THE PRINCIPLES AND PRACTICE OF STEREO-SYNTHESIS WITH MATHEMATICAL APPLICATIONS

by

T.J. Fletcher, M.A.(Oxon.)

Imperial College of Science and Technology
and
Sir John Cass College

A thesis submitted for the degree of Doctor of Philosophy in the University of London.

March 1962

THE PRINCIPLES AND PRACTICE OF STEREO-SYNTHESIS WITH MATHEMATICAL APPLICATIONS

SUMMARY

This thesis describes techniques for making animated stereoscopic diagrams - a problem which dates from the time of Wheatstone but which does not appear to have been solved satisfactorily until the present day. After a historical survey of the development of stereoscopy and of mathematical film animation the classic theory of stereoscopy is used as a foundation for combining standard methods of stereoscopic drawing and film animation to produce an animated stereoscopic film showing three modes of electro-magnetic oscillation in a resonant cavity.

Static and dynamic experiments determine which aspects of the theory are relevant to the problem of producing animated stereoscopy, and it is seen that the affine distortions inherent in stereoscopy must be considered in practical design, but that the more general projective distortions predicted by the usually accepted theory play a much smaller part than is generally assumed because they are only observed to a limited extent.

The relevance of the Luneburg theory is discussed, and also the symmetry groups which are involved in stereo animation. Indication is given of possible future applications in the documentation of mathematics, the preparation of instructional films, and the investigation of the problems of binocular vision.

"Mirrors can be so arranged that, as often as we wish, any object, either in the house or the street, can be made to appear. Anyone looking at the images formed by the mirrors will see something real, but when he goes to the place where the object seems to be he will find nothing. For the mirrors are so cleverly arranged in relation to the object that the images appear to be in space, formed there by the union of the visible rays. And the spectators will run to the place of the apparitions where they think that the objects actually are, but will find nothing but an illusion of the object."

Perspectiva. ROGER BACON (1214-1294).

"Comme Marey et Muybridge devient le cinéma en même temps aux antipodes de la terre, nous étions quelques uns a devine‹ l'ANIMATION dans divers pays, sans nous connaitre. Nous nous sommes rencontrés et nous nous sommes reconnus frères, parce que nous parlons la même langue: UNIVERSELLE"

ALEXANDRE ALEXEIEFF.

CONTENTS

	Summary	(i)
Chapter		page
1.	Introduction.	1
2.	A short history of stereoscopy.	3
3.	The history of mathematical cartoon films.	9
4.	The convergence theory of stereoscopy.	16
5.	Luneburg's theories and their application to films.	41
6.	Static experiments.	57
	I. Experiment "Photofinish".	
	II. Experiment "Guesswork".	
	III. Experiment "Overlap".	
7.	Methods of stereoscopic drawing.	86
	Appendix. Notes on the accompanying stereo slides.	
8.	Stereoscopic animation.	99
9.	The film "Resonant Cavities".	116
10.	Dynamic experiments.	141
	I. Experiment "Jumper".	
	II. Experiment "Drifter".	
	III. Experiments with McLaren's film "Lines".	
11.	Symmetry groups in stereo-animation.	163
12.	Future applications of stereo-synthesis.	173
	Acknowledgements.	178

Supplementary material

Sets of stereo diagrams by T.J. Fletcher.

3 pairs of stereo paintings by N. McLaren.

1 stereo-viewer.

Animated stereo 16 mm film, "Resonant Cavities".

CHAPTER ONE

INTRODUCTION

This thesis is concerned with stereosynthesis, that is with the production of sensations of depth by artificial means; in particular with the production of stereoscopic diagrams which move. A weaker form of the problem can be found more than a century ago in some work of Wheatstone, who attempted to animate pairs of stereo photographs, but here the aim is to produce both appearance of depth and movement completely artificially, and this does not appear to have been done successfully before, in spite of numerous attacks on parts of the problem over the years. The chapters which follow survey the histories of steroscopy and of animated diagrams, and then the theory of stereoscopy. This is classical ground, but many treatments are concerned with stereo-photography, which is not involved here as the parallaxes on stereo pairs are all drawn, so the theory is presented in a form suited to the present problem.

Experiments are described which indicate the strong and weak points of the classical theory, and it is then applied to the problems of animating a film. The subject is "Resonant Cavities", and it depicts the H101, E111 and H111 modes of oscillation in a rectangular cavity - a topic which cannot be illustrated as well in any other medium. The film was photographed by ordinary "flat" methods and projected using the Bolex 16 mm stereo system. The techniques may be applied to all problems of making line illustrations, in colour and in motion, of three-dimensional situations. Resonant Cavities was first presented in public at a meeting

of the Scientific Film Association in London on May 29th, 1958, and nothing like it seems to have been made anywhere else.

The success of the film raised problems on the perception of stereo flicker, and this is considered in relation to the perception of flicker in other systems of stereo projection and to knowledge of eye movements. Two dynamic experiments are described which, as well as contributing to the present enquiry, show how the present techniques can be used to provide experimental situations to investigate problems of physiological optics. The techniques described thus not only depend upon optical theory, but they also offer a means of extending it.

The Luneburg theory of binocular vision has been applied to the design of stereo projection systems, and the relevance of this theory to the present problem is considered; there is a chapter on the symmetry groups of stereo-animation, which is an extension of an earlier paper by the author discussing symmetry groups in "flat" animation, a point of view which gives a method of classifying cyclically repeating diagrams (which are the mainstay of technical animation) and of constructing them in the most economical manner. To conclude there is a brief discussion of the future applications of the techniques.

Throughout there is frequent use of English units of length instead of metric units. This is because the use of feet and inches is standard in animation practice, and it seemed best to continue with them.

CHAPTER TWO

A SHORT HISTORY OF STEREOSCOPY

"To see in relief is to receive by means of each eye the simultaneous impression of two dissimilar images of the same object". This statement, attributed to Euclid, c.280 B.C. (8), shows that the principle underlying stereoscopy has long been understood, and that from the earliest times it has interested mathematicians. Euclid is also credited (7) with the observation that the two eyes see more of a sphere than one eye alone, and similar remarks were made by Ptolemy (A.D.150), da Vinci, Galen, and Gassendus (6). In 1593 Giovanni Battista della Porta produced pairs of stereoscopic drawings (6,8), but it is not known how, or even if, they were viewed. Similar drawings were made by Jacopo Chimenti of Florence round about 1600, and these eventually went to the Wicar Museum in Lille (7a).

Modern work on stereoscopy began with Wheatstone in 1832. As well as preparing pairs of stereoscopic drawings (14) and making the first successful device for viewing them Wheatstone collaborated with Fox-Talbot and made stereo-Talbotyes in 1839 (6). Brewster invented an improved form of steroscope in 1844 and made stereo-Daguerrotypes in 1849. Moser took stereo photographs at Konigsberg in 1841, using two cameras (15) and stereoscopic cameras were commercially available in England by 1857.

About this time other stereoscopic devices were introduced depending on a different principle. Helmholtz had suggested viewing stereo drawings by using complementary colours, and the first anaglyphs were

mathematical diagrams made by Dove and Rollmann in 1853 (6,7). Their pictures were projected in blue and yellow and viewed through spectacles of red and green. D'Almeida projected pictures which were red and green and were viewed through spectacles of the same colours in 1858 (6). The familiar type of anaglyph, printed in red and green, was introduced in 1891 by Ducas du Hauron, and it was only at this time that the name "anaglyph" was first used.

I have described elsewhere (4) the use made of anaglyphs in the teaching of geometry by Richard (13), Perregaux and Weber (9), and Kohler, Graf and Calov (7). This tradition continues in Germany and there are textbooks of geometry incorporating anaglyphs in use there at the present time.

Concurrently with the early experiments in stereoscopy methods of providing an illusion of movement were also being studied (see the following chapter), so it was only natural that an attempt should be made to combine the two techniques, so producing stereoscopic pictures which moved. Wheatstone himself attempted this in a device which may be seen, together with many other pieces of his equipment, in the Science Museum in London. He mounted fourteen pairs of stereo photographs of a small steam engine, photographed in different positions, inside a drum. An epicyclic movement brought the pictures under the lenses of a Brewster stereoscope, and in this way he hoped to obtain a moving picture. The principle, of course, is sound; but the apparatus had no shutter and no provision for registering the pictures accurately and it cannot be said to work satisfactorily. It should be observed that this apparatus is

concerned with synthesising movement artificially since it employs still pictures of the steam engine taken in different positions. The cinematograph camera was not invented until some decades later.

It is difficult to determine the date of the first successful combination of stereoscopy and animation by the persistence of vision. Dubosq experimented with combinations of stereoscope and zoetrope, perhaps as early as 1851 (12). Between 1853 and 1860 there were six patent applications in Britain for inventions to give simultaneous perception of solidity and motion; but a patent application gives no guarantee that the idea has actually been tried, and still less that it has been successfully carried out. Sellers placed stereo photographs in a zoetrope in the U.S.A. in 1861 (5) and made a movie-stereoscope using photographs stacked one behind the other, as in a flicker-book, in 1863. In 1865 Claudet described a combination of stereoscope and phenakistoscope by means of photography (2). Similar work was done by Desvignes (1860), Czugafewicz (1860), Shaw (1856-61), Dumont (1861), Cook and Bonelli (1867) and Hubert de Mollard (1867) (10,12). There are no reports that any of these early experiments produced drawn stereoscopic diagrams which moved, i.e. that they synthesised both movement and depth.

The early history of the cinema has been confused by conflicting claims of priority in which national prestige has played a considerable part; but leaving aside all questions of priority over Edison and Lumiere it is not generally realised that the historic pictures taken at Hyde Park Corner by Friese-Green were taken on a stereo camera. The date of this event is disputed by historians, being placed between 1889 and

1893, and it is likely that some of the dates attached to exhibits in the Science Museum are too early; but it should be emphasised that whilst he took stereo pictures there is no evidence that Friese-Greene had any way of projecting them stereoscopically. Also, to put the event into correct historical perspective, the camera used was of a type which had been invented earlier by Varley and its speed was only three or four frames per second. It was a rapid sequence camera rather than a cinematograph camera.

There are patent specifications for stereoscopic projection systems, one anaglyph and one using polarised light produced by reflection from mirrors, by Grivolas in 1897. Once again, whilst these are sound in principle there is no evidence that satisfactory prototypes were every constructed; and the claim has been made (8) in America that the first successful moving anaglyphs to be projected were produced by Leventhal in 1924.

All the systems of stereoscopic cinematographic projection which are in current use in the Western World derive from the patent for sheet polarizer by Land in 1928. Land projected coloured stereoscopic pictures in 1933. Descriptions of the various present day systems can be found in the references (3, 11, 17), together with accounts of many other devices which have been tried and abandoned.

Commercial developments have centred mainly around 35 mm systems, and less has been done with 16 mm systems. The smaller gauge is standard in educational and academic work, and the only commercially available system with any claims to standardisation is the system developed by

Bolex in Switzerland. This system has been used for the work described in this thesis, and we return to it in Chapter 8.

REFERENCES

1. D'Almeida. Nouvel appareil stereoscopique. Comptes rendus xlv,ii,61. 1858.
2. Claudet. Combination of stereoscope and phenakistoscope by means of photography. Brit.Ass.Report, 1865, Pt.ii,9.
   See also, Laing, Mech.Mag.(2),xiii,190.
3. H. Dewhurst. Auto Precision Stereoscopy. Phot.Journ. Vol.92B.p.2,1952.
   do. Bolex Reporter, 3,3,22-25, Summer 1953.
4. T.J. Fletcher. Solid Pictures. Mathematics Teaching, 2, pp.17-20, May 1956.
5. C. Francis Jenkins. Animated Pictures, Washington, 1893.
6. Judge. Stereoscopic Photography, 3rd edn.
7. Kohler, Graf and Calov. Mathematische Raumbilder, Dresden, 1941.
7a. E.F. Linssen. Stereo-Photography in Practice. London, 1952.
8. Norling. J.S.M.P.T.E., March, 1953.
9. C. Perregaux and A. Weber. Le relief en géometrie. Magron, Bienne, 1914.
10. G. Potonniée. Histoire de la decouverte de la photographie. 1925.
    do. Les origines du cinematographe. 1928.
11. M. Quigley. New Screen Techniques. New York, 1953.
12. G. Sadoul. Histoire générale du cinema. 6 vols. Denël, Paris, 1948.
13. H. Vuibert. Les anaglyphes géometriques. Vuibert, Paris, 1912.
14. C. Wheatstone. Collected works.
15. ------- Bild und Ton, 1954.
16. ------- Polaroid Corporation Report.
17. R. and N. Spottiswoode. The Theory of Stereoscopic Transmission. Univ. of California Press, 1953.

CHAPTER THREE

THE HISTORY OF MATHEMATICAL CARTOON FILMS

Early animation. The Science Museum at South Kensington contains many nineteenth century optical toys which produce an illusion of movement by using the persistence of vision. These go back to the stroboscope which was invented independently by Stampfer and Plateau in 1832. This was immediately applied to the presentation of diagrams of moving mechanisms, and these are certainly the first animated technical diagrams. By coincidence the discovery took place in the very year in which Wheatstone began his researches in stereoscopy. Plateau produced the phenakistiscope in 1833; the familiar zoetrope was invented in 1834 by Horner* (15a).

There are references to even earlier work, Foster (8) describing some diagram illusions goving back to an example due to Roget in 1825. Roget also made an analysis of an interference phenomenon (15) which is probably the first piece of mathematics of a type which is not infrequently involved in film animation. Faraday joined in correspondence on a similar topic. These problems lead ultimately to the theory of the symmetry groups of animation discussed in Chapter 11.

The first animated photographs - sets of photographs showing some object in a set of positions from which a moving picture could, in principle, be synthesised - were made in 1851-2 by Dubosq, Claudet, Wheatstone and Seguin (15a). (This process is currently being revived by avant-garde film makers seeking new artistic effects.)

*Footnote. Horner (1786-1837), an English mathematician, is famous as the inventor of "Horner's method" in the theory of polynomials.

Animated drawings were first projected on to a screen by Uchatius in 1853, and animated photographs were projected by Heyl (U.S.A. in 1870, but he may have been anticipated elsewhere.

The first great master of animated drawing was Reynaud (10,12) whose praxinoscope of 1877 can be seen in the Science Museum in London. It gives a moving cartoon of excellent quality. For many years in the 1880s and 1890s Reynaud's Theatre Optique in Paris displayed an extensive series of animations, until the performance was eventually superseded by the cinema. Reynaud's animations were cartoon films in all but the actual technique employed. He drew his work directly on perforated band film (14,15a), and between 1896 and 1900 he actually used film photographs to assist in the design of his animations, but he never photographed his drawings on film before projecting them. (In the use of the camera to study movement he was following the example of Mueybridge, and in animating from such studies he was anticipating Disney.) In 1903 Reynaud experimented with sets of stereo-photographs mounted in an individual viewing device like a 'mutoscope'; and it is also of interest that his method of drawing direct on film has been revived in recent years by McLaren, whose synthetic work we discuss later.

Many historians have given the distinction of being the first animator who photographed on film to Emile Cohl, but it seems that he must yield priority to Stuart Blackton, an Englishman who worked in America, and who produced Humorous Phases of Funny Faces in 1906 and The Magic Fountain Pen in 1907 (2,10).

The first film by Emile Cohl (a pseudonym for E. Courtet) was shown

in 1908, but was probably made in 1907; and whilst Cohl appears to have seen The Magic Fountain Pen we cannot say whether this was before or after starting his own work. Between 1908 and 1910 Cohl is said to have made 60 films incorporating all kinds of animation techniques. He went to America in 1912 to find many similar films being made there.

Mathematical animation. Whilst all of this work was in the sphere of entertainment it should be remembered that many of the earliest workers on cinematography were principally concerned with the scientific aspects of the subject, and looked on the cine-camera as a means of analysing movement or making scientific records. Münch, of Darmstadt was in this scholarly tradition. In 1908 he saw Blackton's Humorous Phases of Funny Faces and was immediately impressed with the educational possibilities of the technique. By 1912 Münch was able to demonstrate eight mathematical films to a meeting of teachers in Halle, and by 1914 his work ran to no less than 44 separate titles (7). This seems to be a greater quantity of work than anyone has produced since - with the possible exception of Nicolet, in Switzerland - but he was only concerned with depicting certain situations, such as families of coaxal circles, and not in making long films with developing arguments, requiring many different sets of drawings, or a kind which were developed later. Münch's films included much work at an advanced level as well as simpler topics. Film projectors were not then as commonplace in academic circles as they are now, and this is perhaps the reason why his work did not make a bigger impact. It would be wrong to see it merely as a novelty; it was a natural expression of an attitude to geometry and to the problems of teaching it

which we find in such outstanding men as Klein (11) and Hilbert (9), which has never been much in fashion in Britain, but which continues still in Germany in the films on non-Euclidean geometry made recently by the Humboldt University in East Berlin under the direction of Schröder (22).

Before the first World War Detless (1) also made films on similar subjects to Münch. Work recommenced after the war and the professional firms of Deulig and Lichtverlag each made at least three films before 1922, some of which to judge by the scenarios were of outstanding interest. The vigour of the German school of mathematical film making at this time may be judged by a contemporary article by Ebner (3), which as a statement of the pedagogical principles underlying this type of film is still unsurpassed.

This work had very little influence in Britain, and it is doubtful if any of it was ever even shown here; but British developments took place between 1936 and 1939 when four films were made by Robert Fairthorne and Brian Salt. The great contribution which these films made to the art was the portrayal of a continuously developing logical argument. For the first time the conventions of the cinema were given a significance as mathematical symbols; and the film entitled A Hypocyclic Motion in particular showed that a film could not merely realise the dependance of geometrical configurations on a parameter but also conduct a logical, but non-verbal argument in terms of its own particular symbolism.

The chief exponents of the animated technical diagram in France have been Cantagrel (13), as director, and Motard, as animation artist. With the exception of three teaching films for schools made recently they

have done little pure mathematics, their work being mainly in the fields of technology and applied science.

Nicolet has made a number of short, simple films for teaching school geometry in Switzerland, and films of a similar kind have been made in Yugo-Slavia and Belgium. At an advanced level there is the film Gruppen und Phasengesohwindigkeit (18) produced in collaboration by the Physical Institute, Hanover and the Institut fur Film und Bild in Gottingen, as well as a considerable amount of experimental material from the Deutsche Zentralinstitut fur Lehrmittel in East Berlin, including the advanced series on non-Euclidean geometry from the Humboldt University mentioned above.

Mathematical films have been made in America but they have been almost entirely commercial films for elementary teaching whose academic impact has been slight.

The present author has made five films at University level (20) in addition to the three-dimensional work described in this thesis.

REFERENCES

1. Detless. Die Veranschaulichung von veränderlichen Figuren in Unterricht. Unterrichtsblatter f. Math. u. Naturw., 1913, p.39,p.121.

2. Lo Duca. Le Dessin Animé. Paris, 1948.

3. M. Ebner. Mathematische Lehrfilms. Zeits. f. Math. u. Naturw. Unterricht, 53, 1922, p.193. (This contains references to many other early articles.)

4. T.J. Fletcher. Mathematical Documentaries. Science and Film, June,1954.

5. do. Un nouveau langage mathématique. Mathematica et Paedagogia, No.4, 1954-5.

6. do. Les problèmes du film mathématique. Le matérial pour l'enseignement des mathématiques, Neuchatel, 1958.

7. do. The First Mathematical Films. University Film Journal, 17-18, Dec. 1959, 11-15.

8. R.B. Foster. Living Pictures. 2nd edn. 1915.

9. D. Hilbert and S. Cohn-Vossen. Anschauliche Geometrie, Berlin, 1932.

10. Jeanne and Ford. Histoire Encyclopedique du Cinéma, Vol.I, 133-139, Robert Laffont, 1947.

11. F. Klein. Elementary Mathematics from an Advanced Standpoint, New York, 1939. (First published, Germany, 1908.)

12. M. Noverre. Emile Reynaud, sa vie et ses travaux, Brest, 1926.

13. Poncet. Dessin Animé, Art Mondial, p.186, Paris, 1956.

14. M. Quigley. Magic Shadows, Georgetown, U.P., 1948.

15. P.M. Roget. Optical deception in the experience of the spokes of a wheel seen through vertical apertures. Phil. Trans. 1825.

15a. G. Sadoul. Histoire Générale du Cinéma, 6 vols., Denoel, Paris, 1948.

16. Uchatius. Ann. de l'Acad. de Vienne, X, 482, 1853.

17. ------- Zeits. f. Math. u. Naturw. Unterricht, 43, 440, 1912.

FILMS

18. David and Bekow. Gruppen und Phasengeschwindigkeit. Institut für Film und Bild, Gottingen.

19. R.A. Fairthorne and Brian Salt. $\ddot{x} + x = 0$

$\ddot{x} + x = A \sin nt$

A Hypocyclic Motion

Generation of Involute Gear-teeth.

20. T.J. Fletcher. Plucked Strings

The Simson Line

The Cardioid

Four-Point Conics. Sir John Cass College, London.

Four-Line Conics. National Film Board of Canada.

21. Münch. At the time of writing negotiations are in progress for the transfer of some of Münch's surviving work to the National Film Archive, London.

22. K.Schröder. Nichteuklidische Geometrie. 5 parts. DZL, Berlin.

CHAPTER FOUR

THE CONVERGENCE THEORY OF STEREOSCOPY

Depth perception. Stereopsis, with the restricted meaning of the reception of geometrically different images by the two eyes, is only one of the factors involved in depth perception. Depth can be inferred by blocking, foreshortening, change of aspect on movement, aerial perspective (effects of atmospheric conditions, etc.), the relative apparent sizes of familiar physical objects, and differences in accommodation as well as by differences in convergence. It has been suggested even that such factors as the chromatic halo on the retina play a part in depth perception and the list can be extended to a dozen items or more. Gibson (3) attaches great importance to gradations in the texture of surfaces and minimises the effect of stereopsis. In his concern for such problems as flying aircraft he is undoubtedly right; but these things are very different from our present task.

This thesis is concerned with stereosynthesis, which means, in its narrow sense, the use of a pair of diagrams with varying parallaxes to produce an illusion of depth, and the preoccupation with parallaxes is not intended in any way to minimise the importance of these other factors in depth perception, and where appropriate they should certainly be incorporated into the techniques of stereo-illustration. The first three factors listed above are used in the illustrations accompanying this thesis. Aerial perspective can be used in still illustration, but depicting surface texture gradient is a formidable task for the most accomplished artist; and the use of either technique on the moving

elements of animated diagrams would be virtually impossible.

Depth perception also involves psychological factors. The use of the word "psychological" and the related word "rationalisation", which is the psychological process most frequently invoked, can easily be a cloak for ignorance, but in the present context the complex of ideas to which these words point cannot be ignored. The many clues to depth which are normally present are correlated by the observer (usually unconsciously) and he reaches a decision. The ability to correlate depends on his previous conditioning, including conscious and unconscious memory, so it is quite possible for different people to interpret the same clues in different ways, or for the same man to interpret the same clues differently at different times.

When clues point in different directions it is hard to synthesise the information, and the observer can either remain in doubt or reach a decision by ignoring some of the clues which contradict and reconciling the rest. It is quite possible for two interpretations of a situation to alternate in the observer's mind. Nearly all "false" depth perception is "true" in its own more limited context, in the sense that an incomplete set of clues is reconciled consistently. There is also a pronounced tendency for this reconciliation to be as much in sympathy with previous experience as possible. This is so well-known in normal vision that examples are not needed, but two illustrations from less familiar fields may be to the point.

Charnwood (1) describes various stereo situations in which quite different pictures are presented to the two eyes. In the first place

observers synthesise these very unnatural stimuli giving them a meaning of some kind in terms of familiar things (as distinct from suspending judgement); in the second place the meaning given varies from person to person, and (Charnwood observes) when the person is previously committed to some particular theory of binocular vision the interpretation which he arrives at is strangely in sympathy with the theory to which he subscribes. This situation cannot be discussed in terms of physical optics alone.

Another more remarkable example arises when viewing with the image fixed on the retina. It is known that under such conditions the image disappears and reappears spasmodically. The conscious will cannot affect this at all, but it has been demonstrated (6) that recognisable shapes remain visible longer than random, nonsensical ones. Depth perception is not involved here, but the point is that familiarity with shapes leads to a different psycho-physiological reaction - "perceptual elements (as distinct from sensory elements) exist in their own right" - and the old pattern of subconscious mental factors is involved even in these highly unnatural conditions in which the mind has not worked before.

Historical notes. From the early beginnings of the subject in the Sixteenth Century until the completely new ideas introduced in the recent work by Luneburg the theories of perspective and of stereoscopic drawing have all been in one unified tradition. A report of the Air Survey Committee (11) gives ten references on developments in perspective geometry before 1700, beginning with the Institutiones Geometricae written in 1532 by the artist Durer.

Once the principles of perspective are understood it is an easy matter to apply them to the production of drawings of the same object seen from slightly different points of view and so to produce stereo pairs. We have already noted work of this kind by Sixteenth Century artists.

The Eighteenth Century was a period of comparative inactivity, at least as far as the introduction of fresh ideas is concerned, but early in the Nineteenth Century after the labours of Carnot, Poncelet and Monge projective geometry developed into an independent, highly specialised branch of pure mathematics, and most treatments of the theory of stereoscopy make little or no use of its concepts.

There are already extensive modern treatments of the theory, but they are nearly all developed from the point of view of the technician who wishes to take stereoscopic pictures with a camera and later project the results in such a way as to ensure as close a reproduction of the original scene as possible. They are in fact theories of stereo-transmission (9). Excellent as these treatments may be for their own purposes the notations employed are often needlessly cumbersome and overloaded with suffixes and lack the clarity of, say, the classic writings of Helmholtz (4). For the present enquiry we need not so much a theory of stereo-transmission as a theory of stereo-synthesis, so we expound the classical convergence theory of stereoscopy in a notation suited to our purposes. Having done this we will discuss further some of the experimental objections to it, and a later chapter will describe some of the more recent theories, due initially to Luneburg, which seek to

provide a more satisfactory theoretical basis for the phenomena which are actually observed.

The convergence theory. The convergence theory of stereoscopy proceeds from a very simple initial assumption. When a pair of stereo pictures is presented to the two eyes of an observer the theory assumes that he locates a particular point in the scene presented at the point in space where the two lines of sight intersect. The assumption is simply that the two eyes act like a pair of range-finders.

There is considerable evidence against this plausible, but naive, view. There is no evidence that observers can estimate the distance of a point source of light in completely dark surroundings with any accuracy, and on the "range-finder" view they should be able to. The relative distances of two points can, however, be assessed with great accuracy. Also it is common experience that two pictures drawn on convergence principles, or two photographs, may have their relative horizontal positions varied by an amount which, on the convergence theory, is very considerable, and the distortions which the theory predicts are not observed. These observations suggest that it is largely differences in parallax (or in parallactic angles) which cause sensations of depth; i.e. that stereoscopic effects depend on the disparity of the two retinal images. As a descriptive account the retinal disparity theory is probably more accurate than the simple convergence theory, although it does seem that convergence plays some part in depth perception (10).

However the practical implications of the two theories are almost identical in so far as stereosynthesis is involved. The convergence

theory is a completely quantitative theory which gives sufficient (although more than strictly necessary) conditions for the production of undistorted pictures. It is not easy to find a precise statement of the retinal disparity theory, but it seems fair to say that it is not an independent quantitative theory at all, and that in so far as it is quantitative it merely borrows from the convergence theory; but it adds that certain first order distortions which should be observable on the convergence theory are not observed. In other words, in terms of stereosynthesis on the drawing board it gives the draughtsman the reassurance that if convergence principles are followed a certain amount of lack of horizontal registration in the projection system will not be a serious defect. (We cannot be any more quantitative than this because the retinal disparity theory is not quantitative on this vital point.)

For this reason in this thesis we tend to regard the implications of the two theories as practically speaking equivalent. In other investigations, for example in the design of an electro-mechanical model of the human perceptive system, the implications of the two theories are quite different, and they would obviously call for quite separate treatment.

On the convergence theory if we wish to synthesise a solid geometrical figure we must prepare two drawings of it based on two centres of perspective in the same horizontal line and a distance apart equal to that of the observer's eyes, and arrange some viewing device of lenses or filters whereby each eye sees its own drawing from the correct point of view.

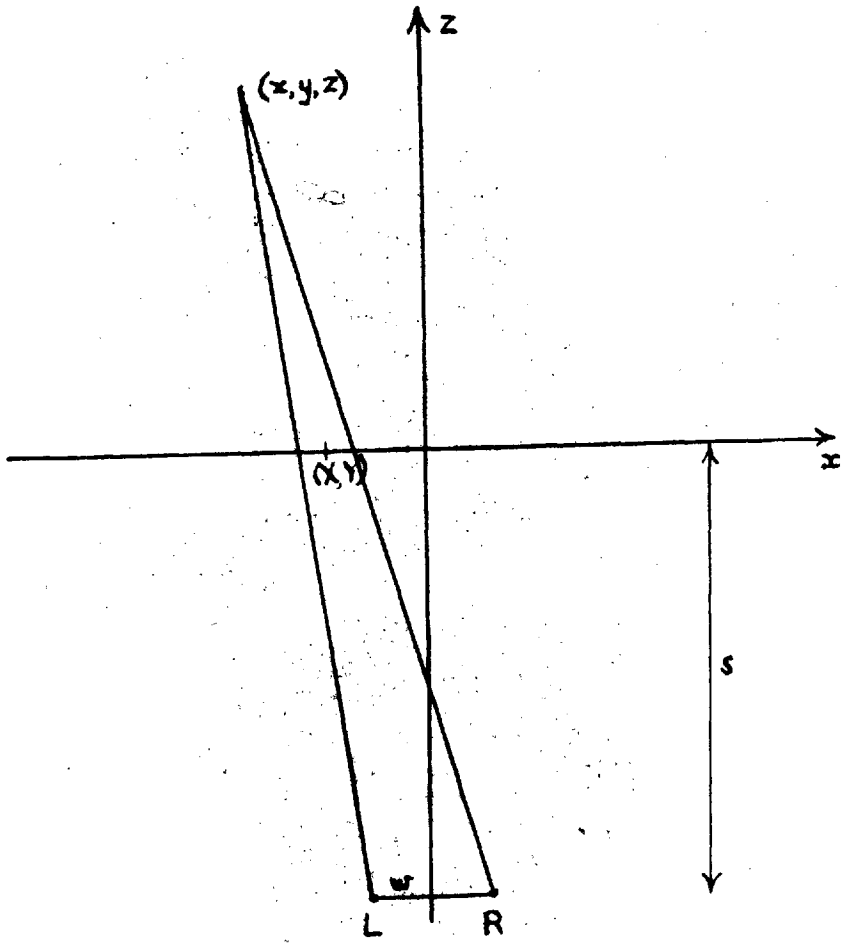


Fig. 4.1

The geometrical notation in Fig. 4.1 will be used throughout. The observer faces a screen. The point mid-way between his eyes, which are a distance 2w apart, is on the normal through the centre of the screen, which we take as origin. The axes Ox and Oy are in the screen, directed to the observer's right and vertically upwards, respectively.

The Oz axis is normal to the screen and directed away from the observer. These axes are the natural ones to take for geometrical animation, where the task is to add the z-dimension to the two-dimensional figures in the xy-plane.

