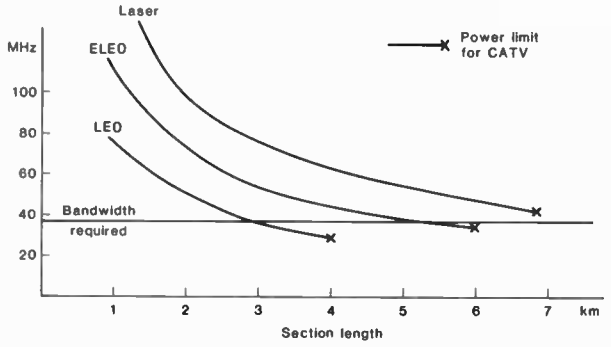


Fig. 12. Range of l.e.d. and laser diodes
Fibre: 4·5 dB and 1·7 ns at 1 km.

The demodulator is of a pulse counting type. Short pulses are generated at both positive and negative going zero crossings. It is also important that this 25 MHz circuitry be kept free of reflections to maintain high linearity.

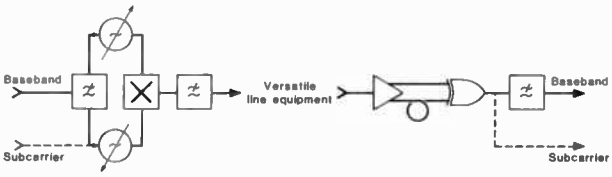


Fig. 13. F.m. terminal.

Tests on this system have confirmed its versatility. Table 2 shows the received power levels to provide transmission of a wide variety of signals. Table 3 gives typical video performance of a basic system; figures of differential gain and phase of 1% pk–pk and 1 V pk–pk respectively have been achieved with an updated design of this type. Figure 14 shows signal/noise performance against optical loss for p-i-n diode and a.p.d. receivers,

Table 2

Receive power level at 0·85 μm

Surveillance television	10 nW	(S/N = 45 dB)
CATV	100 nW	(S/N = 55 dB)
Outside broadcast	1000 nW	(S/N = 65 dB)
60-ch. f.d.m.	500 nW	(10 pW0p)
8 Mbit/s (via f.m.)	25 nW	(10^{-9} b.e.r.)
34 Mbit/s	50 nW	(10^{-9} b.e.r.)

illustrating the typical figures for performance versus loss in analogue systems.

This system is particularly useful in industrial applications where one equipment gives advantages in versatility of handling varying types of signal and giving benefits of maintenance. This type of equipment also has applications in CATV systems, as already demonstrated by its use in a demonstration CATV system set up by British Telecom at their research laboratories.

Table 3

Typical video performance

K-rating	1·5%
Differential gain	3%
Differential phase	3 deg

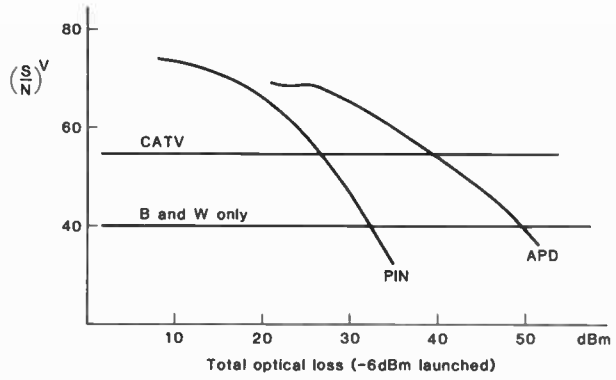


Fig. 14. Noise measurement on p.f.m. link.

4 Conclusion

This paper has described a diversity of applications for optical fibre line systems in p.t.t., military and industrial environments. The selection represents a major subset of the current range of applications. This range is now bound to increase as the level of confidence in the viability and reliability of the technology is proved. Reducing costs of components and fibres, due to better production control and larger numbers, will also increase the range of applications. This could then also permit new services, such as video distribution via fibre optics, to become cost effective.

5 Acknowledgments

I wish to thank many colleagues in STL and STC who have helped by supplying information included in this paper.

6 References

1 Hill, D. R., '140 Mbit/s optical fibre field demonstration system', *Electrical Communication*, **54**, no. 1, pp. 3–13, 1979.
2 Ridler, D. S., 'The changing economics of line transmission', Conference on Telecommunication Transmission—into the digital era, London, March 1981. (IEE Conference Publication 193).
3 Radley, P. E. and Horsley, A. W., 'Technology trends for future optical submerged systems', Conference on Submarine Telecommunications Systems, London, February 1980. (IEE Conference Publication 183).
4 Worthington, P., 'Application of optical fibre system in underwater service', Conference on Submarine Telecommunications Systems, London, February 1980. (IEE Conference Publication 183).
5 Norman, P., Howard, P. J. and Pitcher, P. J., 'An 8 Mbit/s optical fibre system', Conference on Telecommunication Transmission—into the digital era, March 1981, (IEE Conference Publication 193).
6 Farrington, J. G. and Chown, M., 'An optical fibre multiterminal data system for aircraft', *Fibre & Integrated Optics*, **2**, no. 2, 1979.
7 Powell, W. H. and Allsop, B. E., 'Versatile optical fibre transmission system', Conference on Telecommunication Transmission—into the digital era, March 1981 (IEE Conference Publication 193).

Manuscript first received by the Institution on 24th February 1981 and in revised form on 11th May 1981.
(Paper No. 2000/Comm 222)

UDC 535.3/4: 621.372.8: 621.391.63/4
Indexing Terms: Fibre optics, Optical variables measurement

Characterization of single-mode optical fibres

K. I. WHITE, Ph.D., B.Sc., M.Inst.P.,*
S. HORNUNG, D.Phil.,*
J. V. WRIGHT, M.A., M.Inst.P.,*
B. P. NELSON, Ph.D., B.Sc.

and

M. C. BRIERLEY, B.Eng.*

SUMMARY

Experimental techniques for the measurement of total loss and absorption loss, chromatic and monochromatic bandwidth, mode cut-off and refractive index profile are reviewed and illustrated by the results obtained on a particular single-mode fibre. The relationship between theory and experiment is explored.

* *British Telecom Research Laboratories, Martlesham Heath, Ipswich IP5 7RE*

1 Introduction

Two of the most important parameters in any practical communication system are bandwidth and attenuation. They control the rate at which information can be sent and the range over which it can be transmitted. Recent advances in fabrication technology have realized the fundamental low-loss windows that exist in doped silica fibres. Values of 0·65 dB/km at 1·30 μm and 0·20 dB/km at 1·55 μm have been reported[1] in the same fibre. More recent fibres have losses at 1·30 μm which approach the Rayleigh scatter limit of 0·3 dB/km. There are three reasons for high bandwidth. First, single-mode operation clearly removes the inter-modal pulse broadening found in graded index fibres, secondly, the material dispersion of bulk silica has a minimum at 1·27 μm and, thirdly, the total chromatic dispersion of the fibre can be minimized anywhere in the range 1·3 μm to 1·6 μm and beyond by suitably designing the fibre.[2–10] Thus, minimum attenuation and maximum bandwidth can be obtained at the same wavelength. The question of the relative advantages and disadvantages of operating a cabled jointed single-mode fibre link at 1·30 μm or 1·55 μm is receiving considerable attention but is, as yet, unresolved. Nevertheless, it is clear that single-mode fibres are an ideal transmission medium for long-range, high-bandwidth telecommunication systems.

In this paper we review measurements of loss and dispersion on single-mode fibres and, additionally, second-mode cut-off and refractive index profile. The mode cut-off measurement defines the wavelength range over which single-mode operation is possible. The refractive index profile of the fibre not only provides essential information for the fibre fabricator about the index that results from a particular glass composition or set of deposition conditions but also enables predictions to be made about mode cut-off wavelength and pulse dispersion. These techniques are illustrated by results obtained on a particular fibre made at British Telecom Research Laboratories. It is not intended that this is in any way a 'best' fibre.

2 Loss

The total attenuation of a fibre comprises absorption and scatter loss. The scatter loss arises not only from Rayleigh scattering but also from other mechanisms such as waveguide imperfections, bends and microbends which cause light to be radiated from the fibre. While it is common to measure only the total loss, it is nevertheless informative to obtain also at least one of the component parts, the other being derived by subtraction. The total loss is measured over a wide spectral range using a tungsten lamp and monochromator as a light source. The power transmitted through a long (> 1 km) length of fibre is recorded first and then, with launching undisturbed, the fibre is cut near the beginning and a second measurement taken over this short length of a few metres. The loss is obtained from the ratio of these

0033–7722/81/070385+07 $1.50/0

results. Silicon or germanium detectors are used depending on the spectral range covered.

In measuring the loss of multimode fibres it is difficult to obtain very reproducible results. This is because the attenuation measured on a particular length of fibre depends on the relative power levels in, and attenuation rates of, each of the many propagating modes. The application of the results of such a measurement to, for example, a longer piece of the same fibre or that fibre with a joint is non-trivial, as the relative excitation of the modes will vary depending on the particular conditions.

In single-mode fibres the problems are in general less severe. In the spectral range of most interest, that where a system would operate, the fibre is single-mode and so no problems due to differential mode excitation can exist. As the fibre has a radial variation of composition, it may also have radially dependent loss. At shorter wavelengths, where a few modes having different radial power distributions can propagate, the measured loss may thus depend on mode excitation. The loss will increase as cut-off for each mode is approached from shorter wavelengths. Although the fundamental mode in a step index fibre has no cut-off, it is less well guided at longer wavelengths and so becomes lossy due to the bends and microbends experienced by the fibre in its environment. This is in addition to the fundamental infra-red absorption edge of the fibre material.

The absorption loss of the fibre can be measured calorimetrically[11] by detecting the rise in temperature of the fibre when 100 mW of light from a krypton ion or Nd : YAG laser is launched into the fibre. The temperature rise, of the order of a millidegree, is measured differentially between the illuminated and reference samples. Slightly more than 1 m is a convenient length of fibre for the measurement, although only a few centimetres are actually in the calorimeter. The present apparatus has the sensitivity to measure 0·05 dB/km when 1 W is guided by the fibre. The difference between the total and absorption losses gives the scatter loss.

The spectral total attenuation of a single-mode fibre over the range 0·6–1·6 μm is presented in Fig. 1 (a). The increased loss at shorter wavelengths is mainly due to Rayleigh scattering while that at longer wavelengths results from the infra-red absorption edge of the fibre material. The peaks at 1·38 μm and 1·24 μm are OH overtone and combination bands respectively while the feature around 0·86 μm is due to the second mode (LP_{11}) cutting off. The spot measurements of absorption loss (+) and scatter (•), obtained as the difference between total loss and absorption, are also shown. An alternative presentation of the same data is given in Fig. 1 (b), where the three loss parameters are plotted against (wavelength)$^{-4}$. Within experimental error the scatter loss is seen to lie on a straight line passing through the origin. Thus, at least up to 1·32 μm, the longest wavelength of the absorption loss measurement, the

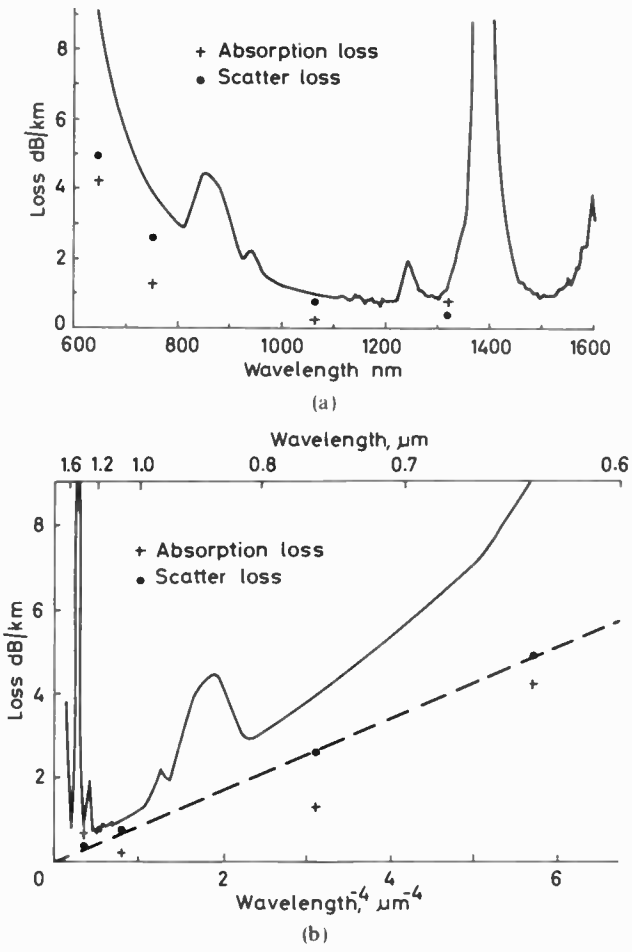


Fig. 1. Fibre loss.

scatter is purely Rayleigh in nature. Scatter losses due to fibre environment or geometric imperfections do not behave as λ^{-4}. It is often assumed that, because fibre losses are now so low, a straight line fit to the base of the attenuation plot on a λ^{-4} graph must be the Rayleigh scatter level. The data show that this is not the case for this fibre. There is a very broad absorption loss feature increasing towards shorter wavelengths without reaching a maximum but still giving a significant contribution to the loss in the region of interest beyond 1 μm. While the fibre used for this set of measurements had a minimum loss of 0·75 dB/km, more recent fibres have losses as low as 0·38 dB/km at 1·30 μm and 1·55 μm.[12]

3 Mode Cut-off

Theoretically, the second mode in a fibre can no longer propagate as a guided wave when, with increasing wavelength, the phase index experienced by that mode falls to that of the cladding. In a step index fibre, the fundamental mode has no cut-off. These limits define the theoretical single-mode regime. Over system lengths of fibre, losses due to bends and microbends (small random fluctuations in the straightness of the fibre) will effectively stop the second mode propagating at wavelengths shorter than cut-off by scattering the

weakly-guided mode out of the fibre. Likewise, the fundamental mode will propagate with unacceptably high losses at longer wavelengths. Thus, for transmission purposes, the theoretical cut-off is not an accurate way of defining the wavelength range over which single-mode operation is possible. However the short lengths of fibre used in transmitter module pigtails need to be assuredly single-mode over their entire length to minimize modal noise problems at the first joint.

Not only is cut-off hard to define, it is also hard to measure. Both bends and microbends cause shifts in cut-off to shorter wavelengths. However, current fibre uniformity (as verified by the Rayleigh nature of the scatter) means that microbend losses need not be large in carefully-laid short straight pieces of fibre. A number of techniques are available for determining the mode cut-off wavelength. These include spectral measurement of:

(a) the power transmitted by the fibre;[13]
(b) the ratio of power transmitted by a fibre when straight and with a bend;[14]
(c) the power transmitted by the fibre when placed between crossed polarizers and excited with an off-axis launch;[15]
(d) the near field;[16]
(e) the power refracted by the fibre when excited by a beam of relatively high numerical aperture;[17]
(f) the far field[18,19] at one wavelength in the single-mode regime.

The first four techniques are all susceptible to microbending effects shifting the cut-off to shorter wavelengths.[20] The use of very short samples to minimize microbend losses often results in another problem—the persistence beyond cut-off of the leaky component of the LP_{11} mode. Far-field methods avoid these problems by making measurements at a wavelength at which the fibre is single moded. Their disadvantage is that they rely on propagation theory to predict a property of the second mode (namely, cut-off) from a measurement on the fundamental mode. The refracted near-field refractive index profiling technique[21,22] makes measurements on lengths of fibre comparable to the fibre diameter and also rejects leaky mode effects. The refracted power cut-off measurement method similarly avoids both leaky rays and also, by virtue of the short length effective in the measurement, microbend effects.

A further complication of cut-off measurements is that the feature observed is often quite broad—even more than 100 nm—and there is not an adequate theory to describe the evolution of that feature in sufficient detail for a unique value of the cut-off wavelength to be predicted.

For convenience, we use the same apparatus to measure mode cut-offs as is used for total spectral attenuation. The transmission of a 1 m length of fibre is measured first when straight and then with a 10 mm radius bend. As the drop in transmitted power at mode cut-off is shifted to shorter wavelengths in the bent fibre, the ratio resulting from these two measurements yields a peak whose long wavelength edge arises from the mode cut-off in the straight fibre. The position at which this edge falls to zero is taken to be the mode cut-off.

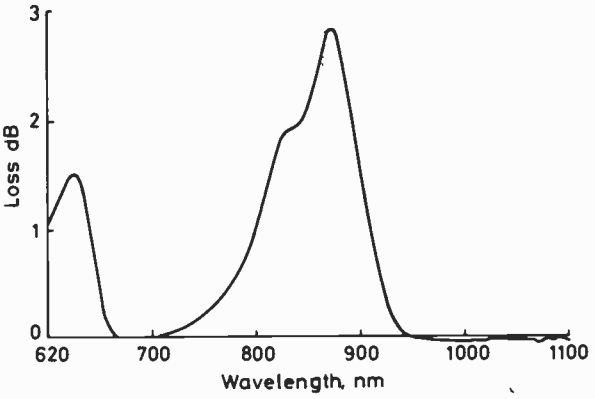


Fig. 2. Mode cut-off.

Figure 2 shows a typical plot giving, for this fibre, the cut-off as 940 nm. Any microbends experienced by the fibre will, however, cause the cut-off to appear at shorter wavelengths. The feature occurring at 800–900 nm on the total loss plot (Fig. 1(a)) is thus identified as the increase in loss of the second mode as it cuts off. The refracted power technique and this approach are said[23] to give similar results if the cut-off is taken as the wavelength at which the gradient of the peak falls to zero at the long wavelength baseline.

4 Dispersion

Pulses of finite spectral width broaden in time during propagation along single-mode fibres because of the wavelength variation of the effective refractive index seen by the mode. This total chromatic dispersion can be resolved into material dispersion, waveguide dispersion and smaller cross-product terms. The fibre dispersion can be obtained from the wavelength variation of group delay for which a pulsed tuneable source is required. The output from a c.w. mode-locked, Q-switched Nd : YAG laser is passed once through a relatively short length of single-mode fibre. Interaction with the fibre via the stimulated Raman effect causes a progressive shift of power to longer wavelengths which stops when the fibre loss rises rapidly at the infra-red absorption edge.[24] The apparatus is illustrated schematically in Fig. 3 while the broadband Raman output is displayed in Fig. 4. Typical source parameters are listed in Table 1. The fibre is a favourable medium for this Raman interaction as the small field width gives good transverse confinement while the long fibre length (relative to the wavelength of light) and low dispersion contain the pump and Raman pulses in synchronism long enough for an efficient shift of power to longer wavelengths to take place. The

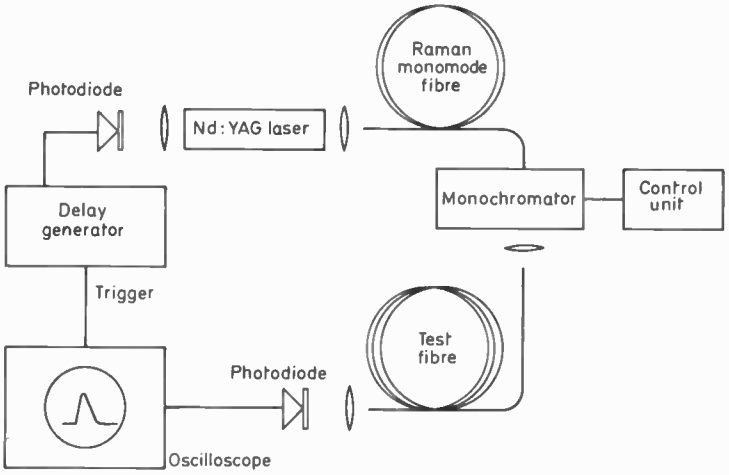


Fig. 3. Schematic diagram of Nd : YAG–Fibre Raman dispersion measurement.

broadband output from the Raman fibre is passed through a monochromator to select a narrow band of light a few nanometres wide which is then launched into the fibre whose dispersion is to be measured. This fibre is several kilometres in length. The pulses emerging from the fibre are detected by a GaInAs photodiode and then displayed on a sampling oscilloscope. This is triggered by pulses, at the pump wavelength, emerging from the back of the Nd : YAG laser and subsequently delayed.

