Cave Science

The Transactions of the British Cave Research Association

BCRA

Volume 19 Number 3 December 1992

Historical Concepts of Speleogenesis

Cave Detection by Resistivity Tomography

Forum

Cave Science

The Transactions of the British Cave Research Association covers all aspects of speleological science, including geology, geomorphology, hydrology, chemistry, physics, archaeology and biology in their application to caves. It also publishes articles on technical matters such as exploration, equipment, diving, surveying, photography and documentation, as well as expedition reports and historical or biographical studies. Papers may be read at meetings held in various parts of Britain, but they may be submitted for publication without being read. Manuscripts should be sent to the Editor, Dr. T. D. Ford, at 21 Elizabeth Drive, Oadby, Leicester LE2 4RD. Intending authors are welcome to contact either the Editor or the Production Editor who will be pleased to advise in any cases of doubt concerning the preparation of manuscripts.

NOTES FOR CONTRIBUTORS

These notes are intended to help the authors to prepare their material in the most advantageous way so as to expedite publication and to reduce both their own and editorial labour. It saves a lot of time if the rules below are followed.

All material should be presented in a format as close as possible to that of *Cave Science* since 1985. Text should be typed double-spaced on one side of the paper only. Subheadings within an article should follow the system used in *Cave Science;* a system of primary, secondary, and if necessary, tertiary subheadings should be clearly indicated.

Abstract: All material should be accompanied by an abstract stating the essential results of the investigation for use by abstracting, library and other services. The abstract may also be published in *Caves and Caving.*

References to previously published work should be given in the standard format used in *Cave Science.* In the text the statement referred to should be followed by the relevant author's name and date (and page number, if appropriate) in brackets. Thus: (Smith, 1969, p. 42). All such references cited in the text should be given in full, in alphabetical order, at the end. Thus: Smith, D. E., 1969. The speleogenesis of the Cavern Hole. Bulletin Yorkshire Caving Assoc., Vol. 7, p. 1-63. Books should be cited by the author, date, title, publisher and where published. Periodical titles should be abbreviated in standard style, or, where doubt exists, should be written out in full.

Acknowledgements: Anyone who has given a grant or helped with the investigation, or with the preparation of the article, should be acknowledged briefly. Contributors in universities and other institutions are reminded that grants towards the cost of publication may be available and they should make the appropriate enquiries as early as possible. Expedition budgets should include an element to help publication, and the editor should be informed at the time of submission.

Illustration: Line diagrams and drawings must be in black ink on either clean white paper or card, or on tracing paper or such materials as Kodatrace. Anaemic grey ink and pencil will not reproduce! Illustrations should be designed to make maximum use of page space. Maps must have bar scales only. If photo-reduction is contemplated all lines and letters must be large and thick enough to allow for their reduction. Letters must be done by stencil, Letraset or similar methods, not handwritten. Diagrams should be numbered in sequence as figures, and referred to in the text, where necessary, by inserting (Fig. 1) etc. in brackets. A full list of figure captions should be submitted on a separate sheet.

Photographs are welcome. They must be good clear black and white prints, with sharp focus and not too much contrast; prints about 15 x 10 cm (6 x 4 inches) are best; if in doubt a selection may be submitted. They should be numbered in sequence, but not referred to in the text, except where essential and then after discussion with the Production Editor. A full list of plate captions, with photographer credits where relevant, should be submitted on a separate sheet.

Tables: These should not be included in the text but should be typed, or clearly handwritten, on separate sheets. They should be numbered in sequence, and a list of captions, if necessary, should be submitted on a separate sheet.

Approximate locations for tables, plates and figures should be marked in pencil in the manuscript margin.

Copyright: If any text, diagrams or photos have been published elsewhere, it is up to the author to clear any copyright or acknowledgement matters.

Speleological expeditions have a moral obligation to produce reports (contractual in the cases of recipients of awards from the Ghar Parau Foundation). These should be concise and cover the results of the expedition as soon as possible after the return from overseas, so that later expeditions are informed for their planning. Personal anecdotes should be kept to a minimum, but useful advice such as location of food supplies, medical services, etc., may be included, normally as a series of appendices.

Authors will be provided with 20 reprints of their own contribution, free of charge, for their own private use.

Manuscripts on disk are welcome, as text may be set directly from them. Please submit a hard copy to the Editor in the normal way, and advise him that you have a disk, which you can submit after any editorial corrections.

If you have any problems regarding your material, please consult either of the Editors in advance of submission.