If the distance from the screen to the observer is s then the observer's eyes, at R and L, are located at the points $(\pm w,0,s)$. If we have an object point with coordinates (x,y,z) elementary geometry shows that the two rays from the point to the observer's eyes cut the screen at two points on the same horizontal level and at a distance $2wz/(s+z)$ apart. Two points such as this are sometimes termed homologous points. Their distance apart we will term the parallax, and this word will always be used in this sense. That is to say a parallax is the distance between two projections of an object-point on to a frontal plane, projecting from the nodal points of the two eyes. It is convenient to call the coordinates of the point midway between the two projections (X,Y) and to denote the parallax by 2Z. Then the object-point (x,y,z) is projected to the two points $(X\pm Z,Y)$, for the right and left eyes respectively, where

$$X = \frac{sx}{s+z} \qquad Y = \frac{sy}{s+z} \qquad Z = \frac{wz}{s+z} \qquad (4.1)$$

which means that

$$x = \frac{wX}{w-Z} \qquad y = \frac{wY}{w-Z} \qquad z = \frac{sZ}{w-Z} \qquad (4.2)$$

These equations can be normalised by writing x, y and z for x/s, y/s and z/s; and X, Y and Z for X/s, Y/s and Z/w. Then they take the form

$$X = \frac{x}{1+z} \qquad Y = \frac{y}{1+z} \qquad Z = \frac{z}{1+z} \qquad (4.3)$$

and

$$x = \frac{X}{1-Z} \qquad y = \frac{Y}{1-Z} \qquad z = \frac{Z}{1-Z} \qquad (4.4)$$

Normalisation in this way amounts merely to measuring in the most convenient units, and it can simplify some theoretical discussions. The equations connecting the parallax with the z coordinate can be written in the form

$$\frac{1}{Z} - \frac{1}{z} = 1 \qquad (4.5)$$

a relation of a form which is extremely common in optical theory.*

The set of points (X+Z,Y) define the picture which must be drawn for the right eye and the set (X-Z,Y) the picture for the left; but for some purposes it is convenient to speak of the (X,Y,Z) space and to regard it as a projective transformation of the (x,y,z) space. The transformation connecting them is given by the equations 4.1 and 4.2; and is, of course, a particular type of bilinear transformation the general study of which is the science of projective geometry as it is known to the pure mathematician.

It is important to note that the capital letters refer to the diagrams as they are projected on the screen. These are not the same as the drawings prepared on the drawing board because they are on a different scale. Whilst it is possible to introduce a further set of coordinates to describe the situation on the drawing board this is not usually necessary.

*FOOTNOTE. Grubb's construction, a geometrical method of realising this algebraic relation, is, of course, a perspectivity.

Indeed, it is a useful economy in notation not to do so. When developing our present techniques we had in mind a final cinematographic projection under certain specific conditions, with a screen measuring 38" by 50" 10 feet away. 7.6" by 10" is a convenient size for drawings, and a technique was devised which was based on this ratio of 5:1 between the linear dimensions of the picture on the screen and the diagram on the drawing board. All work on the drawing board can then be done assuming simply that we are preparing drawings for an observer with eyes $\frac{1}{2}$" apart looking at a screen 2 feet away.

Distortions. We will next consider the distortions which arise within the framework of the theory as it has been developed so far. These can be discussed analytically by altering the appropriate parameters in equations 4.1 and 4.2 and seeing what values (x',y',z') arise instead of the correct values (x,y,z). The space (x',y',z') is then to be regarded as the distortion of the space (x,y,z).

Only one observer in the audience can be in the favoured position which has been used in deriving the basic equations, so we first consider the effect of not sitting in this special position. If the observer sits at a distance s' from the screen instead of s, then s' must replace s in equation (2iii) and he will synthesize a space in which

$$z' = \frac{s'\mathfrak{Z}}{w-\mathfrak{Z}} = \frac{s'}{s} z .$$

This means that all lengths in the z direction are increased in the ratio s'/s, or that the distortion is a simple stretch in depth. This effect is actually observed, and it is a pronounced feature of all stereoscopic systems.

Whilst there is no doubt that this distortion occurs in practice it is much less inconvenient than one might at first think. It is not regarded as a serious drawback in the commercial cinema where the back seats are many times as far away from the screen as the front seats, and even in scientific work where metrical relationships are much more important observers are not unduly worried by this phenomenon. Present day viewers are accustomed to make automatic, unconscious corrections in depth perception at the theatre, the cinema or when watching television and they carry over these habits when watching stereo diagrams.

In spite of this high degree of tolerance on the part of observers this inevitable depth distortion which is inherent in any known stereo system must be borne in mind when designing stereo films, and the chief practical implications are two. In the first place one should not worry unduly about some possible faults in the design of a film because it may be possible to correct them later merely by adjusting the distance from which the observer views. Secondly, it is quite impossible to present correct metrical information to everybody in the audience, so it is a vain hope to try. The designer must concentrate on other spatial features which stereo projection does not distort and see that the fullest possible use is made of them. Where it is essential to present metrical information as accurately as possible the position of members of the audience must be correspondingly restricted.

Just as not all members of an audience can sit at the correct distance from the screen so not all of them can sit directly opposite the centre of it. The consequences of this can be found by replacing X and Y

in equation (2) by X + p and Y + q. This considers the effect of the centre of the screen having x and y coordinates (p,q) relative to the observer instead of (0,0). Then

$$x' = \frac{w(X + p)}{w - Z} = \frac{w[xs + p(z + s)]}{w(z + s) - wz}$$

so $$x' = x + p + \frac{p}{s} z$$

and $$y' = y + q + \frac{q}{s} z$$

$$z' = z$$

This distortion is a shear. Its effect is to prevent an observer from 'seeing round corners' by moving his position in the theatre. Common-sense suggests that it should not be possible to see more in a pair of stereo drawings than they actually contain, and that any attempt to 'see round a corner' must be unsuccessful. This shear is the mathematical expression of this intuitive suggestion. The shear can actually be observed by moving relative to the screen.

These two distortions are both <u>affine</u>. Affine transformations preserve linearity, parallels and metrical ratios which are measured in the same direction. In general they destroy angles and metrical relationships in differing directions.

To sum up, even if the figure is perfectly drawn as an observer moves round the theatre the figure undergoes affine distortions. It is the nature of stereofilms to depict affine geometry, and if we wish them to display metrical geometry we can, in principle, only hope them to do so at one particular place in the theatre.

The next distortions to consider are more serious, and are associated with errors in making the film, or in the optical system; and they cannot usually be altered once the film is complete. There are, however, compensations, as it will be shown that variations in some of the parameters may be used to counterbalance variations in others. Certain writers would appear to deny this, but the special contexts in which such comments are made are important and what is true for one system of projection may not be true for another in which the same parameters are not variable.

In the present system no complications arise from having to photograph with two cameras and none of the parameters of the camera are involved in our calculations. We assume that the camera and projector photograph a flat drawing and project a magnified version of it on to a screen without distortion. Whilst admitting that there may be some degree of distortion in these processes we assume that it is negligible compared with the other, truly stereoscopic, distortions, which can occur at various stages.

Previously the observer moved and the projection system was unaltered, so we employed a system of coordinates with the origin on the screen. As we go on to consider distortions which arise through variations of the parameters in the optical system whilst the observer remains stationary it is convenient to take a system of coordinates with the origin centred on the observer. If we put $\zeta = z + s$ the coordinate system (x,y,ζ) is a Cartesian system with the origin mid-way between the observer's eyes. Equations 4.1 and 4.2 become

$$X = \frac{sx}{\zeta} \qquad Y = \frac{sy}{\zeta} \qquad Z = \frac{w(\zeta - s)}{\zeta} \tag{4.6}$$

and

$$x = \frac{wX}{w - Z} \qquad y = \frac{wY}{w - Z} \qquad \zeta = \frac{ws}{w - s} \tag{4.7}$$

In practice we do not dispense with the variable z, but rather use z or ζ as is most convenient.

A stereoscopic diagram is produced in the following stages. Two drawings are prepared, side by side, assuming given values of s and w; they are then magnified and superimposed on a screen, and from them the observer constructs a three-dimensional picture. Magnifications of the drawings are represented by the transformations

$$X \to X' = mX, \qquad Y \to Y' = mY, \qquad Z \to Z' = m_z Z.$$

In most stereo systems $m_z = m$, but in stereosynthesis this need not be so, and an important degree of freedom is gained by keeping them separate in the theory. Note that all the magnifications which can occur are incorporated in one magnification factor. It is immaterial at the present stage whether these are the intentional magnification when the photographs of the drawings are projected on to the screen or unintentional magnifications or diminutions arising through having the camera or the projector at the incorrect operating distances.

When the two diagrams are superimposed there may be an incorrect registration in the horizontal direction which increases (or decreases) all of the parallaxes by a constant amount, 2p, (which may be negative). This is expressed mathematically by the transformation

$$X'' = X' \qquad Y'' = Y' \qquad Z' \rightarrow Z'' = Z' + p.$$

The observer now constructs the distorted three-dimensional image (x', y', ζ') from X'', Y'' and Z'' using new values s and w; so

$$x' = \frac{w'X''}{w' - Z''} = \frac{w'X'}{w' - p - \zeta'} = \frac{w'mX}{w' - p - m_z\zeta}$$

$$= \frac{w'msx}{(w' - p)\zeta - m_z w(\zeta - s)} = \frac{w'msx}{m_z ws - (m_z w + p - w')\zeta}.$$

The calculation of y' is similar, and

$$\zeta' = \frac{w's'}{w' - Z''} = \frac{w's'}{w' - p - m_z Z} = \frac{w's'}{m_z ws - (m_z w + p - w')\zeta}$$

To summarise

$$\begin{aligned} x' &= \frac{w'msx}{m_z ws - (m_z w + p - w')\zeta} \\ y' &= \frac{w'msy}{Dr.} \\ \zeta' &= \frac{w's'\zeta}{Dr.} \end{aligned} \qquad (4.8)$$

Spottiswoode derives similar distortion equations for his rather different systems.

The most serious distortions, in theory, are those arising from the presence of the ζ term in the denominator. In fact observation and theory part company at this point, but we will first consider the theory on its own merits and see what conclusions can be drawn.

The equations 4.8 specify a group of transformations which are a sub-group of all projective transformations of the space. They are the transformations which leave invariant the origin, all points on the line

at infinity in the ζ plane, and the ζ axis. There is little interest in this aspect of the equations however, the important question is how the parameters may be chosen to restrict the distortions.

If the coefficient of ζ in the denominators is zero then the transformation equations become

$$x' = \frac{w'm}{wm_z}x, \qquad y' = \frac{w'm}{wm_z}y, \qquad \zeta' = \frac{w's'}{wm_z s}\zeta, \tag{4.9}$$

The transformation is affine and furthermore the stretches in the x and y directions are the same. The condition for this is

$$w' = m_z w + p \tag{4.10}$$

The implications of this equation are that an affine projection of the figure desired can be ensured by suitable choice of any of the parameters m_z, w or p. w' is the half-eye width of the observer, and whilst this could be varied by using special spectacles this question does not arise here. The choice of m_z and w is made at the drawing board and is described more fully in Chapter 7. In some forms of projection apparatus there is freedom to adjust p by varying the overlap of the two separate diagrams. This cannot be done in the cinematographic system employed in the present investigation, but it can be done with the various methods of viewing the still slides described in Chapter 7, and it is an adjustment which is essential if they are to be viewed correctly.

The affine transformation in equations 4.9 becomes a similarity transformation if

$$\frac{w'm}{wm_z} = \frac{w's'}{wm_z s} \tag{4.11}$$

i.e. if $$s' = ms, \qquad (4.12)$$

which confirms the earlier conclusion that the screen distance must be correct if the picture is to be seen in correct proportions. This condition is also the condition that each separate picture subtends the correct angles at the eye which is viewing it; and if it is satisfied as well as condition (4.10) then the (x',y',ζ') is simply the (x,y,ζ) space scaled up in the ratio $w'm : wm_z$. In stereo-transmission systems realism is nearly always important. This theory implies that a stereo picture has a definite size - which a 'flat picture' does not have - but for diagrammatic work exact size is unimportant, and there is no need to adjust this scale factor to unity. In this respect stereo-synthesis again has a degree of freedom which is lacking in stereo-transmission in the cinema, of the type discussed by Spottiswoode (9).

These conditions involve w', the semi-eye width of the observer. This is not the same for everyone and some degree of distortion is inevitable on this account, but no writer has suggested that this is a material drawback. It could be overcome if necessary by equipping the observer with suitable spectacles, and we will give this no further consideration.

The distortion due to errors in p alone can be found by putting $w = w'$, $m_z = m = 1$, and $s = s'$ in the equations 4.8. Then, putting $p/ws = P$,

$$x' = \frac{1}{1 - P\zeta}\,x \qquad y' = \frac{1}{1 - P\zeta}\,y \qquad \zeta' = \frac{1}{1 - P\zeta}\,\zeta \ . \qquad (4.13)$$

The effect of varying the overlap of the two images is one of the most

obvious problems in stereoscopy, and it has long been known that the effects which the convergence theory prediots are not found in practice. The overlap of the two images can be varied over quite wide limits without any very pronounced effect on the three-dimensional picture. The 'Overlap' experiment, described in Chapter 6 investigates one aspect of this problem. Practically speaking the divergence between theory and practice is a great convenience because many of the distortions which might be expected to cause difficulties do not appear.

As we remarked earlier, many writers have sought to modify the theory by taking the operative factor in stereoscopic vision to be retinal disparity rather than the convergence of the two lines of sight. That is to say it is assumed that an observer in some way derives information from the fact that the retinal images in his two eyes differ because of the varying parallaxes of objects which are located at different distances. This theory has much to commend it. Rough experiments with stereoscopes, with anaglyphs or with slides projected on a screen all show that the overlap of the two pictures may be varied quite considerably without unwelcome distortions appearing. It might also be relevant to observe that in any case the eyes are continually in motion, performing small saccadic movements, and an observer in some way discounts this and attributes fixity to surrounding objects when there is no such fixity of the image on the retina. Indeed this lack of fixity appears to be an essential feature of normal vision and to be necessary to refresh the nerve endings in the eye (2a,10a). This shows that in another related optical field it is the entire image on the retina that is important and

not a simple, direct one-one relation between a point observed and a point in the eye.

The convergence theory takes no account at all of the finite size of the eye, neither, incidentally, does the differential bipolar theory of Luneburg (see Chapter 5); although this attempts, in its way, to account for the fact that a variety of stereoscopic stimuli give rise to the same stereoscopic sensation. However, the great disadvantage of the retinal disparity theory is that it is difficult to handle, and no quantitative theoretical predictions seem to have been made from it which cannot also be made on the convergence theory. In order to operate within the framework of the theory one obviously has to take a 'gestalt' point of view, emphasising the essential interdependence of the many variables and the importance of considering the situation as a whole without splitting it up into atomic parts. This attitude is important throughout psychology in general, and it is needed in many parts of psychological optics in particular, but time and time again we see that it is grossly unamenable to analysis on the lines of classical physics or mathematics. The 'gestalt' view so often has to be abandoned in order to pursue a quantitative discussion.

Continuing with the theoretical implications of the convergence theory we notice a factor that is likely to give rise to a particularly unwelcome feature in stereoscopic animation. When stereo drawings are prepared the parallaxes involved are always small, and always stretch a draughtsman's technique to the utmost. In a still stereo drawing a small error in parallax causes the corresponding geometrical element to be displaced in

depth. This is a serious practical difficulty. In animation one might expect random errors to cause 'jitter' in those parts of the diagram which are being animated frame by frame. Jitter occurs in any film animation. It is the business of the professional artist to avoid it by tricks whenever he can, and one such obvious trick is to draw cartoons in which the movements are rapid so that the displacements from frame to frame are large compared with the errors. In technical diagrams this is very often impossible, and a certain amount of jitter is inevitable whenever the movements are slow. In stereoscopic work the effect of the unavoidable random errors in drawing parallaxes threatens to be much more serious. How serious, on the convergence theory, is shown by the following considerations. Equation (4.7iii) is

$$\zeta = \frac{ws}{w-Z}.$$

Taking logarithms and differentiating

$$\frac{d\zeta}{\zeta} = \frac{dZ}{w-Z} = \frac{\zeta}{ws}\,dZ, \qquad d\zeta = \frac{\zeta^2}{ws}\,dZ. \tag{4.14}$$

When considering errors in depth which are produced by errors in parallax it has to be decided whether $d\zeta$ or $d\zeta/\zeta$ is the important quantity. But whichever is chosen equation (4.14) suggests that depth jitter will increase with ζ, that is to say it can be reduced by working with the diagram in front of the screen rather than behind. It is questionable if there is any experimental verification of this particular prediction. From equation (4.7i) we have

$$dx = \frac{w}{w - Z}\, dX = \frac{\zeta}{s}\, dX \qquad (4.15)$$

dX and dZ are errors on the screen, and our assumption is that these are uniform magnifications of errors on the drawing board. So comparing equations (4.14) and (4.15) it would appear that errors on the drawing board will ultimately misplace points in the z-direction by ζ/w times as much as in the x-direction. This is very serious because in the system being described at screen distance $\zeta/w = s/w =$ 10 feet/1 inch which is a factor of over 100.

If depth jitter in stereo animation really was a hundred times worse than lateral jitter in ordinary animation the situation would be quite intolerable, and it would be impossible to achieve stereo animation by normal drawing-board techniques. It may be that considerations along these lines have discouraged professional animators from attempting the problem for so long, but at this point once again convergence theory disagrees with experiment. The jitter should also exist on the retinal disparity theory, but the new situation is quite different from the one discussed earlier. There the point was that the distortions produced by a systematic error in parallax over the whole of the diagram do not agree in theory and practice, here the point at issue is the distortion produced by a purely local error in parallax. Later evidence will show that such local errors are very serious if they are static, but in cinematographic projection errors are only on the screen for 1/16 second and whether or not the eye can follow them is a completely new question.

This problem is the subject of Chapter 10, and for the moment we will merely state the conclusion that under the projection conditions

with which we are concerned the eyes are very sympathetic to jitter of this kind, and far from interpreting errors in drawing as random fluctuations in depth they average them out so that binocular animation appears even smoother than monocular.

Three further types of distortion remain to be considered. The film is projected by a twin lens system. It is possible that the focal lengths of these two lenses may differ very slightly and that the two pictures may be magnified unequally. The stereoscopic consequences of unequal magnification have been discussed at length by Ogle (5). His conclusions are that unequal magnification results in an angular displacement of a frontal plane (z = constant) so that one side of the screen, for example, appears nearer to the observer than the other. No distortions of this kind have been observed in the apparatus which was used in the present experiments.

The axes of the twin lenses are 5.6 mm apart, so that if the distance between the screen and the projector is varied the two figures on the screen undergo magnifications about centres which are separated by this amount. This discrepancy is very slight and it is not taken directly into account in the previous theory; but it can be considered indirectly as causing a change in all three of m, m_z and p. So the finite separation of the axes of the two lenses does not introduce any essentially fresh problems.

The final distortion to discuss is the one induced by interchanging the pictures seen by the two eyes. This can be done by appropriately turning round the viewer's polaroid spectacles. The term 'pseudoscopic' is

sometimes used to describe the situation. Elementary calculations show that it is described by the transformation equations

$$x' = \frac{xs}{2z + s} = \frac{xs}{2\zeta - s}$$

$$y' = \frac{ys}{2z + s} = \frac{ys}{2\zeta - s}$$

$$z' = -\frac{zs}{2z + s} \qquad \zeta' = \frac{\zeta s}{2\zeta - s}$$

Points in front of the screen become points behind and vice versa. It is interesting that $x' : y' : z' : \zeta' = x : y : -z : \zeta$ so that the bearing of any points from the observer is unaltered by the transformation, whereas the ray from the point to the x, y, z origin is reflected in the screen.

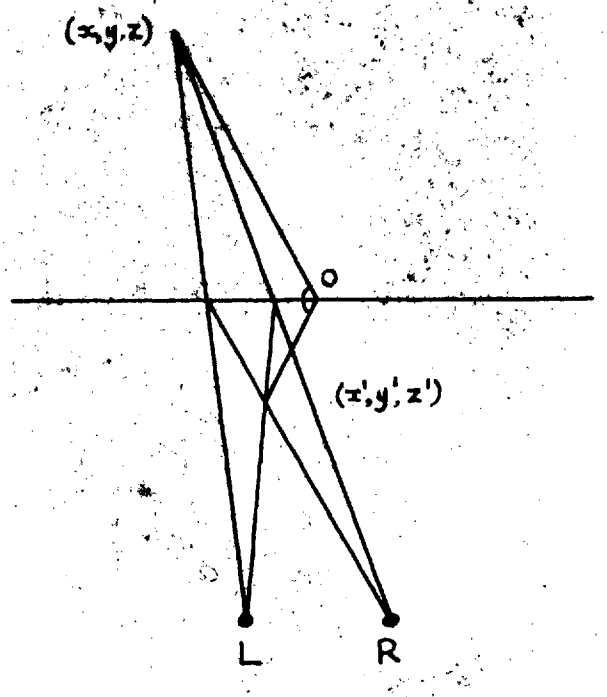


Fig. 4.2

This situation appears at first sight to offer many opportunities for making teaching apparatus to illustrate important principles of projective geometry. For example, in some presentations of the subject many proofs depend upon projecting some particular finite line to infinity. If this line is in the plane half way between the screen and the observer the pseudoscopic transformation sends it to infinity. This means that on the instruction "project the line L to infinity" the audience could do so merely by reversing their spectacles. Once again theory does not agree with experiment. Stereoscopic viewing is a 'gestalt' situation involving many factors and we cannot isolate one factor in this simple way. There is also the point made by Shipley (8) that 'infinity' is an intellectual construction, and "there is nothing in our visual space to correspond with the purely intellectual concept of infinity". Thus observers are usually slow to appraise pseudoscopic transformations whenever the objects depicted resemble familiar, physical objects, or whenever the distances involved are large (on the convergence or disparity theories).

When stereoscopic evidence, in isolation, runs counter to previous experience it is generally rejected. It is not possible to suggest mathematical infinity (as a psychological experience) by stereoscopy any more than it is by getting observers to look at a collimated beam of light.

REFERENCES

1. Charnwood. Essay on Binocular Vision, London, 1950.

2. H. Dewhurst. Auto Precision Stereoscopy. Phot.Journ. Vol.92B, p.2,1952.

2a. R.W. Ditchburn. Eye-movements in relation to retinal action. Optica Acta, 4, 171-176, Feb. 1955.

3. J.J. Gibson. Perception of the Visual World. Boston, 1950.

4. Helmholtz. Treatise on Physiological Optics. 3rd edn. 1910. Trans.J.P.C. Southall, U.S.A., 1925.

5. K.N. Ogle. Researches in Binocular Vision, Philadelphia, 1950.

6. R.M. Pritchard. Stabilized Images on the Retina. Sci.Amer., 204, 6, June, 1961.

7. J.T. Rule. The Shape of Stereoscopic Images. J.O.S.A.,31, pp.124-129, Feb. 1941.

8. Thorne Shipley. Convergence Function in Binocular Visual Space. J.O.S.A., 47, 9, pp.795-821, Sept. 1957.

9. R. and N. Spottiswoode. The Theory of Stereoscopic Transmission. Univ. of California Press, 1953.

10. W.D. Wright. The Role of Convergence in Stereoscopic Vision. Proc.Phys.Soc. B 64, 289-297, 1951.

10a. M.P. Lord and W.D. Wright, Eye Movements during Monocular Fixation Nature, 162, p.25, July 3, 1948.

11. ------- Report of the Air Survey Committee, No.1, 1923, Appendix VI. Stationery Office.

CHAPTER FIVE

LUNEBURG'S THEORIES AND THEIR APPLICATION TO FILM

It has been known since the earliest studies of stereoscopy that the convergence theory was inadequate in a number of ways. The theory is extremely simple and conforms to common-sense notions of how binocular vision might work, and it forms a satisfactory basis for the preparation of stereoscopic drawings and photographs; but some of the effects which it predicts are not observed. Most obviously the projective deformations described in Chapter 4 which are associated with incorrect lateral registration of stereo drawings and incorrect matching of the parameters corresponding to eyewidth are not observable, or at least not observable to anything like the degree which the theory predicts.

Binocular observation of a point source of light gives no indication of its distance, but small differences in the distances of two point-sources can be assessed with great accuracy. In the "Overlap" and "Photofinish" experiments of Chapter 6 observers could make accurate appraisals of many features of the diagrams, but whenever they volunteered information about parts of the picture being 'in front of the screen' or 'behind it' this was usually quite unreliable and bore little relation to the geometrical optics of the situation.

In this connection Martin (10) writes that if the two parts of a stereo picture are moved closer together the whole field appears to draw nearer (as predicted by the theory) "but such apparent movement from or towards the observer ceases immediately the pictures are still, and the whole field looks to be at the same absolute distance as before."

These conclusions are so widely accepted that it is unnecessary to quote further evidence. However until the recent work of Luneburg there does not seem to have been any alternative quantitative theory of stereoscopic vision which could serve as a framework in which to fit observations or as a basis for the construction of stereoscopic drawings. Luneburg's theories have had noteworthy successes in accounting for anomalies in stereoscopic vision which were studied by Helmholtz, Blumenfeld and Ames, and they have been applied to the science of cinematography by Levonian (6) and Hill (5). We will now describe these theories in outline, and then discuss their relevance to the problem in hand.

The following phenomena cannot be accounted for adequately by the convergence theory. Helmholtz observed that if people were seated in a dark room and asked to arrange a set of lights in a straight line across in front of them they did not in fact arrange the lights in a straight line but in curves which he called horopters. These horopters were symmetrical about the centre-line in front of the observer, and they were concave towards him up to a certain distance; and beyond this distance they were convex.

Blumenfeld performed experiments in which observers performed such tasks as the following. They were first given a pair of distant lights and asked to arrange other lights in lines so that they formed a parallel alley; that is a pair of lines parallel to one another and symmetrically disposed about the centre line ahead. It was found that the observers again arranged the lights along curves. In another experiment the

observers were given a pair of distant lights as before, and then lights two by two to compare with them whilst all the other lights of the alley were turned off. They were asked to arrange each pair so that they were the same distance apart as the reference pair. On a Euclidean basis this should of course have led to exactly the same result as the first experiment; but whilst the observers again arranged the lights along curves instead of straight lines they were not the same curves as before, and furthermore the differences were repeated consistently in subsequent experiments. In other words the subjects interpreted parallel and equidistant in such a manner that they led to different results.

More recently in America Ames has conducted a number of experiments on binocular vision which have involved certain types of optical illusion. In one very celebrated experiment he constructed empirically a set of 'rooms' of distorted shapes. The middle-sized room was normal, but the smallest one had concave walls and ceiling whilst the largest one had them convex. When a viewer, who did not know this, was placed at a suitable viewpoint for each room he had the impression that all three rooms were the same size and normal in shape.

Blumenfeld's experiments have been repeated and confirmed by a team working at Columbia University (3), and their investigations also included experiments using stereoscopic sources of light, and it was found that observers treated stereoscopic sources exactly as had been assumed in the convergence theory, in so far as they reacted as if they were seeing just one source of light situated at the point of intersection of the two lines of sight; but there was again a departure from the classical theory since

these stereo-sources also produced Helmholtz's horopters, Blumenfeld's alleys and the related phenomena.

The phenomena are of course 'subjective' in the sense that people are needed to observe them. They are not observable by, say, range-finding instruments. Any theoretical discussion must therefore contain psychological and/or physiological factors, as well as the objective factors of geometrical optics. Luneburg attempted to account for these phenomena by two separate but related theories which may be called the <u>differential bipolar theory</u> and the <u>hyperbolic metric theory</u>.

Throughout this section we adopt Luneburg's notation, and it should be noted that his x, y and z axes do not have the same orientation as the axes in Chapter 4. (No confusion need arise as it is never necessary to use the differing conventions simultaneously.) Every point in space, in front of the observer, has a unique set of coordinates γ, ϕ and θ (see Figs.5.1 and 5.2, which are taken from Reference (9)) which are related to the rectangular cartesian coordinates by

$$\gamma = \pi - \alpha - \beta \qquad \phi = \tfrac{1}{2}(\beta - \alpha)$$

$$x = \frac{\cos 2\phi + \cos\gamma}{\sin\gamma}\cos\theta \qquad \tan\gamma = \frac{2\sqrt{x^2 + z^2}}{x^2 + y^2 + z^2 - 1}$$

$$y = \frac{\sin 2\phi}{\sin\gamma} \qquad \tan 2\phi = \frac{2y\sqrt{x^2 + z^2}}{x^2 + y^2 + z^2 - 1}$$

$$z = \frac{\cos 2\phi + \cos\gamma}{\sin\gamma}\sin\theta \qquad \tan\theta = \frac{z}{x}$$

Luneburg's first, tentative hypothesis is that the absolute values of γ, ϕ and θ are irrelevant, and that the only observables are

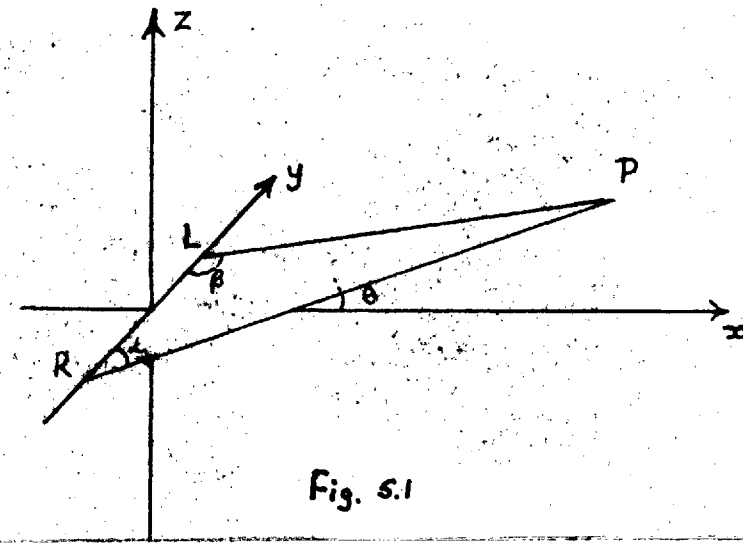


Fig. 5.1

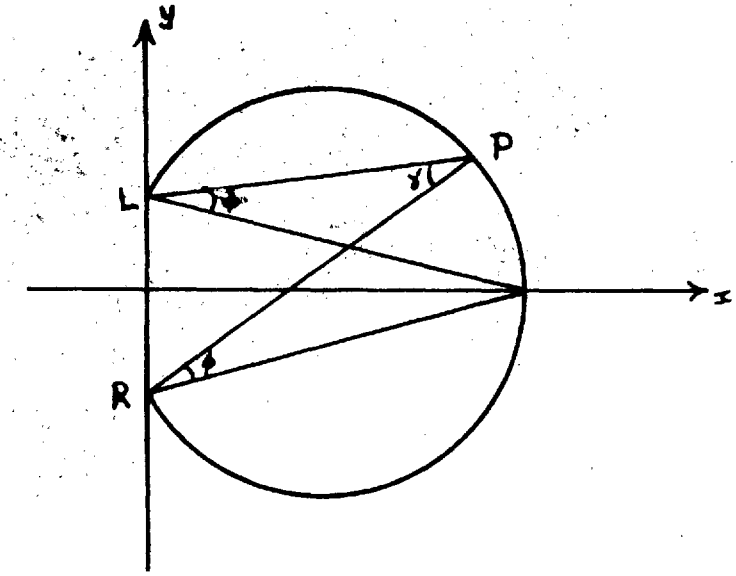


Fig. 5.2.

differences in these coordinates. Putting the matter another way figures related by the transformation

$$\gamma' = \gamma + \text{const.} \qquad \phi' = \phi + \text{const.} \qquad \theta' = \theta + \text{const.}$$

are visually indistinguishable. Such a transformation is called an <u>iseikonic transformation</u>, and figures which are so related are termed <u>equivalent</u>.

These transformations are easily pictured in terms of the coordinate grid. In the horizontal plane in front of the observer the lines γ = constant are Vieth-Müller circles $x^2 + y^2 - 2x \cot \gamma = 1$ and the lines ϕ = constant are Hillebrand hyperbolae $-x^2 + y^2 + 2xy \cot 2\phi = 1$, with asymptotes $y = x \tan \phi$. The three-dimensional grid is produced by rotating the configuration in the xy plane about the y-axis. Iseikonic transformations are translations on this grid. At large distances from the eyes the hyperbolae are indistinguishable from their asymptotes and the Vieth-Müller circles are virtually $x^2 + y^2 - 2x \cot \gamma = 0$, a coaxal system touching the y-axis at the origin. This produces a simplified form of grid, which should not be confused with a simple central projection (although the lines ϕ = const. radiate from the origin) because the two eyes have virtually become a "dipole" at the origin. Grids of these kinds can be superposed on the drawing methods of Chapter 7, so that elements in the drawing can be transformed to equivalent positions.