Table 1

Parameters of Nd : YAG Fibre Raman sources

Nd : YAG laser:		
Operating wavelength	1·064 μm	1·319 μm
Peak power: Q-switched and mode locked	10 kW	2 kW
Pulse duration—full 1/e width	~ 160 ps (Detector limited)	
Pulse repetition rate	100 MHz mode locked burst under 500 Hz Q-switched envelope 0·5 μs in duration	
Fibre Raman source pumped by 1·064 μm:		
Wavelength range	1·064– > 1·6 μm	
Fibre length	200 m	
Fibre loss	0·75 dB/km^{-1} at 1·30 μm	
Fibre Δn	0·0035	
Fibre core diameter	6·7 μm	

Scanning the monochromator varies the wavelength of the pulse and also, because of fibre dispersion, its position on the oscilloscope. The measured variation of group delay, τ, is fitted to a linearized Sellimeier equation of the form:

$$\tau(\lambda) = A\lambda^{-4} + B\lambda^{-2} + C + D\lambda^{2} + E\lambda^{4}$$

This symmetrical form of the equation is used as minimum dispersion is near the centre of the wavelength range under study. Differentiating the fitted curve gives the total chromatic dispersion (Fig. 6). This particular fibre has minimum dispersion around 1·37 μm. This is a useful fibre for Raman generation as the wavelength of minimum dispersion is near that of the higher loss due to the OH overtone. The decrease in the efficiency of the stimulated Raman effect due to the increased loss is countered by the longer interaction length due to low dispersion. Thus, this fibre produces useable output around 1·4 μm. The power falls rapidly after 1·6 μm due to the fibre i.r. absorption edge.

A further source of bandwidth limitation could be the temporal broadening suffered even by monochromatic pulses. A monochromatic pulse is defined as one whose spectral width is small enough for negligible chromatic broadening to occur at the wavelength of operation and over the fibre length under study. This would happen if, for instance, the fibre had geometric fluctuations which allowed the LP_{11} mode to propagate over part of its length, or if deviations from circular symmetry of the refractice index profile lifted the degeneracy of the fundamental mode, thus permitting two orthogonal modes of slightly different propagation constant to be guided.

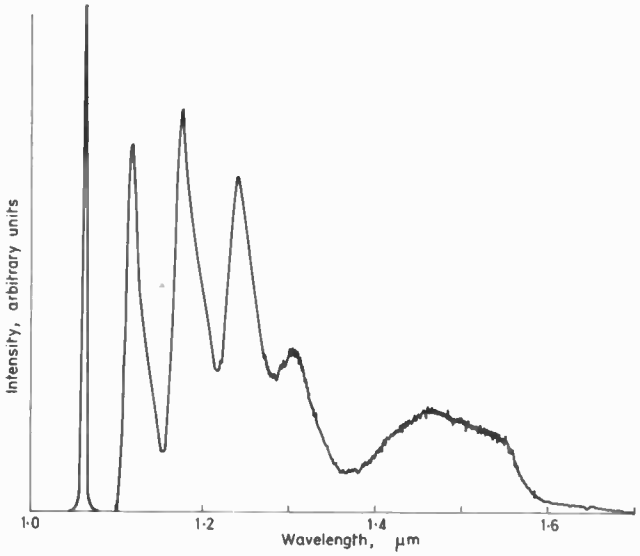


Fig. 4. The spectral output of the Raman fibre.

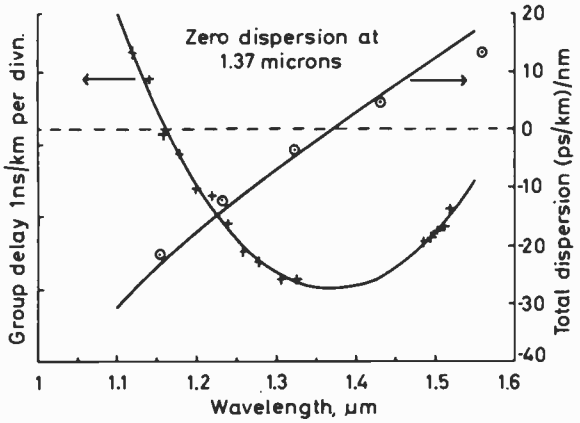


Fig. 5. Fibre group delay and dispersion.

The technique used to measure the monochromatic pulse-broadening is to launch the mode locked output of the Nd : YAG laser (at 1·06 μm or 1·32 μm) into several kilometres of fibre and record the input and output pulse widths. Deconvolution then gives the broadening imposed by the fibre. The spectral width of the Nd : YAG laser pulses is $<0{\cdot}025$ nm. Thus, for a total chromatic dispersion of 20 ps·km^{-1} nm^{-1}, say, the pulse broadening is 0·5 ps·km^{-1}. Pulses having a 1/e duration of 160 ps have been detected using a GaInAs photodiode. This measurement has been performed on a number of fibres up to a maximum length of 27 km. Chromatic dispersion of 0·5 ps·km^{-1} over 27 km would broaden a 160 ps pulse by a negligible 0·6 ps. Fluctuations in mode locking stability cause a $\pm 5\%$ variation in pulse duration. This masks any effect due to monochromatic pulse broadening. A prohibitive number of input and output pulses would be required for significant statistical data to be obtained. Such results that have been obtained indicate a maximum broadening of, at most, a few picoseconds per kilometre. Chromatic broadening from a system laser having a spectral width of several nanometres would be at least an order of magnitude greater. Thus although monochromatic pulse broadening is the fundamental limitation to bandwidth, it is not of consequence in current fibres for presently envisaged systems. However, this parameter must be monitored in case changes in fabrication techniques alter this.

Other work in this area is limited. Negligible pulse broadening very close to the wavelength dispersion has been demonstrated on short (0·76 km) lengths of fibre using a colour centre laser.[25] The output pulses at 1·32 μm were of 5 ps duration and ~ 1 nm spectral width. Pulse broadening measurements over much longer fibres—20 km—have been performed using a double heterostructure InGaAS/InP laser operating at 1·5 μm.[26] The pulse broadening from 380 ps to 400 ps is consistent with the expected chromatic broadening of the 4·8 nm wide source. There was no evidence in either case of monochromatic pulse broadening.

5 Refractive Index Profile

The index profile of a fibre is useful both as a means of relating the deposition conditions during fabrication to the index produced and also as a basis for theoretically predicting propagation properties. The profile may be measured on either the preform or the fibre. Measurements on the preform have the dual advantages that they can be performed as a function of length and are likely to have higher resolution because the preform diameter is two orders of magnitude greater than that of the fibre. The available techniques[27-30] measure, in one way or the other, the deflection of light passing through the preform. Measurements on the fibre are also most useful—particularly to those who do not have access to the preform—as the fibre profile may differ from the preform profile since fibre pulling is a high temperature process, albeit a brief one, and diffusion over a distance of less than 1 μm would change the profile. Reliable evidence on the relation between fibre and preform profiles is so far too limited for firm conclusions to be drawn.

A number of techniques are available for fibre profiling including thin slice interferometry,[31] transverse interferometry,[32] forward scattering,[33] the focusing method[28] and far field.[19] These techniques are generally time consuming and involve computation to obtain the profile. They require lengthy sample preparation (thin slice), extensive data collection (far field, forward scatter) or an expensive microscope for interferometry. Although, in principle, the reflection technique[34] is simpler, end degradation of the fibre sample can be a problem. The refracted near-field technique[21,22] is straightforward in its application and gives the profile directly without any complex computation. It can be applied to both monomode as well as multimode fibres and can also give fibre geometric parameters like core and cladding diameter, ellipticity and concentricity. Circular symmetry of the fibre is neither assumed nor necessary.

Full details of the refracted near-field technique are given elsewhere[22] but an outline of the method is as follows. One end of approximately 1 m of fibre has a flat end-face cut on it. The fibre is then inserted in a cell filled with a liquid of slightly higher index than the cladding. A lens, having a higher numerical aperture than the fibre, focuses a cone of light through the cell entrance window and onto the fibre (Fig. 6.) This focus is scanned across the fibre end-face. Some of this light is guided by the fibre while the rest is refracted out of the side of the fibre. The portion of this light that passes the disk is focused onto the detector. If the intensity distribution in the incident beam is set appropriately, the variation in detected signal is proportional to the index changes encountered as the focus is scanned across the fibre. The instrument can be calibrated to give refractive index differences absolutely.

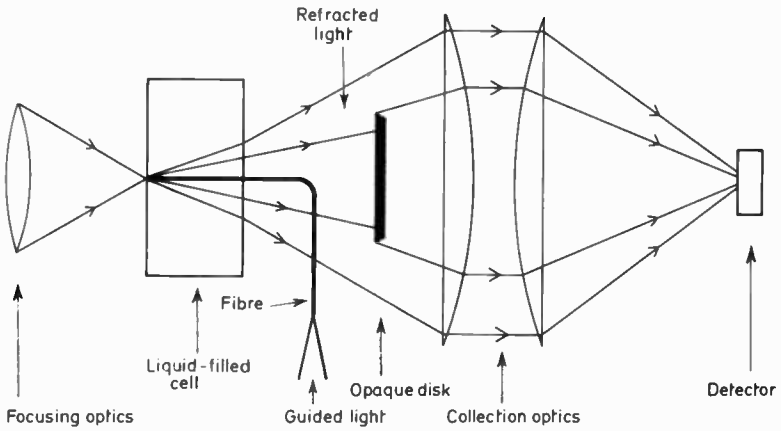


Fig. 6. Schematic diagram of the refracted near-field profiler.

The role of the disk when profiling multi-mode fibres is to reject light that arose from leaky modes and detect only light that comes from the refracting modes—the conditions appropriate to the theory of the measurement. In a single-mode fibre, in the absence of the disk, the measured signal would be all the power that was not guided by the fibre; that is, the exact complement of the near field intensity distribution. The disk ensures that the fibre is profiled by rays entering the front of the fibre at sufficiently high angles that virtually none of that light goes into the guided mode. From the far-field distribution of a single-mode fibre,[18] it can be shown that for a fibre having $V = 2{\cdot}4$ and $\Delta n = 0{\cdot}005$, the far-field intensity has fallen by 20 dB at a half angle of 7°. Typical combinations of focusing lens and disk radius result in the fibre being profiled by a cone bounded by half angles of 22°–30°. Thus, the fibre is clearly profiled by refracted rays.

The profile of the single-mode fibre used in this set of measurements is shown in Fig. 7. The fibre has a cladding diameter of 100 μm, a core diameter of 6·7 μm f.w.h.m. and a maximum index difference of 0·0038. As the entire fibre is profiled, not just the core, both core and cladding diameters can be obtained directly. Two profiles at right angles give estimates of fibre ellipticity and concentricity, while a full raster scan gives the detailed profile and shape of the entire fibre. The measurement is performed at 0·633 μm as a convenient, visible source is available and the relatively short wavelengths aids resolution. The index data can easily be scaled to longer wavelengths using Fleming's data[35] but in any event the effect is small. For a germanium-doped silica fibre, the type of fibre in which the effect is most obvious, the index difference at 1·30 μm is 0·981 times that at 0·633, while at 1·55 μm the figure is 0·993.

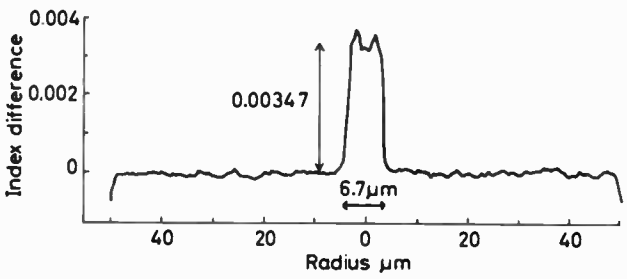


Fig. 7. Single-mode fibre profile.

6 Discussion

In principle, all the fibre parameters could be evaluated from the refractive index profile as a function both of wavelength and also of position along the fibre. In practice, the fibre profile is measured at each end of the fibre and one wavelength. The spatial resolution is about 0·5 μm—approximately 10% of the core diameter. The preform can be profiled as a function of length, but not normally wavelength. As the fibre and preform have different thermal histories, and thus states of strain, the higher resolution obtained in preform measurements is offset by some uncertainty of the applicability of the profile to the fibre. Since fibre dispersion calculations require the second derivative of index with respect to wavelength, this cannot be obtained from profiles obtained at a few wavelengths. The refractive index as a function of wavelength is available[35] but only for bulk samples and a few specific glass compositions. Thus, although it is possible to solve the wave equation using a measured profile,[36,37] the benefits of using a resolution limited profile are offset by the lack of sufficiently accurate data for the wavelength dependence of the refractive index in a particular fibre. The strength of the theory is in pointing out general features of propagation behaviour, for example, the central role played by reducing the core diameter in shifting the minimum dispersion to longer wavelengths.[2–10]

For the fibre under consideration, the profile is not substantially different from a step profile having the index difference and core diameter shown on Fig. 7. Using this equivalent profile and the index data for germanium-doped silica[35] scaled to the dopant level present in this fibre, the total chromatic dispersion—

material dispersion, waveguide dispersion and other cross-terms—has been evaluated. The results are presented on Fig. 5 as the circled dots. There is close agreement with experiment.

The mode cut-off predicted from the profile is around 870 nm. This is somewhat shorter than that measured—940 nm. As cut-off and profile were not measured on the very same piece of fibre, fluctuations in fibre geometry at the ends of a long length of fibre may be responsible.

7 Acknowledgments

Acknowledgment is made to B. J. Ainslie, K. J. Beales, D. Colthorpe and C. R. Day for providing the fibre, to C. J. Todd for supporting discussions, to A. Beaumont for technical assistance and to the Director of the British Telecom Research Laboratories for permission to publish this paper.

8 References

1 Miya, T., Terunuma, Y., Hosaka, T. and Miyashita, T., 'Ultimate low-loss single-mode fibre at 1·55 μm', *Electronics Letters*, **15**, no. 4, pp. 106–8, February 1979.
2 Kapron, F. P., 'Maximum information capacity of fibre-optic waveguides', *Electronics Letters*, **13**, no. 4, pp. 96–7, February 1977.
3 Jürgensen, K., 'Dispersion minimum of monomode fibers', *Applied Optics*, **18**, no. 8, pp. 1259–61, April 1979.
4 Cohen, L. G., Lin, C. and French, W. G., 'Tailoring zero chromatic dispersion into the 1·5–1·6 μm low-loss spectral region of single-mode fibres', *Electronics Letters*, **15**, no. 12, pp. 334–5, June 1979.
5 South, C. R., 'Total dispersion in step-index monomode fibres', *Electronics Letters*, **15**, no. 13, pp. 394–95, June 1979.
6 White, K. I. and Nelson, B. P., 'Zero total dispersion in step-index monomode fibres at 1·30 and 1·55 μm', *Electronics Letters*, **15**, no. 13, pp. 396–7, June 1979.
7 Chang, C. T., 'Minimum dispersion in a single-mode step-index optical fiber', *Applied Optics*, **18**, no. 14, pp. 2516–22, July 1979.
8 Gambling, W. A., Matsumara, H. and Ragdale, C. M., 'Zero total dispersion in graded-index single-mode fibres', *Electronics Letters*, **15**, no. 15, pp. 474–6, July 1979.
9 Tsuchiya, H. and Imoto, N., 'Dispersion-free single-mode fibre in 1·5 μm wavelength region', *Electronics Letters*, **15**, no. 15, pp. 476–8, July 1979.
10 Jeunhomme, L., 'Dispersion minimization in single-mode fibres between 1·3 μm and 1·7 μm', *Electronics Letters*, **15**, no. 15, pp. 478–9, July 1979.
11 White, K. I., 'A calorimetric method for the measurement of low optical absorption losses in optical communication fibres', *Optical & Quantum Electronics*, **8**, no. 1, pp. 73–5, January 1976.
12 Ainslie, B. J., Beales, K. J., Day, C. R. and Rush, J. D., 'Interplay of design parameters and fabrication conditions on the performance of monomode fibres made by MCVD', *IEEE J. Quantum Electronics*, **17**, 1981 (to be published).
13 Worthington, R., 'An apparatus for the measurement of mode cut-off wavelengths of optical fibre waveguides', *J. Phys. E-Sci. Instrum.*, **4**, no. 12, pp. 1052–4, December 1971.
14 Katsuyama, Y., Tokuda, M., Uchida, N. and Nakahara, M., 'New method for measuring V-value of a single-mode optical fibre', *Electronics Letters*, **12**, no. 25, pp. 669–70, December 1976.
15 Kato, Y., Kitayama, K., Seikai, S., and Uchida, N., 'Novel method for measuring cut-off wavelength of He_{21}-, TE_{01}-, and TM_{01}-modes', *Electronics Letters*, **15**, no. 14, pp. 410–1, July 1979.
16 Murakami, Y., Kawana, A. and Tsuchiya, H., 'Cut-off wavelength measurement for single-mode optical fibers', *Applied Optics*, **18**, no. 7, pp. 1101–5, April 1979.
17 Bhagavatula, V. A., Love, W. F., Keck, D. B. and Westwig, R. A., 'Refracted power technique for cut-off wavelength measurement in single-mode waveguides', *Electronics Letters*, **16**, no. 18, pp. 695–6, August 1980.
18 Gambling, W. A., Payne, D. N., Matsumara, H. and Dyott, R. B., 'Determination of core diameter and refractive-index difference of a single-mode fibres by observation of the far-field pattern', *IEE J. Microwaves, Optics and Acoustics*, **1**, no. 1, pp. 13–17, September 1976.
19 Hotate, K. and Okoshi, T., 'Measurment of the refractive-index profile and transmission characteristics of a single-mode optical fiber from its exit radiation pattern', *Applied Optics*, **18**, no. 19, pp. 3265–71, October 1979.
20 Gambling, W. A., Payne, D. N., Matsumara, H. and Norman, S. R., 'Measurement of normalized frequency in single-mode optical fibres', *Electronics Letters*, **13**, no. 5, pp. 133–5, March 1977.
21 Stewart, W. J., 'A new technique for measuring the refractive index profile of graded optical fibres', Paper C2.2, Digest of 'Integrated optics and optical communication', Japan 1977.
22 White, K. I., 'Practical application of the refracted near-field technique for the measurement of optical fibre refractive index profiles', *Optical & Quantum Electronics*, **11**, no. 2, pp. 185–96, March 1979.
23 Bhagavatula, V. A., Private Communication.
24 Cohen, L. G. and Lin, C., 'A universal fiber-optic measurement system based on a near-ir fiber Raman laser', *IEEE J. Quantum Electronics*, **14**, no. 11, pp. 855–9, November 1978.
25 Bloom, D. M., Mollenauer, L. F., Lin, C., Taylor, D. W. and DelGaudio, A. M., 'Direct demonstration of distortionless picosecond-pulse propagation in kilometre-length optical fibers', *Optics Letter*, **4**, no. 9, pp. 297–9, September 1979.
26 Kawana, A., Miya, T., Imoto, N. and Tsuchiya, H., 'Pulse broadening in long-span dispersion free single-mode fibres at 1·5 μm', *Electronics Letters*, **16**, no. 5, pp. 188–9, February 1980.
27 Chu, P. L., 'Nondestructive measurement of index profile of an optical fibre preform', *Electronics Letters*, **13**, no. 24, pp. 736–8, November 1977.
28 Presby, H. M., Marcuse, D. and French, W. G., 'Refractive-index profiling of single-mode optical fibers and preforms', *Applied Optics*, **18**, no. 23, pp. 4006–11, December 1979.
29 Sasaki, I., Payne, D. N. and Adams, M. J., 'Measurement of refractive-index profiles in optical-fibre preforms by spatial-filtering technique', *Electronics Letters*, **16**, no. 6, pp. 219–21, March 1980.
30 Saekeang, C., Chu, P. L. and Whitebread, T. W., 'Nondestructive measurement of refractive-index profile and cross-sectional geometry of optical fiber preforms', *Applied Optics*, **19**, no. 12, pp. 2025–30, June 1980.
31 Presby, H. M., Mammel, W. and Derosier, R. M., 'Refractive index profiling of graded index optical fibres', *Rev. Sci. Instrum.*, **47**, no. 3, pp. 348–52, March 1976.
32 Boggs, L. M., Presby, H. M. and Marcuse, D., 'Rapid automatic index profiling of graded index optical fibers', *Rev. Sci. Instrum.*, no. 4, pp. 867–902, April 1979.
33 Okoshi, T. and Hotate, K., 'Refractive-index profile of an optical fiber: its measurement by the scattering-pattern method', *Applied Optics*, **15**, no. 11, pp. 2756–64, November 1976.
34 Tateda, M. 'Single-mode-fiber refractive-index profile measurement by reflection method', *Applied Optics*, **17**, no. 3, pp. 475–8, February 1978.
35 Fleming, J. W., 'Material dispersion in light guide glasses', *Electronics Letters*, **14**, no. 11, pp. 326–8, May 1978.
36 Cohen, L. G., Mammel, W. L. and Presby, H. M., 'Correlation between numerical predictions and measurements of single-mode fiber dispersion characteristics', *Applied Optics*, **19**, no. 12, pp. 2007–10, June 1980.
37 Okamoto, K., Edahiro, T., Kawana, A. and Miya, T., 'Dispersion minimisation in single-mode fibres over a wide spectral range', *Electronics Letters*, **15**, no. 22, pp. 729–31, November 1979.