Although this theory is still incomplete, it explains quite a lot even at this stage of development. Luneburg gave a theoretical treatment of Ames's rooms the results of which compared well, he claimed, with experiment; and Levonian (6) and Hill (5) applied the theory to the practical problems of stereo camera work for motion pictures. As a result of their studies it was claimed that the technology of filming benefitted because the camera gains an extra degree of freedom when compared with the restrictions which arise from the convergence theory, and this extra freedom was an advantage in overcoming problems of matching cameras, lens spacing, toe-in, distance and "giantising" factors.

When applied to the problems of stereosynthesis the practical benefits are not so great, because, as explained more fully in Chapter 8, camera work in stereosynthesis is very much simpler than in live action photography. However, when synthesising stereo pictures from drawings the variation of the overlap of the two pictures results in a transformation that is virtually iseikonic. That is to say the differential bipolar theory supports the firmly established practical criterion that exact horizontal registration of the two pictures of a stereo pair is not of prime importance. It also confirms that the effects which were sought in the "Overlap" experiment in Chapter 6 should not arise. That the transformations produced by variations in overlap differ from iseikonic ones only in the second order terms can be seen by inspection of Figures 5.1 and 5.2, and does not really call for analytical proof. Variation of the overlap varies angles α and β, but θ and ϕ remain approximately the same and γ is altered by approximately the same amount throughout the figure. We are concerned with viewing conditions in which the pictures at distances of about 10 feet from the observer, and the variations in overlap which may occur do not exceed 3 ins. Therefore the second order terms are quite unimportant.

The differential bipolar theory gives one additional benefit in stereosynthesis since it suggests that elements in a diagram which occur in an inconvenient position can be transferred to some other equivalent position where their appearance will be unaltered. This means, for example, that points whose parallaxes are very small can be moved to some other part of the physical space depicted in the diagram where the

parallaxes are larger. This might be a practical convenience because the difficulties of drawing small parallaxes accurately are the main problems of stereosynthesis; and the bigger the parallaxes the smaller the relative errors. Transformations of this kind can be done on a drawing board by superimposing a grid, as explained earlier, or they may be calculated analytically by the methods used by Luneburg in (7) to discuss equivalent Ames's rooms. However the use of iseikonic transformations is not likely to be of great practical advantage because in general they turn straight lines into curves.

This suggests another possibility. It is in principle possible to construct certain space curves by drawing straight lines and subjecting them to iseikonic transformations. But this comes up against an essential inadequacy of the differential bipolar theory. This theory classes certain related configurations as visually equivalent, but out of the range of configurations which the observer might be seeing he nearly always decides that he is seeing just one - and the theory does not say which interpretation out of the many possible he decides on. We can only say that he decides on the interpretation which is most "reasonable" - meaning the interpretation which is most in accord with his previous experience.

But Luneburg's theories go further without introducing psychological notions with no precise definition. The theory is extended by modifying the hypothesis that the absolute values of γ are of no importance and that only variations are significant. Luneburg introduces a metric to explain how distances are subjectively assessed in terms of γ, ϕ and θ.

Whereas the earlier theory is objective in that the geometry is the same for all observers this part of the theory is subjective in that two parameters σ and μ are involved which are constants for particular individuals. In this way it is possible to aoocunt for the different ways in which different people assess the same situation. The metric is introduced in the following manner. We have at our disposal the coordinates γ, ϕ and θ, of which the first corresponds to the radial variable. These are used to assign further coordinates

$$\xi = f(x)\cos\phi\cos\theta, \qquad \eta = f(x)\sin\phi, \qquad \zeta = f(x)\cos\phi\sin\theta,$$

$f(\gamma)$ is a funotion to be discussed later. These coordinates are to be thought of as being in 'subjective space', and in this space distances are calculated by the Riemannian metric

$$ds^2 = \frac{1}{\sinh^2\sigma(\gamma+\mu)}\left[\sigma^2 d\gamma^2 + d\phi^2 + \cos^2\phi\, d\theta^2\right]$$

These distanoes are subjective distances, i.e. distances as they are perceived by an observer. Most commonly Luneburg puts $f(\gamma) = 2e^{-\sigma(\gamma+\mu)}$, but he is sometimes prepared to work with the function left in a general form and also there is an arbitrary scale factor in his work and he does not always adopt the same normalisation. This space is a hyperbolic space of constant negative curvature, and the references give full arguments why this type of space is necessary, but these do not concern us here. The space can be mapped conformally inside a sphere in a Euclidean space (Poincare's model), in whioh case geodesics are arcs of circles which are normal to the base sphere; and again it can be mapped (but not oonformally) in the Klein model in which the geodesics are straight lines. Luneburg makes most use of the former model.

From this metric it is possible to deduce theories of Helmholtz horopters, alley experiments and Ames's rooms. The notion of equivalent configurations is replaced by that of 'congruent configurations', because in this hyperbolic space there is a six parameter group of transformations which preserve metrical relations, i.e. bodies can be translated and rotated much as in Euclidean space. Much of Luneburg's theoretical discussion leaves the precise nature of the space as an open question, and he deals with the elliptic, Euclidean and hyperbolic possibilities; but he decides finally that the experimental evidence indicates the hyperbolic alternative as the one to select.

There is an analogy between Luneburg's two theories and the special and general theories of relativity. Apart from the similar geometrical ideas underlying both systems, in both cases the "special" theory seems well established and is applicable to a wide variety of phenomena whereas the position of the "general" theory is by no means so certain, and alternative forms of metric have been proposed by others who support the outlook of the theory but who do not accept its details.

Hardy, Rand, Rittler and Blank (3) have proposed another form of metric, which differs in some rather important respects from Luneburg's and gives greater prominence to the most distant point in any configuration - a feature which they think of definite physical importance; and Shipley (11), whilst sympathetic to Luneburg's approach, suggests that parallel geodesics in alley experiments have not be identified properly which means that the question of the sign of the curvature of visual space is still not decided. That is, visual space may be elliptic and not hyperbolic as Luneburg concluded.

We mention these other views to show that the position of this theory is by no means firmly decided, and for this reason alone it certainly cannot be taken uncritically into film technology. There are however greater difficulties. This theory is a truly subjective one in that it allows for different behaviour from different observers. The constants which occur in the theory for different observers have been measured for a number of observers by Zajaczkowska (12) and these show a very wide spread. She reports the following range of values,

σ = 3.3 to 26.4, with 10 as a reasonable average

$K = -e^{-2\sigma\mu} = -0.98$ to -0.05, with -0.5 as a reasonable average.

This range is very wide, and whilst the exact meaning of the spread in terms of differences on the drawing board could only be seen after a tedious calculation it is difficult to see how such a wide range of parameters could be catered for at one and the same time.

Hill has used this theory to confirm some of his recommendations for stereo cinematography, but he does not state what values of parameters have been taken, or how serious the effect of spread is. Indeed at the time at which this paper was written the constants only seem to have been published for one or two observers, and this sample is clearly inadequate.

Therefore the applications of Luneburg's theories to practical technology are doubtful. In any case they only apply to vision at a short range, six feet or less, and with a screen distance of ten feet or the considerably greater distances which are normal in a public cinema all the action of a film takes place outside the critical distance. Also the theories have only been quantitatively confirmed for phenomena which arise

under very special viewing conditions - when seeing points of light devoid of all other indications of depth. The situation when viewing pictures of natural objects is quite different, and it is a most questionable extrapolation to extend the original experimental conclusions to them.

Against this view it can be argued that Luneburg's theories also apply to Ames's rooms, and that the cinema pictures produced with the aid of the theories were satisfactory. The reply to this is in two parts. The 'room' for which Luneburg gave a calculation was only a small, model room, with linear dimensions of less than two metres, and human eyes have shown themselves so accommodating to stereo perspective covering a wide range of parameters in many different circumstances that it is no guarantee of the truth of a theory that acceptable pictures can be made with its aid. Surprisingly satisfactory pictures can be made empirically, especially when many other depth-suggesting factors are present at the same time (see the pictures by McLaren, appended to Chapter 7). Also Luneburg's treatment of congruent Ames's rooms still leaves an unresolved question as the earlier treatment of equivalent rooms did. Congruent rooms form a one parameter system (which arises from the group of translations in hyperbolic space after certain parameters have been eliminated by considerations of symmetry). All rooms of this system are visually equivalent but the observer prefers one interpretation to all of the others. Again it can be suggested that he chooses the one particular interpretation because he rationalises the situation and sees it as much as possible in terms with which he is most familiar. This objection has

to contend with the counter-objection that it has been possible to produce experimentally other types of 'room' in which 'unnatural' things occur in rooms of natural shape - but Luneburg did not discuss these.

Nonetheless there are some extremely special circumstances in which Luneburg's theories could conceivably apply to the design of mathematical films and physical diagrams. It might be possible to use the subjective geometry of the observer to convey information about non-Euclidean situations. For example, the type of configuration occuring in the films from the Humboldt University which display features of non-Euclidean geometry with important applications in the theory of functions might be communicated to an observer by suitably drawn stereo diagrams in such a way that the non-Euclidean features of the configuration became apparent to the observer without being 'objectively there' in the normal interpretation of the film according to Euclidean convergence theory.

If it were possible to do this it would be extremely worth-while, because it would give a direct sensory experience of three-dimensional hyperbolic space. The special conditions under which alley experiments have been performed can scarcely do this to anything like the same extent, because movement has been absent from the situations they have handled and a film offers possibilities of producing configurations which could never be produced with the limited resources of lamps on an optical bench.

However the development of such a film technique presents obvious difficulties and an attempt to do so in the course of the present enquiry would have introduced many new features and would have led too far away from the main problem.

The animation and the projection arrangements would have to be constructed round quite fresh parameters. The entire action of such films would have to take place at distances from the observer of less than ten feet, and the closer the better. Writers on this type of experiment stress the extreme accuracy with which points have to be aligned; "Visual criteria were often so clear as to demand motions as small as 1 mm at the most distant variable points, approximately 2.75 metres distant" (11). It would not be possible to make one film of a subjective non-Euclidean phenomenon that depicted the situation adequately for all observers; at least not until projection apparatus is available which provides for more control over projection parameters than is the case at present.

Finally it should always be remembered that this type of experiment is concerned with a highly unusual notion of parallelism for the "sensory parallelism" studied is not the same thing as physical parallelism. Thus Shipley writes in the instructions to his observers, "The rows are not visually straight and parallel if they seem to converge to some distant vanishing point, as when you look down 'physically straight' railrod tracks. The task is to make the tracks seem to be straight and parallel, even if they would be extended indefinitely into the distance."

These stimulating and challenging theories have a value which is not in any dispute, but here there is yet one more reason to question their relevance to the problems of the stereo cinema in so far as it is trying to reproduce physical reality. But their relevance to the problems of stereo-synthesis, which are ultimately to provide a chain of communication whereby abstract information may be passed from one mind to another, is

immense; and if a projection system of sufficient refinement could be devised quite new possibilities in diagrammatic communication might be opened up, since there would be available a three-dimensional hyperbolic space, directly perceived, with freedom to carry out movements and even other more general transformations.

REFERENCES

1. A.A. Blank. Axiomatics of Binocular Vision. The Foundations of Metric Geometry in Relation to Space Perception. J.O.S.A., 48, 5, May 1958, 328-334.

2. T.J. Fletcher. The Prospects for 3-D Mathematical Films. Univ.Film Jour., 6, Autumn 1954, 1-5.

3. L.H. Hardy, G. Rand, M.C. Rittler, A.A. Blank. The Geometry of Visual Space Perception. Columbian U.C., New York, 1953.

4. H. Helmholtz. Treatise on Physiological Optics. 3rd edn., 1910, trans. by J.P.C. Southall, U.S.A., 1925.

5. Armin J. Hill. A Mathematical and Experimental Foundation for Stereoscopic Photography. J.S.M.P.T.E., 61, 4, Oct. 1953, 461-486.

6. E. Levonian. Stereography and the Transmission of Images. J.S.M.P.T.E., 64, Feb. 1955, 77-85.

7. R.K. Luneburg. Mathematical Analysis of Binocular Vision. Princeton U.P., 1947.

8. do. Metric Methods in Binocular Visual Perception. Courant Anniversary Volume, 1948, 215-240.

9. do. The Metric of Binocular Visual Space. J.O.S.A., 40, 10, Oct. 1950, 627-642.

10. L.C. Martin. Technical Optics, Vol.II, London 1950.

11. Thorne Shipley. Convergence Function in Binocular Visual Space. J.O.S.A., 47, 9, Sept, 1957, 795-821.

12. A. Zajaczkowska. Experimental Test of Luneburg's Theories. Horopter and Alley Experiments. J.O.S.A., 46, 7, July 1956, 514-527.

CHAPTER SIX

STATIC EXPERIMENTS

I. EXPERIMENT "PHOTOFINISH"

Summary. 47 observers were tested on a series of slides showing four vertical lines with different displacements from the screen. The object of the experiment was to see how accurately they could assess their relative positions with different ranges of parallax. After a practice period each observer was shown six pairs of slides with the displacements of the lines from the screen decreasing progressively, and each time he was asked to state the order in which the lines appeared to him.

Three quarters of the observers were able to complete the test with four faults or less (out of a maximum possible of 36). The time required to get accommodated varied greatly from observer to observer. Generally speaking observers had more trouble with the earlier tests where the parallaxes were large than with the later tests where they were small, but the increased numbers of faults towards the end of the experiment showed that the displacements were becoming sufficiently small to cause some difficulty.

The slides were made in the form of a filmstrip by photographing drawings, and by making measurements on the screen it was possible to gain an idea of the errors involved in the drawings and of the errors in the projection system. These were of about the same order of magnitude, and good observers had no difficulty in detecting errors of this size.

The drawings and the projection system. The drawings were made in Indian ink on cellulose acetate, $9\frac{1}{2}$" x $12\frac{1}{2}$". They were photographed on Pan F film

and the negatives projected by two Aldis 1000 projectors, so that the final picture was white against a grey background.

The same drawing was used for the right eye for all the test pictures and six different drawings were used for the left eye. Before the test was performed the observers were given a period of practice for which the same diagrams were used, but interchanged between the eyes. In this way fresh diagrams were presented with the minimum effort.

Each diagram consisted of four vertical lines at different distances. On the original drawings the parallaxes were calculated with the intention of scaling up by a factor of four on projection and having the observer at a distance of 10 feet from the screen. In each test situation one line was in front of the screen by one unit, one line was in the screen, one line was one unit behind the screen, and one line was two units behind. The unit was, in the successive tests, 4 feet, 3 feet, 2 feet, one foot, 6 inches, and 3 inches.

The method of testing. The object of the test was explained to the observer. He was shown the first practice figure and asked how many lines he saw. This showed whether or not he was fusing the four lines correctly. Some observers had no difficulty from the start, but others saw a double image of the line 4 feet in front of the screen. (This corresponded to a parallax on the screen of 1.6".) If necessary the pictures were displaced relatively in and out in order to bring about proper fusion. The same procedure was often necessary with the next figure where the front line was 3 feet in front of the screen; but after this the observers had very little difficulty, and in any case once fusion

was obtained there was usually no further diplopia unless fatique set in much later.

When it was clear that the observer was fusing properly he was asked if any line appeared nearer than the others or any line appeared further off. This was repeated with other diagrams until the observer could announce his decision in the form required; he would say, for example, "The nearest is No.4, then No.2, then No.3, then No.1." When he understood what was wanted and felt easy about the situation he was told that from then on his results would be recorded.

The first test with the very large parallaxes gave most trouble because it was often difficult to fuse properly, and a small number of observers could not even attempt it, whilst doing quite satisfactorily on the later tests. In this test the lines were between half screen distance and twice screen distance, and it is common in writing on stereoscopy to take it for granted that displacements from the screen by amounts even greater than this are quite acceptable. This may well be so for natural photography or for well-trained observers, but if diagram work is to be presented to observers with no special training it may be advisable to keep the diagram well within these limits.

The observers reading of each diagram was recorded and his number of faults computed by counting the number of invertions between the order which he gave and the true order. E.g. if the true order was 2314 and he announced 3124 this counted as two faults.

Results. In view of the difficulties experienced with diagram No.1 the final results are tabulated in two forms, one showing the results on all six tests and one the results on diagrams 2-6 only.

Distribution of faults

No. of faults	No. of observers (6 diagrams)	No. of observers (5 diagrams)
0	6	6
1	14	17
2	5	7
3	8	6
4	2	4
5	1	-
6	2	1
7	-	2
8	-	1
9	-	1
10	-	1
11	1	-
12	3	-
13	-	1
14	1	-
15	-	1
16	-	-
17	3	-

Table I

The table includes figures for three observers who were retested. Two observers out of the total of 47 did not do all six diagrams completely, and were unable to do the test because their stereoscopic vision was inadequate.

The maximum possible number of faults was six on each test, a total of 36. Random guessing would, on the average, give a total of 18. One fault could be attributed to the apparatus rather than to the observers because in diagram No.5 the relative position of two of the lines was in some doubt.

Variations between observers. After a suitable period of practice, in which to become accommodated, 3/4 of the observers could interpret the diagrams quite satisfactorily.

Out of the 47 two could not attempt the test at all as their stereoscopic vision was quite inadequate. (This is not a fair sample of

what may be expected throughout the whole population, for apart from questions of the significance of the size of the sample people who were known to have monocular vision were not invited to take the test in the first place.) Two people could not read the first diagram with large parallaxes, and eight had to be classed as "poor" as they had between 11 and 17 faults. (Random guessing would give a total of 18.)

Three of the poor performers were later retested and they all put up much better performances with 3 faults compared with 14, 4 compared with 12 and 2 compared with 17 respectively. It is ~~most~~ likely that their poor performances on the first test were due in part to being slow to pick up the idea, or what is more likely to being distracted by others waiting to take the test. (This distraction was not normally present.)

These tests by no means approach the limits of stereo-acuitity. The order of discrimination required is very much less than in experiments described in France (1), which are concerned with skilled observers whose powers are being trained to the limit, and by such writers as Helmholtz (2). Nevertheless the distribution of faults on the individual tests (Table II) suggests that when the lines were only displaced from the screen by amounts of the order of 3 inches the observers were beginning to get into difficulties.

Diagram number	Unit of displacements	Total no. of faults made	No. of observers making faults
2	3 feet	37	18
3	2 feet	35	21
4	1 foot	15	13
5	6 inches	23	18
6	3 inches	31	19

TABLE II

Table II shows the total number of faults made by the 48 observers who undertook the test on each particular diagram.

It is quite possible that the performances of most of these observers could be improved by training, but that was not the purpose of the enquiry.

The accuracy of the drawings and of the projection system. The parallaxes which had to be drawn on the drawing board were calculated to .001 inch, but of course drawings cannot be made to this accuracy. When the drawings were ultimately projected on to the screen there was a linear magnification of 4. The accuracy of the drawings was checked by measurement of the projected pictures on the screen. During the course of the experiment it was necessary to realign the apparatus and this opportunity was used to measure the parallaxes on the screen once more. Comparison of the calculated parallaxes with the measured ones, and of the measured ones with the two different settings of the projectors suggested that the distortions arising from faulty drawing and from faulty alignment of the projectors (including optical faults in the lenses etc.) were of about the same order of magnitude.

The detailed results are not recorded here, but they may be summarised by saying that faults in drawing could displace a line by as much as about .1" on the screen and faults in the projectors could do the same. (This displacement is a displacement relative to the rest of the configuration.)

An error of .1" on the screen corresponds to an error of about .025" on the drawing board; and it might be considered that a draughtsman should do better than this. But it must be remembered that such factors as buckling of the cellulose acetate and slight mechanical misalignment on the register pegs all contribute to what are here spoken of as "errors on the drawing board". These faults were just sufficient to make it doubtful if the intended relative positions of the lines had been realised in one case, so on Table I one fault may be attributed to the apparatus rather than to the observers, and on Table II it is possible that the faults on diagram No.5 are due in some measure to the apparatus.

A further check on the apparatus was provided by projecting identical pictures of a piece of graph paper from the two projectors. Measurements on the screen indicated misalignments of up to .1" over the parts of the screen actually used and more at the edges. 18 observers were asked whether or not the screen appeared flat or curved; 13 said that it was cylindrical and concave towards them, 1 said it was spherically concave and 4 said it was flat. Owing to the finite size of the projectors it was necessary to have their optical axes 6.9" apart with the projectors 'toed-in'. Considerations based on the theory of Chapter 4 show that this should result in a distortion of the type observed. This 'toe-in' could well account for the greater part of the misalignments attributable to the projection system.

Conclusions

i) Only two people out of 47 were unable to do the test and only eight had to be classified as 'poor', and of the poor ones the three who

were re-tested did very much better at a second attempt. Therefore the information which the draughtsman attempted to convey was communicated with considerable success.

ii) The large parallaxes corresponding to displacements of 4 feet in front of the screen and 8 feet behind were a definite distraction and displacements as large as this should be avoided if possible.

iii) Displacements of down to 3" were read with ease by many observers, but the best results were achieved with displacements of about 1 foot. However the smallest displacements used were beginning to tax the observers and although they had by no means reached the limit of their skill their performances were beginning to deteriorate.

iv) Inaccuracies on the drawing board and in the projection system can each produce parallax errors which correspond to displacements of up to 6" (or more) so the production of diagrams of this type taxes drawing techniques to the full. The draughtsman's hand is not as accurate as the observer's eye.

REFERENCES

1. H. Helmholtz. Treatise on Physiological Optics. 3rd edn. 1910. trans. J.P.C. Southall, U.S.A., 1925.

2. ------- Notice provisoire sur le stéréoscope de démonstration et d'entraînement. Institut Geographique National, France, Feb. 1942.

II. EXPERIMENT "GUESSWORK"

Summary. Observers saw projected stereoscopic diagrams of two lines, one vertical line in the plane of the screen and one line in a horizontal plane inclined at various angles to the screen. The two lines formed a letter "L". The vertical line was a reference and the observer tried to estimate the length of the oblique line. The relations between the estimates and the lengths of the test line when calculated by the convergence were then investigated.

There was only the one slanting line with the one parallax in the whole diagram, and there were no clues of depth by line tapering, aerial perspective or blocking, but there was a slight perspective clue which tended to assist the observer in the first part of the experiment and to hinder him in the second. The test revealed wide variations of ability, and it was not possible to assess this spread in any satisfactory numerical way. Nevertheless the information obtained is of definite value in the design of stereoscopic diagrams and this experiment suggests further tests of a similar kind which might yield results more susceptible to numerical analysis.

Procedure and Results. The observer was shown a projected stereoscopic diagram of two lines, a vertical one in the plane of the screen whose length was always 20", and another of varying length in a horizontal plane and at varying inclinations to the plane of the screen. He was then asked to locate the position of the test line, and to estimate its length in space taking the length of the vertical line as ten units. Before the results were recorded he was shown two test situations and

asked to describe what he saw. This was to see that he was capable of undertaking the test properly, and also to give him an opportunity to get his vision accommodated. He was then shown about five of the frames that were used later in the test for practice, and his task was explained to him. He did not make estimates while this was done. A standard set of instructions was given, although a little experience was necessary before its form became finalised.

The drawings used were drawn in black on a white background on a large scale so that errors could be minimised as much as possible. The negative was projected so that the observer saw white lines on a grey background - a format that is common in mathematical films - so that the viewing conditions were very like the conditions which would actually occur when looking at a stereo film diagram in black and white.

When the observer had seen the practice diagrams the test began, and his estimates were recorded. In the first half of the test he saw ten patterns. In the first two the test line was in the plane of the screen, in the next two inclined backwards at 20°, in the next two at 40°, the next two at 60°, and the last two at 80°.

Nearly every observer managed this part of the test satisfactorily, and could explain where the test line was. It was normal for observers to underestimate the length of the test line when it was foreshortened, and the greater the foreshortening the worse was the estimate. However when an observer's stereoscopic ability was good his estimates were often very good. J.W. was an excellent observer, and in spite of the foreshortening his estimates of the length of the test line were as often

above the theoretical value as they were below. On the other hand subjects who clearly had difficulty in seeing the picture stereoscopically greatly underestimated the length of the test line, and in a number of cases (J.V., R.B.M., D.W.) the observer gave estimates which were much closer to the projection of the test line on the screen than to its length in space.

The observer A.H. was interesting because his first estimates were well below the correct target, but as the test went on he clearly began to see the diagrams in depth better and better. When the earlier frames were repeated at the end of the experiment he had no difficulty in seeing them properly in depth, he greatly increased his estimates and came very close to the correct values.

The diagram contained a slight perspective clue, since the test line was below eye-level, and as it turned further back so its projection on the screen rose above the vertical. A monocular observer could have interpreted what he might have recognised as a perspective drawing of a horizontal line tilted back, and he could have made an allowance for this and made a guess of its length in space. Some of the observers probably did this, but it is most noticeable that their estimates were considerably below those of the observers whose comments made it clear that they really were appreciating the stereosoopic clues properly.

The second half of the test used the previous drawings pseudoscopically. Having made the drawings for the first half of the test it seemed a reasonable economy of effort to use them as much as possible, but this did mean that the conditions during the second half of the test

were not quite the same as in the first half. The test line started in the plane of the screen and came forward by steps of 20° in accordance with the theory of pseudoscopic transformations discussed in Chapter 4, but as it came forward it also tilted up slightly. This meant that in the second half of the test observers tended to fall into two fairly well defined classes. The first class saw the diagram stereoscopically, and they tended to make very good estimates. When an observer could manage the second part of the test the estimates which he gave were nearly always better than those which he gave for the first part (A.E.K.,G.G.,D.L.J.). This may have been due in some measure to practice, but the remarks which the observers made nearly all indicated that their estimates were better because they definitely considered that the test conditions were easier; and this is also the case with A.H. who re-did the first part of the test after doing the second. The situation was summed up by A.E.K. who said, "The second set seemed far more stereoscopic."

It is difficult to account for this, but the agreement was so pronounced that some explanation is called for. Meanwhile this observation that objects in front of the reference plane are "more stereoscopic" than objects behind should definitely be borne in mind when stereoscopic illustrations are designed.

When assessing these results it should be remembered that certain important evidence is lacking. We do not know what estimates these same observers would make if they were viewing physical objects, such as pieces of wire, in three-dimensional space under comparable lighting conditions. It would be possible to conduct an experiment to decide this

by making models of wire and displaying them on a suitably equipped stage; but this would take us rather far from the present enquiry because the aim is primarily to convey geometrical information by means of stereoscopic diagrams and not to simulate the physical conditions of the three-dimensional space in which we live.

Other observers found great difficulty in doing the second part of the test at all. Many of them were not conscious of the difficulty and they took the perspective clue as indicating that the line was in the same position as before, behind the screen, and made their estimates on this assumption (i.e. they rationalised by ignoring the parallax clue entirely). As these estimates were made on a totally false hypothesis they cannot be compared quantitatively with the previous ones, but they must be interpreted in a different way. Observers such as B.H. and E.H. thought at first that the test line was behind the screen, but as the parallaxes increased they were forced to modify their views, and once they realised what was happening they had no further difficulty in locating the test line in its correct place even on frames where they had been in error before. H.W. was of this type, but he required a comparatively large parallax before he overcame the suggestion made by the false perspective clue, and when the angle of inclination of the test line to the screen was reduced to 40° he suffered a relapse and declared once more that it was behind the screen.

Some people never saw the line come forward at all, but it is striking that in so many cases the observers' estimates were consistently below what they were in the first part of the test (which is as they

should be on the convergence theory) which seems to indicate that even in spite of the belief that they were seeing the same thing as before the stereoscopic factor was influencing their judgement even if the conscious interpretation which they put upon it was the wrong one.

The results of each observer were plotted graphically. Calling the theoretical length of the test line l in each case, the estimated length l', and the inclination of the test line to the screen θ the ratio $(l'/l)^2$ was plotted against sin θ. This was done to see if an observer's inaccuracies could be explained by the hypothesis that he was flattening the diagram on to the screen in such a way that it underwent a uniform strain. Swenson has suggested (1) that stereoscopic diagrams are seen compressed to about 3/4 of their theoretical depth. If this hypothesis is tenable it is possible to associate with each observer a particular value of k, the ratio of estimated depths to true depths. Elementary calculation shows that on this hypothesis

$$\frac{l'^2}{l^2} - 1 = (k^2 - 1) \sin^2\theta,$$

so that if the observations are plotted as described the points should, on the 'uniform strain' hypothesis, lie on a straight line of slope $k^2 - 1$. If the observer saw and estimated everything correctly the plots of his observations would lie on the horizontal line $\frac{l'}{l} = 1$, whereas if he was at the other extreme and saw each diagram as quite flat on the screen, and estimated accordingly, the points would lie on a curve which is very close to, but slightly below, the diagonal line joining the unit points on each axis. Other systematic errors affect the graph in their

own characteristic ways. If, for example, the observer underestimates each time by one (a situation which is quite possible) the points which represent his estimates are moved differently in each case but they lie distributed about the horizontal line $\frac{l'}{l}^2 = 0.8$.

Sixteen observers were tested; the following being of special interest.

A.E.K. (See Fig.6.1) A consistent observer who was confident about his observations. The results in the second part of the test were noticeably better than for the first part. The 'uniform strain' hypothesis applies reasonably well to these results with values of $k = 0.46$ for the observations behind the screen and $k = 0.74$ for those in front.

J.V. This observer did not see the diagrams properly in the second part of the test and it seems that he was systematically underestimating in the first part. The slope of these points suggested a value of $k = 0.46$ or rather greater.

J.W. (See Fig.6.2) His estimates showed a very wide spread but he seemed to see the diagram more deeply than anyone else. The points were far too scattered to be placed on a line, but this observer was perhaps the only one to whom one would be inclined to attach a value of k as high as 1.

D.W. (See Fig.6.3) A plot typical of a person who was not seeing the diagram deeply at all. His remarks confirmed this - but this observer performed excellently on the "Photofinish" test making only one error.

H.W. This observer took some time to realise that the test line was coming out of the screen in the second half of the experiment, and even then he only saw it for large values of θ. But when he had decided that the test was emerging from the screen his estimates were astonishingly good.

x denotes readings in front of the screen; ⊙ readings behind.

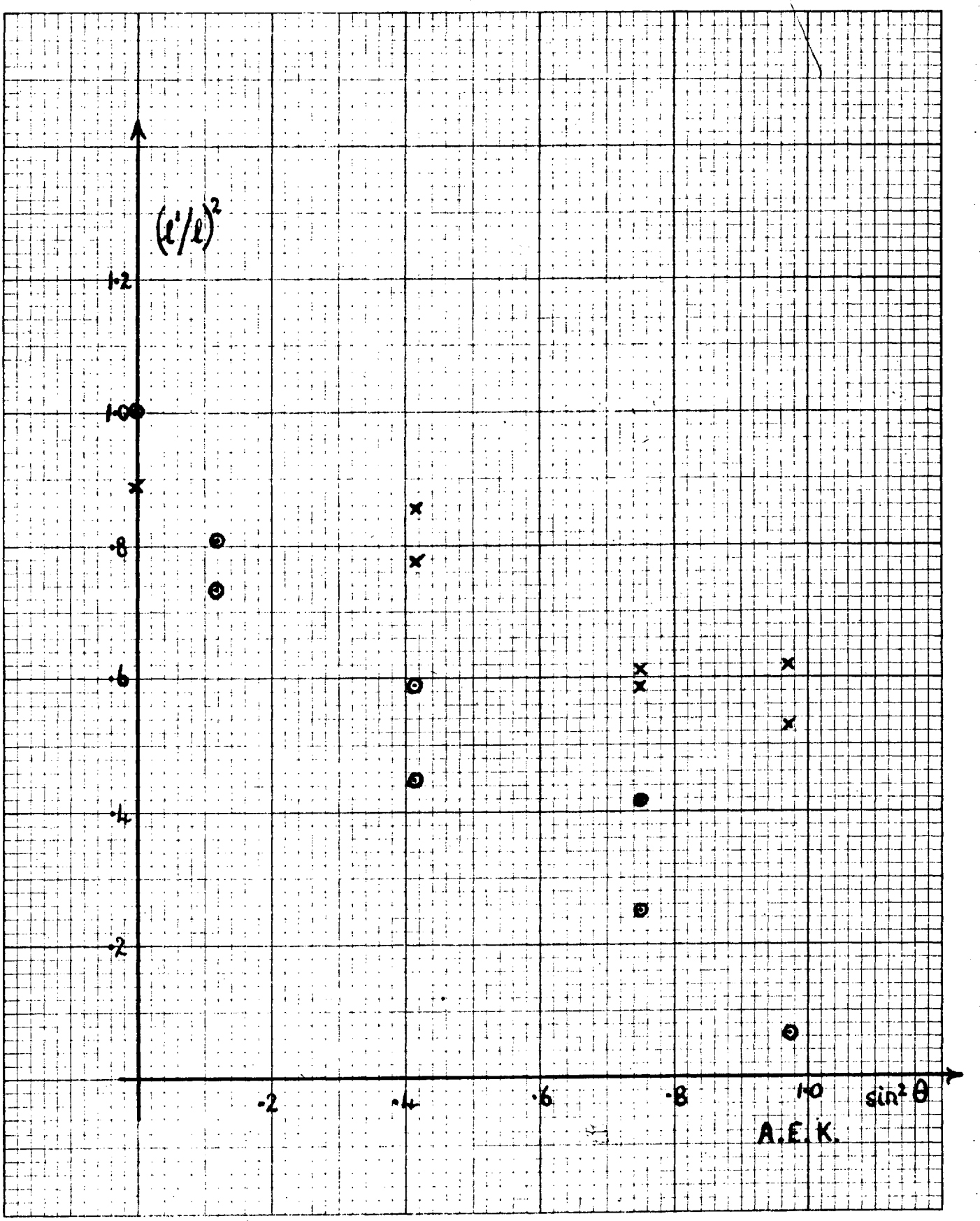


Fig. 6.1

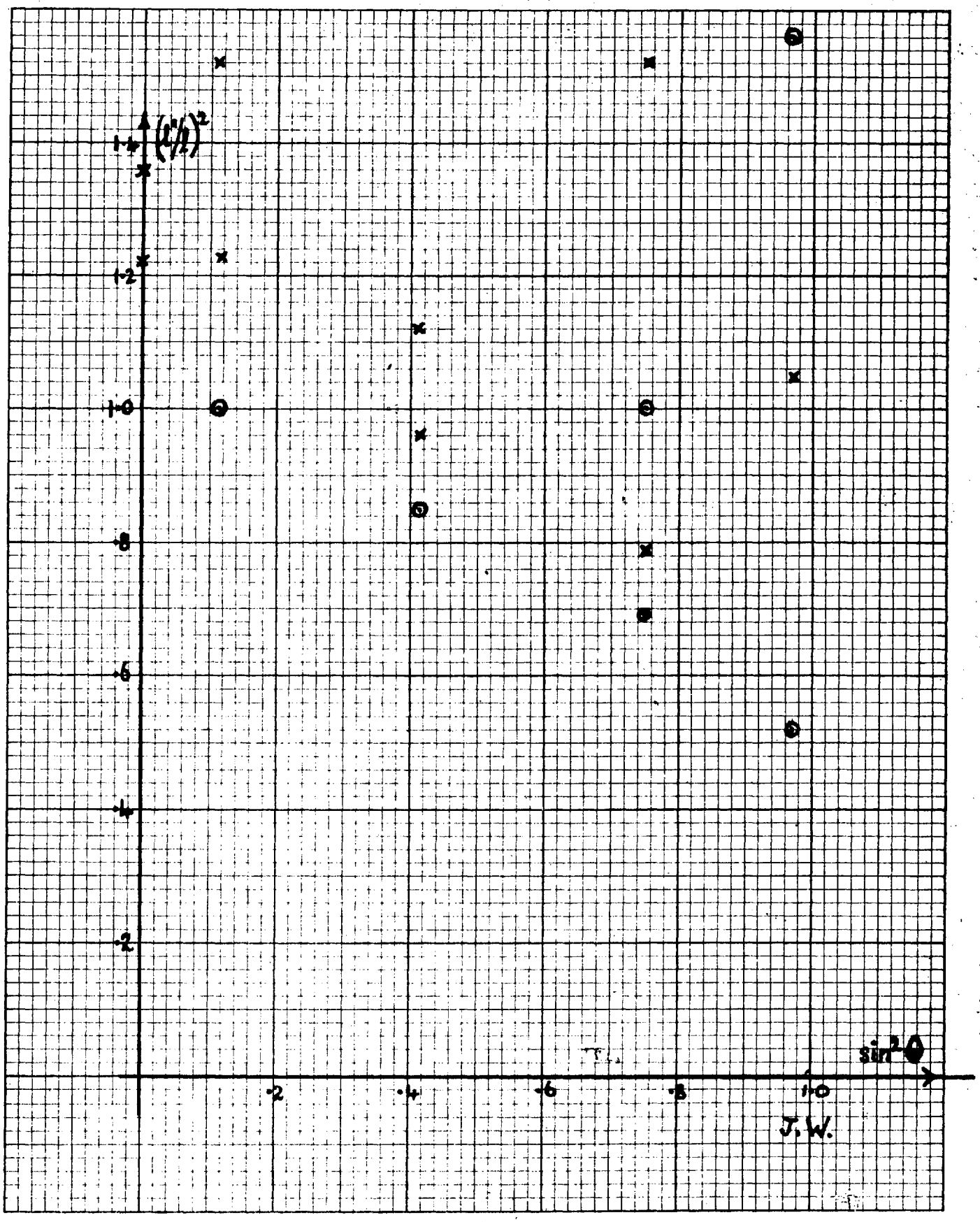


Fig. 6.2

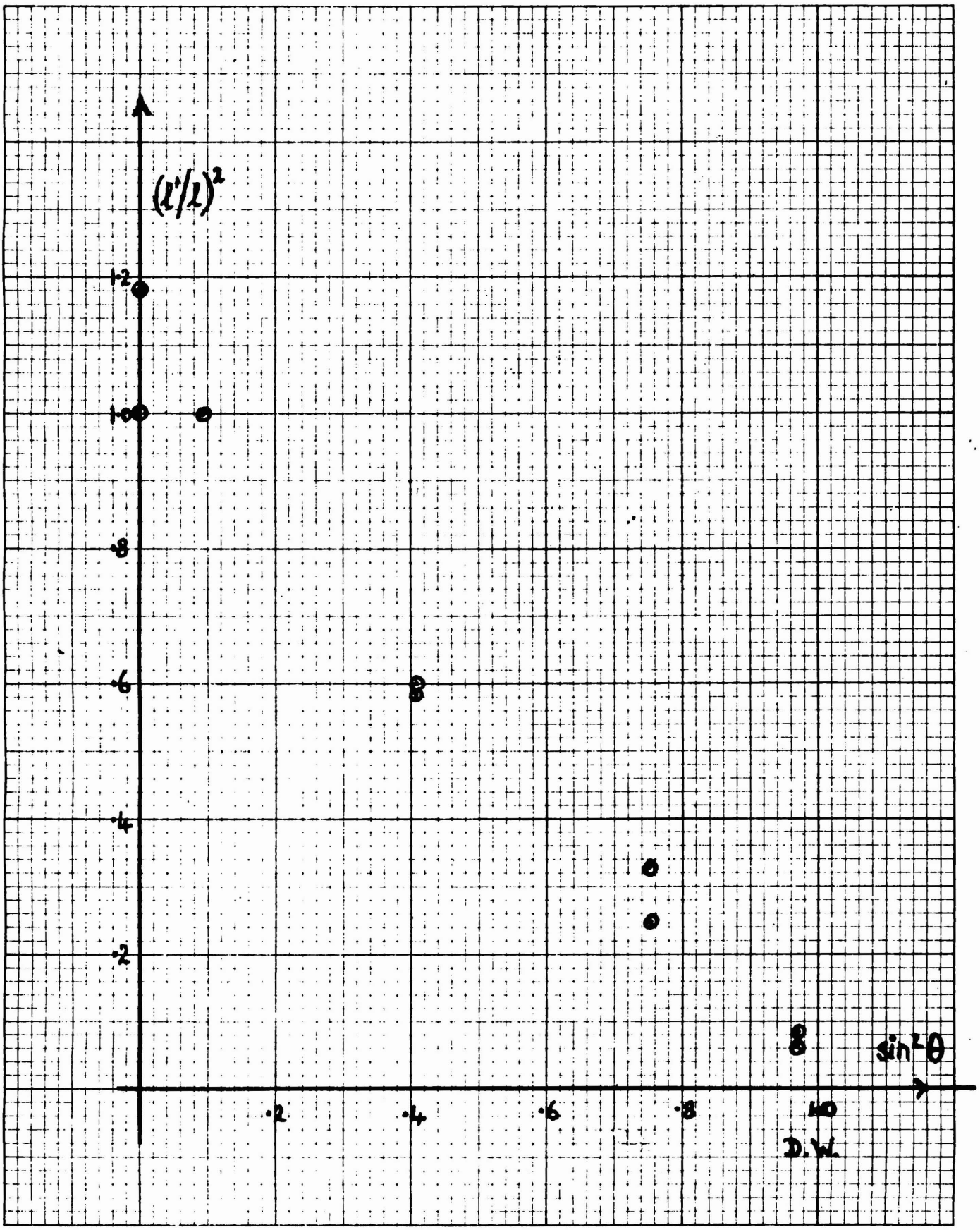


Fig. 6.3

Photographic details. The original drawings were half the linear dimensions of the projected diagram, and were drawn in Indian ink on cellulose acetate. They were photographed on Pan F; $\frac{1}{4}$ second, f16 (slight under-exposure). The negative was projected on two Aldis 1000 projectors, with a 12 ft 6 in. throw giving a picture 38 in. x 50 in.

The observer was 10 ft. from the screen, his eyes opposite the middle of the screen with the projectors behind and slightly above his head.

The line width in the original drawings was 0.02 in., and on projection this became 0.07 in.

Conclusions. i) Observers estimates varied widely; the test was difficult and only about 40% were definitely locating the test line in space properly and estimating accordingly. Underestimates were far more common than overestimates; i.e. with the very small clues given depth was insufficiently appreciated.

ii) It is quite possible that observers could be trained to give better performances under these test conditions, but that was not attempted as stereoscopic films will have to be shown to untrained audiences.

iii) Lines emerging from the screen were harder to locate than lines going in, but when located their lengths were definitely estimated with greater accuracy. To some extent the difficulty arose in this experiment from the "false" perspective clue, but other experiments also suggest that negative parallaxes are more aggressive, and in consequence harder to accept but more suggestive once they have been accepted.

iv) Since nearly all the estimates were too small it might be advisable to exaggerate parallaxes in diagram work even by as much as 50%. This makes drawing easier and increases the relative accuracy. But when harmonising parallax clues with other indications of depth the situation which the viewer is asked to appraise is very much simpler and so depth exaggeration should not be overdone. Remembering the affine distortions which are inevitable when stereoscopic material is projected the sounder solution is probably to sit those with poorer stereoscopic vision in the back seats!

v) This experiment was an overall test of a communication system, from drawing board to the observer's personal estimate. The errors could have been caused in part by errors in drawing and projection as well as by defects in the observers' vision, and we do not know how the same observers would make estimates on three-dimensional wire models viewed under comparable conditions, nor how they normally compare vertical and horizontal distances. These are interesting questions, but they are not relevant to the present enquiry, because diagrammatic stereo work is concerned mainly with conveying abstract relations to the observer and only secondly with duplicating his experience of the physical world.

REFERENCES

1. H.A. Swenson. The relative influence of accommodation and convergence in the judgement of distance. J.G.Psychology, 7, pp.360-380, Oct. 1932.

III. EXPERIMENT "OVERLAP"

The convergence theory suggests that if a stereoscopic pair of pictures is viewed with incorrect horizontal registration very serious distortions result. It is generally agreed that whilst some distortion may arise it certainly does not do so to the extent which the theory predicts. This appears to be true whatever the method of viewing employed, and this is the main factor which has led many people to abandon the convergence theory for the retinal disparity theory.

Whilst it is easy to speak of "the retinal disparity theory" it is not quite so easy to decide exactly what the theory is, for apart possibly from Luneburg (neither of whose theories is really a theory of retinal disparity at all) no one seems to have put this theory on to a quantitative basis. In so far as the theory is quantitative it would seem to be merely a re-statement of the convergence theory with the additional qualification that observations are unaltered by a translation of the image across the retina of either eye, and in consequence the observer loses one degree of freedom and is unable to assign an absolute distance to the objects viewed if he works on stereoscopic data alone.

Observations of this type are notoriously dependent on the particular viewing conditions employed so it seemed desirable to investigate the point using our own particular projection conditions. With this in mind a set of still stereoscopic slides was prepared in which the observer saw four vertical lines. The situation is shown in plan in Fig.6.4.

With correct adjustment the four lines were the edges of a prism whose cross-section was a parallelogram; in particular the two near lines

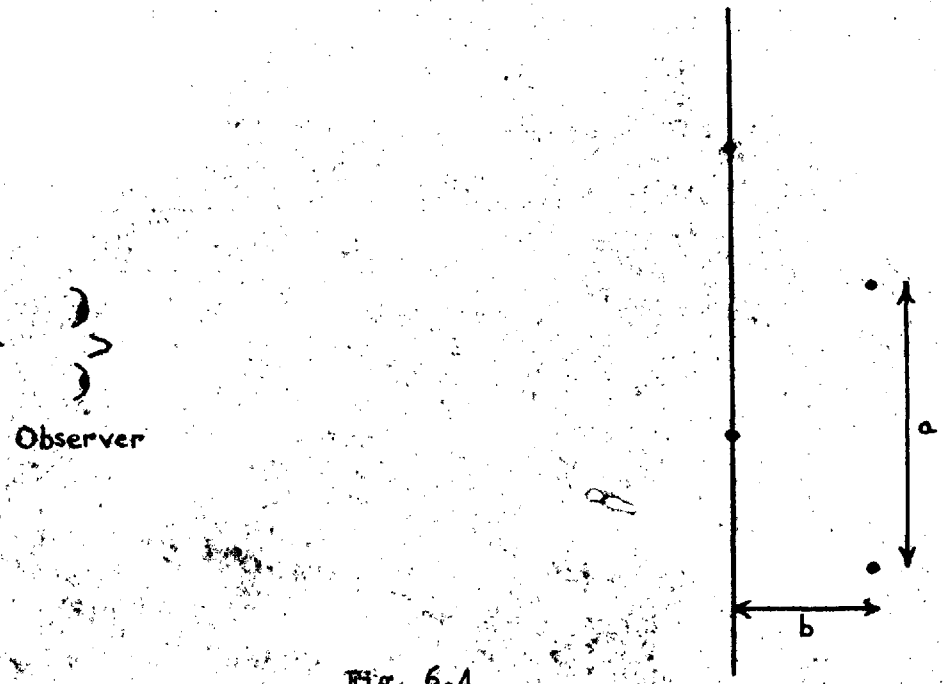


Fig. 6.4

and the two far lines were an equal distance apart. The convergence theory indicates that this equality should disappear if the overlap of the two pictures is varied. One of the pictures carried a scale and the other a pointer which enabled the operator to control the overlap. The aim of the experiment was to investigate the distortions as the overlap was varied; in particular an attempt was made to see if the observer could instruct the operator to produce an alignment in which the front and back lines were equally spaced and to compare this with the setting which was theoretically correct.

The results gave no support to the convergence theory and they did not fit into any coherent theoretical pattern. This is not say that the

results of the experiment were entirely negative, for in the present context the only apparatus available for the projection of moving stereoscopic diagrams, the Paillard-Bolex system, has no control over this paramater; so if the effect of varying the overlap of moving diagrams required extensive study the existing apparatus would be quite inadequate. The experiments confirm the opinion of the manufacturers that control of this parameter is not important, and one preset value suffices.

The experiments also confirm some of the conclusions of the "Guesswork" experiment that the metrical information which a stereoscopic diagram can convey is very limited and that it is a vain hope to expect diagrams of this type to bring out the finer details of metrical relationships.

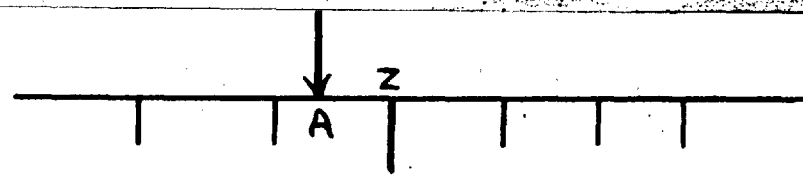
Apparatus. The drawings were photographed on 35mm film and projected using two Aldis 1000 projectors. The original drawings were black on white and the negatives were projected, giving white lines against a grey background. Six situations were drawn in all, but two were photographed twice to provide the observers with additional material for practice.

The observer viewed the screen at a distance of 10 feet. The magnification from the drawing board to the screen (which was 38" x 50") was 4 times (linear). The dimensions of the six scenes are given in Table III, together with the theoretical parallaxes required on the drawing board. Scenes iii) and ii) were the ones which were also used at the beginning for practice.

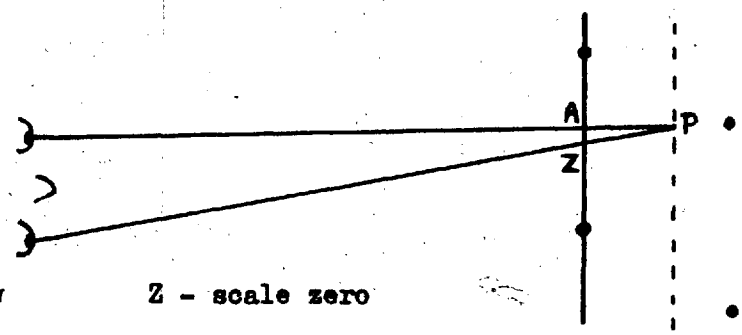
	a (inches)	b (inches)	Full parallax on drawing board (ins.)
i)	20	24	.104
ii)	18	-12	-.0695
iii)	16	18	.0816
iv)	16	-18	-.110
v)	14	-24	-.1565
vi)	20	36	.1445

Table III

The right eye slides all carried an arrow which moved over a scale on the left-eye slides to enable the operator to measure the overlap of the two. The observer was instructed to ignore this scale so that his judgement was not prejudiced. This warning proved unnecessary because a quite unexpected effect was observed. Over a very wide range of overlap observers saw the arrow in line with the zero of the scale and interpreted the depth of the scale in space accordingly (Fig.6.5). A most remarkable example of the determination we have to make sense out of stereoscopic nonsense.



Combined arrow (left eye) and scale (right eye) as seen on screen



A - arrow Z - scale zero

The scale and arrow are located in a plane through P.

Fig. 6.5.

Theory. In Chapter 4 we investigated the effect of projecting stereo pictures with a systematic alteration to all of the parallaxes, and showed that if all the parallaxes on the screen are increased by an amount 2p the following distortions are predicted by the convergence theory:-

$$x' = \frac{x}{1 - P\zeta} \qquad y' = \frac{y}{1 - P\zeta} \qquad \zeta' = \frac{\zeta}{1 - P\zeta} \quad \text{where} \quad P = \frac{p}{ws}$$

Only the values of x concern us here, and we will call the abscissae of the two lines in the screen x_1 and x_2 when the figure is correctly aligned, and the abscissae of the other two lines x_3 and x_4. Then

$$x_1' - x_2' = (x_1 - x_2)/(1 - P\zeta_1) \qquad (\zeta_2 = \zeta_1)$$

and

$$x_3' - x_4' = (x_3 - x_4)/(1 - P\zeta_3) \qquad (\zeta_4 = \zeta_3)$$

Now $x_1 - x_2 = x_3 - x_4$ so

$$\frac{x_1' - x_2'}{x_3 - x_4'} = \frac{1 - P\zeta_3}{1 - P\zeta_1} = \frac{ws - p\zeta_3}{ws - p\zeta_1}$$

$$= \frac{ws - p(z_3 + s)}{(w - p)s} \qquad \text{since } \zeta = z + s, \text{ and } z_1 = 0,$$

$$= 1 - \frac{p}{w - p}\,\frac{z_3}{s}$$

For most practical purposes $p < w$.

If $z_3 > 0$ and $p < 0$ the ratio is < 1, and $x_3' - x_4'$ appears too big.

If $z_3 < 0$ and $p > 0$ the ratio is > 1, and $x_3' - x_4'$ appears too small.

Hence whether we are dealing with the case when the lines in the screen are the front pair or the back pair (as the observer sees them) if $p > 0$ the back pair should appear further apart than the front pair. By "appear further apart" we refer to their position in space as the observer should

assess it if he conforms to the convergence theory - we do not refer to the angle which they subtend at the observer's viewpoint.

The Experiment. 16 tests were made on observers who, for the most part, were known to be competent as they had performed well previously on other tests of this series.

Many of them made the unsolicited comment that as the overlap of the diagrams was increased the entire figure seemed to approach, and as the overlap was diminished the figure appeared to recede, but all found great difficulty in making the quantitative judgements that were demanded.

With the first few observers the attempt was made to get the overlap adjusted so that the front and back lines appeared equally spaced by adjusting the overlap in accordance with the observations, increasing the overlap (i.e. decreasing p) if the back lines appeared too far apart and decreasing it if they appeared too close together. It soon became clear that this approach was taking too much of the theory for granted, and so the procedure was modified. Each scene was presented to the observer with a variety of overlaps arranged at random, and he was asked if the spacing of the back pair of lines was more or less than the spacing of the front pair.

An observer who conformed to the theory would have answered "more" whenever $p > 0$, "less" when $p < 0$ and "equal" when $p = 0$; but this type of behaviour was not to be found. Most observers gave replies which were random, although a few appeared just a little better. Precise use of the word 'random' in this context would presuppose a certain statistical procedure, but in the circumstances it was not felt that the labour of a

detailed statistical analysis was justified. Such an analysis would, in any case, only be worthwhile if at the same time an attempt was made to relate the results to the performance of observers on similar tests using physical objects, and to their opthalmic deficiencies (if any).

The estimates made about transverse lengths in different frontal planes were most unreliable. It may well be that the same holds true for similar situations in which observers are looking at physical objects, but we have no knowledge of experiments covering this point. For such experiments it would be necessary to erect a 'stage' on which observers saw vertical rods which were evenly lit in such a manner that all but the strictly stereoscopic clues to depth had been removed, and arranging such conditions would call for very special care.

Six observers (out of 15) consistently said that the distance between the back lines was 'less' or 'equal' however the overlap was adjusted. This indicates a decided tendency not to allow sufficiently for the apparent decrease in size with distance.

Whilst the front two lines were always intended to be at equal distances from the observer, and the back two also, slight, unavoidable errors in drawing altered this slightly and presented a situation that was rather like that in experiment 'Photofinish'. A number of observers commented on this; and this confirms the very high degree of accuracy which observers can attain under the conditions of that experiment.

One observer (J.C.) did noticeably better than the others on this test, and did better still when he was retested later. Therefore it might be possible to train talented observers to perform comparatively well,

but this was outside the scope of the enquiry. Such talent is rare, and it must not be expected in general audiences.

Conclusions. Varying the overlap of stereoscopic drawings does result in an observeable approach and regression of the entire picture, but over a wide range of variation the secondary effects on the transverse dimensions which the convergence theory predicts are not found.

The absence of such a control on the Paillard Bolex projection system is therefore not a serious drawback.

When information is presented stereoscopically the comparisons which observers who are proficient but not specially trained make of transverse dimensions in differing frontal planes are quite unreliable. There is no evidence of how this compares with similar observations on physical objects; but in this connection the following question arises:-

"Do we wish stereoscopic diagrams to simulate physical reality, or do we wish them to convey abstract information?"

The experimental techniques to be followed must vary according to the answer given. Experiments have been done in U.S.A. relating 'objective geometry' and 'subjective geometry', but there is not as yet any generally accepted body of knowledge on this topic.

Footnote. In the three experiments described in this chapter the structure (grain etc.) of the screen played no part. This was not allowed for in the theory in any way and no comments by observers indicated that they were at all influenced by any awareness that there was a screen at a particular physical distance. Indeed, the very few references made to objects being "in front of (or behind) the screen" were as often as not quite incorrect.

CHAPTER SEVEN

METHODS OF STEREOSCOPIC DRAWING

The theory of stereo drawing was described in Chapter 4, and using the same notation we now describe how this theory may be embodied in a practical scheme. There are three principle methods, orthodox perspective as practised by artists, the well-established method of constructing a perspective drawing from side view and plan, usually known as 'the architects' method', and methods depending on calculating the position of key points and plotting their positions on graph paper. The calculation may be done on a machine (1).

With all of these methods it is convenient to take the drawings for the right and left eye from a common master by tracing; and whilst this is not essential it is also in accordance with animation practice, and throughout the subsequent discussion it is assumed that the final drawings will be produced this way.

Conventional methods of central perspective use a vanishing line, corresponding to the line at infinity in the horizontal plane, with a range of vanishing points on it corresponding to the points at infinity of parallel pencils of lines in the horizontal plane. If perspective drawings for the right and left eyes are to be prepared from the same master then there must be a separate range of points on the vanishing line for each eye, the two ranges having a uniform displacement from one another of $2w/m$. This method suffers from the disadvantage that it requires an excessively wide drawing board - at least five times as wide as the finished drawing. This can be avoided by having a mechanical

attachment to extend the vanishing line artificially, but such an attachment is inconvenient, and it must be very rigid and accurately made if it is to be suitable for stereoscopic work.

The architects' method has been described by Rule (2). In this method the drawings for the two eyes are constructed from a side view and plan as shown in Fig. 7.1. Again the two eye positions have to be separated by a distance $2w/m$. This method requires much less space than the previous one; and the greatest waste of space now comes from the distance to the two eye positions. This has to be s/m and is once more rather larger than may be desirable on a drawing board. However this time these points can quite easily be situated on a mechanical extension to the drawing board as only the two fixed points are ever used and the extension can be by a thread and does not have to be rigid, as before. The previous method calls for a complete range of points, one for each direction in the horizontal plane of the space being depicted. In addition, proceeding as we do here it is possible to save space on the drawing board by suitable choice of the factor m, and to achieve the same results in considerably smaller space than Rule.

Rule gives many practical rules for getting a realistic effect in drawings, and these should be followed wherever possible. The comments on the appended set of stereo illustrations bring out many of the rules and the illustrations show the effect of working both with and without them.

A disadvantage of Rule's method is that errors in parallax are additive, and since these are far and away the most important errors in stereo illustration it is convenient to adopt a procedure which avoids this unfortunate feature.

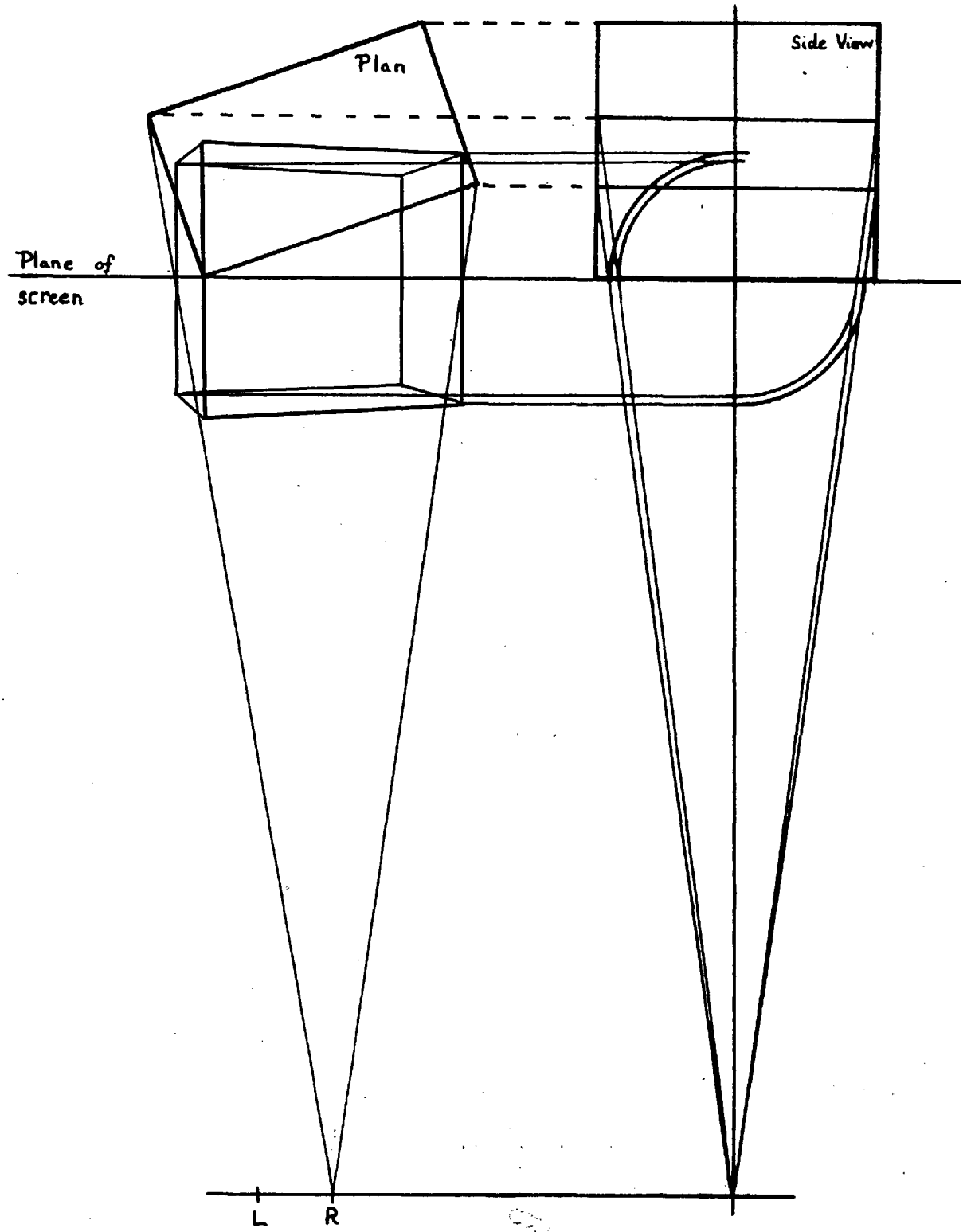


Fig. 7.1

The equations of Chapter 4 may be used to calculate the positions of key points in a diagram, which can be plotted on graph paper to give the master drawing. This is slower than purely geometrical methods, but errors do not accummulate. Therefore the best method of all is perhaps a combination of the two. Rule's method can be supplemented by adding a scale of parallaxes to the plan drawing - it will be remembered that all points at the same depth have the same parallax. These can either be listed on a numerical table or plotted on a graph with an exaggerated scale. It is obviously best to standardise on a convenient set of parameters, and the values $s = 10$ ft., $w = 1\frac{1}{4}$ ins. and $m = 5$ can be recommended. The drawing for one eye can be constructed by the architects' method and Rule's extension of the method can be used to locate the main outlines of the drawing for the second eye, but to ensure the greatest possible accuracy the parallaxes should then be read off the scale, transferred to callipers and marked on the drawings. A scale of this kind, used for stereoanimation, is given in Fig. 8.3.

Graph paper has inaccuracies which are obvious to the eye, and whilst it is quite satisfactory for setting up X and Y coordinates it is unwise to use graph paper calibrations to measure parallaxes. The ideal instrument would be callipers with a calibrated vernier head, but such instruments are not readily available. The head of a standard pair of callipers can be graduated so that accuracies of better than .01 in. can be obtained, and as subsequent errors in positioning points and lines on the drawing can scarcely be reduced below this it is probably sound enough practice to work only to .01 in. on callipers, and if this is done

then they can be set with sufficient accuracy on a diagonal scale.

Errors in parallax can always be made relatively smaller by increasing the size of the drawing. (In this connection it is, of course, immaterial whether the final pictures are to be projected or viewed with a stereoscope since it is the angle which the drawings subtend at the eye which matters.) However when animation is involved as well it is not practical to use very large drawings, since so many are needed and they have to be handled on the animation rostrum. Questions of cost and convenience of storing also limit the size. Therefore it is good practice to work with a field of something like 11 ins. x 8 ins. which is established as a standard size in ordinary animation. Two drawings like this, side by side, mean that the field to be photographed is 11 ins. x 16 ins., and it is not convenient to work with anything much bigger.