Manuscript first received by the Institution on 24th March 1981.
(Paper No. 2001/MI 21)

Some Recent Optical Developments—2

P.F.M. Fibre Optic Module

New pulsed frequency modulation (p.f.m.) fibre optic modules have been designed by ITT Opto-Electronics especially for use in cable television applications where the benefits of optical transmissions are considerable. Advantages include longer transmission distances with lower noise intrusion than is possible with coaxial cable, and the use of p.f.m. permits higher quality links without the need for good source linearity.

The transmitter module source may be either a high radiance l.e.d. or e.l.d. (edge emitting l.e.d.) and this can be interchanged without the need for re-characterization. There is minimal distortion from either fibre dispersion or modal noise. The receiver module employs an avalanche photodiode detector.

Normally a 50 μm core graded index silica fibre optic cable is employed between the modules and the optical path loss is typically 21 dB at 57 dB signal-to-noise ratio. Optical input/output correction is possible using the OCN101 connector in conjunction with an ITT jewelled ferrule. The operating temperature range is 0–50°C.

These transmitter and receiver modules are currently being tested by British Telecom for possible inclusion in a central antennae system in Milton Keynes.

Fibre Optic Transmission System

McMichael's 2000 series of real time multiplexers now includes a wide range of interface modules in which output data are provided by coaxial transmission up to 200 metres or optional fibre optic transmission of up to 2 km. Cost effective savings can be made installing multi-fibre cable as opposed to installing a number of larger coaxial cables.

The standard system is available with speech, analogue-to-digital and digital-to-analogue, sub-multiplexer and RS 232 interface modules. Up to 32 input channels can be accommodated in input data rates of up to 38·4 kilobaud. The system has the flexibility to alter framing and clocking rates so as to minimize the overall system channel jitter when used in an asynchronous mode.

Fibre Optic Modems

Cossor Electronics have produced the FDS 1003A optic modem which matches requirements for RS 232 transmission systems. It is suitable for use with large core plastic-coated silica or glass fibre, and has a very high data integrity, being immune from extraneous electromagnetic interference. The maximum transmission distance is 3000 metres. With a maximum data rate capability in excess of 75 kb/s, the FDS 1003A modem can also be used for general data systems.

Evaluation of Short-range Optical Fibre Systems

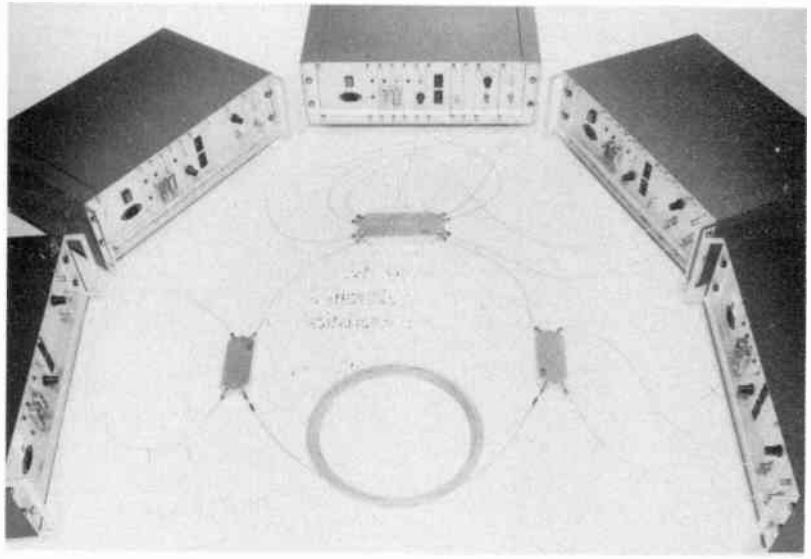
ERA Technology, the independent contract research organization, is proposing to evaluate fibre optic communication systems designed for use on industrial sites or in large business centres.

Fibre optic communication systems offer immunity from electromagnetic interference and intrinsic safety in hazardous conditions. They also provide greater security in the transmission of confidential data. However, there is very little information available on the reliability of actual fibre optic networks, methods of installation or maintenance requirements. Potential users need to know how cost-effective fibre optics are by comparison with conventional links and in what circumstances the new technology represents the best possible choice.

ERA proposes, therefore, to investigate existing short-range applications of fibre optics in industrial process control and inter-connected business communications, including computing and management information systems. The fibre optics investigation has been designed to produce data which will be of great value to users and potential users of optical systems, to suppliers of cable, control components and transducers, and to companies involved in the design and implementation of site-communication systems.

ERA's external research will be supplemented by the company's own experience in the propagation of optical modes in fibre optic cables.

Fibre Optic Data System



Complete STL optical fibre multiterminal data system comprising five terminals, single fibre optic cables, and three transmission star mixers (As referred to in the paper 'System applications of optical fibre transmissions' by Dr P. E. Radley on pp. 379–84 of this issue)

UDC 621.372.8: 621.375.826: 658.382
Indexing Terms: Lasers, Fibre optics, Accident prevention

Optical fibre transmission systems—the problems of formulating optical radiation safety criteria

J. D. TOPPING, B.Sc.*

and

J. C. NORTH†

SUMMARY

Although national and international standards regarding laser safety either already exist or are in the course of preparation, they do not specifically cover the use of lasers in optical fibre transmission systems.

It is necessary to relate the safety requirements for lasers as published in the national and international standards to optical sources as used in optical fibre transmission systems. This paper attempts to do this and to stimulate discussion of the problems involved. The paper concludes that under normal operation optical fibre systems are intrinsically safe, since no exposure to optical radiation is possible. In cases where a fibre end is exposed, minimum viewing distances are proposed.

Cautionary Note:

The following paper is written in an attempt to stimulate discussion on ways in which the possible hazard of optical fibre systems can be evaluated. As no standard specifically related to such systems presently exists (BS4803: 1972 Guide on protection of personnel against hazards from laser radiation, is due to be published in revised form later this year), much of what is given below can only be regarded as the authors' own interpretation of the situation. No attempt should be made at hazard assessment solely on the basis of this paper.

* *Network Executive, British Telecom, 2–12 Gresham Street, London EC2V 7AG*
† *British Telecom Research Laboratories, Martlesham Heath, Ipswich.* (Mr North died in February 1981.)

1 Introduction

Optical fibre transmission systems may use, as the optical source, either semiconductor lasers or light-emitting diodes (l.e.d.) depending upon the particular application. The semiconductor laser used on optical transmission systems is only one of the many types of laser used throughout industry, commerce, medicine and other fields. Many of these lasers can be hazardous owing to the high optical power density in the beam. The human eye is particularly at risk in this context, since it is capable of magnifying the laser light intensity many times by its focusing power. Other hazards associated with certain types of laser include possible skin damage, atmospheric contamination and collateral radiation.

As a result of these hazards, a number of national and international standards have been or are being produced which define the radiation safety of laser products and equipment. The relevant British standard, BS4803, is currently under revision and is expected to be similar in content to an emerging international standard under the auspices of the International Electrotechnical Commission. An American National Standard for the safe use of lasers (ANSI Z136.1—1976) also exists.

Consultation of these documents is thus the duty of all designers and users of optical fibre systems incorporating lasers. The position of systems based on l.e.d.s is however less clear, but with the constant improvement in performance of these devices and in the absence of any other suitable safety criteria, the requirements of the laser safety standards should be adhered to. In attempting to apply these standards, system designers are likely to encounter a number of aspects which may present difficulties.

2 Laser Classification

Lasers are classified according to various output parameters including emission level, emission duration, wavelength etc. For pulsed lasers details of pulse width, pulse repetition frequency, pulse energy etc. are required. The classifications may be summarized as follows:

Class 1: Intrinsically safe because the output power or energy levels are within safe limits, or by virtue of the engineering design.

Class 2: The maximum permissible power/energy levels are as for Class 1 but with a relaxation in the spectral region 400 to 700 nm.

Class 3A: This represents a further relaxation on Class 1 in the spectral region 400 to 700 nm.

Class 3B: Direct viewing of lasers in this class may be hazardous but this will not generally be the case for diffuse reflections.

Class 4: This is the classification for high power lasers. They may produce hazardous diffuse reflections and cause skin damage.

0033–7722/81/070393+04 $1.50/0

Classification of a laser product is the responsibility of the manufacturer or his agent. Any person or organization performing subsequent modification to the product is also obliged to perform this duty.

It is important to know the classification since this affects the safety features to be incorporated, the labelling and the administrative controls.

Despite the fact that most semiconductor lasers used on optical fibre systems will be in Class 3B, the systems themselves, once operational should, in the authors' opinion, be regarded as falling into Class 1. Such systems will be closed systems (when all cables are properly terminated) with the optical cable forming part of the protective housing. No human access to optical radiation should be possible except under circumstances involving disconnection of the optical path, such as

(i) The transmission fibre could be broken at some point on the route as the result of deliberate or accidental damage.
(ii) An optical connector could be unmated.

In both cases the transmitter may or may not continue to function depending upon the system design.

Access to exposed fibre ends is also possible during installation and repair procedures but in such cases the terminal equipment associated with those fibres should be inoperative. Certain test equipment may however be in use on these occasions and this aspect is considered later.

In order to assess the potential hazard arising from cases (i) and (ii) above, it is necessary to refer to the maximum permissible exposure limits given in the relevant standards documents.

3 Maximum Permissible Exposure

The maximum permissible exposure (MPE) is defined as that level of laser radiation to which, under normal circumstances, persons may be exposed without suffering adverse effects and is based on information from experimental studies. The MPE levels represent the maximum level to which the eye or skin can be exposed without consequential injury and are related to the wavelength of radiation, the pulse duration or exposure time, the tissue at risk, and for radiation in the range 400 nm to 1400 nm, the size of the retinal image. They should be used as guides in the control of exposures and not regarded as precisely defined lines between safe and dangerous levels.

Table 1 presents information on the MPE values at the cornea for direct ocular exposure to laser radiation at wavelengths 700 to 10^6 nm. The situation is complicated by the fact that for repetitively pulsed lasers, as is the case with optical fibre systems, the MPE is determined by the most restrictive of the following requirements:

(a) The exposure from any single pulse within the train shall not exceed the MPE for a single pulse
(b) The average irradiance for a pulse train of duration T shall not exceed the MPE given in Table 1 for a single pulse of duration T.
(c) If the individual pulse duration is less than 10^{-5} s, for a simple pulse train having a uniform pulse repetition frequency (p.r.f.) greater than 1 Hz, the MPE applicable to each pulse shall be taken as the MPE of a single isolated pulse reduced by the p.r.f. correction factor C_5 which varies with the p.r.f. N:

For N between 1 Hz and 278 Hz, $C_5 = 1/\sqrt{N}$

For $N = 278$ Hz, $C_5 = 0{\cdot}06$

(d) If the individual pulse duration is greater than 10^{-5} s, the following formula shall be used to evaluate the MPE applicable to each pulse:

$$\text{MPE (single)} = \frac{\text{MPE}(nt)}{n}$$

where n = number of pulses in the train
t = individual pulse width; and
MPE (nt) = MPE applicable to a pulse width nt s.

There are additional criteria for pulse trains having a non-uniform or variable p.r.f., or with regular patterns.

During normal operation of most optical fibre systems, the most restrictive condition is likely to be the

Table 1

Maximum permissible exposure at the cornea for direct ocular exposure to laser radiation (intra-beam viewing)

Wavelength (nm)	Exposure time (s)						
	$\leqslant 10^{-9}$	10^{-9} to 10^{-7}	10^{-7} to 18×10^{-6}	18×10^{-6} to 50×10^{-6}	50×10^{-6} to 10	10 to 10^3	10^3 to 3×10^4
700 to 1050	$5C_4\times10^6$ W·m^{-2}	$5C_4\times10^{-3}$ J·m^{-2}		$18C_4\,t^{0.75}$ J·m^{-2}			$3{\cdot}2C_4$ W·m^{-2}
1050 to 1400	5×10^7 W·m^{-2}	5×10^{-2} J·m^{-2}			$90t^{0.75}$ J·m^{-2}		16 W·m^{-2}
1400 to 10^6	10^{11} W·m^{-2}	100 J·m^{-2}	$5600\,t^{0.25}$ J·m^{-2}			1000 W·m^{-2}	

$C_4 = 10^{[(\lambda-700)/500]}$ where λ is the wavelength from 700 to 1050 nm.

Table 2

Maximum permissible irradiance at cornea

Wavelength (nm)	Maximum permissible irradiance at cornea ($W \cdot m^{-2}$) for exposure times of:			
	1s	10s	10^3 s	10^4 s
850	35	20	6·38	6·38
900	45	25	8·0	8·0
1050 to 1400	90	50	16·0	16·0

average irradiance consideration (criterion b) and for easy reference Table 2 gives the MPE values of Table 1 in terms of irradiance at the cornea for different exposure times and for selected wavelengths commonly encountered on such systems.

For low bit rate systems, the mean power consideration may not be the most severe. Caution is therefore required in the application of Table 2.

Bearing in mind that the radiation emerging from a fibre end or from the unmated half of a non-lensed connector does so in the form of a divergent beam (Fig. 1) it should be possible to calculate a critical viewing distance closer than which viewing is unsafe. The nominal ocular hazard distance (NOHD) may be defined as the range at which the beam irradiance or radiant exposure falls below the appropriate MPE. It has been shown by Timmerman[1] that for multimode fibre the NOHD is given by the following formula:

For a step index fibre $$\text{NOHD} = \left[\frac{P}{\pi NA^2 I}\right]^{\frac{1}{2}}$$

For a graded index fibre (α profile parameter = 2) $$\text{NOHD} = \left[\frac{2P}{\pi NA^2 I}\right]^{\frac{1}{2}}$$

where P = mean power irradiated from the fibre
NA = numerical aperture of the fibre
I = permissible irradiance at the cornea

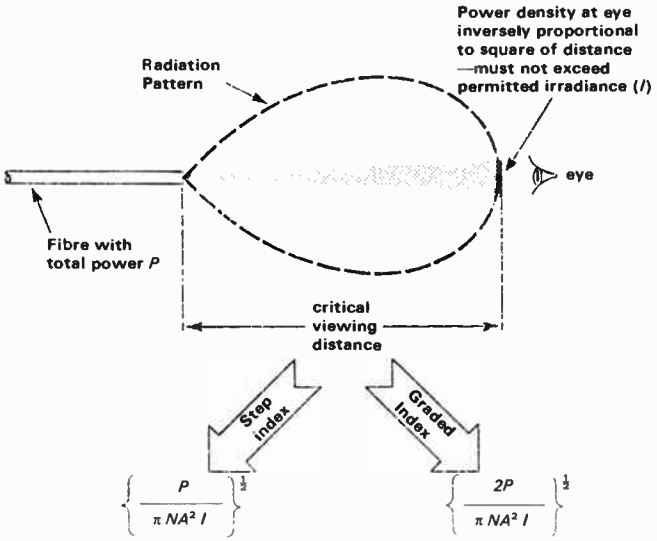


Fig. 1. Critical distance from a multimode fibre.

Table 3 gives examples of the nominal ocular hazard distance at wavelengths of 850 nm and 1050 to 1400 nm for multimode fibre in order to meet the MPE limits and assume that no special viewing optics are used.

As can be seen from Table 3 the eye has to be very close to the fibre end for any hazard to occur. Such distances appear much closer than any person other than extraordinarily myopic individuals would attempt to view an object.

The classic paper on the subject of eye accommodation is that by Alexander Duane[2] which, although written as long ago as 1922, is still regarded as the authoritive reference for such information. This paper reports on the testing of 4200 individuals for their monocular and binocular accommodations and suggests that the probability of being able to focus on an object closer than about 60 mm is very small provided that persons requiring spectacles do in fact wear them (Fig. 2). Any attempt by an individual to view a fibre end closer than

Table 3

Critical viewing distance for multimode fibre

Wavelength (nm)	Mean power in the fibre (mw)	Critical distance (mm) for continuous viewing			
		Step index fibre		Graded index fibre	
		NA = 0·13	NA = 0·20	NA = 0·13	NA = 0·20
850	0·1	17	11	24	16
	0·3	30	19	42	27
	1·0	54	35	77	50
	3·0	94	61	133	87
	10·0	172	112	243	158
1050 to 1400	0·1	11	7	15	10
	0·3	19	12	27	17
	1·0	34	22	49	32
	3·0	59	39	84	55
	10·0	108	71	153	100

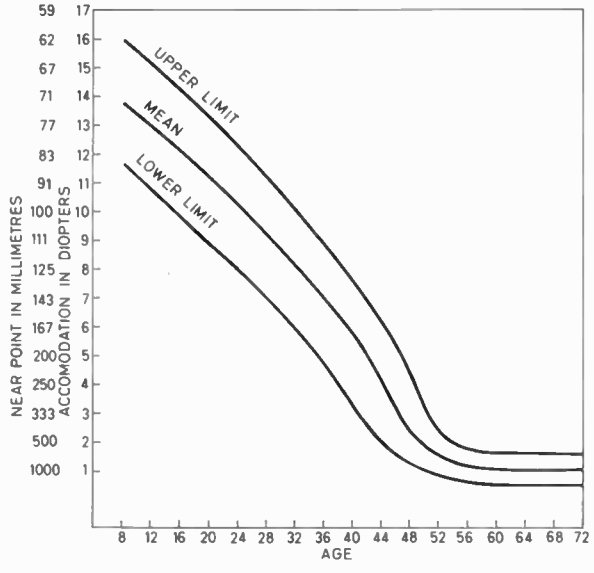


Fig. 2. Monocular accommodation curves.

their minimum focusing distance will cause the image of the fibre end formed on the retina to be defocused. Though the potential hazard is not affected until the retinal spot has a diameter greater than 0·4 mm, it does mean for example, that for a person with a minimum focusing distance of 60 mm, the hazard will be reduced by this defocusing at viewing distances less than 50 mm.