Rule is chiefly concerned with black and white drawings, and his recommendations sometimes need modification if colour is used. For example it is not necessary to make the razor cut which he recommends where two skew lines cross if the lines are of different colours; but this point and a number more are covered adequately by the accompanying collection of still slides. As we often emphasise, the designer of stereo diagrams has to consider all the factors involved in the preparation of ordinary technical illustrations as well as the problems peculiar to stereoscopy. Depth perception depends on many simultaneous factors of which retinal disparity is only one, and non-stereoscopic factors on their own can give very clear indications of three dimensional relations. The accompanying illustrations use perspective, tapering lines and other

firmly established devices of the draughtsman's art to supplement the stereoscopy, and the slides show the effect of incorporating these other factors to varying extents.

Two further methods of drawing should be mentioned. The crystal lattice illustrations by von Mieses and von Laue (3) show that excellent stereo illustrations can be made by parallel projection. This method is obviously much easier than central projection as parallaxes are directly proportional to the distances of their associated points from the picture plane. It is so simple to adapt conventional methods of drawing by parallel projection that further explanation is unnecessary, also the techniques which we have described earlier are very easily modified for parallel projection.

Parallel projection has the disadvantage that drawings based on it are necessarily deprived of the normal perspective clues, but in many technical fields parallel projection is a firmly established medium of communication, and in these it would be quite natural to construct stereo diagrams by the same technique. It is however a convention that parallel projection is suited for drawings of things which are 'small' (crystals, small pieces of wood or metal work) but not for things which are 'big' (architecture). If the convergence theory is assumed for stereocinematography then the objects which concern us are 'big'; but many writers on stereoscopy consider that there are not proper grounds for considering that a stereo picture has a definite size, and since our concern is in any case primarily with abstract diagrams there is no doubt that diagrams constructed by parallel projection would be perfectly adequate for many

purposes. This is especially true when the properties are affine; and we give elsewhere the general recommendation that stereo diagram films should be based on the affine group of transformations.

The last drawing technique to consider is an extension of another established technique of the architect's office - the use of a specially constructed perspective grid, which functions rather as a set of three-dimensional coordinates transferred to the paper. Such grids are easily constructed using the principles which we have described and they could well be useful in a large programme of stereo work.

APPENDIX

Notes on the accompanying stereo slides

The drawings for these slides were made on cellulose acetate. This enables crossing lines to be drawn on different sheets and avoids any tendency for colours to run at points of intersection, but it means that opaque colours must be used. The drawings were made on punched sheets for accurate registration, with a field size of 11 ins. x 8 ins.

Some of these slides illustrate the use of some depth clues to the exclusion of others, and some are deliberately included as examples of mistakes to avoid, rather than as examples to follow.

The slides in sets of three labelled L, M and R (left, middle and right) have parallaxes in the ratio 1:2, and different combinations produce the same drawing with three different parallaxes.

The slides were photographed on 35 mm Kodachrome on an Agfa 'Silette' camera, with a close-up lens attachment; and they are mounted in standard 2" x 2" cover glasses. They may be projected on two standard projectors fitted with polaroid filters or inspected on the accompanying viewer. (The slides were not jigged when they were mounted, and may require a little fresh alignment each time.)

General Conclusions.

i) Although the parallaxes involved are very small drawing techniques which lend themselves to animation if employed with care can produce satisfactory pictures.

ii) Tapering lines, colour, blocking and perspective all increase the effectiveness of the diagram.

iii) Whilst relative parallaxes should be as correct as possible the absolute scale of parallaxes is not as important as the simple convergence theory indicates. It may therefore be wise to exaggerate parallaxes a little so that the errors which are inevitable on the drawing board are proportionately less.

iv) The slides are equally acceptable when viewed on the accompanying viewer or by projection; the condition for good viewing being merely that the drawings subtend the correct angle at each eye. This depends on s, which may be varied when the slides are projected on a screen, but not when they are viewed on a viewer. However, the value is not critical.

Detailed Commentary on the slides.

In the notation of Chapter 4 these diagrams were constructed with a standard value of m = 5 but with varying values of s, the theoretical viewing distance in projection. A projector with 100 mm lens requires a 13 ft. throw to produce the theoretically correct viewing conditions.

Some of the slides are described as "twin-pairs". These are pairs of slides taken from the same drawing, with one of the slides reflected right-left on mounting. In this way a stereo picture of a symmetrical object can be produced with only one drawing.

Set No.1. Serial Numbers 1/1 to 1/21.

Shot on Kodachrome A. Photoflood lighting was not available and so this set is under exposed and not very suitable for projection, but all the slides can be seen satisfactorily on the viewer. Unless otherwise stated s = 8 ft.

1/1-2. Twin pair. Rectangular box. Green drawing ink on Ethulon.

Parallaxes calculated. Range on drawing board from 0 to -0.25 ins.

1/3-4. Twin pair. A set of circles on a cylinder normal to the screen.

1/5-6. Twin pair. Latitude circles on a sphere viewed along its axis. m = 5, s = 100 ins. Rose red and green drawing ink on Ethulon. Coordinates calculated and plotted on graph paper. Range of aprallaxes -0.14 ins. to +0.11 ins. This picture is confusing, and was expected to be so. The aspect is very unfavourable, and the red and green circles are liable to appear relatively displaced. (Some observers confirm this.)

1/7-8. L and R. Hyperboloid of one sheet. Fifteen generators, twist of 144°, viewed down through the middle. Scarlet lines on Ethulon. This picture was later redrawn with tapering lines, and this increased its comprehensibility.

1/9-10. L and R. Hyperboloid of one sheet. Fifteen generators, twist of 144°, viewed obliquely. Parallax range 0 to -0.25 ins. Cobalt blue ink on Ethulon. One line in this diagram looks very wrong. This does not appear to be a fault in the drawings, but the impression arises from the extremely oblique angle at which the line is seen.

1/11. A spare copy of 1/12.

1/12-13-14. L, M and R, spaced 2:1. Stellar Octangular (i.e. two interlocking tetrahedra). s = 10 ft. Scarlet, green and cobalt ink on Ethulon. A 'solid drawing' in a parallel projection. Parallax ranges from 0 to +0.15 ins. with M and R, and correspondingly greater with the other combinations. The theoretical parallax range is 0 to +0.16 ins., so M and R should give a picture which is very slightly flat, and others should appear considerably distorted. Note the effect of the parallel projection.

1/15-16-17. L, M and R, spaced 1:2. Desargues' Theorem in space. Parallax range, -0.11 ins. to +0.15 ins., with L and M, and correspondingly higher with the other pairs. Cerise and veridian ink on cellulose acetate. Thin lines, not tapered. The combination of M and R seems to give the best result. This configuration is invariant under projective transformations.

1/18-19. L and R. Rectangular box. s = 6 ft. Drawn by Rule's method and traced in black ink on cellulose acetate. Parallax range 0 to +0.14 ins. No razor cuts where lines cross. Thin lines which did not taper. A version with tapering lines appears later. It is instructive to view these slides pseudoscopically.

1/20-21. Twin pair. Rectangular box, 24" x 30" x 21". s = 8 ft. Cobalt blue ink on cellulose acetate. Parallax range 0 to -0.14 ins. The aspect is very poor in this diagram.

Set No.2. Serial Numbers 2/1 to 2/21.

Shot on Kodachrome A. Photoflood lighting. Weston meter II reading 50-75. 1/5 second at apertures varying from f5.6 to f16, but mostly at f8.

2/1-2-3. Colour charts.

2/4-5. L and R. Hyperboloid of one sheet. A redrawn version of 1/7-8. Parallaxes increased to a range of 0 to -0.30 ins. Cobalt blue on Ethulon, with tapering lines. Three of the lines at the top left have razor cuts at the crossings. The parallaxes are quite large, and the picture appears very deep. Even more taper could profitably be used.

2/6-7. L and R. Rectangular box. 1/18-19 redrawn with tapering lines; which is definitely helpful. This tapering can be 'overcome' if the slides

are viewed pseudoscopically.

2/8-9. L and R. Modification of the Desargues' configuration on 1/15-16-17. Bigger cuts have been inserted where the lines cross. No tapering.

1/10-11-12. L, M and R, spaced 1:2. Hyperboloid of one sheet seen from one side, slightly on the slant. Parallax range +0.062 ins. to -0.058 ins. on L and M. Trox green on cellulose acetate, with tapered lines. A very effective diagram, but close examination reveals one line that is out of place.

2/13-14-15. L, M and R, spaced 1:2. The previous figure with the addition of extra lines intended to show the generation of the hyperboloid by two intersecting places. This is a complicated figure, and it would be clearer if it was redrawn with a more distinctive colour scheme.

2/16-17. L and R. 2/4-5 with the addition of extra lines. This diagram shows that where lines of a different colour, or even of the same colour intersect the point of intersection should be marked in some way. The mere stereoscopic presentation of the two lines without some indication of the point of intersection is not in itself enough.

2/18-19. Twin pair. A redrawn version of some of the circles of 1/3-4. It is instructive to view these slides pseudoscopically, when the circles are located on a cone instead of a cylinder. One of the circles is noticeably displaced because its parallaxes are inaccurate.

2/20-21. Colour charts.

Stereoscopic Paintings by Norman McLaren

These three stereo paintings were made by McLaren in the late 1940s and slides taken from them are included here because of their great historical and artistic interest. The originals were paintings, about 11" x 8", which were viewed on a Brewster stereoscope. Seeing these paintings led Spottiswoode to invite McLaren to produce his two stereo films for the stereo cinema at the Festival of Britain in 1951 (see Chapter 8). The methods employed in making them were largely empirical.

These paintings have not been previously reproduced or published anywhere, and I am indebted to Mr. McLaren for permission to include the reproductions here.

REFERENCES

1. A.R. Curtis, J.G.L. Michel and E.D. Sully. Central Projection and Stereoscopic Diagrams on the Differential Analyser. Math.Gaz.XXXIV, No.310, p.276, Dec. 1950.
2. J.T. Rule. Stereoscopic Drawings. J.O.S.A., 28, 313-322, Aug. 1938.
3. Von Mieses and von Laue.

CHAPTER EIGHT

STEREOSCOPIC ANIMATION

The stereo system to be employed. Unless it is to involve a completely new range of equipment there is no doubt at all that a stereoscopic film projection system for academic use must be built round the 16 mm gauge which is standard in all educational institutions. This means that the only system which is readily adaptable to existing projection apparatus is the system developed by Bolex in Switzerland. The taking part can be fitted to the Bolex H.16 camera and consists of a lens, followed by a system of prisms; ultimately two images are recorded side by side on standard 16 mm film. For projection there is another twin lens assembly provided in a variety of mounts which will fit most 16 mm projectors. This is composed of two Petzval type lenses of 20 mm focal length with a working aperture of f 1.6. Their optical axes are set parallel and spaced 5.6 mm apart. Two polarizing filters are set at 90° to each other and both at 45° to the horizontal, an orientation which is standard in cinema work. The film must be projected on to a metallized screen and viewed through polaroid spectacles.

Since no live action photography is involved in this programme we are only concerned with the projection system, and the pairs of stereo drawings which have to be photographed can be taken side by side on the animation rostrum with an ordinary animation camera used in the normal way.

This system has proved extremely convenient, and it has been possible, for example, to send copies of a film to Berlin and to the U.S.A. and to have it projected with apparatus available there; and also to arrange

performances in various parts of Britain, the projection on some occasions being done by operators with no previous experience of the system at all. Whilst other 16 mm systems of projection have been developed* they do not appear to have any technical advantages over the one adopted, and in any case they are not generally available. There are reports from Russia of a stereo system in Leningrad for which the audience do not have to wear spectacles, and such systems have also been studied in the West, but no apparatus of this type is currently available in Britain.

Apart from economy a one projector system has other advantages over a system using two. Synchronisation of the projectors is a problem when two films are used, and there is the possibility of relative movement between the two when taking, printing and projecting. Levonian has discussed the magnitudes of these errors in a commercial system in the U.S.A. (5) and on his figures the relative movement between the two films is much the worst at the camera stage, which causes an intermittent variation in the angular convergence of the eyes thirty times as great as that caused by irregular movement on projection. No measurements have been made on the steadiness of the gate of the animation camera employed for this programme - which was a Kodak Model A with special conversions - but since the film passes through the gate of an animation camera at only one twenty-fourth of the speed involved in live action photography, and since the claws of the animation camera had been specially built up to do, as far as possible, the work of register pins there is every reason to suppose that the photography was steadier than that which Levonian measured - or at least no worse. The figure for intermittent variation in

* See references in Chapter 2.

convergence which Levonian quotes, 0.00082 radians for each camera, corresponds to an error on the animation table with the present system of about 0.02 ins. Now it is not unreasonable to aim to produce drawings to this accuracy, so while the difficulties of working to the small parallaxes involved in stereo-animation have long been recognised it is possible that they have been exaggerated; and indeed it would seem that with care a draughtsman should be able to produce stereo-animation of about the same order of accuracy as live action stereophotography.

In making this estimate we have in mind the stereo-animation of technical drawings, as it is obviously a very different problem to animate, say, a cartoon animal which is to appear solid and not merely a cardboard figure located in a plane at a certain distance. Also, in cartoons the intermittent variations in convergence angles are variations in separate parts of the picture, whereas the variations which Levonian discussed were variations over the entire field as a whole. Only experiment can decide whether or not the two cases are comparable.

Since the Bolex system employs 16 mm film with two pictures side by side the resulting stereo picture should be compared for quality more with 8 mm "flat" films than with 16 mm. In the present programme a screen size of 38 ins. wide and 50 ins. high was the standard, which corresponds to a projector throw of 17 feet. Since the axes of the two projection lenses are parallel, and since we have seen that the exact amount of overlap of two stereo pictures is not of major importance it follows that the exact distance of the projector from the screen and the corresponding size of picture is not oritical. The same is certainly not true for some other

systems, especially when live action is involved.

All of this makes it very simple to arrange a theatre for projection by the present system. Polaroid filters cut down the brightness of the available illuminations but even so the film Resonant Cavities has been shown to audiences of over one hundred. When the audience is as big as this the affine distortions at the back of the theatre are larger than is desirable, the films being designed for a viewing distance of 10 feet, but the system can certainly be used adequately with audiences of over fifty, and this is large enough for many academic purposes.

It is a feature of the Bolex system that the screen is narrower than it is high. Long association with cinema conventions makes this an unfamiliar format, but for technical illustration it is just as convenient as the customary film picture shape. It is a shape which is standard in technical books, and the accompanying Fig. 8, taken from Gibson (2) giving the field of binocular vision is an argument that ordinary films should be made with a "horizontal" format and stereo films with a "vertical". However the last point is academic when, as here, the screen occupies only a comparatively small, central part of the total binocular field.

The general techniques of stereosynthesis. The first three-dimensional cartoon films were three entertainment films which were made for showing at the stereo cinema at the Festival of Britain in 1951. They were Now is the Time and Around is Around by Norman McLaren, and The Owl and the Pussycat by Halas and Batchelor. McLaren's account (7) is the only available paper on the problems of stereo animation. He summarises six basic techniques:-

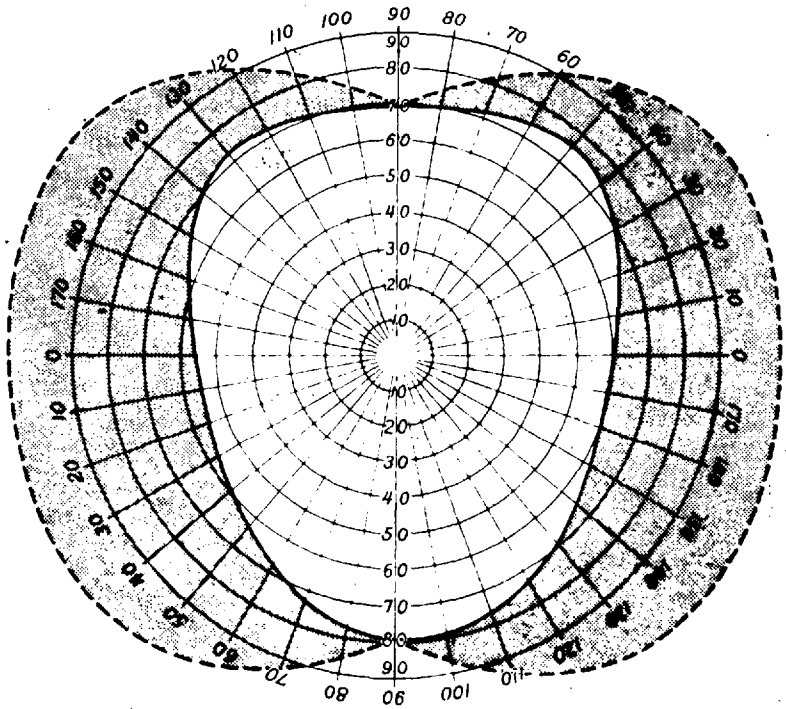


Fig. 8.1
~~FIGURE 47.~~ The Visual Fields of Each Eye and the Binocular Field, as Measured with a Perimeter

i) use of stereo pairs of drawings, with parallaxes drawn in,

ii) double punching of single cards or cells,

iii) the use of movable cut-outs,

iv) movements of the horizontal panner under the animation camera,

v) horizontal panning or lens shifting in the optical printer,

vi) frame stagger on horizontal action shots.

vi) is the method often described as autostereoscopy (9), that is to say the same picture is presented to the right and left eye at different times. This has certain advantages as it enables a stereo picture to be produced from one ordinary piece of photography. In this way, for example, a stereo picture has been produced from a film of a city photographed from an aircraft - a technique which is just an extension of the method of producing stereo aerial reconnaissance pictures with a single camera. This type of stereo-photography is usually very 'rule of thumb', for in

entertainment work as long as the picture has depth it does not matter very much how accurately it is rendered. For serious work the method is very restricted, but we describe it a little more fully later in the chapter.

Methods ii) to iv) depend in various ways on the principle that all objects in a frontal plane at a given distance from the observer carry the same parallax on the screen. They are based on the same principle as a drawing instrument consisting of a frame with a screw thread for lateral displacement which was invented by Rood as long ago as 1861 (8). These methods were the main ones employed in the three films named above, and none of these films employed method i). But for serious mathematical work i) is almost the only method which can be employed at all systematically, and the future of scientific and mathematical stereo illustration depends upon its development. Methods ii) to iv) produce a type of stereo picture which may be described as 'multiplane'* rather than as truly stereoscopic. The picture has no gradation in depth, but consists of a series of stage flats. It is possible to vary the depth of any given plane containing some of the action, but this is very different from building up a picture in which there are continuous variations in depth all the time. These multiplane methods have a restricted application to technical illustration but they are only of minor importance, not being anything like flexible enough to give the range of depth variation that is required for even the simplest stereo films which are likely to be of mathematical interest.

*FOOTNOTE. The use of the 'multiplane' to describe certain animation techniques employed by Disney is quite a different matter, as that multiplane technique is not for stereoscopic films at all.

The mainstay of technical illustration must be the production of stereo pairs with the parallaxes drawn in - the problem that was not tackled in the films for the 1951 exhibition.

Those films all used two films and two projectors. Now is the Time was made with cut-out figures on an animation field of 10" x 14". The range of parallaxes corresponded to distances from infinity to half-screen distance. The technique of varying the parallaxes of a single drawing at the printing stage was also used, the operator varying the parallax "in sympathy with the linear perspective of the flat drawing"(3).

Around is Around used parallaxes produced by double punching of the location holes on art work. The scene for each eye was shot separately. Frame stagger was also used, and with especially striking effect on Lissajou figures photographed from an oscilloscope. Special techniques were neoessary in order to get a steady controlled picture. Frame stagger is a particularly good method of producing abstract designs since synchronisation of the projectors is a problem, and if the synchronisation slips it does not matter in this type of work as it merely produces another abstract diagram with a different gamut of parallaxes.

Halas and Batchelor's The Owl and the Pussycat* was a drawn, multi-plane cartoon, achieving its parallaxes by similar methods to McLaren. The parallaxes were calculated by Spottiswoode, and were worked out to 0.001 in. in some cases and 0.01 in. in others. Independently the author found it advisable to work in a similar way on the present programme which

*FOOTNOTE. An attempt was made to borrow The Owl and the Pussycat to accompany the premier performance of Resonant Cavities in London in 1958; but unfortunately no prints were available by then.

is on a rather different scale, since here $m = 5$ whereas Spottiswoode worked for a much bigger theatre with $m = 25$. This agreement on the degree of accuracy required in spite of the many differences between the two systems indicates strongly that future work should be standardised around accuracies of this order. 0.001 in. is attainable mechanically on the drawing board, and 0.01 in. can be attempted when drawing by hand.

Suggested drawing techniques. The cinemaphotography for the present programme was done on a Kodak Model A camera, especially modified by the addition of a 180° fading shutter and an 8:1 gear with an electrically operated clutch. The apparatus is capable of taking one picture per second with a half second exposure, either single shot or continuously, forwards or backwards. It is mounted on a rostrum providing a vertical movement of three or four feet, with two panning bars on the animation table which give independent horizontal movements calibrated to 0.001 in.

For stereo animation the two pictures are placed on the rostrum side by side. Each picture is approximately 7.6 ins. wide by 10 ins. high. After some preliminary experiments the spacing for congruent points was fixed at 10.7 ins. When this spacing had been standardised the peg bars on the animation table were fixed at this distance, the layout being shown in Fig. 8.2. Note that the drawing for the right eye has to be photographed on the left, because the film is inverted as it passes through the projector. The full field size of the camera under these conditions is approximately 14" x 19", so it can be seen that the system is rather inefficient in that only about 60% of the available area on the film is used to record the two pictures of the stereo pair. (This inefficiency is

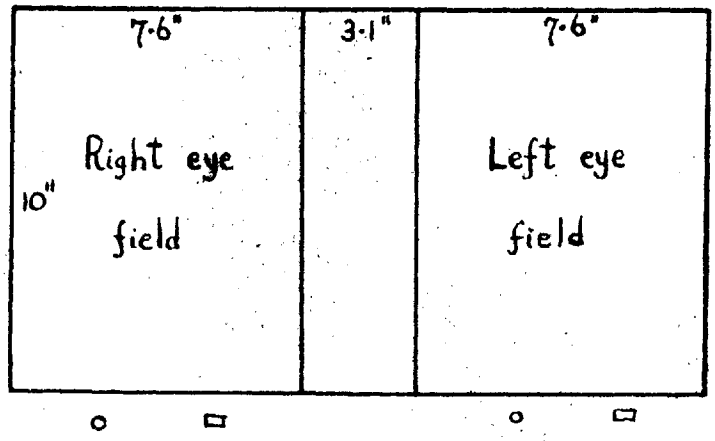


Fig. 8.2

inherent in the optical system, and is not due merely to the way in which it is used here.)

This standardised field size is adequate for a wide range of work, and it would be a simple matter to fit the peg bars to panning bars so that the spacing could be varied if it was desired to work with varying field sizes. This is not really necessary because in stereo work the difficulty of drawing parallaxes accurately rules out using very small fields, and since a double field is in use all the time the camera has to be high on its stand (which does not make for convenience in operation) so there is very little room for variation in either direction. Also the technique of tracking (or zooming) which is a normal part of 'flat' cinematography is quite unsuited to stereo as it introduces a variety of distortions simultaneously.

The drawing board carried only one peg bar, as it is not necessary to

have the drawings side by side at this stage. The general drawing procedure was to prepare a master drawing using the modifications to Rule's techniques described in Chapter 7. This master drawing carried the static backgrounds for both pictures and the positions of a certain number of animation cells. The number used was usually one in eight or one in four. Conforming to standard animation practice the remaining cells were interpolated by repeated bisection; i.e. after preparing numbers 0, 4, 8 etc. these were used to draw numbers 2, 6, 10 etc., and these were then used to construct 1, 3, 5 etc.

In order to maintain close control over the parallaxes these were read off a depth scale beside the plan drawing (Fig.8.3), and repeated checks were made with callipers, the screwhead of which had been calibrated to 0.01 in. and which could be read to one quarter or one fifth of a division. The drawing accuracy was not as great as this, the limitation being on the location of the nib of the pen to coincide with the calliper reading. Relative accuracy can be increased if the scale of the parallaxes as a whole is magnified. We give elsewhere evidence that this procedure does not result in distortions as serious as might be supposed, and in situations where the information to be conveyed is not primarily metrical some artificial increases in parallax are easily tolerated, and they may be increased by as much as 50% above their theoretical values. However where it is important to render metrical values as accurately as possible they should not be increased by as much as this. In these cases a 10% or 20% increase might compensate for the observer's tendency to see a stereo picture slightly flattened and produce

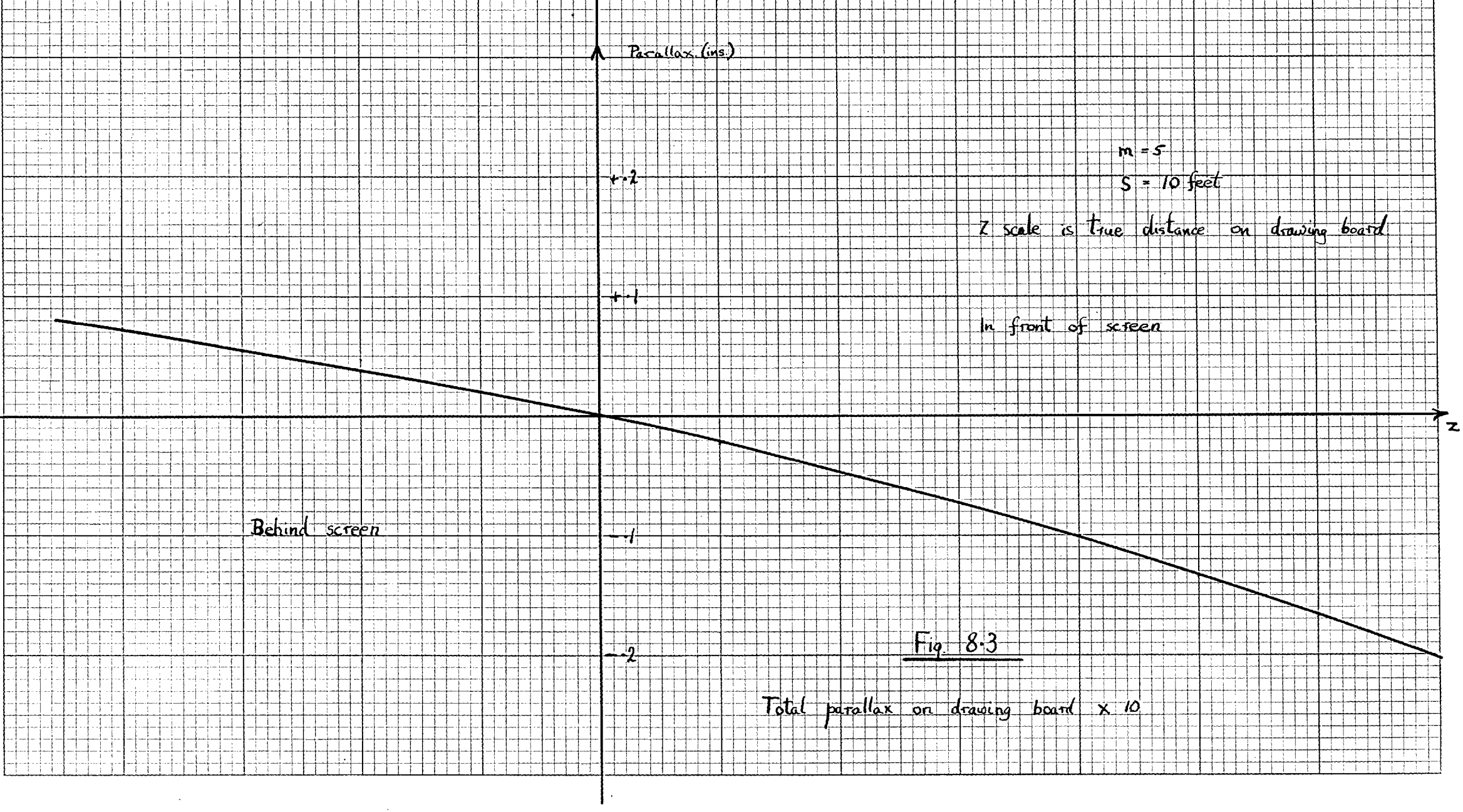

Parallax (ins.)
+.2
+.1
m = 5
S = 10 feet
Z scale is true distance on drawing board
In front of screen
Z
Behind screen
-.1
-.2
Total parallax on drawing board x 10

Fig. 8·3

a result a little better than the theoretically correct parallaxes. Only a statistical survey could decide this point conclusively.

The drawings were prepared on 0.003 in. cellulose acetate, and the finished drawings were checked for smoothness by superimposing the action for each eye separately on the peg bar. They were also checked for correct parallaxes by comparing each pair of corresponding drawings on the peg bar. This two-way check is essential for good stereo animation, and it is this stage which is much more tedious than ordinary, monocular animation; because if an error is detected and corrected empirically it may then be found that the cell no longer runs smoothly on the other check. This means that a simple error may not be redeemed by a simple correction but it may trigger off a chain of empirical adjustments on neighbouring cells. Careful checking at this stage is most important, and the final quality of the animation depends on it more than anything else.

The effect of errors is discussed elsewhere, but this is a convenient place at which to tabulate the various parameters involved in the overall communication system, drawing board - camera - projector - screen - observer. It is noticeable that there is a close parallel between the smallest movements producible on the animation rostrum and the maximum binocular acuity of trained observers, between the normal accuracy to which a draughtsman can work and the resolving power of the eye, and between random errors in animation and in the projection system and the saccadic flicks which are a feature of normal vision. (See the appended table.)

Drawing board	Projection system (m=5, s=10 ft.)	Eyes
		2" max. binocular acuity of trained observers (4)
.001 in. calibrations on animation rostrum panning bars	.005 in. on screen = 9" approx.	5" - 15" horizontal fine tremor (1)
		20" practical limits of binocular acuity
		40" diam. of foveal cones
		45" vertical fine tremor (1)
	.033 ins. = 1' approx.	1' max. resolving power (3)
.01 in. draughtsman's accuracy	.05 ins. total parallax for object $2\frac{1}{2}$ ins. in front of screen	
	2' intermittent variations in convergence in a two projector stereo system (5)	2' standard deviation of convergence (4)
.02 in. approx. line width	5' approx. line widths on screen	
	6' steady errors of horizontal registration in a two projector stereo system (5)	3' - 14' saccadic flicks (6)
		25' horizontal width of Panum's fusion areas
.25 in. total parallax at distance 2s		
	2.1 ins. = 1° approx.	1° diameter of fovea
.5 in. total parallax at $\frac{1}{2}$s		
7.6 ins. width of drawings	15° width of screen	
10 ins. height of drawings	20° height of screen	

Comparative table of parameters

Model techniques. It would be possible to produce technical stereo-animation by photographing models. (The movement would then be 'synthetic' but the stereopsis would not.) It would take us too far from the present line of enquiry to discuss these methods, we merely observe that they do not seem ever to have been attempted. The technique would be extremely laborious, more like the techniques of puppet animation than anything else, but with their aid it might be possible to achieve effects which would be impossible by drawing alone. For example, moving planes which both reveal their position by texture or by surface reflection and which are also transparent so that objects behind them are visible are impossible to animate by drawing, but they might be possible to model. They are needed to produce effective diagrams of some situations in three-dimensional projective geometry, differential geometry, stellations of regular solids and the like.

Autostereoscopy. We now consider more fully the production of depth effects by using only one series of drawings and presenting them to the two eyes with a difference in phase. Consider a set of drawings in which a point moves in a horizontal line, and let its x coordinate in frame n be ξ_n. Now suppose that frame n is presented to the left eye at the same time as frame n+r is presented to the right eye. In the notation of Chapter 4

$$X_n = \tfrac{1}{2}(\xi_{n+r} + \xi_n)$$

$$Z_n = \tfrac{1}{2}(\xi_{n+r} - \xi_n)$$

$$x_n = \frac{w(\xi_{n+r} + \xi_n)}{2w - \xi_{n+r} + \xi_n}$$

$$z_n = \frac{s(\xi_{n+r} - \xi_n)}{2w - \xi_{n+r} + \xi_n}$$

This means that using the methods of drawing previously described a point can be made to describe an arbitrary curve $z = f(x)$, y = constant, by a construction of the type shown in Fig.8.4. The diagram, which is for the case $r = 1$, is self-explanatory.