It should also be noted that the figures obtained by Duane refer to visible radiation (400 to 700 nm wavelength). Most laser devices used on optical systems operate in the infra-red region and at this wavelength the eye is comparatively long-sighted, introducing an added safety factor.

It thus appears that the launched power levels on present day optical fibre systems are not of sufficient proportion to present a serious threat of eye damage in the event of accidental exposure to an exposed fibre end. This will also be true of many currently available connectors and while human access to optical radiation is more likely in this case, prudent design of equipment could ensure that it was impossible to bring an eye any closer than the NOHD. However where any form of lens arrangement is incorporated into a connector, the NOHD can be expected to be considerably greater and special safeguards will have to be incorporated to prevent excessive accidental exposure.

In cases where laser devices are incorporated directly into the connector, hazard assessment must take into account the output beam geometry of the device modified as necessary by any focusing optics.

4 Test Equipment

During installation, fault location and maintenance, certain items of test equipment such as fibre attenuation meters and time domain reflectometers will be employed. These present somewhat more of a problem since during their operation the system may not necessary be closed. Furthermore, many of the procedures followed involve visual examination of the fibre including the possible use of viewing optics, e.g. fibre jointing, connector fitting etc.

Classification of test equipment employing laser devices is the responsibility of the manufacturer. If the equipment falls into Class 1 then clearly there is no problem. If however it is assessed as a Class 3B product then strict safety procedures must be followed. Whilst it is possible to adhere to these safeguards at the launch end it may be more difficult or even impossible to control the situation at some point on the route such as attempting to locate a cable fault using an optical time domain reflectometer. As with normal system operation it is possible to establish the NOHD from an exposed fibre end when test equipment is in use. It should be noted that in establishing the MPE limits the most restrictive criteria may be the single pulse, or pulse train limitation rather than the mean power consideration assumed in Table 2 above.

5 Conclusions

We have explained how optical fibre system using laser sources must adhere to certain safety standards. L.e.d.-based systems should, in the absence of any other standard adopt the same criteria.

Under normal operation, optical fibre systems should be regarded as Class 1 products, i.e. intrinsically safe, since no exposure to optical radiation will be possible. It is only when a connector is unmated or a cable is intercepted that any risk of exposure to excessive radiation arises. Since a fibre end will act as a divergent beam source, a nominal ocular hazard distance can be established beyond which the level of optical radiation is below the maximum permissible exposure levels. With the power levels presently employed such distances are small and are likely to be well below those distances likely for normal viewing.

Connectors using lens techniques or which incorporate laser devices may present a more serious threat and should be separately assessed.

Additional safeguards will be required during installation and maintenance procedures particularly when certain test equipment is in operation.

6 Acknowledgment

Acknowledgement is made to the Director, Transmission Department, British Telecom, for permission to publish this paper.

7 References

1 Timmermann, C. C., 'Handling optical cables; safety aspects', *Applied Optics*, **16**, no. 9, pp. 2380–2, September 1977.

2 Duane, A., 'Studies in monocular and binocular accommodation with their clinical applications', *Am. J. Opthalmol.*, 5, pp. 865–77, 1922.

Manuscript first received by the Institution on 5th February 1981 and in final form on 1st April 1981 (Paper No. 2002/Comm 223)

UDC 535.3: 621.3.049.77: 621.372.8
Indexing Terms: Integrated optics

Integrated Optics: a tutorial review

P. J. R. LAYBOURN, M.A., Ph.D.*

and

Professor J. LAMB, D.Sc., Ph.D., C.Eng., M.I.E.E.*

SUMMARY

Integrated optics involves the use of thin transparent dielectric layers on planar substrates as optical waveguides. Various methods can be used to couple light into and out of the guiding layers, including butt and transverse coupling, prisms and diffraction gratings. Optical components are defined by photolithography, and may be passive (such as strip waveguides with bends, junctions and directional couplers) or active (such as switches, modulators and beam deflectors). The substrate and film materials may be glass for passive circuits, or electro-optic or semiconductor single crystal for active devices. Hybrid integration of active and passive devices may be used to produce complex integrated circuits. Alternatively, the use of semiconductor materials offers the possibility of monolithic integration of all types of component on a single substrate.

* *Department of Electronics and Electrical Engineering, University of Glasgow, Glasgow G12 8QQ.*

1 Introduction

Light guided by fine glass fibre waveguides is rapidly becoming a major communications carrier. The extension of guided wave optical technology into the development of miniature optical waveguide components and circuits made by planar technology on rigid substrates has been called 'integrated optics', by analogy with semiconductor integrated electronics to which it bears some similarity in manufacturing processes. However, integrated optical circuits are by no means scaled-down versions of integrated electronic circuits, and are much closer in concept to microwave stripline waveguide circuits.

Integrated optical devices are formed from thin layers of transparent dielectric on a flat glass or crystal substrate. Light energy propagating in a dielectric multilayer structure parallel to the layers will tend to be confined within the layer of highest refractive index or smallest velocity of propagation. An optical waveguide formed by such a structure will have a high-refractive-index guiding layer, the thickness of which is of the order of the wavelength of light; the thick supporting substrate will be of lower index. The guiding layer confines the light beam in one dimension, but if a thin strip of guiding material is used the light can be confined in two dimensions, and may be directed by curves or branches in the guiding strip. One- or two-dimensional thin-film optical waveguides form the basis for integrated optical devices in which several optical guided-wave components are formed on a substrate by thin-film technology and optically connected by thin light-guiding films.

A wide range of materials has been used for thin-film optical circuits, ranging from glasses and plastics to ferroelectric crystals and compound semiconductors. Techniques for device fabrication have been derived in a large part from the microelectronics industry, but have had to be developed to accommodate the various materials and produce the higher edge definition required. To achieve true monolithic integration, in which light sources are fabricated with other optical components on a single substrate, compound semiconductors consisting of group III and group V elements appear to be the most promising materials, but there are many applications in which monolithic integration is not a necessity and other materials may be employed.

Much research and development effort has hitherto been directed towards communications applications, but there are also potential uses for integrated optical systems in many other fields, such as the sensing of physical parameters, particularly in hazardous environments; processing wide-band information; radio-frequency spectrum analysis; matched filtering; and signal delay and storage. It is the aim of the authors

0033–7722/81/070397 +17 $1.50/0

to present a basic understanding of the physical principles of thin-film optics, together with an appreciation of the techniques used to fabricate optical circuits in some of the more important materials.

2 Basic Principles of Propagation

Consider a plane electromagnetic wave travelling in an isotropic medium 1 of refractive index $n_1 = c/v_1$, where c is the velocity of light in vacuum and v_1 the corresponding velocity in 1; the wave encounters a boundary with a different medium 2 of lower index n_2. In

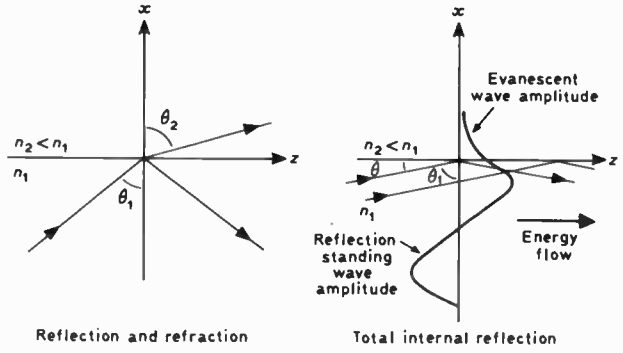


Fig. 1. Reflection at a boundary.

general, there are transmitted and reflected waves making angles θ_2 and θ_1 to the normal, as shown in Fig. 1. The angles are related by Snell's Law (matching of velocities tangential to the boundary):

$$\sin\theta_1/\sin\theta_2 = n_2/n_1 \qquad (1)$$

As θ_1 increases and the wave approaches glancing angle, θ_2 reaches $\pi/2$: there is then no wave transmitted into medium 2, and writing $\theta = \pi/2 - \theta_1$ as the angle between the wave normal and the boundary:

$$\sin\theta_1 = n_2/n_1 = \cos\theta = 1 - 2\sin^2\theta/2$$

Hence at the critical angle:

$$\sin^2\theta/2 = \frac{1-n_2/n_1}{2} = \frac{n_1-n_2}{2n_1} = \frac{\Delta n}{2n_1} \qquad (2)$$

For θ and Δn sufficiently small, equation (2) reduces to:

$$\theta_{crit.} = (2\Delta n/n_1)^{\frac{1}{2}} \qquad (3)$$

In order to satisfy the conditions imposed upon the fields at the boundary, an evanescent field must be present in medium 2. The energy flow is tangential to the surface whilst the amplitude decays exponentially with distance, normal to the surface. The combination of the normal components of the incident and reflected waves produces a standing wave in amplitude normal to the boundary which propagates in a direction tangential to the boundary.

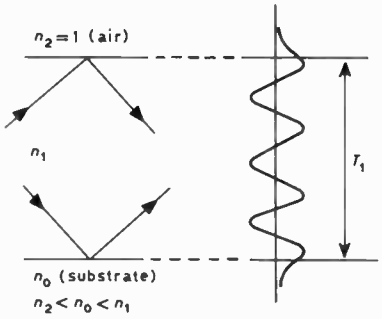


Fig. 2. Higher-order mode.

Let a second boundary be now introduced so that medium 1 becomes a planar slab waveguide supported

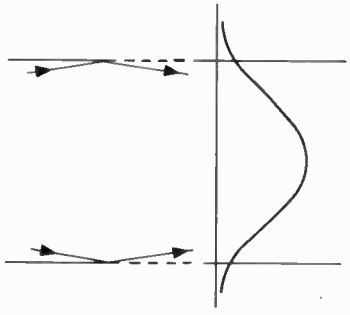


Fig. 3. Lowest-order mode.

on a substrate of refractive index n_0 with the air above ($n_2 = n_a$). Then, provided that critical internal reflection takes place also at the boundary between n_1 and n_0, there will be a certain value of θ for which the same standing wave pattern normal to the boundaries is maintained by both upper and lower reflections: this constitutes a higher-order propagated mode (Fig. 2). If the thickness, T_1, of the slab is maintained constant and the angle θ reduced, we finally achieve a standing wave pattern corresponding to the lowest order TE_{01} mode (Fig. 3). If T_1 is now reduced, the angle θ required to sustain a particular mode will progressively increase until it exceeds the critical value for the lower boundary when energy for that mode is transmitted into the substrate. Thus, the higher order modes are progressively eliminated as T_1 is reduced, leading eventually to monomode propagation (Fig. 4). Further reduction of T_1 causes all this wave energy to be transmitted into the substrate.

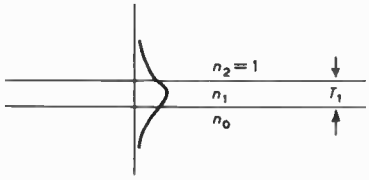


Fig. 4. Monomode.

The subscript 'a' is used to denote air (or vacuum) and if ω ($= 2\pi f$) is the angular frequency of the wave, the wave number $K_a = \omega/c = 2\pi/\lambda_a$. In medium n_1, $K_1 = \omega/v_1 = 2\pi/\lambda_1 = n_1 K_a$. The relation between these values for an unbounded medium and the propagation constant β and guide wavelength λ_g in a bounded medium can be illustrated by drawing wave-fronts and wave-normals for reflection at the guide boundary as in

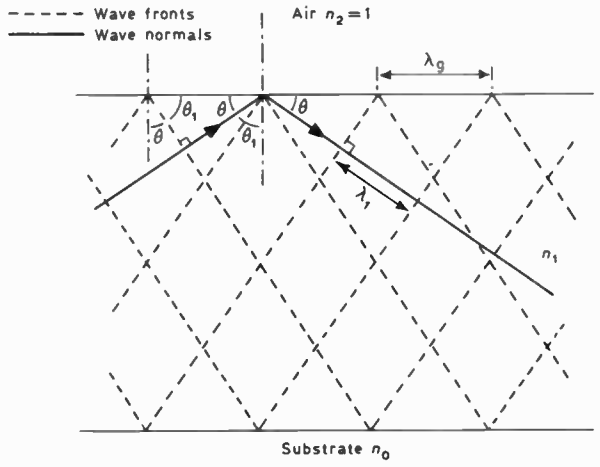


Fig. 5. Reflected wave fronts.

Fig. 5, from which:

$$\lambda_1 = \lambda_g \cos\theta \tag{4}$$

and

$$\beta = K_1 \cos\theta = n_1 K_a \cos\theta \tag{5}$$

As the primary interest is with propagation in solid media, the small difference between the velocity of light in air and in vacuum is negligible, so that $K_a = K$, $\lambda_a = \lambda$, $n_a = 1$. Then from equations (4) and (5), substituting $n_1 = c/v_1 = \lambda/\lambda_1$:

$$\beta/K = n_1\lambda_1/\lambda_g = \lambda/\lambda_g \tag{6}$$

The asymmetrical planar guide comprises a substrate, n_0, of thickness, T_0, carrying the guiding layer, n_1, of thickness T_1 with air above n_1 and below the substrate. Solving Maxwell's equations for this composite structure and applying the appropriate boundary conditions leads to the expressions for transverse electric, (TE) and transverse magnetic (TM) modes, from which the curves

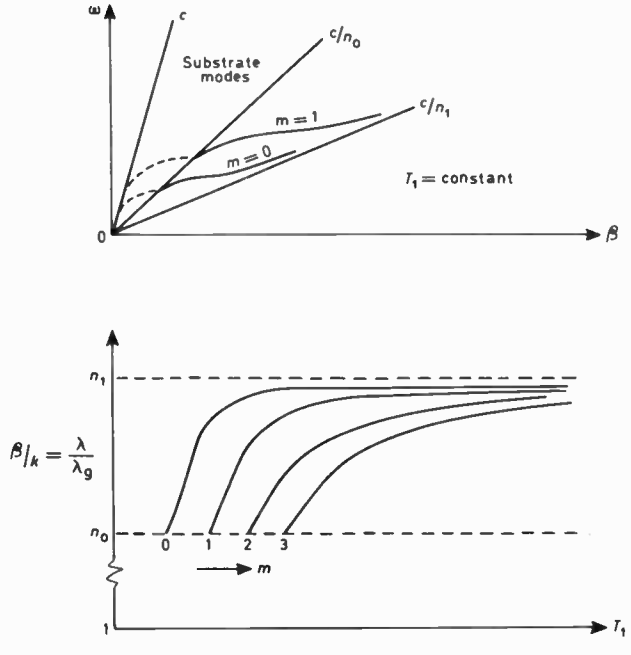


Fig. 6. Modes in asymmetrical guide.

of Fig. 6 have been drawn showing ω versus β and β/K versus the thickness of the guiding layer, T_1, for each mode. For $\beta/K < n_0$, propagation is essentially in substrate modes and, therefore, consideration of guided modes in medium 1 involves the region between $\beta/K = \lambda/\lambda_g = n_0$ and $\beta/K = \lambda/\lambda_g = n_1$. As T_1 increases, propagation in all modes approaches asymptotically the value $\beta/K = n_1$, $\lambda_g = \lambda_1$ appropriate to plane wave propagation in an unbounded medium, n_1.

3 Coupling between Planar Guides

The foregoing discussion of mode propagation in a planar light guide is based upon the condition that critical internal reflection occurs both at the upper boundary between the guiding layer and the surrounding air and at the lower boundary with the substrate. Although electromagnetic fields exist in the ambient media, these are associated with evanescent waves, the amplitudes of which decay exponentially with distance, x, normal to the direction of propagation. The decay constants in the exponential term $e^{-\gamma x}$ are:

$$\text{in air} \qquad \gamma^2 = K^2[(\beta/K)^2 - 1] \tag{7}$$

$$\text{in the substrate} \quad \gamma_0^2 = K^2[(\beta/K)^2 - n_0^2] \tag{8}$$

Since energy does not propagate into the surrounding media, it is important to enquire into the methods by which energy can be coupled from one or other ambient medium into the guiding structure and vice versa.

Beam couplers for planar guides can be classified as (a) transverse, 'end-fibre', couplers in which the incident beam is focused on an exposed edge of the guide and (b) longitudinal couplers in which the beam is incident obliquely to the upper or lower surface of the guide. The former include 'direct', 'head-on' or 'butt' couplers, which require critical alignments between the incident beam and the cleaved or polished end surface of the guide. Field contour-matching of the transverse Gaussian shape of an input laser beam to the fundamental TE_0 mode of the planar guide can produce efficient coupling but longitudinal coupling is frequently preferred due to its practical advantages.

3.1 Prism Coupling

A very convenient method of exciting a given mode in a planar waveguide is to make use of the evanescent field external to a surface at which critical internal reflection occurs. This can readily be achieved using a prism coupler, as shown in Fig. 7. Light enters the diagonal face of the prism of refractive index, n_p, and is totally internally-reflected at the lower face, which is pressed close to the guiding layer leaving an air gap (T) of a fraction of a wavelength. There is an exponential decay of the amplitude of the evanescent field in the narrow air gap, expressed by equation (7), with $\beta_p = 2\pi \cos\theta_p/\lambda_p$, θ_p being the angle between the wave-normal and the lower surface of the prism and λ_p the wavelength of a plane wave in the bulk prism material. This situation is akin to

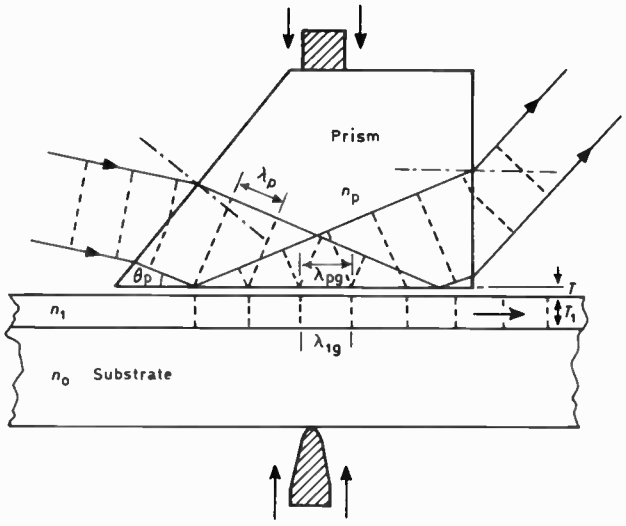


Fig. 7. Prism coupler.

that shown in Fig. 5 with the guide, n_1, replaced by a prism, n_p, which is in effect infinitely thick. The wavelength in the prism, resolved parallel to the boundary is

$$\lambda_{pg} = 2\pi/\beta_p = \lambda_p/\cos\theta_p \propto 1/\cos\theta_p \qquad (9)$$

By varying the angle θ_p, from zero to the critical angle (eqn. (3)), λ_{pg} can be controlled over a wide range. If now λ_{pg} is adjusted so that it is equal to the wavelength, λ_{1g}, of a given mode in the planar guide, this mode will be energized through evanescent field coupling. Essentially, resolved propagation parallel to the boundary in the prism together with its associated evanescent field in the air gap are both phase-matched to the guided mode in the planar waveguide.