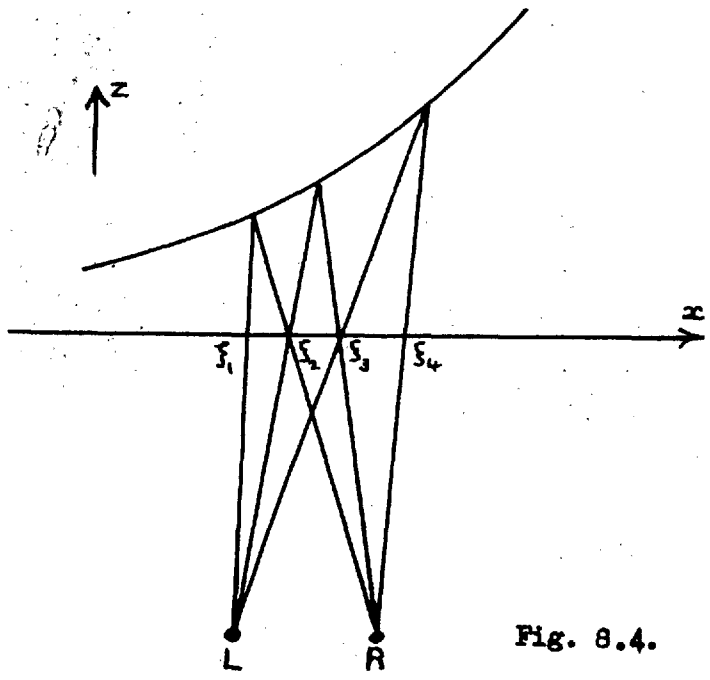


Fig. 8.4.

This type of animation is only of very limited use in technical illustration as it can only be used for points which are moving horizontally (or for points whose vertical movement is very small) and there is no simultaneous control over depth and speed. Once the phase difference has been decided there is no further control of the speed at which the point moves along its locus. The algebraic expression for the speed of movement is too complicated to be of value, but to the first order the angular displacement of the point from frame to frame, as seen by the observer, is $2wz/s(s + z)$.

REFERENCES

1. R.W. Ditchburn. Eye movements in relation to retinal action. Optica Acta, 4, Feb. 1955, pp.171-6.
2. J.J. Gibson. Perception of the Visual World. Boston, 1950.
3. A.C. Hardy, F.H. Perrin. Principles of Optics. McGraw Hill, 1932.
4. J. Krauskopf, T.N. Cornsweet, L.A. Riggs. Analysis of Eye Movements during monocular and binocular Fixation. J.O.S.A., 50, 6, June 1960, pp.572-8.
5. E. Levonian. Stereography and the Transmission of Images. J.S.M.P.T.E. 64, Feb. 1955, 77-85.
6. Mary P. Lord, W.D. Wright. The investigation of eye movements. Rep.Prog.Phys. XIII, 1950, pp.1-24.
7. N. McLaren. Stereographic Animation. J.S.M.P.T.E. Dec. 1951.
8. O.N. Rood. Amer.Journ.Sci. and Arts, XXXI, June 1861, 71.
9. R. and N. Spottiswoode. The Theory of Stereoscopic Transmission. University of California Press, 1953.

CHAPTER NINE

THE FILM "RESONANT CAVITIES"

In this chapter we give the mathematics involved in the design of the stereoscopic film Resonant Cavities. This film, showing the H_{101}, E_{111} and H_{111} modes of oscillation in a rectangular resonant cavity is the first animated stereoscopic diagram to be made by the techniques described. It is a six minute, colour film on 16 mm, consisting of three two minute sections each containing a steadily repeating cycle of the oscillation concerned. This arrangement was used so that the time spent on animation could be reduced by shooting a few repeated cycles and printing duplicates from them after they had been cemented into closed loops. The preparation of a film with the usual cinematic devices of continuity would have been more expensive and the gain would have been small.

The lines of electric force in the cavity are shown in red and the magnetic lines in green. The cavity itself, a rectangular box in shape, is shown in white, the whole diagram being against a black background. This is especially convenient for technical diagrams in stereo as it avoids edge effects where there is only one image, a feature which can cause difficulties in live action stereo.

A line of force in an electric or magnetic field is a line such that the tangent at every point is in the direction of the electric (or magnetic) vector. If lines are drawn through every point of a small closed curve they lie on a tubular surface called a tube of force. The customary convention is that when suitable units are chosen unit tubes of

force terminate on unit charges, and the density of the tubes of force measures the intensity of the electric field. When diagrams are drawn there are two possibilities. If the diagram is only being used to give a qualitative picture it is sufficient to draw a few representative lines of force; but if a quantitatively accurate diagram is required the tubes must be spaced correctly, and then it is usually convenient to represent tubes by lines along their axes. The term 'line' will be understood in this sense from now on. The correct spacing of the lines when the diagrams move involves many considerations which the standard works on electro-magnetic theory do not discuss. Tubes of force are a convenient fiction and one can easily speak of a fractional number of tubes; but one can only use an integral number of lines in a diagram, and so it is not always possible to construct diagrams which are quantitatively accurate.

Even when the number of tubes of force is integral it is not always possible to draw a diagram showing them in a satisfactory manner. Consider, for example, the tubes of force which leave a charged sphere. The numbers of lines which can be drawn radiating equally in all directions from a sphere are bound up with the point symmetry groups in three-dimensional space. Therefore even in the simplest cases the design of line of force diagrams which are as accurate as possible is difficult, and involves mathematical considerations far beyond the purely electrical phenomena which they illustrate. If the diagram has to move, and if we wish the correct quantitative relationships to be maintained throughout the movement the questions are even more complicated.

For these reasons the design of animated diagrams of even such a

standard topic as cavity resonators demands a considerable amount of quite fresh mathematical investigation. We give below the calculations which were necessary to design the film Resonant Cavities, giving the calculations in a general form so that it would be possible to produce versions of these diagrams containing larger numbers of lines if one wished to do so.

Line of Force Illustrations of Rectangular Electro-Magnetic Resonators

Notation. Throughout we omit the factor $e^{i\omega t}$, as is customary, and we write $\pi x/a = \theta$, $\pi y/b = \phi$, $\pi z/l = \psi$, and $\omega t = \tau$.

The standard forms of equations for E and H waves are (1):-

E waves

$$
\begin{aligned}
E_z &= \sin m\theta \sin n\phi \cos \nu\psi \\
E_x &= -\frac{1}{k^2} \cdot \frac{m\pi}{a} \cdot \frac{\nu\pi}{l} \cos m\theta \sin n\phi \sin \nu\psi \\
E_y &= -\frac{1}{k^2} \cdot \frac{n\pi}{b} \cdot \frac{\nu\pi}{l} \sin m\theta \cos n\phi \sin \nu\psi \\
H_z &= 0 \\
H_x &= \frac{i\omega\varepsilon}{k^2 c} \cdot \frac{n\pi}{b} \sin m\theta \cos n\phi \cos \nu\psi \\
H_y &= -\frac{i\omega\varepsilon}{k^2 c} \cdot \frac{m\pi}{a} \cos m\theta \sin n\phi \cos \nu\psi
\end{aligned}
\qquad (9.1)
$$

where
$$
\frac{\omega}{2\pi} = f_{mn\nu} = \frac{c}{2\sqrt{\varepsilon\mu}} \left[\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2 + \left(\frac{\nu}{l}\right)^2 \right]^{\frac{1}{2}}
$$

$$
k^2 = \left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2 .
$$

The only modes we need to consider are E_{110}, which is H_{101} reoriented, and E_{111}.

H waves

$$
\begin{aligned}
H_z &= \cos m\theta \cos n\phi \sin \nu\psi \\
H_x &= -\frac{1}{k^2} \cdot \frac{m\pi}{a} \cdot \frac{\nu\pi}{l} \sin m\theta \cos n\phi \cos \nu\psi \\
H_y &= -\frac{1}{k^2} \cdot \frac{n\pi}{b} \cdot \frac{\nu\pi}{l} \cos m\theta \sin n\phi \cos \nu\psi \\
E_z &= 0 \\
E_x &= \frac{i\omega\mu}{k^2 c} \cdot \frac{n\pi}{b} \cos m\theta \sin n\phi \sin \nu\psi \\
E_y &= -\frac{i\omega\mu}{k^2 c} \cdot \frac{m\pi}{a} \sin m\theta \cos n\phi \sin \nu\psi
\end{aligned}
\qquad (9.2)
$$

The only modes we need consider are H_{101} and H_{111}.

All of the above equations can be multiplied by an arbitrary amplitude factor A. The E and H components can be multiplied by separate amplitude factors B and A, since this amounts to choosing the ratio $\varepsilon : \mu$ to suit our convenience.

Note that all lines have the same profile when viewed in the x, y or z directions, since $H_z : H_y$ is independent of z, etc.

The form of the E and H lines

E waves E lines

$$
\frac{dx}{\frac{m}{a} \cos m\theta \sin n\phi} = \frac{dy}{\frac{n}{b} \sin m\theta \cos n\phi}
$$

$$
\frac{(\cos m\theta)^{a^2/m^2}}{(\cos n\phi)^{b^2/n^2}} = \text{const.}
$$

$$
\frac{dy}{-\frac{1}{k^2} \cdot \frac{n\nu\pi^2}{bl} \cos n\phi \sin \nu\psi} = \frac{dz}{\sin n\phi \cos \nu\psi}
\qquad (9.3)
$$

$$(\cos n\phi)^{\frac{a^2}{m^2}+\frac{b^2}{n^2}}(\cos \nu\ )^{a^2/m^2} = \text{const.} \qquad \text{(9.3 cont.)}$$

(assuming $m \neq 0$ and $n \neq 0$, which is so in the cases which concern us.)

H lines

$$\frac{dx}{\frac{n}{b}\sin m\theta \cos n\phi} = \frac{dy}{-\frac{m}{a}\cos m\theta \sin n\phi}$$

$$\sin m\theta \sin n\phi = \text{const.} \qquad (9.4)$$

$$z = \text{const.}$$

H waves E lines

$$\cos m\theta \cos n\phi = \text{const.} \qquad (9.5)$$

$$z = \text{const.}$$

H lines

$$\frac{(\sin m\theta)^{a^2/m^2}}{(\sin n\phi)^{b^2/n^2}} = \text{const.} \qquad (9.6)$$

$$(\sin m\theta)^{\frac{a^2}{m^2}+\frac{b^2}{n^2}}(\sin \nu\psi)^{b^2/n^2} = \text{const.}$$

(assuming that $m \neq 0$ and $n \neq 0$).

We are also concerned with the case $n = 0$; and the solution then, which is not given by the equations above, is $y = \text{const.}$, $\sin m\theta \sin \nu\psi = \text{const.}$

Currents in Walls

The surface current density, K, is perpendicular to the tangential magnetic field, and is of a magnitude given by

$$H_{tan} = \frac{4\pi}{c} K.$$

The scalar multiplier depends on the system of units employed, which for our purpose is irrelevant, and we later drop this factor.

E Waves

The wall currents all flow parallel to the axis of the waveguide, and behave correspondingly in a resonator.

On the walls

$$x = 0 \qquad J_y = -H_z \qquad J_z = H_y$$

$$y = 0 \qquad J_x = H_z \qquad J_z = -H_x$$

$$z = 0 \qquad J_y = H_x \qquad J_x = -H_y$$

These equations become

$$x = 0 \qquad J_y = 0 \qquad J_z = -\frac{i\omega\epsilon}{k^2 c} \cdot \frac{m\pi}{a} \sin n\phi \cos \nu\psi$$

$$y = 0 \qquad J_x = 0 \qquad J_z = -\frac{i\omega\epsilon}{k^2 c} \cdot \frac{n\pi}{b} \sin m\theta \cos \nu\psi \qquad (9.7)$$

$$z = 0 \qquad J_y = \frac{i\omega}{k^2 c} \cdot \frac{n\pi}{b} \sin m\theta \cos n\phi \qquad J_z = \frac{i\omega\epsilon}{k^2 c} \cdot \frac{m\pi}{a} \cos m\theta \sin n\phi$$

In $x = 0$ and $x = a$ the flow lines are $y = \text{const.}$,

$y = 0$ and $y = b$ the flow lines are $x = \text{const.}$,

$z = 0$ and $z = 1$ the flow lines are as E lines, see above.

H waves

$$x = 0 \qquad J_y = -\cos n\phi \sin \nu\psi \qquad J_z = -\frac{1}{k^2} \cdot \frac{n\pi}{b} \cdot \frac{\nu\pi}{l} \sin n\phi \cos \nu\psi$$

$$y = 0 \qquad J_x = \cos m\theta \sin \nu\psi \qquad J_z = \frac{1}{k^2} \cdot \frac{m\pi}{a} \cdot \frac{\nu\pi}{l} \sin m\theta \cos \nu\psi \qquad (9.8)$$

$$z = 0 \qquad J_y = -\frac{1}{k^2} \cdot \frac{m\pi}{a} \cdot \frac{\nu\pi}{l} \sin m\theta \cos n\phi \qquad J_x = -\frac{1}{k^2} \cdot \frac{n\pi}{b} \cdot \frac{\nu\pi}{l} \cos m\theta \sin n\phi$$

The surface charge density is $\sigma = E_n$.

It is a useful convention, when necessary, to draw planes at $x = \frac{1}{2}a$, $y = \frac{1}{2}b$, and $z = \frac{1}{2}l$ and to imagine a magnetic charge on these fictitious boundaries. Then, always working on the side of the plane nearest the origin, the magnetic charge density is $\rho = H_n$.

Also 'magnetic currents' are given by a set of equations like the above, but with all the signs reversed.

Henceforth we multiply the E and H vectors by arbitrary coefficients A and B, which we later choose in order to give a suitable scale to the diagram.

The Design of the Diagram for H_{101}

The diagram is constructed with the particular values $a = 24$ ins, $b = 30$ ins. and $l = 18$ ins.

The equations needed for this particular case, still omitting the factor $e^{i\omega t}$, are

$$H_z = Al \cos\theta \sin\psi \qquad E_z = 0$$

$$H_x = -Aa \sin\theta \cos\psi \qquad E_x = 0$$

$$H_y = 0 \qquad E_y = -iB \sin\theta \sin\psi$$

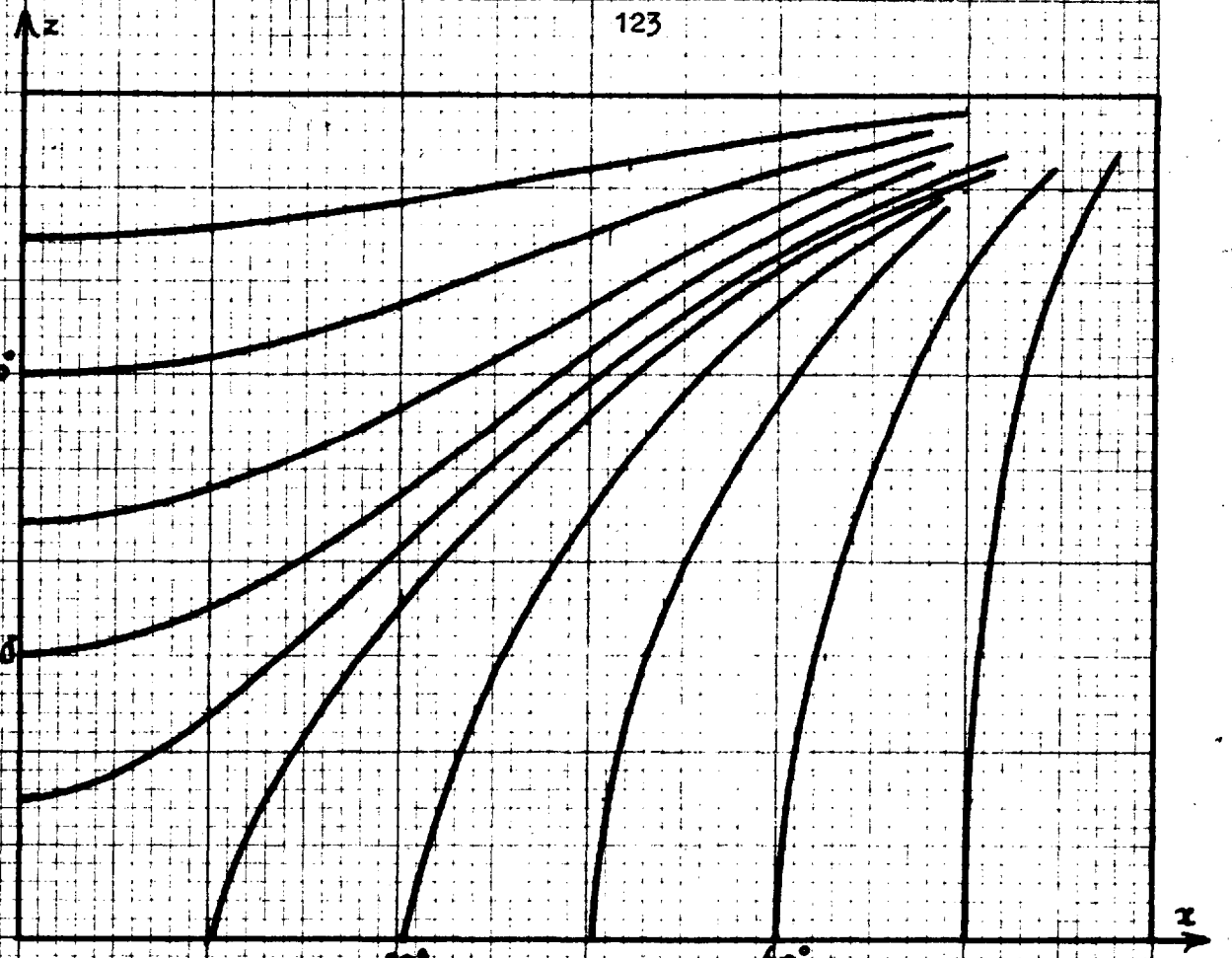
The electric currents in the plane $y = 0$ are

$$J_x = Al \cos\theta \sin\psi \qquad J_z = Aa \sin\theta \cos\psi$$

The electric charge density in the plane $y = 0$ is $-iB \sin\theta \sin\psi$. In the plane $y = 0$ the electric charges flow along the lines

$$\frac{(\cos\psi)^{l^2}}{(\cos\theta)^{a^2}} = \text{const.} \qquad (9.9)$$

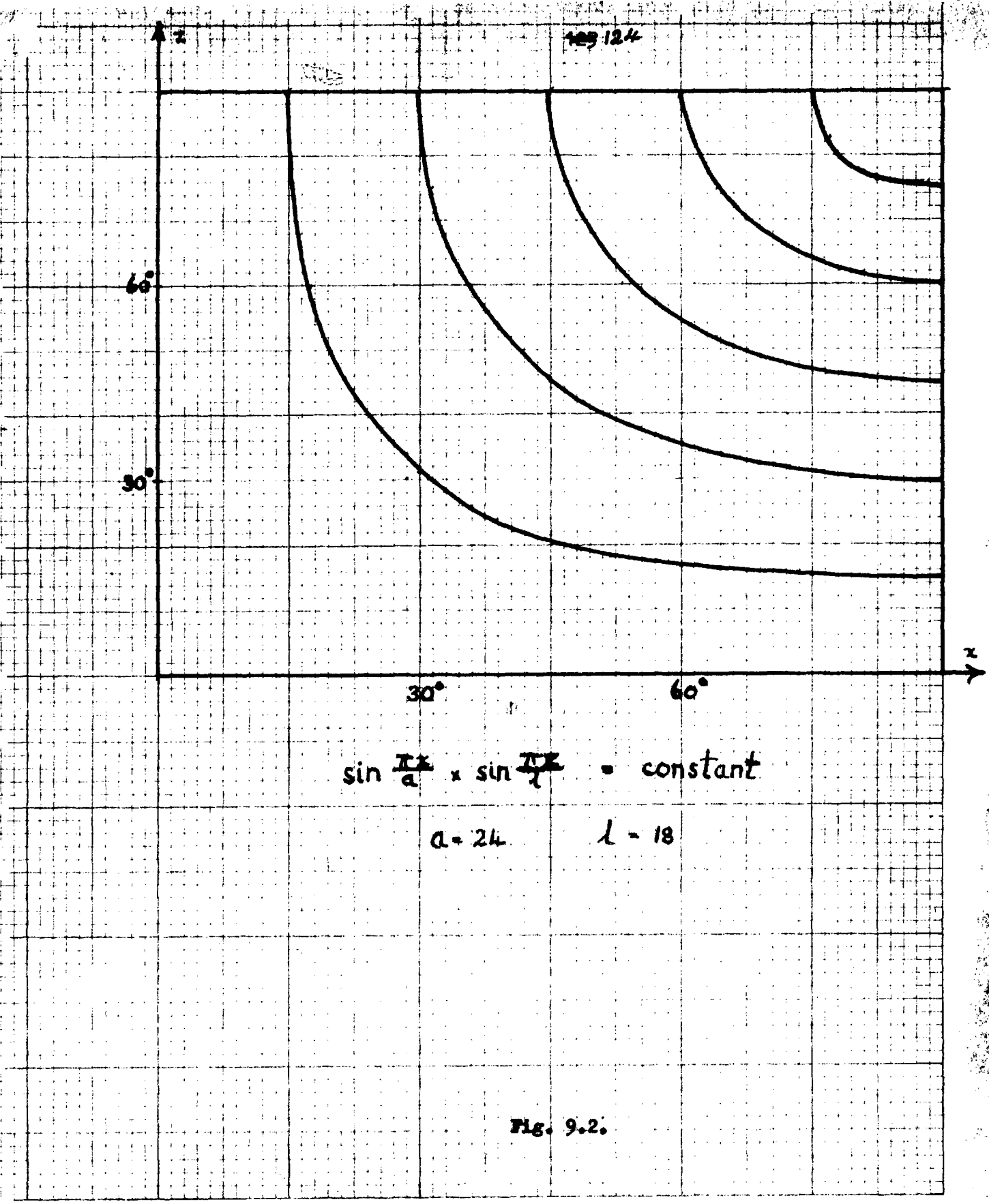
See Fig.9.1



$$\frac{\left(\cos\frac{\pi z}{l}\right)^{l^2}}{\left(\cos\frac{\pi x}{a}\right)^{a^2}} \doteq \text{constant}$$

$a = 24$ $\qquad$ $l = 18$

Fig. 9.1.



$$\sin \frac{\pi x}{a} \times \sin \frac{\pi z}{l} = \text{constant}$$

$a = 24$ $\quad l = 18$

Fig. 9.2.

The magnetic charge density in the planes

$x = \frac{1}{2}a$ is $A a \cos 4$

and $z = \frac{1}{2}l$ is $-Al \cos \Theta$.

The magnetic currents are, in the planes

$x = \frac{1}{2}a$ $M_z = -iB \sin 4$

and $z = \frac{1}{2}l$ $M_x = -iB \sin \Theta$.

The magnetic charges flow along the lines $y = \text{const.}$ in each case.

The E lines have equations $x = \text{const.}$ and $z = \text{const.}$

The H lines have equations $y = \text{const.}$ and $\sin \Theta \sin 4 = \text{const.}$

See Fig. 9.2.

The Flow of Electric Charges in the case H_{101}.

The lines of electric force are all parallel to the y axis, and they terminate on charges in the planes $y = 0$ and $y = b$. The problem is to space the lines in such a way that throughout the cycle they give a quantitative picture of the field strength inside the resonator and the charge distribution on the end walls.

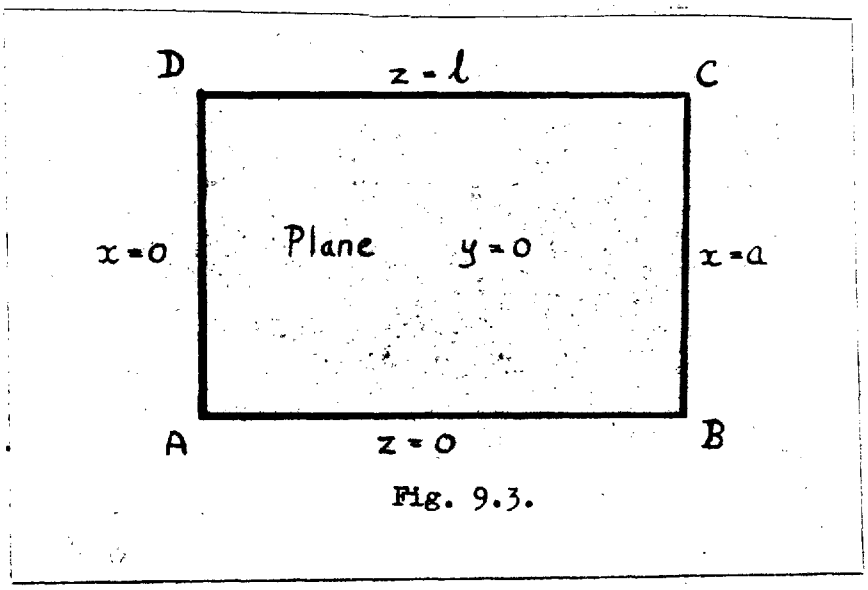


Fig. 9.3.

The instantaneous current density across the edge $z = 0$ at the point x is

$$J_z = a \sin\theta \cos\tau$$

and the instantaneous total flux across the segment AP of this line, which is obtained by integrating with respect to x is

$$\frac{a^2}{\pi}(1 - \cos\theta)\cos\tau$$

Similar relations hold along AD.

The instantaneous current densities across $z = 0$ and $x = 0$ are in the ratio a:1, and the total currents over any interval of time are in the ratio $a^2:1^2$. Therefore to give a reasonable quantitative picture one should divide the lines up into qa and ql segments and let ra and rl lines cross each segment in the course of each quarter cycle. (q and r are integers, and the choice of them determines how many lines are to be shown in the diagram.)

The line AB must be divided into qa segments across which the flux is equal, and the ra charges which cross each segment follow a track through the corresponding centre of current density. That is to say the charges cross AB at points given by

$$1 - \cos\theta = \frac{p}{qa} \qquad p \text{ odd}, \quad 0 < p < 2qa,$$

$$\cos\theta = 1 - \frac{p}{qa}.$$

The instantaneous current is proportional to $\cos\tau$, so that the total flux since the beginning of a cycle is proportional to $\sin\tau$.

The ra charges have to be introduced with the correct phase angles. To achieve this a quarter cycle must be divided up into ra divisions and the lines introduced after charges $\frac{1}{2}$, $1\frac{1}{2}$, $2\frac{1}{2}$ etc. have flowed across the segment in question. This means that the lines appear at intervals given by

$$\sin \quad = \frac{p}{2ra} \qquad\qquad p \text{ odd}, \quad 0 < p < 2r.$$

The total number of charges crossing AB in each quarter cycle is now qra^2 and the total number of lines of force which appear in the diagram is $2qr(a^2 + l^2)$. Therefore to produce a diagram with enough on it to be 'interesting' which is at the same time quantitatively accurate by our present criteria it would be reasonable to work with, for example,

$$a = 3, \qquad l = 2, \qquad q = r = 1.$$

This would give a diagram with 26 vertical lines all of which move at varying depths. The preparation of a diagram with as many lines as this is clearly very tedious, so it seemed better at this early stage of development to produce a diagram with fewer lines which gave a reasonable qualitative picture without aspiring to strictly quantitative accuracy. But to illustrate the principles involved we continue the general discussion.

With the above values of a, l, q and r the division along the line AB are at the points given by

$$\cos \theta = 1 - \frac{p}{5} \qquad\qquad p = 1, 3, 5$$

$$\theta = 48°12', 90°, 131°48',$$

and the charges appear with phase angles given by

$$\sin \tau = \frac{p}{6} \qquad p = 1, 3, 5$$

$$\tau = 9°36', 30°, 56°26'.$$

The division along AD are at points given by

$$\cos \psi = 1 - \frac{p}{2} \qquad p = 1, 3$$

$$\psi = 60°, 120°,$$

and the charges appear with phase angles given by

$$\sin \tau = \frac{p}{4} \qquad p = 1, 3$$

$$\tau = 14°30', 48°36'.$$

This is shown in the accompanying diagram.

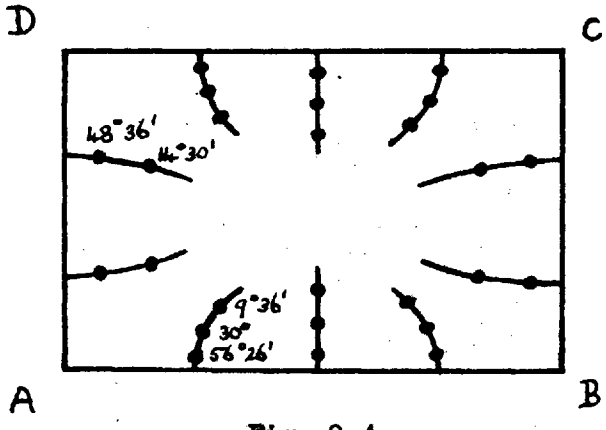


Fig. 9.4.

A convenient scheme would be to animate using a 19 frame quarter-cycle, varying τ by steps of 5°. This would give a full cycle of 76 frames which would last $4\frac{3}{4}$ seconds at the normal silent projection rate of 16 frames per second. But the diagram we envisage does not discriminate between positive and negative charges, so to the eye it would appear to recur after half this time. If a slower oscillation was required it would be convenient to vary τ by steps of 3°. 5° steps lead to the following schedule:-

Phase angle	Number of lines in picture	Number of frames
0	0	2
9°36'	6	1
14°30'	10	4
30°	16	3
48°36'	20	1
56°26'	26	8
90		—
		19

Before the analysis of the E-lines is complete two further problems have to be considered:-

i) whilst the tracks which the charges follow throughout the cycle are known we do not yet know where each charge is on its track at any particular time,

ii) we do not know how far in towards the centre of the face the charges move. That is, we do not know their positions at the time $\tau = \frac{\pi}{2}$.

The answer to i) is required in order to produce the drawings for the successive frames of the films, and the answer to ii) is required to settle an artistic point of the design. The oscillations have to be given a suitable amplitude so that the movement of the lines covers a reasonable range.

The mathematical answers to both of these questions can be derived from the equations

$$\frac{dx}{dt} = Al \cos\theta \sin\psi \cos\tau$$
$$\frac{dx}{dt} = Aa \sin\theta \cos\psi \cos\tau \qquad (9.10)$$

which come from equations 9.8.

Confining attention to the lines of flow which commence on the segment AB we also have from equation 9.9

$$(\cos\theta)^{a^2} = (\cos\psi)^{l^2} \times \text{const.}$$

where the constant is easily seen to be $(\cos\theta_0)^{a^2}$ where θ_0 is the abscissa of the point on AB at which the line of flow commences. These equations can be solved in the sense that they can be reduced to integrals; but the integrals do not seem to be expressible explicitly, and in any case they are too complicated to be of very much value in numerical computation. For film purposes as good a procedure as any is to use the equations to give approximate finite differences between one frame and the next. A film is in any case a succession of discrete pictures and a finite difference method is always very much in the spirit of film animation.

It is easy to solve the equations explicitly for the particular lines of flow for which $x = \frac{1}{2}a$ or $z = \frac{1}{2}l$, and this solution is sufficient to

settle the design question implicit in problem ii).