At first sight, this might appear to be an inefficient method of coupling because of the marked decay of evanescent field amplitude across even the very narrow air gap. Nevertheless, coupling efficiencies in excess of 70% can be achieved since the coupling region extends over many wavelengths. The extent of the coupling region is controlled by the clamping pressure which deforms the materials, giving rise to a domed area of contact. If, however, the length of the contact region is allowed to become too great, energy will begin to couple back in the reverse direction from guide to prism, thereby decreasing the efficiency of coupling. Design details are available in the published literature[1,2] but, in practice, the simple expedient is followed of gradually increasing the contact pressure until maximum coupling is observed. In addition, the prism may be truncated, as in Fig. 7, and the coupling region is terminated by experimental adjustment of the incident beam position.

It is necessary to choose the prism material so that its refractive index, n_p, is greater than that of the planar guide, n_1. Reference to Fig. 6 shows that in order to couple to any selected mode in the planar guide $\lambda/\lambda_{pg} = n_p \cos\theta$ must be capable of being varied between n_1 and n_0. Since $\cos\theta < 1$, this requires that $n_p > n_1$. In practice, $n_p = n_1$ implies propagation parallel to the boundary ($\theta = 0°$) which is to be avoided because of the critical tolerance conditions imposed upon the value of θ—hence the choice of a value for n_p somewhat in excess of n_1. This can be a severe restriction on the choice of prism material required for use with guiding layers of relatively high refractive index, e.g. CdS ($n = 2{\cdot}32$), $LiNbO_3$ ($n = 2{\cdot}22$), GaAs ($n = 3{\cdot}36$).

For given conditions of operating wavelength, guide thickness and prism index, each propagated mode has associated with it a specific launching angle, θ_p, from which the guide wavelength or propagation constant can be computed. The prism coupler is used extensively in the laboratory to evaluate the characteristics of an unknown waveguide by varying θ_p and observing mode coupling over a number of modes.

By converse argument, such a prism can be used to couple energy out of the guide. Although the prism-coupler is a useful laboratory tool, it is not suitable for device applications.

3.2 Grating Coupling

Consider a light beam incident normally from air on to a regularly corrugated surface of a transparent material of index n_1, as in Fig. 8(a). Depending upon the grating spacing, $D(> \lambda_1)$, light will be diffracted into orders, m, for each of which the angle, θ_m, of the diffracted beam to the normal is given by the grating formula:

$$\sin\theta_m = m\lambda_1/D \qquad (10)$$

If the beam is incident at an angle θ_i (Fig. 8(b)), then

$$\sin\theta_m = \frac{1}{n_1}\sin\theta_i + m\lambda_1/D \qquad (11)$$

If, now, region 1 constitutes a planar guide of thickness T_1 supported on a substrate of index n_0 ($< n_1$) then one or more of the diffracted orders may suffer total internal reflection at the boundary between 1 and the substrate. This does not necessarily build up into a propagated mode in the guide unless the diffracted angle is such that a transverse standing-wave pattern is established by critical internal reflection at *both* boundaries of the planar guide. The condition for coupling into a guided mode is therefore:

$$\sin\theta_m = \lambda_1/\lambda_{1g} = \beta/K_1 = \frac{1}{n_1}\sin\theta_i + m\lambda_1/D$$

or

$$\beta = 2\pi/\lambda_{1g} = K\sin\theta_i + mK_D \qquad (12)$$

where $K_D = 2\pi/D$ is termed the grating wave number.

Hence, for suitable grating period, D, rotation of the input angle, θ_i, can cause the mth diffracted order to couple synchronously to a given mode of propagation in the planar guide. In practice, θ_i is adjusted so that use is made of the lowest possible diffraction order. Coupling

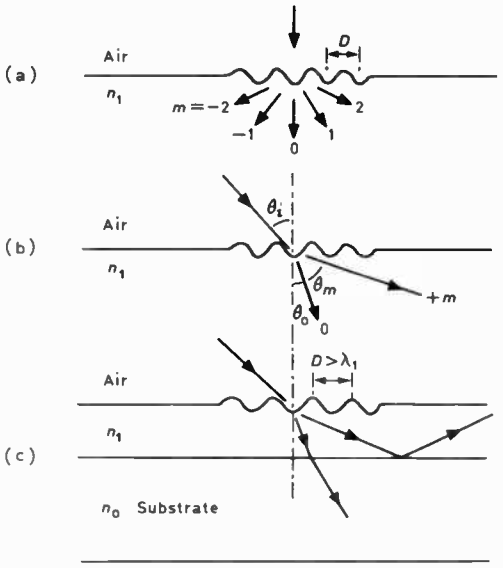


Fig. 8. Diffraction grating coupler.

efficiencies realized in the structures of Fig. 8(c) do not generally exceed 40% but if the grating period, D, is made less than λ_1 ($= \lambda/n_1$), it is possible to achieve efficiencies approaching 80% utilizing reverse coupling with the input beam incident through the substrate, as shown by Dalgoutte,[3] (Fig. 9). Only in the reverse coupling direction is there sufficient path difference to achieve a diffracted order with $D < \lambda_1$. Again, the angle of incidence is chosen to provide phase-matching to a guided mode. Coupling through the substrate is effected via a prism of refractive index similar to that of the substrate cemented on to the under surface.

It is implied that grating couplers have to be carefully designed to meet the particular conditions of the experimental arrangement. In contrast with the prism coupler, they form an integral part of the guiding structure and the only experimental variable is the angle of the incident beam. By reciprocal action both forward and reverse grating couplers can be used to extract light energy from a waveguide mode in the form of a radiated beam.

3.3 Directional Coupling

Consider two identical planar waveguides as in Fig. 10, the upper guide being inverted so that there is a very narrow air gap ($T < \lambda$) between the two. Suppose that

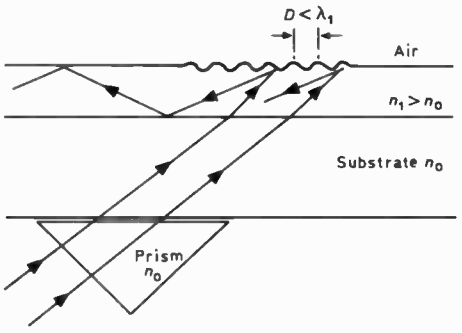


Fig. 9. Reverse coupler with prism.

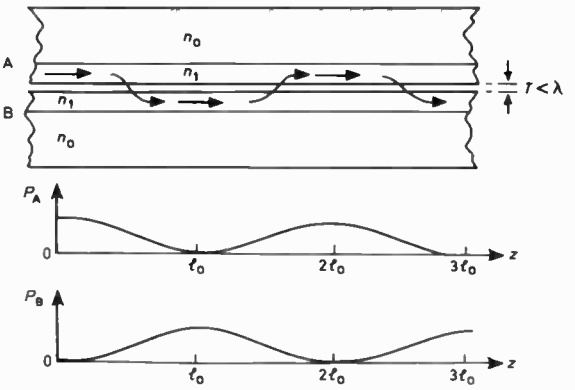


Fig. 10. Directional coupling.

energy in a single mode arrives in waveguide A: interaction of the evanescent fields in the air gap causes progressive transfer of energy into waveguide B over many wavelengths, as already described for the prism coupler. Over the coupling length, l_0, assuming no loss, all the incident energy from guide A will be transferred to guide B. The process is then reversed, with all the energy reappearing in guide A after traversing a further length l_0. If a and b are the wave amplitudes of the modes in the respective guides, the coupled mode equations may be written:

$$\frac{da}{dz} = -j\beta a + jCb \tag{13}$$

$$\frac{db}{dz} = -j\beta b + jCa \tag{14}$$

where β is the phase constant in each guide and C is a coupling coefficient per unit length.[4,5] It can be shown that the coupling length for complete energy transfer, l_0, is given by $\pi/2C$. Provided that the phase constants β of the two coupled modes are identical, there will be complete energy transfer even in physically dissimilar waveguides. The calculation of the coupling constant C depends on the product of the fields associated with both waveguides over a plane normal to the direction of propagation and is given in detail by Marcuse.[5]

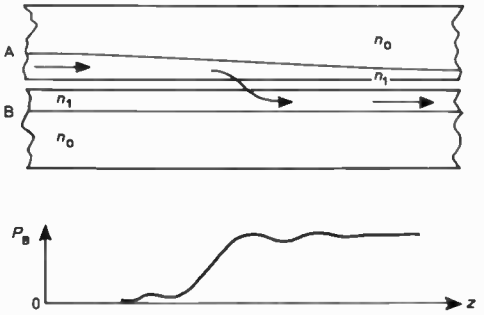


Fig. 11. Taper velocity coupling.

In the case of dissimilar phase constants, the energy transfer is not complete and will be extremely small for a large dissimilarity. The taper coupler[6] makes use of the fact that coupling will only occur where phase constants

are closely matched to relax the need for waveguide matching during manufacture. Instead of using guides with matched phase constants, one of the guides has a tapering cross-section (Fig. 11) so dimensioned that, at some part of the taper, the local phase constant is equal to that of the other waveguide. With the guides positioned as before, coupling *will* occur around the point of exact phase match, when energy will transfer from one guide to the other, but the energy will then be trapped in the second guide, due to the mismatch beyond the matched region. Dimensions are not critical, and coupling efficiencies in the region of 100% have been predicted and measured.[6,7]

4 Rectangular Waveguides

So far the propagation of light in planar or slab waveguides has been considered, with the light confined in only one dimension. In many cases it is necessary to confine the light to a particular path on the surface of the substrate. By defining the high-index guiding region as a thin strip, total internal reflection will prevent the light beam spreading out across the substrate, and the strips can be curved or branched to direct the light beam as required. The exact solution of electromagnetic propagation in a rectangular waveguide is not analytically possible, although approximations can be made,[8] or numerical techniques employed. Usually, only small changes in propagation constant occur when a planar slab waveguide is reduced to a rectangular strip and the mode structures of the two types of waveguide are similar. The strip waveguide may be formed either as a ridge on the surface of the substrate, or by diffusion as a region of higher refractive index below the substrate, or as a rib of increased thickness within a thin planar slab (Fig. 12). All the methods described in Section 3 for coupling light into and out of planar waveguides may be employed with rectangular waveguides. Particular advantages to be gained from the use of rectangular waveguides are the ability to guide or deflect the beam along predetermined paths, their suitability for electro-optic modulators and deflectors with the consequent reduction in the voltage necessary to achieve a given field strength, and the possibility of increasing the number of optical paths on a given substrate. Despite these

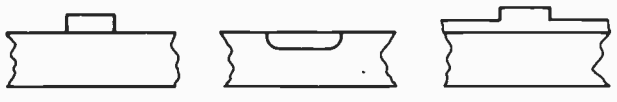


Fig. 12. Rectangular or strip waveguides.

advantages, there remain specific applications for planar waveguide devices, such as the microwave spectrum analyser and surface acoustic wave beam deflectors: examples will be described subsequently in Section 6.

5 Technology and Materials for Integrated Optics

The ability to fabricate usable devices for applications in integrated optics is principally governed by considerations of materials technology: deposition, ion-exchange, diffusion, photolithographic mask-making, selective etching, electron-beam writing and lithography. In order to obtain guiding of a light beam in a predetermined path it is necessary to increase the phase constant along the path relative to its surroundings, which can be achieved either by purely geometrical means or by a combination of geometry and a change in material properties. These processes will be illustrated in turn by considering a number of selected examples which have been successfully used to achieve light guiding and deflection in thin-film structures.

5.1 R.F. Sputtering

The most common application of r.f. sputtering techniques has been the deposition of thin films of Corning 7059 glass on to glass substrates of lower refractive index. In conventional direct-current sputtering, a negative voltage is applied to a target of the material to be sputtered, which in a low-pressure d.c. glow discharge attracts positive ions accelerated across the Crookes dark space with sufficient energy to erode the target material. The ejected material is collected on the anode formed by the substrate on to which the film is to be deposited. This process cannot, however, be used with insulators, due to the fact that the sputtering process is inhibited by the build-up of positive charge on the target. Instead, an r.f. discharge is employed: during the half-cycle when the voltage on the target is negative with respect to the plasma, sputtering occurs due to positive-ion bombardment. The build-up of positive charge is then neutralized in the succeeding half-cycle by the arrival of electrons at the target. Operating frequencies in the region 10–20 MHz effect a compromise between allowing the number and energies of positive ions to be sufficient to cause erosion and preventing build-up of positive charge which would inhibit further bombardment. The plasma is confined by a magnetic field, and the electrodes have to be water-cooled. An atmosphere of argon and oxygen is commonly employed, at a pressure of between 10^{-2} and 10^{-3} torr. The sputtering process modifies the glass composition, so that films with different refractive indices are deposited, depending upon the r.f. power density at the sputtering cathode.[9] The rate of film growth is of the order of 10 nm per minute.

Other optical thin films which have been prepared by sputtering include barium silicate on glass,[10] lithium niobate on sapphire[11] and zinc oxide on sapphire or lithium niobate.[12,13]

Since sputtering through a mask is impracticable, only planar films are produced: these must subsequently be etched through a mask to form strip waveguide structures.

5.2 Ion Exchange

An alternative to growing a thin guiding-film on the surface of a substrate is to modify the substrate surface itself by raising the refractive index of the surface layer, thus creating an integral optical guide within the substrate (Fig. 12(b)). A frequently-used technique is the exchange of silver ions from a molten silver nitrate bath with the sodium ions in the surface of a soda-lime glass substrate.[14] (Similar processes have been used to toughen glassware by putting the surface layer under compression.) Silver nitrate is molten at 220°C and may be used at temperatures up to 350°C without decomposition. Soda-lime glass substrates (commonly microscope slides with 14% Na_2O by weight) are dipped into the molten silver nitrate bath for diffusion times ranging from a few minutes to a few hours.

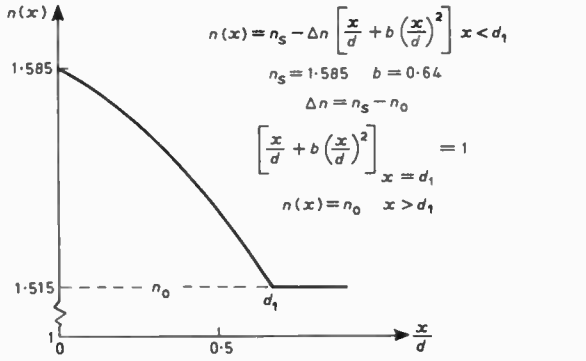


Fig. 13. Second-order polynomial refractive index index distribution for Ag-Na ion-exchanged waveguides.

The silver is distributed in a depth profile such as that shown in Fig. 13,[5] markedly different from the error-function complement profile arising from the diffusion of a single ionic species. The exchange and counter-diffusion of two species sets up an electric field, which, in turn, affects the diffusion rates. The original substrate refractive index of 1·51 is raised to about 1·60 at the surface, and below the surface is proportional to the silver-ion concentration. The high concentration of silver at the surface can give rise to staining by free atomic silver; by diluting the silver nitrate melt with sodium nitrate the surface silver-ion concentration, refractive index and liability to staining are reduced.[16] A typical molar ratio of 4×10^{-3} (silver nitrate to sodium nitrate) in the melt halves the increase in surface refractive index over the substrate index.

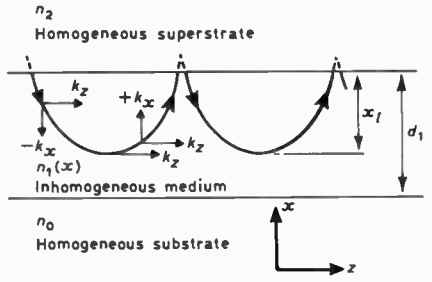


Fig. 14. Ray optics representation of a guided wave in an inhomogeneous medium.

Propagation in diffused waveguides with index grading in depth differs from that in confined guides of uniform index. Thus, after reflection from the solid-air interface, the wave direction changes gradually until it eventually becomes parallel to the axis (Fig. 14). It transpires that the refractive index at this depth is the effective refractive index of the propagating mode. A surface at this depth, x_l, is the virtual lower waveguide boundary for that particular mode: only evanescent fields exist at greater depths.

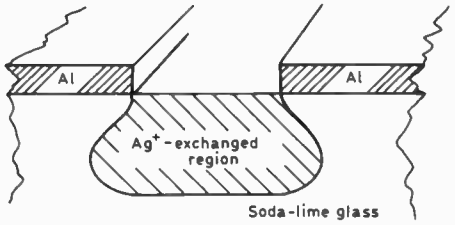


Fig. 15. Ion exchange through a metal mask.

The manufacture of light-guiding films in glass by silver-ion exchange is rapid and straightforward, although the accurate control of diffusion depth and mode propagation constant requires that the diffusion bath temperature be maintained to within ±0·1°C.[17] Strip ion-exchange waveguides may be produced by masking the substrate beforehand with an evaporated aluminium film with windows through which ion exchange will take place. Although lateral diffusion would occur directly beneath a non-conducting mask the equipotential surface created by the metal mask inhibits this diffusion, as shown in Fig. 15. In consequence, strip waveguides formed by ion diffusion using an evaporated metal mask are laterally well-defined with enhanced width control in the fabrication process. Losses in ion-exchange guides are relatively low, and approach the bulk loss of the substrate material.

5.3 Titanium Diffusion into Lithium Niobate

Lithium niobate and lithium tantalate may be prepared as large transparent single crystals, with very high electro-optic and piezo-electric coefficients. They are therefore of some importance as the basis of electro-optic and acousto-optic modulators, both in bulk optical devices and in thin-film guided systems. The confinement of the light beam in an optical waveguide is particularly advantageous in the case of electro-optic modulation, as the electrode spacing and hence the applied voltage may be much reduced from the values used in bulk modulators.

Several methods exist for the creation of optical waveguides on the surface of a lithium niobate (or similar) crystal, including outdiffusion of lithium and indiffusion of metals. Metal indiffusion is suitable for photolithographic definition of the guiding regions and titanium indiffusion has become the most common method of preparing strip waveguide modulators on $LiNbO_3$. A thin layer of titanium metal (~ 200 Å) is

evaporated on to the surface of the lithium niobate substrate, on which a negative of the required waveguide pattern has been created as a photoresist film. The 'lift-off' technique is used to remove titanium from all but the waveguiding areas, leaving the waveguide pattern defined in titanium on the substrate surface. The substrate is put into a sealed alumina tube together with lithium niobate powder, and heated in a furnace to about 1000°C. The titanium is first oxidized, and subsequently slowly diffuses into the lithium niobate surface, taking approximately 24 hours for complete diffusion. Although inward diffusion of titanium occurs, the final depth of the strip which remains above the surface of the lithium niobate exceeds the initial depth of the deposited titanium. It is believed that this is due to lithium diffusion from the substrate into the titanium film leading to the formation of mixed oxides. The addition of titanium dioxide to lithium niobate raises the refractive index by a small amount, of the order of 0·01, but during the heating process it is necessary to prevent the simultaneous outdiffusion of lithium. This is achieved by the presence of lithium niobate powder, which saturates the atmosphere inside the sealed alumina tube with lithium vapour, preventing its outdiffusion from the substrate.[18]

5.4 Photolithography

The essential feature of photolithography stems from the fact that different molecular weights of a given polymer species exhibit different solubilities in certain solvents. Indeed, solvents can be found which preferentially dissolve the lower molecular weight component, having a negligible influence on the constituents of higher molecular weight. These principles have been applied successfully for solution fractionation of a polymer having a wide distribution of molecular weights using the 'solvent–non-solvent' method. Since it is possible to degrade a high molecular weight polymer by irradiating it with incident photons, for which the photon energy, hv, exceeds that required to fracture a carbon–carbon bond, it follows that, by employing a suitable solvent, the degraded (lower molecular weight material) can be dissolved, leaving the higher molecular weight material *in situ*. This is the principle governing positive photoresists but, by suitable choice of polymer–solvent systems, a negative resist can be used in which the original higher molecular weight material is dissolved, leaving the degraded material behind.

The polymer is deposited on the substrate by first dissolving it in a solvent which rapidly evaporates. A few drops of solution are applied to the substrate which is then spun rapidly about an axis normal to its surface. Subsequent evaporation of the solvent leaves a thin coating of polymer on the surface; this is then exposed to incident light which has been focused through a mask formed in the desired pattern in chromium on a glass substrate. The pattern of the mask is replicated in the layer of photoresist remaining on the substrate after removal of the parts on which the photons have been allowed to impinge (positive resist).

These techniques are well-known from their extensive application in the formation of masks and surface patterns for microelectronic devices and they likewise play a vital role in the fabrication of both passive and active optical waveguide structures. However, for optical waveguide applications very smooth edges (better than, say 500 Å) are required in order to reduce scattering losses and the necessary length of the optical device compared with its width places severe restrictions on the optical camera-lens system required. For example, if the mask for a waveguide 2 mm long and 5 μm wide is to be made by photographic reduction of a line 2 mm wide, then the line would need to be 80 cm long. Extremely good lenses are required to cope with the inherent problem of aspect ratio.

Although replication of the pattern on a mask which has itself been formed by photographic reduction is a standard application of photolithography, an extension of these techniques provides the opportunity to form diffraction gratings directly in the photoresist material. Thus, if a monochromatic laser beam is split into two separate beams of equal intensity by means of a dielectric-coated beam splitter, one of the divided beams can be caused to traverse a longer path than the other and, by suitable arrangement of reflecting mirrors, the two beams can then be recombined to form an interference pattern on the surface of the photoresist film as in Fig. 16. If each beam is incident at an angle ψ to the normal (Fig. 17), the fringe spacing, D is given by $D = \lambda/2 \sin \psi$, where λ is the wavelength of the incident light. Thus, the minimum theoretical grating spacing is

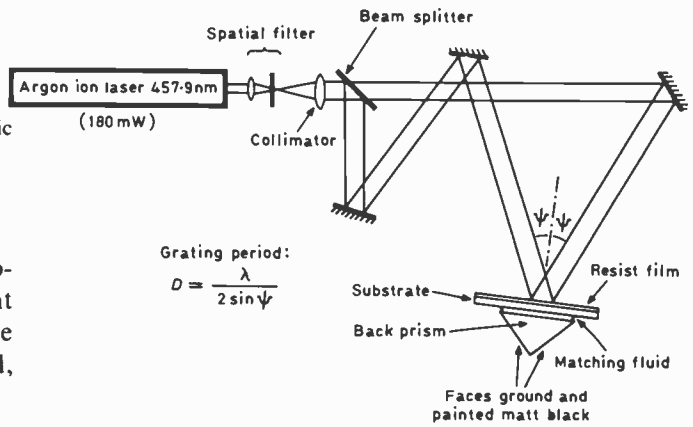


Fig. 16. Three-mirror interferometer for the production of holographic diffraction gratings.

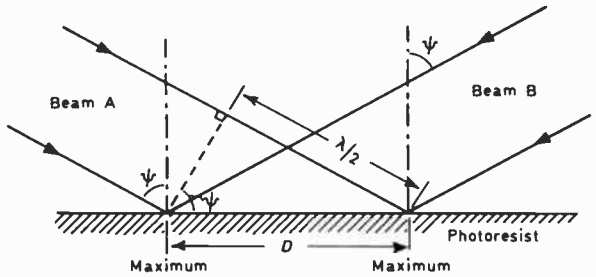


Fig. 17. Interference of two beams on photoresist film.

$\lambda/2$, when $\psi = \pi/2$, and the two beams then arrive at grazing incidence but travelling in opposite directions. This is not, however, a practical proposition since, for values of ψ approaching $\pi/2$, spurious modulations occur on the desired pattern due to increasing reflectivity at the interfaces with increasing angle of incidence. In addition, for large values of ψ the beams expand laterally and any optical imperfections in the wavefronts become amplified. A suitable compromise is to work with $\psi \simeq 60°$. Since D is proportional to λ the use of blue or ultra-violet light to which the photoresist is sensitive produces a finer grating pitch. The effective wavelength of the incident light can be further reduced by placing a prism of higher refractive index, n_p, against the exposed surface of the photoresist using a matching liquid, in which case $D = \lambda/2n_p \sin \psi$.

Unwanted reflections from interfaces parallel to the substrate surface establish standing-wave patterns with nodal planes parallel to the substrate. Thus, at some plane within the film of photoresist, there may occur a field minimum inhibiting the exposure of the photoresist through its entire thickness. The developed pattern then remains embedded in the photoresist but without 'cutting through' to the surface of the substrate. This may be overcome by using the technique of simultaneous exposure and development.[19] The photoresist-covered substrate is immersed in the developer solution contained within a hollow prism. As the positive photoresist is exposed, development occurs immediately, the top layers of material dissolving away to allow the underlying layers to be exposed and developed in turn. This precludes the use of a front prism but permits the fringe pattern to be cut right through to the surface of the substrate.

Having produced such a grating pattern in the photoresist film, the pattern can then be etched chemically or by ion-beam into the surface of the underlying substrate, which will normally comprise a light-guiding layer formed on its own base substrate. Embedded waveguide gratings of 190 nm period have been prepared in this way in glass layers.[20]

A similar technique of grating preparation can be employed on materials of high refractive index. This finds particular application in the manufacture of gallium arsenide and other semiconducting lasers with distributed feedback or distributed reflectors. Gratings designed to reflect at the laser frequency are incorporated into the multilayer laser structure, thus obviating the need for conventional cleaved-face mirrors. In principle, this allows semiconducting laser sources to be formed as an integral part of semiconductor optical waveguide structures, thus realizing the ultimate aim of truly integrated optical devices.

One of the most useful features of photolithography is that of metal lift-off, which is especially useful in the fabrication of strip (3-dimensional) waveguides. A pattern is generated in a positive photoresist deposited on top of a planar guide surface and the regions which have been exposed to the incident radiation coming through the gaps in the mask are removed by solution in the developer. A metal coating is then applied to the entire surface followed by subsequent immersion in a solvent (e.g. acetone), which dissolves the original resist material and washes away the photoresist-and-metal covering on those regions which replicate the lines on the original mask. This 'lift-off' is facilitated by ultrasonic agitation and what finally remains on the guide surface is a metal pattern replicating the gaps in the original mask which can then be used as a barrier in a further etching process in which all but the wanted portions of the slab guide are removed. In the case of ion-exchange waveguides, the metal pattern allows ion exchange with the glass substrate to take place in those regions which are not covered by the protective metal barrier (Sect. 5.2).

6 Devices as Components of Integrated Optical Systems

The essential principles underlying a multichannel communication system are those of combining information from the separate channels, transmitting the combined signals over a single transmission link and separating the individual channels at the receiving end, routing these to their selected destinations. Present-day applications of optical communications utilize fibre optic transmission lines but continue to employ standard electronic switching techniques. The challenge of integrated optics is to provide optical methods for signal multiplexing, demodulation and routing. These various functions involve optical beam splitters, both passive and active, filters, selective couplers, switches, modulators and detectors. Various means of achieving these functions have been described in the literature and an outline is given here of some of the more important techniques for practical realization.

6.1 Passive Waveguide Junctions (Beam Splitters)

A basic element of any optical system is the Y-junction, by means of which signals from separate sources are combined, or received power may be divided between two or more channels. Consider, for example, a basic Y-junction beam splitter (Fig. 18) fabricated in $LiNbO_3$. Calculated and experimental performances of such a junction have been compared by Sasaki and Anderson,[21] who show that the total power transmission (sum of the powers in each arm of the junction) decreases markedly

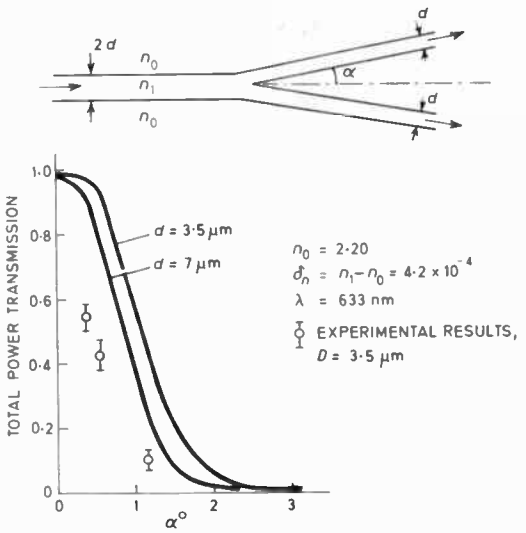


Fig. 18. Passive Y-junction.

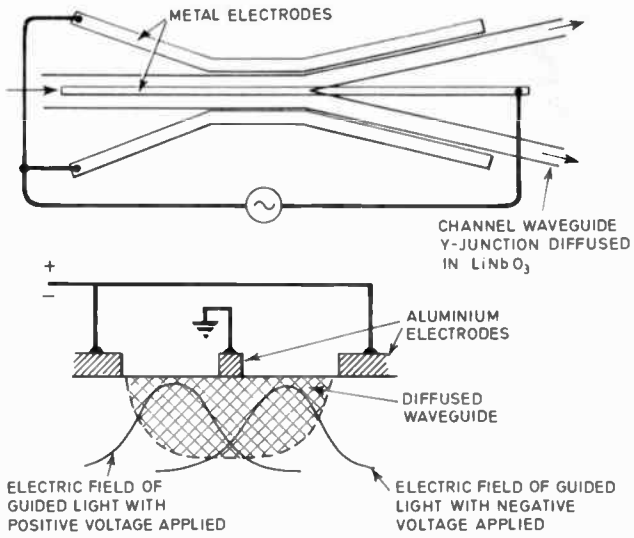


Fig. 19. Electro-optic Y-junction switch.

with increasing half-angle, α. The power lost is radiated into the substrate and the fact that the measured power transmission decreases even more rapidly than the predicted behaviour is attributed to imperfect fabrication of the sharp tip of the division between the output guides. It has been shown that the total power transmission depends critically upon the half-angle which, in the present instance, must not exceed 0·5° for acceptable insertion loss. In practice, this implies that the length of the junction must be many times the width of the waveguide in order to provide effective separation of the output arms so that access is possible to each. Some 3000 wavelengths are required in order to give a separation of about 30 μm between the output arms. Thus, although the actual division of the incident beam into the two arms occurs over a relatively short distance, the device itself is necessarily very much longer.

6.2 Active Y-junction Switch

Although there are applications where approximately equal power division of an incident beam is required (3dB coupler) the major interest in the Y-junction is in its use as a switch which, in turn, implies its fabrication in an electro-optic material. Single crystal $LiNbO_3$ is almost universally employed because of its relatively low loss combined with large values of electro-optic coefficients. Such a structure is shown in Fig. 19.[22] A titanium-diffused waveguide structure in the form of a Y-junction is fabricated in single crystal $LiNbO_3$, the crystallographic orientation of which is selected to give maximum electro-optic coefficient in the transverse z-direction. A metal electrode pattern is superimposed with biasing applied so that one side of the input waveguide has an enhanced value of refractive index at the expense of the other side. This causes the beam to be deflected towards the region of higher refractive index so that, when it arrives at the Y-junction, it is positioned to follow the corresponding output arm with continued assistance from the biasing field, since the electrode structure extends beyond the junction. When switched with 30V applied between centre and outer electrodes, the overall efficiency of transmitted power can be much greater than for the passive Y-junction and, moreover, practically all the energy can be contained in a single arm (Fig. 20). Indeed, the efficiency of the electro-optic deflection is such that much larger junction angles can be tolerated than is possible with a passive Y-junction (see Fig. 18, with $\alpha = 1{\cdot}15°$). The structure of Fig. 20 with $\alpha = 0{\cdot}57°$ had a measured insertion-loss of less than 1 dB, whilst that with $\alpha = 1{\cdot}15°$ was 3 dB.

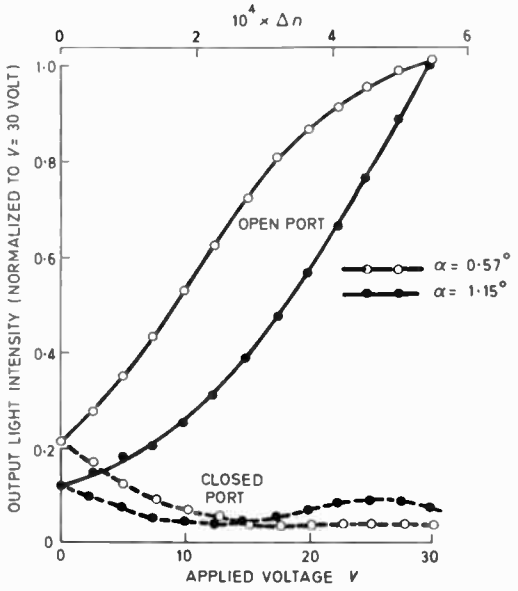


Fig. 20. Switched Y-junction in $LiNbO_3$ Ti-diffused = 633 nm.

Clearly, efficient switching with high extinction ratios is a practical possibility with a potential bandwidth of 10 GHz.[22] Despite the somewhat larger junction angle permitted, the physical length of the active Y-junction switch remains several hundred wavelengths.

6.3 Grating Filters and Beam Deflectors

The planar guide with corrugated surface, forming a diffraction grating coupler, was discussed in Section 3.2. The angle to the normal of the incident beam, θ_i, for synchronous coupling to a mode is given by equation (12):

$$\sin\theta_i = \beta/K - m\,K_D/K = \lambda/\lambda_{1g} - m\,\lambda/D \qquad (12)$$

where D is the grating period and m is the diffracted order. If $D = \lambda_{1g}/2$ then $\sin\theta_i = -\lambda/\lambda_{1g}$ and since $\lambda > \lambda_{1g}$, as a condition for a propagated mode, then $|\sin\theta_i| > 1$, which applies that even for $m = 1$ (first-order diffracted beam) no coupling can occur from an incident beam via such a grating nor, reciprocally, can outward coupling take place to a radiated beam. However, with $D = \lambda_{1g}/2$, a mode already established in the guide travelling in the axial direction normal to the grating lines will be reflected back along its path of incidence since reflections from successive grating corrugations will differ in optical pathlength by $2D = \lambda_{1g}$. In general, a guided wave of guide wavelength λ_{1g} will experience maximum reflection from a grating of spacing D when incident at an angle ϕ to the grating normal given by (Fig. 21):

$$\lambda_{1g}/2 = D\cos\phi$$

or

$$D = \lambda_{1g}/2\cos\phi \qquad (15)$$

Given the angle of incidence, ϕ, and the grating period, D, Bragg reflection occurs only for the wavelength λ_{1g} given by equation (15). Hence, the grating acts as a rejection filter: wavelengths other than $\lambda_{1g} = 2D\cos\phi$ are transmitted along the direction of the incident beam.

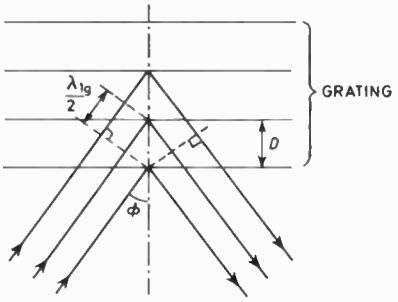


Fig. 21. Reflection from several grating corrugations.

In practice, the bandwidth of the reflection filter is governed by the finite length of the grating, the non-zero width of the lines and the loss in the transmission medium. This filtering effect is similar to that obtained with a multilayer dielectric filter.

The grating period given by equation (15) may be less than half the free space wavelength of the incident light wave ($D < \lambda/2$). Since for a glass film waveguide of refractive index n_1 supported on a substrate of refractive index n_0, λ_{1g} lies between λ_0 and λ_1 (Fig. 6) and $n_0 \simeq n_1 \simeq 1{\cdot}5$, we can take $\lambda_{1g} \simeq \lambda/1{\cdot}5 \simeq 0{\cdot}7\,\lambda$. Thus, a grating space of about $0{\cdot}35\,\lambda$ ($D = \lambda_{1g}/2$) will be required for Bragg reflection of the incident guided wave. Such grating corrugations are too closely spaced to be obtained by standard photographic reduction: recourse is had to the technique of holographic grating formation as described in Section 5.4.

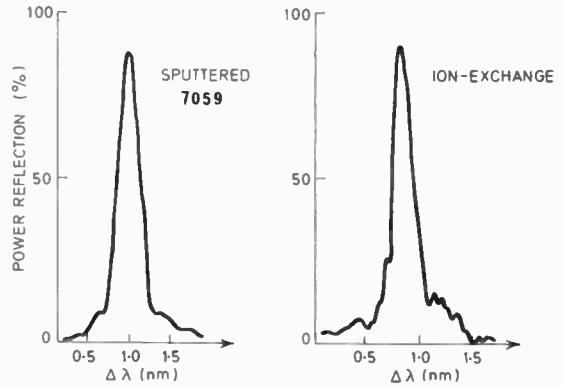


Fig. 22. Bragg rejection filter response in planar waveguide.

Grating filters fabricated by these means have been etched into sputtered Corning 7059 planar waveguides formed on Fischer microscope slides and into ion-exchange guides on similar substrates.[20] Measured results taken at normal incidence ($\phi = 0°$) for filters formed in each of these two types of waveguides are shown in Fig. 22. In each case, half-power bandwidths of less than 0·5 nm (5 Å) are achieved at a centre wavelength in the region of 610 nm. These results demonstrate clearly the possibility of achieving narrow bandwidth filtering for applications in optical wavelength demultiplexing. For example, successive grating filters, each of differing periodicity, will each reflect its associated band of frequencies and, by arranging these obliquely to the incident beam, it is possible to achieve spatial channel separation between the incoming frequencies. A particular example of this is provided by the crossed grating demultiplexer of Fig. 23, in which two gratings of equal periodicity are formed in a photoresist layer deposited on a 7059 planar guide at an angle of 40·5° to each other.[20] If the incident beam

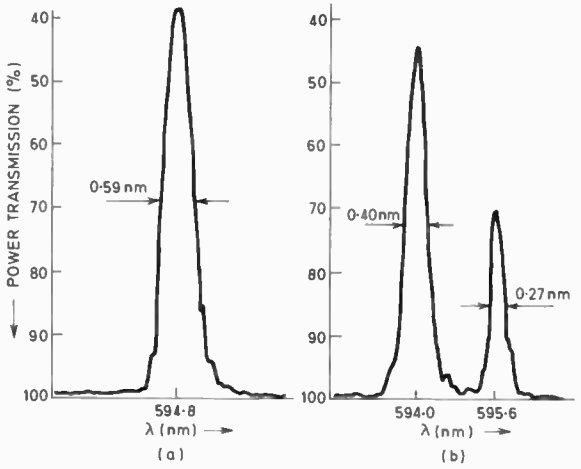


Fig. 23. Crossed grating demultiplexer response.

bisects the angle between the normals, some 60% of the incident power is divided between two reflected beams, each at 40·5° to the incident beam (Fig. 23(a)). Rotational displacement of the gratings by 0·25° with respect to the incident beam changes the respective angles of reflection to 41° and 40° so that different frequencies are then required to achieve maximum power in each of the reflected beams (Fig. 23(b)). Although the frequency of the incident beam has here been swept in order to demonstrate this response, the practical application would be the simultaneous separation of two frequencies, both of which are contained in the incident beam.