When $x = \frac{a}{2}$ $\qquad$ $\sec \psi \, dz = Aa \cos\tau \, dt$

so $\qquad \frac{1}{\pi} \log \tan\left(\frac{\psi}{2} + \frac{\pi}{4}\right) = A \frac{a}{\omega} \sin\tau + \text{const.}$

The value of the constant is such that $z = 0$ when the charge appears, so we may write

$$\tan\left(\frac{\psi}{2} + \frac{\pi}{4}\right) = \exp\left[\frac{Aa\pi}{l\omega} (\sin\tau - \sin\tau_0)\right],$$

has its maximum value when $\tau = \frac{\pi}{2}$, so

$$\tan\left(\frac{\psi_{max}}{2} + \frac{\pi}{4}\right) = \exp\left[\frac{Aa\pi}{l\omega} (1 - \sin\tau_0)\right].$$

The value of θ_{max} for the charge flowing along the line $z = \frac{1}{3}l$ is given by a similar expression, with a and l interchanged; but direct comparison of the two is complicated by the fact that the values of τ_0 are not the same in the two cases. Numerical work is greatly assisted by drawing the graph of θ and K (Fig.9.5) for the relation $\tan\left(\frac{\theta}{2} + \frac{\pi}{4}\right) = \exp K$. Using the particular values of parameters chosen earlier ψ_{max} is given with $\sin\tau_0 = 1/6$, and θ_{max} with $\sin\tau_0 = \frac{1}{3}$. We have

$$\tan\left(\frac{\theta_{max}}{2} + \frac{\pi}{4}\right) = \exp\left(\frac{5}{8} \cdot \frac{A\pi}{\omega}\right)$$

and $\qquad \tan\left(\frac{\psi_{max}}{2} + \frac{\pi}{4}\right) = \exp\frac{A\pi}{\omega}.$

A suitable value for A may be decided by using the graph. ω is also involved in the equations, and no numerical value has been assigned to it yet;

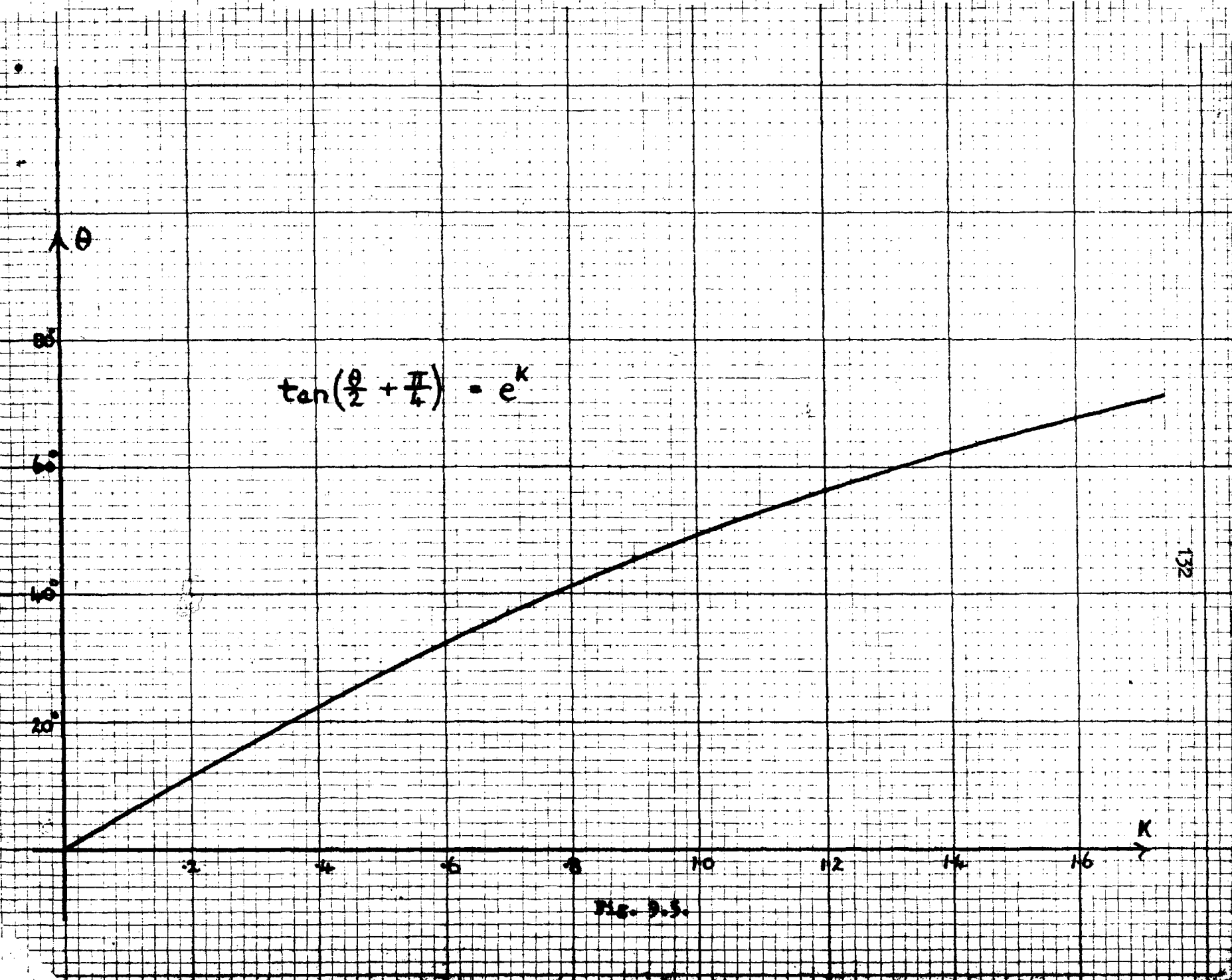

θ
80°
60°
40°
20°
$\tan\left(\frac{\theta}{2} + \frac{\pi}{4}\right) = e^{K}$
K
.2
.4
.6
.8
1.0
1.2
1.4
1.6

Fig. 9.5.

but choice of a numerical value for ω is linked with the time scale we choose to adopt. We have fixed the number of frames which a complete cycle will take and if we adopt one frame as the unit of time then a value of ω is automatically decided; but one may prefer to leave the time scale arbitrary at this stage and to choose the units elsewhere.

When the constants to be used have been decided equations 9.10 when put into the form

$$\delta x = Al \cos\Theta \sin\Psi \cos\tau \quad \delta t$$

$$\delta z = Aa \sin\Theta \cos\Psi \cos\tau \quad \delta t$$

enable the positions of the lines in successive frames to be calculated.

This is a convenient time at which to normalise the time scale, taking $\delta t = 1$, the interval between frames. Fig. 9.6 shows how Θ_{max} and Ψ_{max} are connected for charges appearing at the same epoch (i.e. for the same value of τ_0), and it may also be useful when the above design questions are being settled.

It may be observed that the original set of equations (1) involved four physical constants. The velocity of light, c, and the electric and magnetic constants of the dielectric ε and μ occur in the equations, and we are free to introduce an arbitrary multiplying factor for the amplitude. In the design of the film there are four corresponding decisions. There are choices of q, r, A and the time scale.

The Flow of Magnetic Charges in the case H_{101}

We have explained earlier that it is convenient to imagine partitions introduced at the planes $x = \frac{1}{2}a$, $y = \frac{1}{2}b$, $z = \frac{1}{2}l$ and to imagine

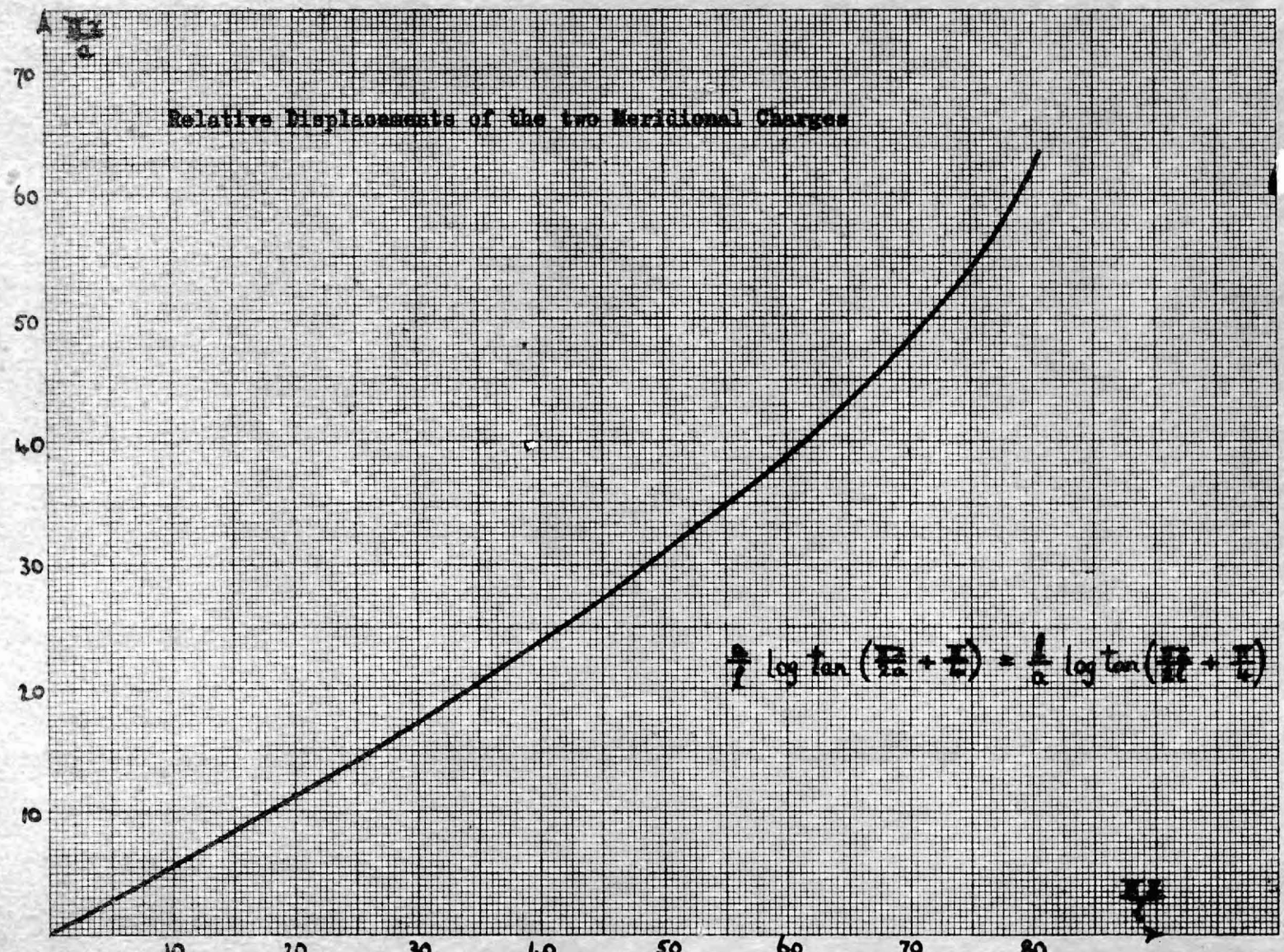

Relative Displacements of the two Meridional Charges
70
60
50
40
30
20
10
10
20
30
40
50
60
70
80

Fig. 9.6.

fictitious 'magnetic currents' in these partitions. The magnetic lines of force are related to these currents just as the electric lines of force are related to the electric currents in the boundary walls in the previous section. (There is actually a difference in sign, but this does not concern us as this sign is not going to be shown in the final diagram.) Fig. 9.7 shows the plane $x = \frac{1}{2}a$.

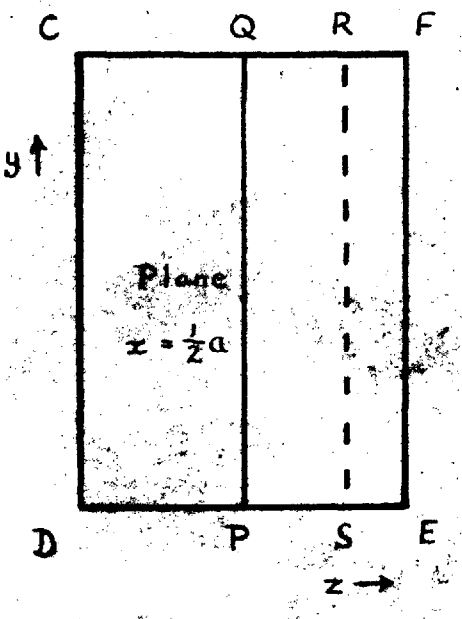


Fig. 9.7.

Denoting the components of magnetic current by M_y and M_z we have

$$M_y = 0$$

$$M_z = B \sin \psi \sin \tau$$

The instantaneous current across PQ is $B \sin \tau$.

As before we now divide PQ into q segments and let r magnetic

charges flow across each segment in each quarter cycle. The magnetic charges move along equally spaced lines parallel to DE, with coordinates

$$y = \frac{pb}{2q} \qquad p \text{ odd}, \quad 0 < p < 2q,$$

and they emerge from PQ at times given by

$$\cos\tau = \frac{p}{2r} \qquad p \text{ odd}, \quad 0 < p < 2r.$$

The charge density is equal to H_z, so integrating with respect to z the instantaneous total charge in the region PQRS is $qr \sin\psi \cos\tau$, and we require the values of ψ for which this is $\frac{p}{2r} \cdot qr$, i.e.

$$\frac{p}{2r} = \sin\psi \cos\tau$$

$$\sin\psi = \frac{p}{2r} \sec\tau \qquad p \text{ odd}, \quad 0 < p < 2r.$$

Similar considerations apply in the plane $z = \frac{1}{2}l$. The magnetic lines are plane curves, approximately elliptical in shape which join the magnetic charges in the planes $x = \frac{1}{2}a$ and $z = \frac{1}{2}l$. The amplitude of the oscillation and the position of particular charges at particular times may be discussed as the similar questions were discussed for the electric lines earlier, but the discussion is a little simpler this time because the lines along which the charges flow are straight.

The above treatment is sufficiently detailed to show how the necessary data can be calculated to draw diagrams of any required degree of accuracy. The number of lines needed in this theoretical treatment is rather too large to be practically convenient in drawing, so a slightly simplified version was adopted in order to make the film. Nevertheless when stereo diagrams are projected the lines do really appear "more

spread out"; and so when technical diagrams are designed it is advisable to include more elements than one might include in corresponding "flat" diagrams.

Practical Details of Construction for the H_{101} Diagram

The scene was designed inside a rectangular box with dimensions a = 24", b = 30", l = 18". (These dimensions refer to the size of the figure in space, as finally projected, and not to dimensions on the drawing board.) Six electric lines were used and two magnetic lines. The time scale was moved in steps of 5°, which meant that 19 sets of drawings had to be prepared for a complete quarter-cycle.

The calculations showed that there was a rather quick movement as the H-lines appeared and disappeared, so to ease the animation a slight empirical adjustment was made to the movement at this point.

Practical Details of Construction for the E_{111} Diagram

By re-orienting the diagram by interchanging the y and z axes, and by altering the length b to 22½" it was possible to use the drawings already prepared for the H-lines of the H_{101} mode again. The E lines had to be calculated according to the principles already given, and they presented extra difficulties as in this mode of oscillation they are twisted curves moving in space.

Practical Details of Construction of the H_{111} Diagram

It is possible to derive the diagram for this mode of oscillation from the diagram for the E_{111} mode by interchanging the E and H lines and translating by $\frac{1}{2}\pi$ in each direction in space. This meant that a lot of

the previous calculations and drawings could have been used again, but there was the disadvantage that if this plan had been carried out to the last detail there would not have been enough H lines in this diagram for it to have been very informative, therefore some additional H-lines were calculated.

The diagram was constructed using the box that had previously been used for the H_{101} diagram.

When preparing drawings in perspective a translation is not as simple a matter as it might appear, because translation alters all the details of perspective. Therefore whilst it was possible to use some of the material over again a lot of quite fresh drawing was necessary as well.

The film was shot on Kodachrome A stock with Photoflood illumination. Three cycles of animation (half-cycles of the electro-magnetic oscillation) were shot in each case and lengths of them subsequently printed off. The title were shot separately. In the printing the green, which was a good colour in the original, was washed out to a near-white: but this was only a minor defect. Colour materials which have more recently become available should avoid any repetition of this.

Performances of the film. The film Resonant Cavities has either been seen by representatives of the following organisations or been shown to audiences at the institutions:-

Scientific Film Association

Shell Film Unit

G.E.C. Research Laboratories

The Royal Society

The National Academy of Sciences and the National Research Council, U.S.A.

Deutsches Zentralinstitut fur Lehrmittel, East Berlin, East Germany

University College of North Staffordshire

Oxford University

Various technical colleges.

Electrical engineers and mathematicians have agreed that the film accomplishes what it sets out to do and that it gives a fresh understanding of the principles which it illustrates. The nature of this 'understanding' is (almost by definition) difficult to convey in words; but it is common experience that the purely verbal treatment of many parts of mathematics, or the treatment by algebraic symbols alone, leaves a gap which remains to be filled. And that when words have reached their limit comprehension may be enhanced, often in a vivid way, by some other means of communication which appeals more directly to the imagination. Klein, Hilbert and Polya, to name but three prominent mathematicians, have held this view and described it graphically.

The film has also been seen by experienced professional animators who have commented favourably on the quality of the animation, and agreed that it is certainly as smooth as conventional non-stereo animation done by similar drawing techniques.

The road is open for the wider use of stereo films for technical instruction and for the documentation of pure and applied mathematics but for the "chicken-and-egg" situation which hinders further development. Showing such films requires special equipment. Few people have the

equipment because there are so few films to show, and it is hardly worth making more films when there are so few facilities for showing them. A similar situation operates with the art films of Norman McLaren. Encouraged by the success of his films for the Festival of Britain McLaren would like to make more; but the very few outlets for such work makes it scarcely practical to devote time to it which could be spent on other work, with an importance of its own, which it is certain will be widely used.

Some of the showings of Resonant Cavities listed above were arranged by projectionists with no previous experience of stereoscopy, and audiences of nearly one hundred were accommodated. The system is by no means ideal for audiences as large as this, but the technique is sufficiently well developed to be used with comparative ease with academic audiences of a reasonable size. Were there more films of this type to enable a programme of some length to be shown then it seems quite certain that they would be well used.

REFERENCES

1. H.R.L. Lamont. Wave Guides. Methuen.

CHAPTER TEN

DYNAMIC EXPERIMENTS

I. EXPERIMENT "JUMPER"

Summary. The success of the film Resonant Cavities raises the question, "Since the inevitable errors in the animation were of a size which is easily noticed in static diagrams why did they not lead to disturbing fluctuations in depth?". To investigate this further a special stereo film was made, showing three vertical white lines against a black background. These lines were in the plane of the screen, but the central line was given small displacements forward or back for one or two frames to simulate the errors which can arise through inaccuracies in drawing.

Observers were asked to report on these "jumps", with results that were completely negative. Jumps of this order of magnitude are often lost in the random motions of the picture due to irregularities of the positioning of the film in the gate of the camera and the projector, and even when the jumps on the pictures for the right and left eyes were clearly visible separately they were not visible when viewed stereoscopically.

In other words, one frame is too short a time for slight errors in drawing to produce the serious errors in depth which application to the simple convergence theory predicts. This may seem a little surprising in view of the observations of a number of workers of flash stereoscopy, but the viewing conditions are by no means identical and this failure to observe sudden jumps is in accord with Levonian's measurements on live action stereo-cinematography (reviewed in Chapter 8) (6), with recent

work by Julesz (3) and with current knowledge on eye-movements (2, 4, 8).

Pilot Experiment. A preliminary experiment was made in which the observer saw three vertical white lines against a grey background. Every few seconds the picture for one eye only was altered for single frames, or sometimes for a double frame, by moving the middle line by amounts of the order of a line width (0.02 ins. on the drawing board). This corresponded to theoretical movements in depth of 3 ins. or more when the film was viewed under normal projection conditions. Experiment "Photofinish" had shown that variations in depth of this order are seen very easily indeed under static conditions.

Observers were quite unable to say whether the line jumped forwards or backwards. The sideways movement was of the order of 0.1 ins. on the screen (certainly one tenth or less of the theoretical movement in depth). This sideways movement was seen equally well under monocular conditions or under binocular conditions without polaroid glasses. This suggested two things:-

i) if the smoothness of stereo-animation is subjected to the normal controls which are exercised on ordinary two-dimensional animation the results may well be quite satisfactory.

ii) it would be of interest to make a similar test film in which a one frame "jump" was given in opposite directions to the images for the right and left eyes - which gives, of course, a stereo jump in depth with no transverse component - and to see if this was noticeable under normal projection conditions.

Quite a crude, empirical displacement by normal drawing methods was

adequate for the pilot experiment; but in order to match the displacements for the two eyes accurately in the main experiment the animation rostrum was equipped with two panning bars controlled by screw threads which were calibrated down to 0.001 ins.

Main Experiment. Three vertical white lines were shown against a black background. The film was shot on 50 ft. of Kodachrome A. A code letter which was of no concern to the observer was changed every ten seconds, and somewhere in the course of each interval a deliberate jump was given to the central line. The jumps were produced by having the two drawings of the central line on separate panning bars.

There were twelve jumps in all, six forward and six back.

Eight jumps lasted for one frame, and four lasted for two frames.

The lateral displacements on the animation rostrum varied between 0.003 ins. and 0.024 ins., which corresponds on projection to changes of depth, on the convergence theory, of between 1.5 ins. and 12 ins. seen at a distance of 10 ft.

The same drawing was used for the two outer lines for the right and left eyes so that they were certain to appear at equal depths. To use the same drawing for each eye it was necessary to double expose the film. It was possible for an error in the location of the drawings on the animation table to displace the central line from the plane of the outer two by a small, constant amount. No comment to this effect was made by any observer, and in any case this error was immaterial as it was only the relative movement forward or backward that was being studied.

The width of the lines on the drawing board was about 0.04 ins.

This was magnified on the screen by a factor of 4.34. (Not the conventional 5 because the throw of the projector was restricted.)

The film was projected at 16 frames per second, with the usual projector shutter, i.e. the light was modulated twice each frame to reduce flicker. The waveform of the light output from a projector is usually taken to be square, with equal periods on and off.

Results and Discussion. Tests on five observers were quite conclusive. They were so bad at reporting jumps that any statistical analysis of the results was quite unnecessary. The two frame jumps were not seen any better than the one frame jumps, with projection at 16 frames per second. When projected the picture made random movements all the time due to the irregular movements of the film through the gate of the camera and the projector, and irregular jerks of at least 0.1 in. (i.e. as big as some of the artificially imposed jumps) were clearly visible if the screen was examined at close quarters. The deliberately imposed jumps could be seen by an observer without polaroid spectacles, as they could in the pilot experiment, and they were identifiable because there was a simultaneous movement in opposite directions of the two images of the centre line which were normally superimposed. But this was not noticed by stereo observers - as far as they were concerned the 'signal' did not stand out above the 'noise'.

There was some softness of focus, due mainly to the projection lenses, which resulted in the final line widths being about twice their widths on the drawing board multiplied by the scaling up factor. This might have increased the difficulties of accurate depth perception,

but a similar factor was involved even in the very accurately observed "Photofinish" experiments.

Some lateral drift was introduced by the double exposure of the outer lines. It was visible both with and without polaroid spectacles, and whilst it was irrelevant as far as this experiment was concerned it does indicate another source of error in stereo animation which is as serious as errors in draughtsmanship.

Experiments in flash stereoscopy have been carried out by many workers, including Dove, Langlands (5), Ogle (9) and Wright (10). These experiments all indicate that varying degrees of depth information can be obtained when a configuration is illuminated by a flash of short duration. The classical experiments of Langlands showed that some slight sense of depth was observed even when the test objects were only visible for a hundredth of a second or less, with little further loss when the period of illumination was reduced still further. As the period of visibility was increased, a sudden change in sensitivity was observed in the neighbourhood of one-tenth of a second, but after one second's exposure the rate of increase was again very slow. (We quote this from Charnwood's account (1), where he adds that this evidence has been used to support widely differing points of view.)

This reduction in sensitivity when the time of illumination is reduced below one-tenth of a second is encouraging from the point of view of a film-maker, but in view of the agreement of all of the observers that some degree of depth discrimination does remain, even for the shortest tested periods of illumination, it may cause some surprise that

a jump of short duration did not stand out in a similar way. However in the present circumstances there is the minimum of monocular clues, and the observer saw one set configuration before and after the jump, so the conditions differed materially from those in the flash stereoscopic experiments.

It is a common feature of stereo observation in other circumstances that it may take a long while for the picture to "click" in the observer's mind. This is difficult to measure, and it is even more difficult to get reproduceable results as peoples' reactions vary greatly from time to time, and in any one set of conditions an observer learns and his reaction time is usually cut down with practice. When stereo photographs of physical objects are viewed a three-dimensional impression is given immediately, if only by the monocular clues. However the slowness of an observer's reactions to purely stereoscopic information can be seen if pictures are viewed pseudoscopically. Geometrical diagrams of cylinders and cubes are often interpreted as such even when viewed pseudoscopically, and it may take some time (20-30 seconds or more) for an observer to realise that the picture should be interpreted as the frustum of a cone or as a wedge-shaped solid. Pictures of landscapes or people may never take on their theoretical pseudoscopic distortions because the other clues to depth are altogether too strong.

Some recent work by Julesz (3) is also relevant. He constructed stereo patterns of dots by means of a computer and a CRT transducer. This produced a lattice of dots of random brightness. Stereo pairs were constructed in which a square in the middle of the field was raised up

or lowered. This configuration was completely devoid of monocular clues, and it often took observers some minutes before they interpreted the pictures correctly. Julesz discusses his results in terms of the schema in Fig. 10.1. (In this I dot some lines when Julesz draws them all in full.)

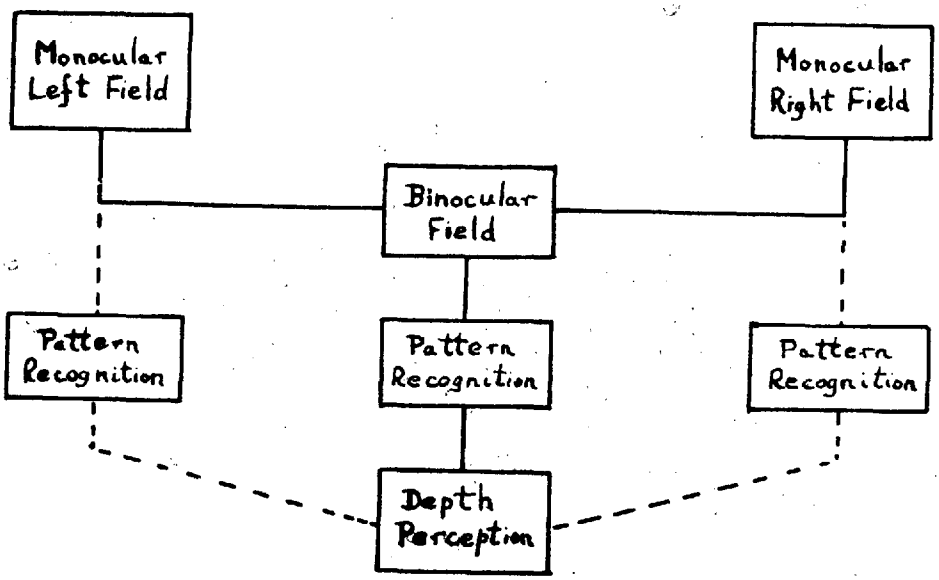


Fig. 10.1.

He decides that, under the highly unusual conditions of his experiments, depth perception is closely related to pattern recognition in the binocular field; i.e. information flows principally along the lines shown in full in the diagram and the flow along the dotted lines is less important. (I have already given reasons above why the reverse may be the case under other viewing conditions.) However, when both monocular

and binocular factors are involved in depth perception it does seem that the "flow" along the solid lines in Julesz's scheme is slower than the flow along the dotted lines (perhaps by a factor of 100 or even 1000).

The observation on the present experiment accorded with this; it being particularly noticeable how a pair of jumps in opposite direction were clearly visible without polaroid spectacles, but were most usually missed by an observer with them.

The present results should also be considered in relation to knowledge of eye movements.* These are of three main types, "fine tremors", "slow drifts" and "sudden flicks" (2, 4, 8). The first are of the order 10" (for horizontal movements) at a frequency of 30-80 cycles per second. The slow drifts are corrected about every $\frac{1}{2}$ second by sudden flicks of 3'-14'.

The function of these movements in refreshing the nerve fibres in the eye is not our concern here, but they may also have some function in increasing visual acuity, for it is remarkable that stereoscopic acuity goes down to 10" or below for trained observers (4) whereas the foveal cones subtend an angle of about 40".

Errors on the drawing board, when translated into parallaxes on the screen are far above the limits of binocular acuity, but they are below the magnitude of the sudden involuntary flicks which the eye is continually making, since 0.01 ins. on the drawing board corresponds to about 1' at the eye of the observer when he sees the final picture projected on the screen. There is some mechanism by which the brain

* FOOTNOTE. Refer back to the Table in Chapter 8.

ignores saccadic jerks in the eye, and whilst there is admittedly a very great difference between movements in the eye and movements of the object observed it is not surprising that jerks of a similar magnitude are ignored under stereoscopic projection conditions.

The degree of importance of convergence in depth perception is undecided. Wright says that "convergence and disparity both contribute to the stereo sense, and the wider the field of view the more important does convergence become". The fields of view with which he was concerned are considerably wider than here. Ogle does not agree (9).

Krauskopf et al. show that the function of the saccadic jerks is to bring the eyes "back on target" after each slow drift away. Saccades are always correlated in the two eyes, and a secondary effect of this is to correct convergence errors at the same time. The standard deviation of the convergence of the eyes is 2' (the same order of magnitude as the saccadic jerks) whereas, they report, stereo acuities of the order of 2" have been reliably obtained. In view of this it is difficult to see that convergence (as such) is of major importance, since it has such a surprisingly high standard deviation and it is only corrected every $\frac{1}{2}$ second or so. In any case this does confirm that the convergence factor requires longer to act than some other factors in depth perception.

Levonian (6) has discussed the related problem of the relative movement of the two pictures which is inevitable in stereo systems employing two cameras and two projectors. He gives figures which show that under the least favourable circumstances this can result in unsteady registration of the two pictures on the screen with errors of about 6' of

arc horizontally, and 3.5' of arc vertically. This, he says, gives no trouble, because "the reaction time of the extraocular muscles is about one-tenth of a second, so the eyes cannot possible follow the 24 cycles per second image movement due to lack of registration". (The difference between projection speeds of 16 frames per second and 24 frames per second is obviously slight.) Whilst Levonian's conclusions support the view that small variations in drawings will not be detected the variations which occur in the present experiment are definitely more critical than those which Levonian discussed; because he was concerned with relative movements of particular elements of the pictures, while the picture as a whole remains steady.

Krauskopf et alia do not discuss fine tremor as it is too small to be significant in the problems which concern them; but it is noteworthy that just as bad errors on the drawing board are of about the same order of magnitude as the eyes' saccadic jerks so the limits to which the author's animation equipment will work are very much the same order as the eyes' fine tremor. The panning bars are calibrated in 0.001 ins. and will work to a quarter of a division or less. This corresponds in the finished film to about 2" at the observer's eye, but of course fine tremor of this magnitude is lost completely on projection in the vibrations of the projector. However the frequency of the fine tremor is interesting, because maximum binocular acuity (2") being so far below the limits of tremor (about 40" in the horizontal direction) it is clear that our depth perception mechanism ignores at least one type of oscillation of a frequency of 30-80 cps. Film projection rates are not so far outside this range.

The conclusion is therefore that the failure to observe jumps of a magnitude which is clearly visible in static tests is not surprising, and it is understandable when considered in relation to the large number of interacting factors which are involved in even the simplest stereo-experiments which may be devised.

REFERENCES

1. Charnwood. An Essay on Binocular Vision, London, 1950.

2. R.W. Ditchburn. Eye movements in relation to retinal action. Optica Acta, 4, p.171-176, Feb., 1955.

3. B. Julesz. Binocular depth perception of computer-generated patterns, B.S.T.J., XXXIX, 5, Sept. 1960, 1125-1161.

4. J. Krauskopf, T.H. Cornsweet and L.A. Riggs. Analysis of eye movements during monocular and binocular fixation. J.O.S.A., 50, 6, June 1960, 572-8.

5. Langlands. Trans. Opt. Soc. XXVII, 45, 83.

6. E. Levonian. Stereography and the physiology of vision, J.S.M.P.T.E., 62, Mar. 1954, 199-207.

7. do. Stereography and the transmission of images, J.S.M.P.T.E., 64, Feb. 1955, 77-85.

8. M.P. Lord and W.D. Wright. Eye movements during monocular fixation. Nature, 162, p.25, July 3, 1948.

9. K.N. Ogle. Stereoscopic acuity and the role of convergence. J.O.S.A., 46, 4, p.269-273, April, 1956.

10. W.D. Wright. The role of convergence in stereoscopic vision. Proc. Phys. Soc. B, 64, p.289-297, 1954.

11. Ishak, Hefzalla and Badbury. A letter. J.O.S.A., 46, p.303, April, 1956

II. EXPERIMENT "DRIFTER"

Summary. A stereo film was made showing only a vertical line segment, white against a dark background, which moved intermittently, to and from the observer at varying speeds. The objects of the experiment were to see whether the fluctuations in depth were correctly perceived, and to see if the segment appeared to vary in length according to the predictions of the convergence theory.