6.4 Mach–Zehnder Interference Modulator

Modification of the refractive index in an electro-optic light-guiding material by application of an electric field provides the opportunity to control the phase velocity of a guided wave. This, combined with equal power division of an incoming light beam at a Y-junction permits differential changes to be made in the relative phase between the signals in the two arms. Subsequent recombination of the two beams gives rise to constructive or destructive interference in the single output guide, converting the essential process of phase modulation into intensity modulation at the output. Such a Mach–Zehnder type of interferometer is illustrated in Fig. 24(a).[23] Using the process of titanium diffusion into lithium niobate, parallel single-mode waveguides, separated by sufficient distance to reduce cross-coupling to a minimum, were joined at either end by symmetrical Y-junctions, with a half-angle of 0·57°. Electrodes were only deposited along the parallel waveguide region so that equal power division took place at the incoming Y-junction. The outer electrodes were connected, so that the changes in refractive index in the two waveguides were of opposite sign. Experimental results for the intensity modulation observed by applying d.c. pulses at 1 kHz rate with a width of 20 μs are shown in Fig. 24(b). The half-wave voltage from peak to minimum is 2 V, with an extinction ratio of −18 dB at an applied voltage of 2·2 V. The insertion loss of the device is estimated to be approximately 5 dB. With an electrode capacitance of 6·2 pF an attainable bandwidth of 1 GHz is indicated for a 50 Ω load.

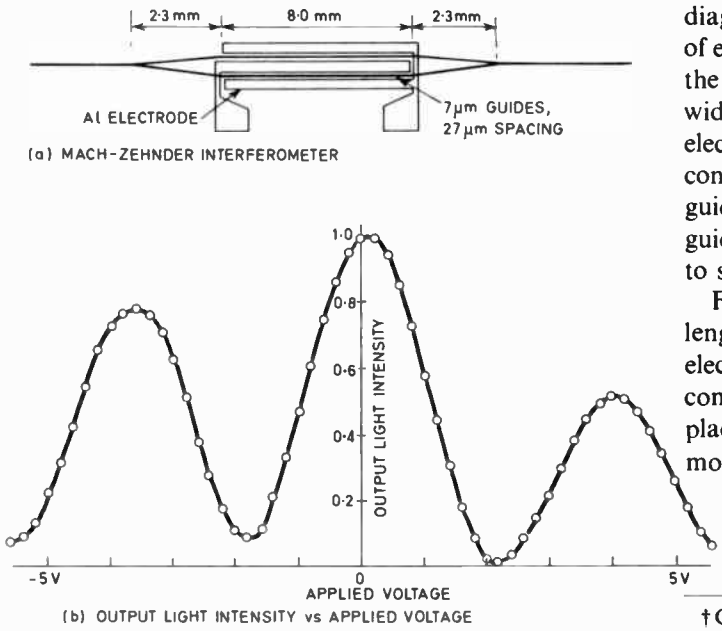


Fig. 24. Mach–Zehnder interferometer and response.

6.5 Directional Couplers as Switches

The evanescent fields outside the guiding regions of optical dielectric waveguides may be exploited to allow transverse coupling between adjacent waveguides, as described in Section 3.3. If the waveguide modes have equal propagation constants, the energy transfer is complete, but if the propagation constants differ, only a part of the energy in one guide will be coupled to the other, and the coupled energy will subsequently be recoupled back into the first guide. The spatial period of the coupling process depends not only on the guide separation but also on the degree of mismatch between the mode propagation constants. So for guides of fixed separation the coupling length l is reduced from the matched value l_0 as the mismatch is gradually increased. If the mismatch is sufficient to reduce l to $l_0/2$, an interesting situation will exist, as illustrated in Fig. 25(b). Energy launched in one guide will completely couple to the second guide over a length l_0 if the two guides are exactly matched, but for the particular mismatched case when $l = l_0/2$, over the same distance l_0 the energy will all be recoupled into the first guide.

An optical switch constructed on this principle has been called the 'COBRA'†.[24] Titanium-diffused lithium niobate waveguides were formed in the pattern shown diagramatically in Fig. 25(a) with electrodes laid on top of each waveguide. The coupling length l_0 which defined the device length was 1 mm for identical guides 3 μm wide separated by 2 μm. The fringing fields beneath the electrodes act in opposition to raise the propagation constant β in one guide while lowering β in the other guide. With zero applied field, complete coupling from guide 1 to guide 2 was observed, while 6 V was necessary to switch all the input light back into guide 1.

Fabrication of the device is critical, since its coupled length must be exactly l_0 (or an odd multiple of l_0). No electrical adjustment can be made to ensure that complete transfer of power from guide 1 to guide 2 takes place when the guides are exactly matched. However, by modifying the electrode structure so that one electrode

† Commutateur Optique Binaire Rapide

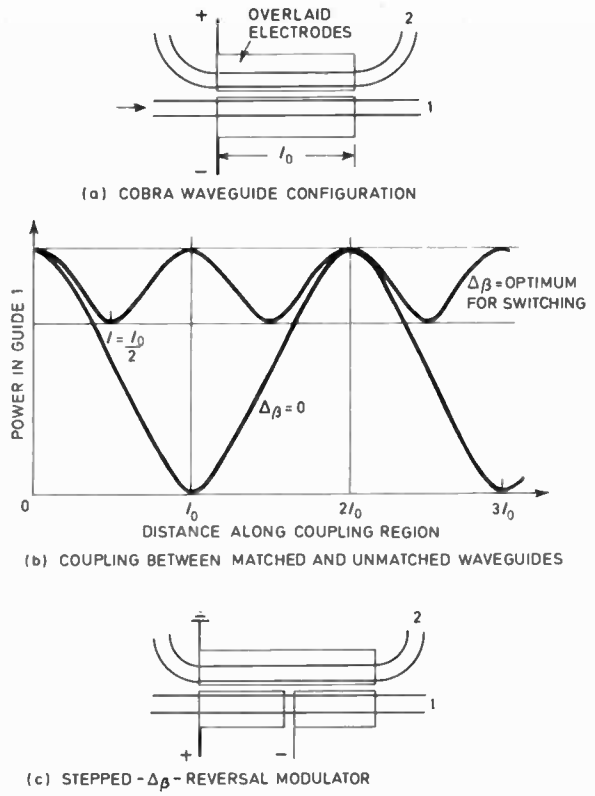


Fig. 25. Directional-coupler electro-optic switches:
(a) Cobra waveguide configuration;
(b) Coupling between matched and unmatched waveguides.

is divided into two halves as in Fig. 25(c), with opposite polarities on each half, the effective length l_0 can be adjusted to equal the physical length of the modulator.[25] For straight-through operation the divided electrode is reconnected and the device is operated as described previously.

6.6 Acousto-optical Thin Film Beam Deflectors

When an acoustic (i.e. elastic) wave travels through a bulk solid, the material experiences a periodic variation of density along the path of the wave. In a transparent solid the density variations will be accompanied by corresponding refractive index changes, which constitute a moving phase-diffraction grating. A light-beam crossing the path of the acoustic wave will be scattered by the phase grating, but the interaction will only be of significant magnitude if the conditions of Bragg diffraction are met:

$$\sin\theta_B = \frac{\lambda_1}{2\Lambda} \qquad (16)$$

where θ_B is the angle between the light beam normal and the acoustic beam wavefronts (Fig. 26(a)), λ_1 is the wavelength of the light within the solid medium and Λ is the acoustic wavelength. The light beam will be deflected by $2\theta_B$ from its original path.

By launching the acoustic wave at and parallel to the surface of the solid, a 'surface acoustic wave' (s.a.w.) will result, most of the wave energy being concentrated within a depth of one acoustic wavelength. Such a wave can be generated by an interdigital electrode system comprising two or more parallel electrodes (Fig. 26(b)) deposited on the surface of a piezo-electric material. When a voltage is applied, alternate pairs of electrodes will be attracted and repelled. An alternating applied voltage will set up s.a.w.s travelling across the surface away from the electrodes. Such a s.a.w. transducer will be most effective when the electrode separation is equal to one half of the generated s.a.w. wavelength.

If a thin optical waveguiding film is formed on the surface of the same piezo-electric solid, an optical surface wave may be made to interact efficiently with the s.a.w. since both the light and the acoustical energy may be confined in the same thin surface layer. Use of the Bragg condition of equation (16) with λ_1 replaced by the planar guide wavelength λ_{1g}, produces an efficient light-beam deflector. Lithium niobate has a large piezo-electric coefficient, and has been used for optical surface wave deflection with optical guides formed both by titanium indiffusion and by lithium outdiffusion.[26] A crystalline quartz substrate with a superposed guiding film of higher refractive index glass has also been used for a s.a.w. beam deflector.[27]

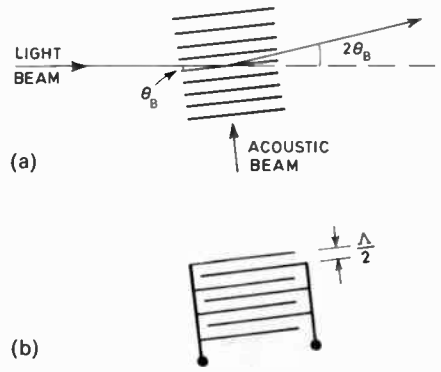


Fig. 26. (a) Bragg diffraction by acoustic waves;
(b) S.a.w. transducer electrode pattern.

From equation (16) it is clear that the s.a.w. wavelength Λ must be within an order of magnitude of the optical guided wavelength, λ_g, for significant deflection of the optical beam. For 2° deflection of the 633 nm He–Ne laser output in $LiNbO_3$, ($\lambda_g \simeq 290$ nm), $\Lambda \simeq 8$ nm, and the transducer finger spacing must be 4 nm. The proportion of the incident beam deflected depends on the generation efficiency of the s.a.w. and on the acoustic beam width (which defines the interaction length). Diffraction efficiencies are generally low, up to 20%, although the diffracted on/off ratio may be extremely high, thereby providing effective switching action. The acoustic input power is limited by the electrical breakdown characteristics of the transducer.

6.7 Microwave Spectrum Analyser

It is apparent from equation (16) of the previous Section that, if θ_B is small, the beam deflection angle $2\theta_B$ of a

thin-film s.a.w. beam deflector is proportional to the acoustic input frequency, f:

$$2\theta_B \simeq \lambda_{1g} \cdot f/V \qquad (17)$$

where V is the s.a.w. velocity ($3{\cdot}6 \times 10^3$ m s^{-1} in lithium niobate). Thus, within the bandwidth limits of the acoustic transducer, the diffracted output beam may be deflected by sweeping the acoustic input frequency. The fractional bandwidth of the transducer is approximately equal to the inverse of the number of electrode pairs. To optimize the Bragg diffraction of the light beam for a range of frequency f, extending over the bandwidth, it will also be necessary to maintain the angle θ_B given by equation (17), between the lightbeam and the acoustic wavefronts. This requires an angular adjustment of the acoustic beam with input frequency, which may be accomplished by using an array of tilted transducers of decreasing finger spacing, as shown in Fig. 27.[28] Each transducer accepts a limited frequency band and, at its centre frequency, the angle of the acoustic diffraction grating to the optical beam is optimized for Bragg diffraction. By combining the outputs of the transducers the acoustic beam is automatically rotated as the frequency changes while the overlapping frequency responses of the transducers ensure efficient s.a.w. generation over a wide band. Such a device has been used as a beam deflector over an acoustic range of 240 MHz to 400 MHz.[29]

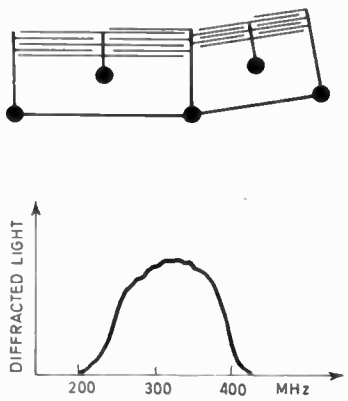


Fig. 27. Beam-steering s.a.w. transducer.

The number of deflected beam positions that can be separately identified is governed by the angular width of the diffracted light beam. If the light beam width is D, the diffraction width of the beam $\Delta\theta$, will be approximately λ_{1g}/D radians. Hence the number of resolvable spots is given by:

$$N \simeq \frac{2\theta_B}{\Delta\theta} = \frac{D\Delta f}{V}$$

where Δf is the total acoustic bandwidth. In the example quoted earlier, the expected number of resolvable spots was 129, whilst the measured number was 125.

By focusing the deflected output beam on to a row of suitably-spaced photodetectors each resolvable beam position may be detected. If an electrical signal containing several frequency components is now fed to the transducer array, each component will diffract a beam through a different angle to a corresponding photodetector. Thus, the device analyses the spectrum of the electrical input signal. In order to improve the resolution of the spectrum analyser,—that is, to increase N—the width D of the light beam must be increased. Prism input and output couplings have been used in experimental beam deflectors,[29] but the optical beam width is limited by the small coupling region to less than 1 mm. The use of geodesic lenses has been suggested to expand the input beam and focus the deflected light. Such a lens consists of a depression in the surface of the substrate, over which an optical guiding film is formed. As the optical path across the centre of the depression is increased, the depressed region of waveguide has a focusing effect on a guided beam.

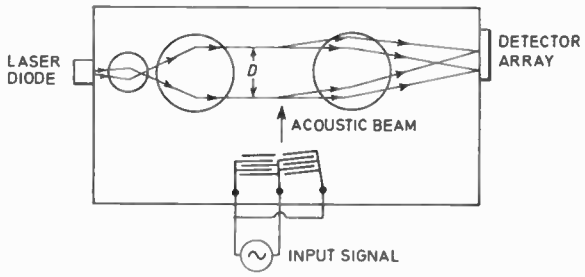


Fig. 28. S.a.w. microwave spectrum analyser.

The schematic layout of a microwave spectrum analyser is shown in Fig. 28. It is expected that such spectrum analysers will be able to operate over a 1 GHz swept frequency range with a resolution of 1 MHz, which implies an optical beam width D of at least 3·6 mm.

6.8 Coupling of Devices

The thin-film planar or three-dimensional optical devices described in the preceding sections have been mainly constructed as individual units, working in isolation. These have been formed in various materials and, at the present stage of the technology, the essential problem is to devise ways of intercoupling such devices to build a complete system. This can be accomplished either by 'hybrid' means, in which the separate devices are linked by passive optical bridges or eventually by true integration, in which the active structures and associated passive linkages are formed on a common substrate which may include the laser source.

The requirement to employ hybrid interconnection of separate devices stems from the practical necessity to work with relatively long optical structures, in which coupling and interaction occur over several hundred optical wavelengths. For strip waveguides, the Y-junction is at present the basic means for beam splitting and/or switching and, owing to the very small permissible junction angle (Sect. 6.1.) it is physically

necessary to extend the length of the Y in order to provide access to the separate signals in each arm. Inevitably, this precludes cascading many devices on the same substrate, so that it is necessary to introduce some form of optical interconnection between discrete devices. One way of achieving this is to employ the 'sandwich ribbon' as a coupling waveguide.[30] This consists of a high index light-guiding filament bonded to one side of a flat ribbon substrate of lower refractive index. Evanescent fields are available for directional coupling at the exposed surface of the ridge (Fig. 29) whilst the reverse side of the flexible supporting substrate can be handled with relative impunity. Optimum coupling can be obtained by adjustment *in situ*, followed by permanent bonding using epoxy resin or other adhesive.

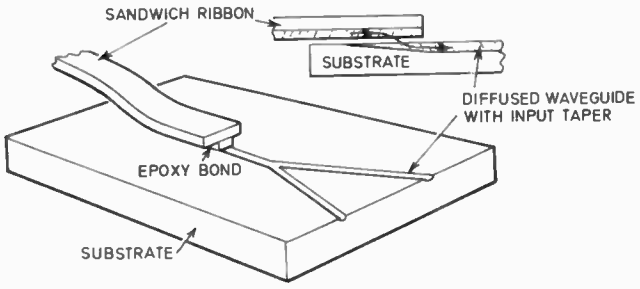


Fig. 29. Coupling to optical device by sandwich ribbon.

Alternatively, bridging waveguides of higher refractive index can be formed as overlay tapered connections between devices on the same substrate.[31] In principle, this is the beginning of true optical integration, in which sources and detectors, along with switches and modulators, can be incorporated into the same monolithic structure.

7 Applications of Semiconductors in Integrated Optics

Because of their potential application as laser sources, attention has been focused on group III-V semiconductor compounds for applications in integrated optics. Gallium arsenide lasers have already become the most important light sources for optical communications, emitting at around 0·9 μm wavelength, although the improvement in optical fibre loss by moving to longer wavelengths has prompted the development of indium phosphide and similar compounds for injection lasers.

Early GaAs homojunction lasers, consisting of a simple junction between p- and n-type GaAs, were inefficient and could not be operated continuously at room temperature because of the high lasing-threshold current density. Although the diode junction region, where optical gain occurs, was slightly raised in refractive index above the surrounding material by the combined effects of impurity concentrations, free-carrier concentration and gain, the optical waveguiding effect was minimal and absorption loss in the passive p and n regions was high. By creating strong waveguiding

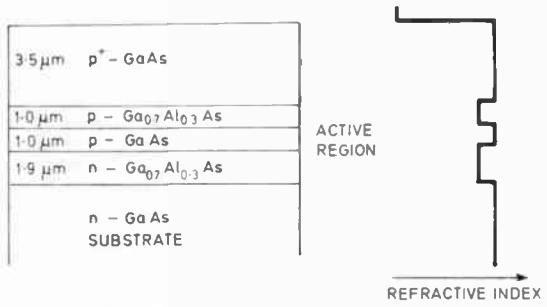


Fig. 30. Heterostructure diode laser.

around the junction it is possible to confine the light to the gain region of the laser (as is done in heterostructure lasers), reduce the lasing threshold, and obtain c.w. room temperature operation. The construction of a heterostructure laser and its resulting refractive index profile are shown in Fig. 30. Lower-index layers of p- and n-type $Ga_{0.7}Al_{0.3}$ As surround a thin p-type GasAs region of high refractive index, forming an effective planar optical waveguide.