Theory. The theory of Chapter 4 applies; and so we have:-

the line segment $(0, y, z)$ is mapped on the drawing board by $X = 0$, $Y = sy/(s + z)$, $Z = wz/(s + z)$. The distortions to be expected are described by

$$y' = \frac{w'sy}{m_z ws - (p + m_z w - w')\zeta} ,$$

$$\zeta' = \frac{w's'\zeta}{m_z ws - (p + m_z w - w')\zeta} ,$$

so

$$\frac{dy'}{y'} = \frac{(p + m_z w - w')d\zeta}{m_z ws - (p + m_z w - w')\zeta} ,$$

$$= \frac{V}{s - V\zeta}\, d\zeta ,$$

where $$V = \frac{p + m_z w - w'}{m_z w} .$$

So $$\frac{dy'}{y'} = V \cdot \frac{d\zeta}{s} \quad \text{approx.}$$

This means that if the parameters p, m_z, w and w' are incorrectly balanced the line segment will appear to change in length, increasing or decreasing as it recedes as $V > \text{or} < 0$.

Practical design. This movement can be synthesised on the animation rostrum by moving a vertical line segment behind a black mask, the angle of the mask depending on the eyewidth of the observer. Fig. 10.2 shows the arrangement for the right eye. The arrangement for the left eye is similar.

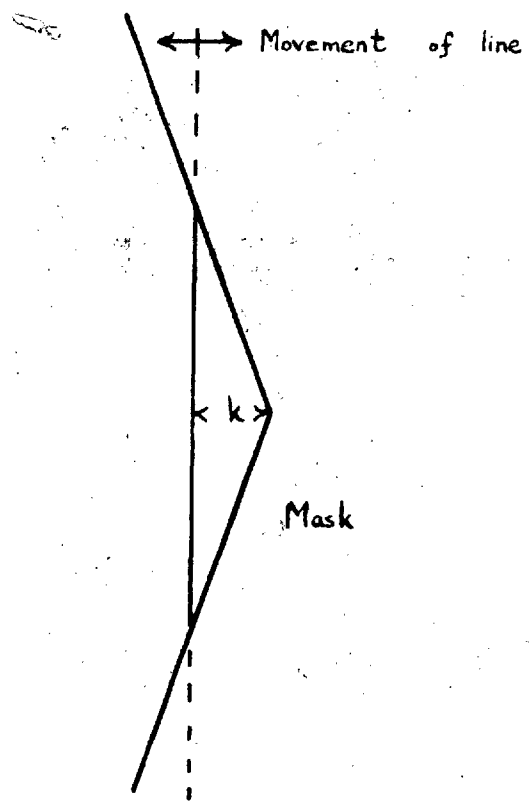


Fig. 10.2

For a theoretically correct projection the length k should be equal to w on the screen, so on the drawing board it should be w/m.

The forward and backward movement of the line was restricted to the range $\frac{1}{2}s$ to 2s, where s was the screen distance of 12 feet. Earlier static experiments had shown that it is desirable to limit the picture range in this way. The speeds of movement used were such that the line moved through this range in a time of some four or five seconds, which meant that the parallaxes had to be controlled accurately and varied very slowly. This was achieved by mounting the moving line segments on panning bars which were moved by screw threads calibrated to 0.001 in. The movement was punctuated by occasional pauses. There seemed no theoretical objection to this, and it reduced the time spent on animation.

Three scenes were shot with different values of k which amounted to working with values of w (the semi-eye width) on 1.24 ins., 1.04 ins. and 1.44 ins. This covers the range of eyewidths measured by French (1). The effective value of V could also be varied by varying the throw of the projector, since this gives a direct variation of m.

Results and Discussion. The distortions predicted by the convergence theory were not observed. Observers tended to see the picture in one of the two states. Either the line appeared to stay in the plane of the screen (this is how it was described, but we know that such assessments of absolute distance are quite unreliable) and to pulsate in length, or else it seemed to move backwards and forwards while its length remained constant.

Observers rather vacillated between the two states. It is possible

that training might encourage the second state, but it was not the object of the experiment to assess the effect of training.

No observer reported any significant difference between the three scenes with different values of k, or observed the even greater distortions produced by decreasing the throw of the projector from 15 feet to 11 feet. This is in complete accord with results obtained in other parts of the programme. The distortions which were sought were projective distortions, i.e. they were outside the group of affine distortions which are significant in stereo projection, and they were part of the group of distortions involved in such experiments as "Overlap" which are not significant - at least to the first order.

The observers were set a rather difficult task, since it involved comparing unlike things - a length seen with a length as it was previously - and judicious balancing of lengths in two different directions perceived by different mechanisms. Experiments "Guesswork" and "Overlap" had both revealed the difficulties of length comparisons of this type, so it is not surprising that the observers rationalised the situation, and either declared that they were seeing something that corresponded closely with everyday experience of moving physical objects, or discounted one feature of the diagram (changing parallax) and accounted for the other as simply as possible (changing size).

This complete discounting of parallax clues and a rationalisation of the remaining phenomena can be seen elsewhere. For example in the refusal to see stereo pictures of familiar physical objects pseudoscopically, and in some experiments by Ames on balloons (2). In these

observers were shown two balloons at a distance of about 10 feet. The sizes of the ballons could be varied by working a pump, and the brightness of the illumination was also variable. There was a pronounced tendency to think that the brighter balloon was nearer, and to think that variations in size were to be accounted for by the targets coming and going. Parallax clues, of course, conflict with these interpretations, but in Ames's experiments the observers were very willing to ignore such clues in order to arrive at a consistent everyday interpretation of the remaining depth clues.

Variations in eye-width are also, theoretically, relevant in this experiment, but the observations of three observers were in such complete agreement, even when the value of V was altered by reducing the throw of the projector, that it was quite unnecessary to measure the eye-widths of the observers (as had been originally intended) in order to account for any small differences. Differences due to naturally occurring variations in eye-width should be quite significant in situations such as this on the convergence theory. For example, taking the equation above for $\frac{dy'}{y'}$ and putting $m = 1$, $w = 1.24$ ins., and $w - w' = 0.2$ ins., (corresponding to the range reported by French) and $d\zeta / s = 3/2$ (the variation in this experiment) we have $dy'/y' = 1/4$. This is to say that in the course of its movement between the far and near points the length of the line segment should alter by about 1/4. One would have expected an effect of this magnitude to be noticeable, at least by experienced observers. Further, it should be present in all systems of this type, since the relevant parameter is associated with the observers and is quite independent

of the particular system involved. However the literature does not seem to record any unwelcome effects in the projection of stereo films due to variations in the eye-widths of observers, although the situation with binocular precision instruments may well be different.

Conclusions. When faced with a configuration as simple as the one in the experiment observers do not necessarily interpret the parallax information correctly, but are prepared to ignore it and to rationalize the remaining information.

Observers cannot detect variations due to the distortion factor 'V', at least for the restricted range of values likely to occur if normal care is taken with the projection conditions.

This experiment reinforces the conclusions of other experiments that it is difficult to compare vertical lengths with lengths receding into the picture.

REFERENCES

1. J.W. French. Trans. Opt. Soc. XXIII, 44 and 220, 1921-2.

2. Scientific American Reader. p.578. New York, 1953.

III. EXPERIMENTS WITH McLAREN'S FILM "LINES"

Here we describe some experiments performed by the author and McLaren, using the latter's film "Lines". The main experiment was not strictly concerned with stereosynthesis as it involved presenting quite distinct images to each eye, but it seems reasonable to describe it here as the film had been made by synthetic techniques and was projected on a stereo system. It shows how the combined use of these techniques gives an experimental situation in which fresh aspects of vision can be investigated.

McLaren's film Lines is an abstract animated film with a musical accompaniment. It is in two parts. In the first part vertical, or slightly sloping, white lines which extend from the top to the bottom of the screen perform sideways movements against a plain background of varying colour. The second half of the film is an optical rotation of the first by 90°, with a changed colour scheme, the moving lines being black and the background tints differing from those used before.

Two experiments were done with this material. The section with vertical lines was projected autostereoscopically; and whilst this had never been considered while the film was being made it was found that a two frame stagger gave a most pleasing sensation of depth to many of the scenes just as it had done in "Around is Around", which McLaren had made for the stereo-cinema at the Festival of Britain.

Fresh ground was explored by projecting the two halves of the film simultaneously (through two synchronised projectors fitted with polaroid) - one version to each eye. This gave rise to some most surprising phenomena, because whilst one eye was shown a set of vertical lines going

from the top to the bottom of the screen, and the other eye a set of horizontal lines going right across the combined effect was always of a mesh of lines, but with gaps which were continually changing. (See Fig.10.3)

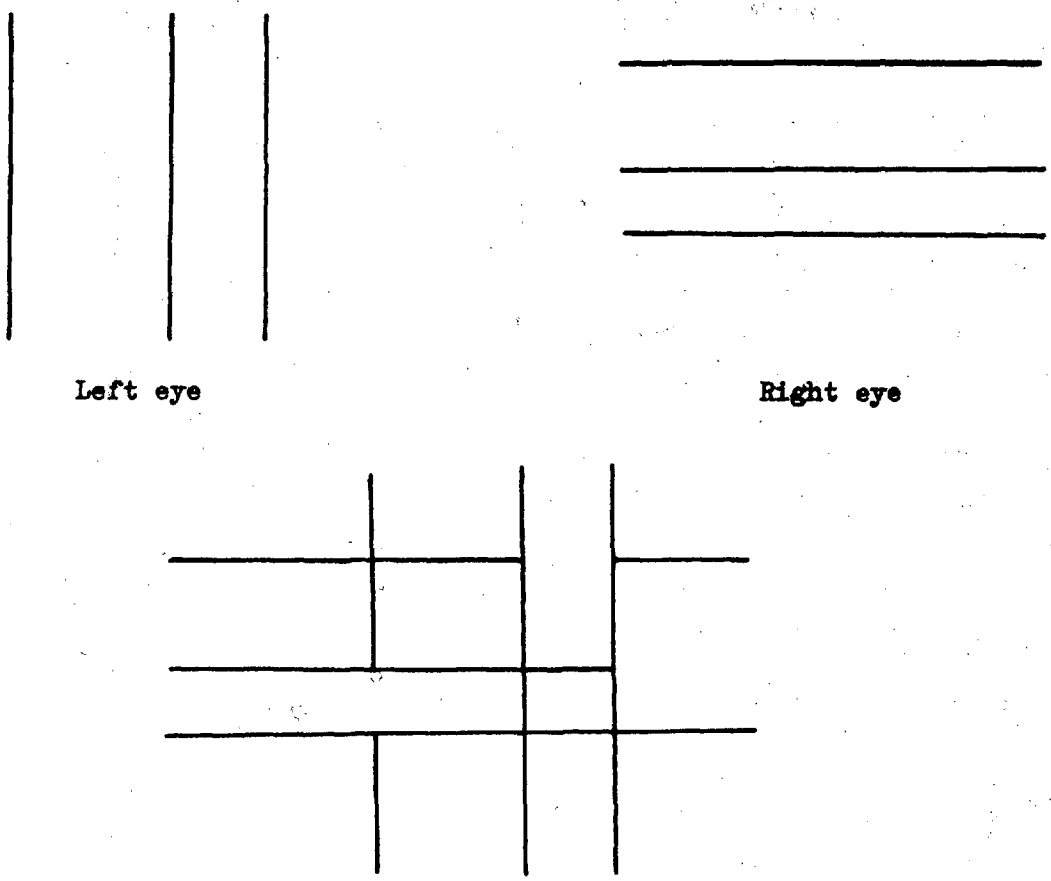


Fig. 10.3.

Observations on totally different static pictures, viewed in a stereoscope, have been widely reported and it is well known that features of the sensed configuration come and go; but this addition of movement

to the situation increased the effects which have been described to a startling extent. The horizontal lines frequently appeared broken, but they were always broken at points where they were crossed by vertical lines, and the movement of the vertical lines greatly encouraged the illusion, closing up gaps and opening out new ones as they moved to and fro. Some gaps were also to be seen in the vertical lines, but these were much less frequent. The reason for this was presumably that observers tended to interpret the configuration as a material structure in which it was mechanically necessary that the vertical lines should not be interrupted.

A subsequent experiment showed that similar effects can be observed with still pictures on a Brewster stereoscope, but the effects are not so remarkable in static configurations.

These results are very reminiscent of Pritchard's results (1) on viewing under conditions in which the position of the image was fixed on the retina. He reports that parts of the image disappear and reappear in a random fashion, but that familiarity with the shapes involved affects the parts which come and go. (Pritchard's viewing was monocular, and an essential feature of the experiment being described was that each image was seen by one eye only. There was no question of stereosynthesis.) Pritchard remarks that "perceptual elements (as distinct from sensory elements) exist in their own right" and that in fixed image conditions entities that are recognisable shapes are visible longer than meaningless ones. The same is true in the present situation, and the two main points which the experiment emphasises are

i) the condition of fixed image viewing is not essential for this type of phenomenon,

ii) movements in the field of view of the other eye can induce parts of the image to appear and disappear.

It might be worth while to make an animated film, in structure very like some of McLaren's entertainment work, in which different configurations of the type used by Pritchard are presented separately to the two eyes and where movements in the two fields of view overlap one another.

This experiment differs from Pritchard's in one other important respect. Whilst the effects described were most striking to some observers others could not see them at all; whereas when viewing with the image fixed on the retina parts of the image come and go for all observers without exception.

REFERENCES

1. R.M. Pritchard. Stabilized images on the retina. Scientific American, 204, 6, June, 1961.

CHAPTER ELEVEN

SYMMETRY GROUPS IN STEREO-ANIMATION

The author has shown elsewhere (1) how a re-interpretation of the known symmetry groups associated with three-dimensional configurations gives groups which describe the symmetries in space and time of cyclically repeating two-dimensional diagrams. In particular 194 of the 230 crystal groups apply to film animation. Appreciation of these groups is of practical importance because every n-fold symmetry offers in principle the possibility of dividing the labour of animation by n. Symmetries which are purely spatial only offer a reduction in drawing at the expense of making multiple exposures at the photographing stage and this is not usually any gain, but the glide symmetries which involve using the same drawings in different positions at different times in the animation cycle are definitely useful. In stereo-animation symmetries in space alone have an extra importance since multiple exposure is a more accurate way than drawing of reproducing parallaxes in different parts of the diagram.

We will now consider the groups which arise in stereo-animation. The stereo-animator is not concerned with a truly three-dimensional domain, but with a pair of two-dimensional projections, which we will henceforth call L and R. Hence there is no question of discussing the dynamic groups in three-dimensions which are related to four-dimensional static configurations as the two-dimensional animation groups are related to the three-dimensional static groups. The number of truly three-dimensional dynamic groups must be very large indeed, our problem is rather more simple, but we still do not attempt a complete enumeration. The symmetry

operations involved are the spatial ones which were involved in the two-dimensional case, coupled with the interchange of L and R, and the superposition of one on the other in various ways. By definition we are concerned with diagrams which repeat cyclically; these may be reversible in time (which is another symmetry) and there may be glide symmetries in which the overall diagram is invariant under a change of phase by a half-period coupled with a spatial operation of some kind.

These symmetries can reduce the labour of stereo-animation in various ways. They occur first, historically, in the earliest animations in the stroboscope, by Plateau and Stampfer, and in the recognition by Wheatstone and Brewster that the pair of stereo drawings of a symmetric object, dead ahead, are reflections of one another.

The symmetry groups which arise in this context do not appear to have been studied before, but the number of different cases which can arise is very large and a systematic enumeration would be tedious (at least as complex as the enumeration of the 230 crystal groups). Knowledge is not therefore as advanced as in the case of 'flat' animation where it would be possible to draw up programmes to fit all possible cyclic animation routines, so that, for example, an animation camera could be controlled by a computer having a large measure of auto-programming. The list of operations given below is probably complete, and the number of groups can be overestimated by saying that all the groups which occur are the direct product of two groups of the 'flat' type and the cyclic group of order 2, or sub-groups of one of these product groups. This would give a systematic enumeration if it were known which combinations are to be

rejected; for there is the additional requirement that L and R must always be composed of points in homologous pairs, and it is not clear exactly how this restriction limits the groups which actually arise.

The practical importance of these symmetries should not be over-emphasised as there are two reasons why they are rather false friends to the animator. Many of them involve pseudoscopic transformations, the applications of which are rather limited, when a true reflection in the plane of the screen would be extremely useful; and the basic equations of stereoscopy in Chapter 4 show that parallaxes not only play a strong role in determining z coordinates but also play a weaker part in determining x and y coordinates. This means that many animations achieved by symmetry operations are accompanied by projective transformations which it is often desirable to avoid.

The simple autostereoscopic operation described in Chapter 8, whereby a depth effect is obtained by presenting the same pictures to the two eyes in different phase, is, in general, asymmetric and does not possess the group property. But there are more elaborate systems of symmetry which contain a superficial element of autostereoscopy where it will be found that the group arises from other considerations, the autostereoscopy being present because there is a glide symmetry in time.

Symmetry groups arise when the symmetry operations which are permissible in the context have been defined. The following operations are involved in stereo-animation, and they may produce, destroy or change stereopsis in various ways.

i) interchange of L and R,

ii) simultaneous reflection of L and R, each in a line in its own plane,

iii) simultaneous rotation of L and R, each about a point in its own plane,

iv) simultaneous translation of L and R,

v) reflection of one of L or R only,

vi) rotation of one only,

vii) translation of one only, with or without a glide reflection,

viii) shift in time of L or R, the other remaining fixed,

ix) reversal of the time scale,

x) time shift of a half period together with a spatial transformation of some kind.

Some of these operations will now be considered in detail.

Translational symmetries play a comparatively unimportant role in stereo-animation because the stereo observer is tied to a fixed point in the space which he inhabits. Spatial symmetries are destroyed by central projection. Simple translations of a configuration in stereo space cannot be produced by simple panning movements of the drawings, as in flat animation, because the aspect is continually changing. (In the author's film Plucked Strings a travelling wave is animated by a mechanical panning movement. It would have been useful to have employed a similar device in Resonant Cavities to show that stationary waves in the resonator are produced by the interference of travelling waves. But in stereo this cannot be done.)

It is possible to draw stereo pairs by parallel projection, instead

of conical projection. The crystal lattices of von Mieses and von Laue are done this way. The theory of this process is very simple since parallax is directly proportional to depth, and the representation is effective enough if the field is limited. In this system it is possible to produce rigid translations by panning movements and so the symmetry groups involving translations are of far greater importance.

In stereo-animation by conical projection the distinction made in crystallography between space groups and point groups applies, and the most important symmetry groups are the point groups in which the origin is fixed.

The relative movement of L and R poses problems peculiar to the stereo situation, and useful animations can be produced in this way.

EXAMPLE. Let L consist of the set of points

$$Y_p = 0, \qquad X_p = p \qquad p = 0, \pm 1, \pm 2 \ldots$$

and R consist of the set

$$Y_q = 0, \qquad X_q = q + \delta_q \qquad q = 0, \pm 1, \pm 2 \ldots$$

where the δ_q are all less than the distance apart at which two homologous points can be fused.

Now animate with L fixed and R moved one step to the right at each frame. An animation results in which the depth of the points (and to a lesser extent their x-coordinates) varies from frame to frame. It is possible to animate a wave motion by arranging that $\delta_q = \sin kq \times$ constant.

Similar considerations apply to a vertical movement of one drawing when they are composed of the sets

$$L: \qquad X_p = 0 \qquad Y_p = p,$$

$$R: \qquad X_q = \delta_q \qquad Y_q = q.$$

This procedure is restricted by the inter-dependence of the z-coordinate and the others, but if strict quantitative accuracy is not required animations of such subjects as vibrating crystal lattices can be obtained in this way.

Some configurations are invariant under a shift in time of L or R, the other remaining fixed. This can arise if homologous points oscillate horizontally, one having a period n times the other.

Again there are three-dimensional configurations which are invariant under rotations of L and/or R by $2\pi/n$, with or without an interchange. 'n' need not be the same for the two.

EXAMPLE. In Fig. 11.1 the dot and the cross denote the points seen by the right and left eyes, and the diagrams show configurations which are invariant under rotations of $2\pi/n$ for $n = 2,3$ and 4; or which are invariant under rotations of π/n with interchange. (The scale of the configuration is assumed small enough for the separation of homologous points not to exceed the fusion distance.)

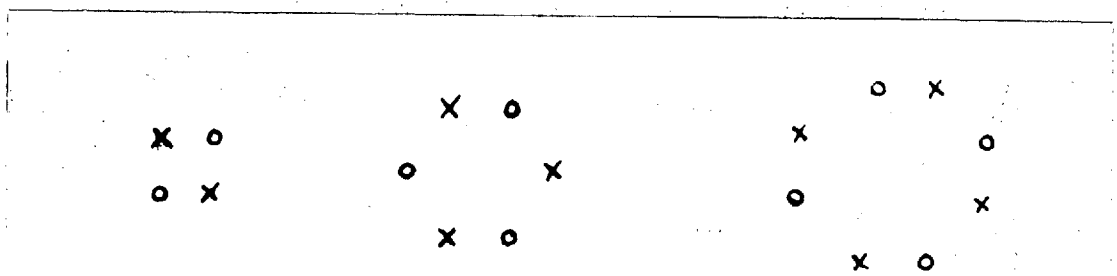


Fig. 11.1.

EXAMPLE. L and R may consist of two regular n-gons, with different values of n but both of the same parity. The n-gons have vertical axes of symmetry and the same vertical height. Then if the separations of homologous points is not too large these diagrams are a stereo pair for a twisted polygon in space. One of these may even be a circle.

These two examples show very clearly that the symmetries of the two-dimensional drawings may be very different from the symmetries of the associated three-dimensional configuration. The theory of the prismatic class of symmetry groups which occurs in two-dimensional animation (1) could be extended to deal with these stereo cases, but the question will not be taken any further here. The only rotational symmetries considered henceforth will be the rotations of π which inevitably result from a combination of reflections in the x and y axes.

A wide class of groups involves the following operations, for which a special notation will be introduced. Let

A, B and C denote reflections of the drawings in the X-axis, Y-axis and O,

P denote an interchange of R and L,

T denote a reversal in time,

subscripts r and l denote that an operation is performed on R or L only,

superscript *g* denotes than an operation is accompanied by a glide in time of half a period.

A reflects the three-dimensional configuration in the xz plane,

B reflects it in the yz plane, and simultaneously pseudo-reflects it in the xy plane,

C reflects it in the Oz axis, and simultaneously pseudo-reflects it in the xy plane.

The transformations with subscripts r and l can only be performed in special cases, when there are still pairs of homologous points in R and L to fuse after the transformation has taken place.

P is the only operation which does not, in general, commute with the others; and there are numerous identical relations among the operators, such as

$$A_r = PA_lP, \qquad A = A_rPA_rP \qquad \text{etc.},$$

the product of any two of A , B and C gives the third, where

$$A_s = A, A_r \text{ or } A_l, \text{ etc.},$$

the product of any two of A_sg, B_sg and C_sg gives the third, without the g.

Ignoring the glides the full group containing all the operators is of order 64, and the other groups are sub-groups of this. The other possible groups of this class arise by replacing one or two of the simple symmetries by glides. These groups have two aspects, the operations of the group may be used to generate the complete figure from a portion of its elements, or the emphasis may be on the invariance of the entire figure under the group transformations. The following examples bring out the two points of view.

EXAMPLE. L and R each consist of four points, one in each quadrant. The eight displacements from the x axes are all equal; whereas the eight displacements from the y axes are different, those on L being $\frac{1}{2}$, $\frac{3}{2}$, $\frac{5}{2}$ and $\frac{7}{2}$, and those on R being 0, 4, 8, 12.

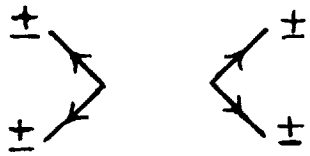
There are 32 ways in which the diagrams may be registered under the camera and this produces parallaxes $\pm(2n + 1)/2$; $n = 0, 1, \ldots\ldots 15$. These can obviously be taken in sequence so that an oscillation of period 64 is produced in any one quadrant and in fact an arrangement can be found which produces oscillations in two of the quadrants simultaneously, 180° out of phase. It is doubtful if a system can be found which will oscillate all four points in a coherent manner, but this can be done by appropriate masking and double exposure. Animators would be pleased to find methods of producing a 64 frame animation from only two drawings more often.

The remaining examples show particular configurations belonging to groups containing various operations. An ad hoc notation is employed which is self-explanatory. Dots and crosses denote points seen by the different eyes (it does not matter which), and arrows denote a movement. + and - denote displacements behind and in front of the screen.

EXAMPLE	P, A, B^g, C^g	See Fig. 11.2
EXAMPLE	T, A, B^g, C^g	See Fig. 11.3
EXAMPLE	A_r^g, B_l^g, C	See Fig. 11.4
EXAMPLE	A_r, T, A_l^g	See Fig. 11.5

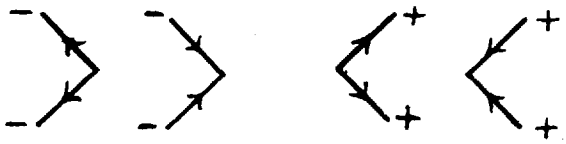
REFERENCE

1. T.J. Fletcher. Film Groups, Math.Gaz., XL, 331, p.15-19.



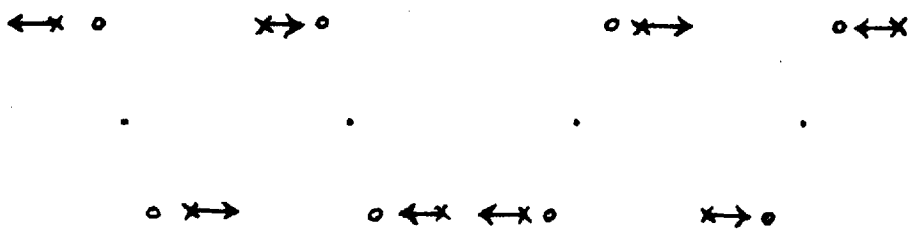
Two half cycles

Fig. 11.2.



Four quarter cycles

Fig. 11.3.



Four quarter cycles

Fig. 11.4.



o fixed. × oscillate with different amplitudes

Fig. 11.5.

CHAPTER TWELVE

FUTURE APPLICATIONS OF STEREOSYNTHESIS

Introduction. The invention of alphabetic script as a means of communication was a major achievement, but it is not the only form in which logically connected information may be conveyed. In some respects the emphasis on the printed word in current scientific communication is a result of the historical accident that other aids to communication were invented much later than the printing press, just as in earlier times the reason for the Greek insistence on verbal reasoning in geometry was the material circumstance that it was not possible at that time to produce non-verbal, or non-literal, structures of the flexibility required for a lengthy mathematical argument. There have been many occasions in the history of mathematics when the solution of a problem has been principally a matter of finding an appropriate notation. Likewise work which was discovered by the great men of the past is very often made available to a far greater public by the invention of some more suitable notation.

The film is firmly established as a means of documentation in medicine; it also has an unquestioned place in the biological and social sciences, and in the physical sciences where change of the time scale can make visible previously unsuspected phenomena (3, 4). In the hands of Marey and Mueybridge the embryo scientific film was a research tool and a means of exposition even before the rapid sequence camera gave birth to the cine-camera proper. The mathematical film has been slower to develop, but it has provided the twentieth century not merely with an illustrative adjunct to conventional methods but also with the possibility of entirely

new mathematical and scientific notations which are potentially capable of presenting logical trains of deduction within their own system of conventions (1, 2).

The philosophy of communication by film will not be discussed here, we will merely observe that the stereo film is potentially richer than the printed book because it presents information in a three-dimensional continuum instead of a one-dimensional sequence. This potential richness will not be realised until the medium has developed a wealth of conventional symbols and conventional associations to use in addition to the naive pictograms which are graphical illustrations at the moment.

Practical Techniques. The techniques of design developed in this thesis are adequate for the production of moving diagrams in three dimensions. The usual pieces of advice on stereo drawing (to be found in the classical references) should be followed, and if reasonable care is taken to control parallaxes the techniques of film animation and stereoscopy may be combined successfully.

Stereosynthesis has the advantage over stereo-photography that it is possible to avoid the region of monocular vision at the sides of the picture.

The animated stereo film is, practically speaking, restricted to line illustration since it has not been demonstrated that there are any effective air brushing techniques to produce subtleties of shading and of aerial perspective in the moving elements, and there is as yet no solution to the problem (fundamental in three-dimensional illustration) of indicating a plane which moves in space but which at the same time permits

other planes to be seen through it. However line illustration will depict a wide range of phenomena in wave and field theory, crystallography, projective and differential geometry, and in the theory of functions of two variables. The medium is seriously restricted by the inability of the observer to "look round corners" or to turn an illustration as he can a spatial model, but it does permit the use of elements which materialise and vanish, and of distortions which are impossible mechanically. It is an excellent medium in which to depict transformations of all kinds. The limitations of the medium become its great glory in the documentation of affine geometry, for the stereo world is an affine world where the motion of the observer induces deformations of the figure which are part of the information which it communicates.

The medium should not be trusted to convey subtleties of Euclidean geometry, for even when these are correctly realised in theory experiment shows that they are often not perceived correctly. This is not necessarily the fault of the medium, but stereo illustrations must be designed for a complete communications chain, from the mind of the designer through the drawing board, camera, film, projector, screen and polaroid spectacles to the eyes and mind of the observer. The capacity of the channel is no stronger than the capacity of the weakest link, and psycho-physiological factors must be considered in the design of stereo illustrations every bit as much as geometrical optics.

The stereo film is also an experimental tool as it can produce situations in which to experiment on the phenomena of binocular vision. It can produce configurations which do not have mechanical limitations,

it can divorce stereo clues from others so that they may be studied independently, and it can provide accurately controlled movements which can be reproduced indefinitely under the same conditions. When using stereo film in this way the chief obstacles are the amount of residual movement in the picture, due mainly to the unsteadiness of the projector gate, and the photographic difficulties of getting a really pronounced black and white image with the currently available 16 mm film stock. There is also a certain softness of the image when projecting with the only available 16 mm stereo twin lens. Softness of focus is much more noticeable in line diagram than it is in ordinary photography. Even good quality 35 mm still projectors show some softness on line diagram, and many 16 mm projectors which give a quite acceptable picture on live action cinematography show deteriorations on line diagram.

There is no evidence that phenomena requiring a Luneburg metric are actually observed under the normal conditions of projection of stereo films; particularly when recognisable physical objects are involved and the observer is strongly inclined to rationalise, which means in this context to ignore certain specific pieces of information while integrating the rest into a consistent whole. But this should be investigated more thoroughly by the production of stereo diagram to be viewed at short ranges (six feet or less), and containing very few elements as in the three point and four point alley experiments. If Luneburg phenomena could be produced reliably by stereo animation then the way would be open to quite novel methods of documenting hyperbolic geometry. Also moving stereo diagrams would be a useful experimental technique to investigate

the incorporation of time into Luneburg's theories - about which enigmatical hints were given in his last, posthumous paper.

As it is the range of possibilities opened up by this new technique are wide enough to provide work for many years, and give a challenge to document large areas of mathematics and related scientific disciplines in a new way.

REFERENCES

1. R.A. Fairthome. Film and scientific symbols. Chapter in "Science and Films", ed. B. Lloyd. London, 1948.

2. T.J. Fletcher. Les problemes du film mathematique. Chapter in "Le material pour l'enseignement des mathematiques." Delachaux et Niestle, 1958.

3. A.R. Michaelis. Research Films in Biology, Anthropology, Psychology and Medicine. New York, 1955.

4. " In the Beginning was the word. Univ. Film Journ. I, p.1-6, 8, p.10-15, 1955.

ACKNOWLEDGEMENTS

Grants from the Nuffield Foundation made possible the purchase of a camera and the making of a number of copies of the film for experimental showing, and I would gratefully acknowledge this assistance. Also I would thank the Principal and Governors of Sir John Cass College for accommodating the programme; many friends in the film industry for giving me their time and professional knowledge when discussing my problems; Professor W.D. Wright and Dr. A.E.R. Church who supervised my work (the latter until his retirement); and Dr. B. Spain for his continued encouragement.