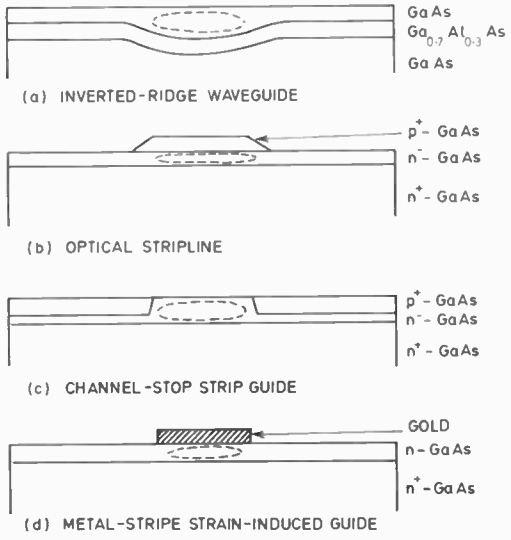


Fig. 31. Semiconductor strip waveguiding structures.

A wide range of thin-film optical components, apart from lasers, can be formed on GaAs substrates. Passive waveguides may be made by a variety of methods. As in the heterostructure laser, compositional changes may be used to alter refractive index. Figure 31(a) shows an inverted-ridge waveguide,[30] a groove is first etched into the GaAs substrate and after an optical buffer layer of $Ga_{0.7}Al_{0.3}As$ has been grown over the surface, the groove is filled by a further layer of GaAs. A more conventional ridge waveguide may be formed by etching the top layer of a multilayer structure. In the structure shown in Fig. 31(b), the p^+ ridge is of *lower* refractive index than the n-type layer beneath but loads it sufficiently to confine the light beneath the rib.[33] Proton bombardment of the crystal reduces the free-carrier concentration in the surface layer, thus producing a high-index film since the plasma contribution to the refractive index increases with decreasing electron

concentration.[34] Ion-implantation with beryllium can be used to *lower* the refractive index of the surface layer,[33] a photoresist mask is sufficient to stop the beryllium ions and allow strip waveguides to be defined (Fig. 31(c)). Finally, strain in the surface of GaAs under an evaporated strip produces enough refractive index change (by the photoelastic effect) to guide light beneath the strip even when no electrical bias is applied to the metal strip[35] (Fig. 31(d)).

The evaporated gold strip of the last example may be used as a Schottky barrier. By reverse-biasing the Schottky barrier a high electric-field depletion region will be formed, which may be made to extend throughout the entire waveguiding region. The electro-optic effect in GaAs is considerable; by applying 8 V across a 0·6 μm thick depletion region an optical phase retardation of 60° has been observed over 0·4 mm waveguide length.[36] By using an interference technique, the phase modulation could be converted to intensity modulation. As with lithium niobate, an alternative approach to intensity modulation is to use a directional coupler, the coupling coefficient of which may be altered.

Directional-coupler switches have been made of two closely-spaced parallel Schottky-barrier ridge waveguides.[37] The change in phase constant induced by an applied bias on one arm is sufficient to alter the coupling coefficient. Over a coupled length of 8·5 mm with guides spaced by 8 μm, 30 dB extinction ratio was observed at an applied bias voltage of 28 V.

Thus, with lasers and photodiodes, a wide range of optical components can be made on GaAs substrates, so that the use of semiconductor-based optical system offers the best possibility of true monolithic integration. Some special fabrication techniques will have to be used. Isolated semiconductor lasers are generally provided with mirrors by cleaving the end faces of the crystal chip. Such a technique is not possible in an integrated circuit, and optical feedback must be provided by diffraction gratings incorporated in the optical waveguiding region, at either end of the laser gain region (distributed Bragg reflectors) or throughout the lasing region (distributed feedback). Beyond the laser source, connected optical waveguides and components must be made of material with a shorter-wavelength bandgap than that of the laser, since otherwise the laser emission wavelength will be strongly absorbed by the external circuitry. Many compositional changes will thus be needed over a single integrated circuit chip. Molecular beam epitaxy (MBE) of thin semiconductor films provides an adaptable fabrication method. MBE consists of the separate evaporation, under ultra-high vacuum, of the elemental constituents of the layer to be grown, the molecular beams being collimated and directed at the substrate where the required epitaxial layer will then grow. Masking can be used to define areas of growth of a film of particular composition. Films of GaAlAs have been grown by MBE with optical losses as low as liquid phase epitaxially grown films.[38] Composition variations are particularly easy using MBE, and interest in longer-wavelength lasing semiconductors is increasing. Indium phosphide films are being grown at Glasgow University and elsewhere, and it is hoped that a number of compounds of In, P, Ga and As, with a wide range of bandgaps and hence of lasing wavelengths, may be grown by MBE to form monolithic integrated optical circuits. Moreover, it may prove advantageous for associated integrated electronic circuits to be formed by diffusion in the same chip as the optical circuits, to achieve the highest modulation and switching speeds.

Progress into full integration has as yet been limited. Small numbers of components have been grown on a single substrate; perhaps the most spectacular experiment has been the formation of six distributed feedback laser diodes of different output wavelengths, the outputs combined by strip waveguides and Y-junctions.[39] It does seem at present that monolithic integration in semiconductors holds out a strong promise of success, aided by the consequent elimination of optical coupling problems between heterogeneous components and devices.

8 Conclusion

Because of the many-wavelength dimensions of the components, it is unlikely that an integrated optical circuit will contain as many components as a large-scale electronic integrated circuit, so that hybrid integration of optical circuit components fabricated on diverse substrate materials is still a practical possibility. However, monolithic integration using III-V semiconductor substrates and passive and active guiding layers holds out the attraction of including sources and detectors with all other types of optical component, and with the associated integrated electronic circuitry.

Improvements to fabrication techniques are still required; present lithographic techniques place limits on strip waveguide attenuation (because of boundary roughness) and device size. Tuned waveguide elements cannot be made by photolithography; however, the improved definition of electron-beam writing should allow the use of such components, with a consequent reduction in device length.

9 References

1 Harris, J. H., Shubert, R. and Polky, J. N., 'Beam coupling to films', *J. Opt. Soc. Am.*, **60**, pp. 1007–16, 1970.

2 Ulrich, R., 'Optimum excitation of optical surface waves', *J. Opt. Soc. Am.*, **61**, pp. 1467–76, 1971.

3 Dalgoutte, D. G., 'Periodic couplers for the excitation of optical guided waves in thin films', Ph.D. thesis, University of Glasgow 1973.

4 Yariv, A., 'Coupled mode theory for guided wave optics', *IEEE J. Quantum Electronics*, **QE-9**, pp. 919–33, 1973.

5 Marcuse, D., 'The coupling of degenerate modes in two parallel dielectric waveguides', *Bell Syst. Tech. J.*, **50**, pp. 1791–1816, 1971.

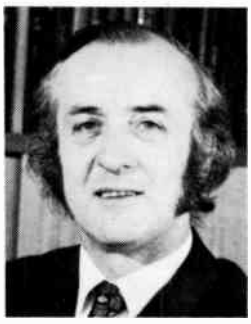
6 Wilson, M. G. F. and Teh, G. A., 'Tapered optical directional coupler', *IEEE Trans. on Microwave Theory and Technique*, **MTT-23**, pp. 85–92, 1975.

7 Milton, A. F. and Burns, W. K., 'Tapered velocity couplers for integrated optics design', *Applied Optics*, **14**, pp. 1207–12, 1975.
8 Marcatili, E. A. J., 'Dielectric rectangular waveguide and directional coupler for integrated optics', *Bell Syst. Tech. J.*, **48**, pp. 2071–2192, 1969.
9 Pitt, C. W., 'Sputtered-glass optical waveguides', *Electronics Letters*, **9**, pp. 401–3, 1973.
10 Goell, J. E., 'Barium silicate films for integrated optical circuits', *Applied Optics*, **12**, pp. 737–43, 1973.
11 Takada, S., Ohnishi, M., Hayakawa, H. and Mikoshiba, N., 'Optical waveguides of single-crystal $LiNbO_3$ film deposited by r.f. sputtering', *Applied Physics Letters*, **24**, pp. 490–1, 1974.
12 Hammer, J. M., Chanin, D. J. and Duffy, M. T., 'Fast electro-optic waveguide deflector modulator', *Applied Physics Letters*, **23**, pp. 176–7, 1973.
13 Westwood, W. D. and Ingrey, S. J., 'Sputtered ZnO optical waveguides', *Wave Electronics*, **1**, pp. 139–51, 1975.
14 Giallorenzi, T. G., West, E. J., Kirk, R., Ginther, R. and Andrews, R. A., 'Optical waveguides formed by thermal migration of ions in glass', *Applied Optics*, **12**, pp. 1240–5, 1973.
15 Stewart, G., Millar, C. A., Laybourn, P. J. R., Wilkinson, C. D. W. and De La Rue, R. M., 'Planar optical waveguides formed by silver ion migration in glass', *IEEE J. Quantum Electronics*, **QE-13**, pp. 192–200, 1977.
16 Stewart, G. and Laybourn, P. J. R., 'Fabrication of ion-exchanged optical waveguides for dilute silver nitrate melts', *IEEE J. Quantum Electronics*, **QE-14**, pp. 930–4, 1978.
17 Hutchins, R. H. and Millar, C. A., 'Manufacturing tolerances for silver-sodium ion-exchanged planar optical waveguides', *J. Phys. D: Applied Physics*, **11**, pp. 1567–76, 1978.
18 Esdaile, R. J., 'Closed tube control of out-diffusion during fabrication of optical waveguides in $LiNbO_3$', *Applied Physics Letters*, **33**, pp. 733–4, 1975.
19 Tsang, W. T. and Wang, S., 'Simultaneous exposure and development technique for making gratings on positive photoresist', *Applied Physics Letters*, **25**, pp. 415–8, 1974.
20 Yi Yan, A., 'Frequency selective grating filters for integrated optics', Ph.D. Thesis, University of Glasgow 1978.
21 Sasaki, H., and Anderson, I., 'Theoretical and experimental studies on active Y-junctions in optical waveguides', *IEEE J. Quantum Electronics*, **QE-14**, pp. 883–92, 1978.
22 Sasaki, H. and De La Rue, R. M., 'Electro-optic Y-junction modulator/switch', *Electronics Letters*, **12**, pp. 495–6, 1976.
23 Sasaki, H., 'Efficient intensity modulation in a Ti-diffused $LiNbO_3$ branched optical waveguide device', *Electronics Letters*, **13**, pp. 693–4, 1977.
24 Papuchon, M., *et al.*, 'Electrically switched optical directional couplers: Cobra', *Applied Physics Letters*, **27**, pp. 289–91, 1975.
25 Schmidt, R. V. and Kogelink, H., 'Electro-optically switched coupler with stepped $\Delta\beta$ reversal using Ti-diffused $LiNbO_3$ waveguides', *Applied Physics Letters*, **28**, pp. 503–6, 1976.
26 Schmidt, R. V., Kaminow, I. P. and Carruthers, J. P., 'Acousto-optic diffraction of guided optical waves in $LiNbO_3$', *Applied Physics Letters*, **23**, pp. 417–9, 1973.
27 Kuhn, L., Dakss, M. L., Heidrich, P. F. and Scott, B. A., 'Deflection of an optical guided wave by a surface acoustic wave', *Applied Physics Letters*, **17**, pp. 265–7, 1970.
28 Tsai, C. S., Nguyen, L. T., Yao, S. K. and Alhaider, M. A., 'High performance acousto-optic guided-light beam device using two tilting surface acoustic waves', *Applied Physics Letters*, **26**, pp. 140–3, 1975.
29 De La Rue, R. M., McK. Trantor, G. S. and Wilkinson, C. D. W., 'Light beam deflector using surface acoustic wave-guided optical wave interaction', Proceedings of the International Meeting on Electron Devices, Washington D.C., pp. 617–9, 1975.
30 Millar, C. A. and Laybourn, P. J. R., 'Coupling of integrated optical circuits using sandwich ribbon fibres', *Optics Communications*, **18**, pp. 80–2, 1976.
31 Tien, P. K., Martin, R. J. and Smolinsky, G., 'Formation of light-guiding interconnections in an integrated optical circuit by composite tapered-film coupling', *Applied Optics*, **12**, pp. 1909–16, 1973.
32 Tsang, W. T. and Wang, S., 'Mode properties of GaAs–$Ga_{1-x}Al_xAs$ heterostructure inverted-ridge optical waveguides', *Applied Physics Letters*, **28**, pp. 665–6, 1976.
33 Leonberger, F. J., Donelly, J. P. and Bozler, C. O., 'Low-loss GaAs $p^+n^-n^+$ three-dimensional optical waveguides', *Applied Physics Letters*, **28**, pp. 616–9, 1976.
34 Garmire, E., Stoll, H. and Yariv, A., 'Optical waveguiding in proton-implanted GaAs', *Applied Physics Letters*, **21**, pp. 87–8, 1972.
35 Westbrook, L. D., Robson, P. N. and Majerfeld, A., 'Strain-induced optical waveguiding in GaAs epitaxial layers at 1·15 μm', *Electronics Letters*, **15**, pp. 99–100, 1979.
36 Cheo, P. K., 'Electro-optic properties of reverse-biased GaAs epitaxial thin films at 10·6 μm', *Applied Physics Letters*, **23**, 439–41, 1973.
37 Kawaguchi, H., 'GaAs rib-waveguide directional-coupler switch with Schottky barriers', *Electronics Letters*, **14**, pp. 377–8, 1978.
38 Merz, J. L. and Cho, A. Y., 'Low-loss $Al_xGa_{1-x}As$ waveguides grown by molecular beam epitaxy', *Applied Physics Letters*, **28**, pp. 456–8, 1976.
39 Aiki, K., Nakamura, M. and Umeda, J., 'Frequency multiplexing light source with monolithically integrated distributed—feedback lasers', *Applied Physics Letters*, **28**, pp. 506–8, 1976.

Manuscript received by the Institution on 13th May 1981.
(Paper No. 2003/CC 344)

The Authors

Peter Laybourn took a B.A. degree in mechanical and electrical sciences at Clare College, Cambridge, in 1963, and a Ph.D. in microwave harmonic generation at Leeds University in 1966. He was a Research Fellow at the University of Southampton from 1966–1971, working on optical fibre communications. Since 1971 he has taught in the Department of Electronics and Electrical Engineering at the University of Glasgow, where he is now a Senior Lecturer and member of the integrated optics research group.



John Lamb has been James Watt Professor of Electrical Engineering at Glasgow University since 1961, and was Vice Principal of the University from 1977–1980. He took an honours degree in electrical engineering at Manchester University in 1943 and was Fairbairn Prizeman; his subsequent higher degrees of M.Sc., Ph.D., and D.Sc. also granted by Manchester. Dr Lamb lectured at Imperial College from 1946 to 1961 where his principal research interests were in ultrasonic properties of liquids and gases, and in viscoelastic properties of liquids. Since 1969 he has led a large research group at Glasgow in integrated optics.

New and Revised British Standards

Copies of British Standards may be obtained from BSI Sales Department, 101 Pentonville Road, London N1 9ND. Non-members should send remittances with orders. Subscribing members will be invoiced and receive 50% discount.

SOFT SOLDER WIRE

BSI has revised BS 441 **Purchasing requirements for flux-cored and solid soft-solder wire** (£3.60) which covers soft solder of circular cross-section, with or without one or more continuous cores of flux. The standard specifies wire dimensions, flux content, inspection procedure, test requirements of solder and flux, and marking requirements. Whereas the 1954 edition contained certain details of flux and solder compositions, the revised version refers the user to the new standard for fluxes (BS 5625) and the recently revised standard for solders (BS 219). The scope of the new document has also been widened to take into account the much larger variety and versatility of flux-cored solders now available. It was considered appropriate to include 'solid' solder wire since this is usually supplied in a form and packaging similar to those associated with flux-cored types.

PERFORMANCE OF RADIO RECEIVERS FOR SOUND BROADCASTING

BS 4054 **Methods of measuring and expressing the performance of radio receivers for sound broadcasting** is intended to standardize the conditions and methods of measurement to be used for determining characteristics of a radio receiver so as to make possible the comparison of results from different observers. The standard was originally published in 1966, and is now revised and issued in a number of parts, based, where appropriate, upon a general set of conditions laid down in Part 1.

The standard constitutes a catalogue of selected measurements, recommended for assessing the essential properties of radio receivers. Neither mandatory nor limiting, the standard provides a choice of measurements for each particular case. The methods of measurement are generally conceived to permit analysing of the overall performance of the receiver, considered as a quadripole (two-port), without endeavouring to study its elements separately.

Part 1 **General conditions** (£17.50) is used, where applicable, in conjunction with other parts, describing general conditions for measurements to be applied to various types of radio receivers, including particularly those for broadcast transmissions. It also describes methods of measurement which are deemed to be applicable to several types of radio receivers and which therefore do not appear in other specialized parts.

Part 2 **Measurements related to the audio-frequency part of the receiver** (£14.50) gives details of the audio-frequency input signal and input arrangements to be used for the measurements, and methods are included for measurements of distortion, sensitivity, acoustic and electrical response characteristics, stereophonic characteristics and other miscellaneous properties.

Part 3 **Radio frequency measurements on receivers for amplitude-modulated emissions** (£17.50) is applicable particularly to receivers of the full-carrier double-sideband type. Methods are given for the measurement of distortion, r.f. selectivity, r.f. sensitivity, the susceptibility to interference entering through ways other than the normal aerial input circuit, and of interference due to internally generated unwanted signals. Part 4 of the standard, in course of preparation, will provide equivalent measurements applying to f.m. receivers.

Part 5 **Measurement on frequency-modulated receivers of the response to impulsive interference** (£6.20) is identical with IEC 315–5 and describes conditions and methods of measurement for evaluating the protection afforded in f.m. receivers against impulsive interference entering through the normal aerial and produced by the ignition systems of engines etc. The impulsive interference is simulated in this measurement by a generator giving short pulses at 100 Hz. The method of measurement applies most particularly to radio receivers designed for f.m. sound broadcasting between 41 MHz and 108 MHz. It is applicable to both mains-operated and to battery-operated receivers.

CONFORMAL COATING MATERIAL

BS 5917 **Conformal coating materials for use on printed circuit assemblies** (£6.20) specifies general requirements and tests for electrical insulating materials suitable for application to printed circuit assemblies as conformal coatings, that is insulating coatings enveloping the board and the components mounted on it. Additionally, it defines the test specimens required and, in an appendix, gives information about the general characteristics of the various resins and recommendations on the procedure of application. The standard does not consider questions of compatibility between coatings and the components of printed circuit assemblies. There is no international standard for this material, but a proposal has been made by the United Kingdom to the International Electrotechnical Commission based on this BS.

RECHARGEABLE SINGLE CELLS

The new British Standard BS 5932 **Sealed nickel–cadmium cylindrical rechargeable single cells** (£2.60) is related to IEC Recommendations 285–1, 285–1A and 285–2. It specifies performance requirements and type tests, closely following the proposals so far established for the revision of IEC Publication 285. The cells specified can be used over a wide range of industrial, and consumer applications, wherever compact and lightweight power is needed. Four of the cell designations are physically interchangeable with four of the round batteries specified in BS 397.

POWER ELECTRICAL EQUIPMENT STANDARDS

New publications prepared for the electrical industry deal with power transformers and potentiometers.

BS 5953 **Guide to power transformers** Part 1 **The application of power transformers** (IEC 606) (£10.40) applies to transformers complying with BS 171 and gives information to assist in the determination and selection of transformer characteristics. The subjects covered include the specification of tapping quantities, selection of winding connections and parallel operation of transformers for three-phase systems, loading capability of the neutral point and calculation of voltage drop (or rise) for a specified load condition.

BS 5981 **D.C. potentiometers** (IEC 523) (£10.40) applies to voltage-measuring instruments for laboratory use in which the voltage to be measured is balanced against a known voltage. The specification includes limits of intrinsic error, permissible variations with influence quantities, marking and symbols