Cave Science

TRANSACTIONS OF THE BRITISH CAVE RESEARCH ASSOCIATION

Volume 19 Number 3 December 1992

Contents

Cover: The main streamway in the low level of Tiencuan Dong, in the Xingwen karst of Sichuan, China. Initial development of a phreatic tube, following the bedding and linking between major joint fissures, emphasizes the geological control of cave inception; progressive recognition of these early geological influences has prompted the evolution of concepts concerning speleogenesis, as traced by David Lowe in this issue of Cave Science. Photo: Tony Baker, China Caves Project.

Editor: Dr. T. D. Ford, 21 Elizabeth Drive, Oadby, Leicester LE2 4RD.

Production Editor: Dr. A. C. Waltham, Civil Engineering Department, Nottingham Trent University, Nottingham NG1 4BU.

Cave Science is published three times a year by the British Cave Research Association and is issued free to all paid up members of the Association.
1992 subscription rate to Cave Science is £16.00 per annum (postage paid).
Individual copies, back issues and details of annual subscriptions to Cave Science, and of Association membership, can be obtained from the BCRA General Administrator, 20 Woodland Avenue, Westonzoyland, Bridgwater TA7 0LG.
The Association's permanent address is: BCM BCRA, London WC1N 3XX.

ISSN 0263-760X

CAVE SCIENCE Vol. 19, No. 3, December 1992
Transactions of the British Cave Research Association

A Historical Review of Concepts of Speleogenesis

D. J. LOWE

Abstract: During more than one hundred years of research into the genesis of limestone caverns many contributions and theories have appeared, most of which have added to the understanding of how caves develop. Few workers have adequately addressed the question of how caves originate. Relatively recent advances, including the recognition of the types of dissolutional process active at the interface between fresh water and salt water beneath young carbonate land masses, have allowed reassessment of the potential timescales of speleogenesis and of the dissolutional processes and driving mechanisms which might be involved in cave inception. This review attempts to examine the history of cave development theories by reference to the work of a selection of authors, and to assess the validity of these ideas when compared with a modern inception horizon hypothesis of the origin of limestone caverns.

A PhD thesis *[The origin of limestone caverns: an inception horizon hypothesis]* submitted by the present author (Lowe, 1992) attempted to provide a new working hypothesis which would accommodate and inter-relate a variety of formerly diverse elements potentially pertinent to speleogenesis in its broadest sense. To achieve this end it was necessary to re-examine and re-assess much that had been written, and either accepted or disregarded, in the past. Lateral thought indicated that potential alternative interpretations could be placed upon some earlier observations or results, on a purely theoretical level, if some of the commonly accepted parameters of speleogenesis were adjusted or ignored. Much that existed within the then current wisdom was found to remain more or less applicable to processes and timescales of cave *development* but inapplicable to the timescales and processes of the earliest cave growth, which were referred to by Lowe as cave *inception.* A quantum jump, both philosophical and actual, is required to progress from the theoretical existence of a carbonate terrain which carries no underground drainage to a carbonate terrain within which groundwater movement and, by definition, cave growth has been conceived. Allowance for this quantum jump necessitates the expansion of the definition of a cave to include *all* voids within a rock mass which are capable of transmitting groundfluids. Whilst some workers in the past have been able to demonstrate the nature of the quantum jump as a syngenetic process (or set of processes) closely related to diagenesis, there has been only limited success in describing possible mechanisms to accomplish the transition within mature rock sequences.

By "moving the goalposts" that the then current wisdom had erected, it was possible to follow a thought process which linked a number of pre-existing but generally isolated concepts and to demonstrate, among other things, that on a theoretical level:

a. Cave inception in mature carbonate sequences is probably related not to the purer parts of the succession, but to beds which are either themselves impure or are adjacent to impure/non-carbonate beds. Such dissolutional foci are termed *inception horizons.*

b. The driving mechanism for speleo-inception in mature carbonate sequences has little in common with the hydraulic situation pertaining in explorable cave systems.

c. The dissolutional mechanisms responsible for speleo-inception in mature carbonate sequences are unlike those normally considered to dominate cave development.

d. The earliest solvent motion and dissolution within mature carbonate sequences is probably infinitesimally slow and yet effective over enormous distances, on a regional or basinal scale, thus bearing no relation to modern landscapes and current hydraulic gradients.

e. Dissolutional processes active today at the freshwater/saltwater interface have probably been similarly active and affecting syn-depositional carbonates in similar environments throughout (at least) Phanerozoic time.

f. During timescales of spelegenesis which span millions, tens of millions or even hundreds of millions of years, processes similar to those deduced to be active today have been active within syn-depositional and mature carbonate sequences. The relict features of any such episodes can be preserved to different degrees and become involved in subsequent phases of speleogenesis.

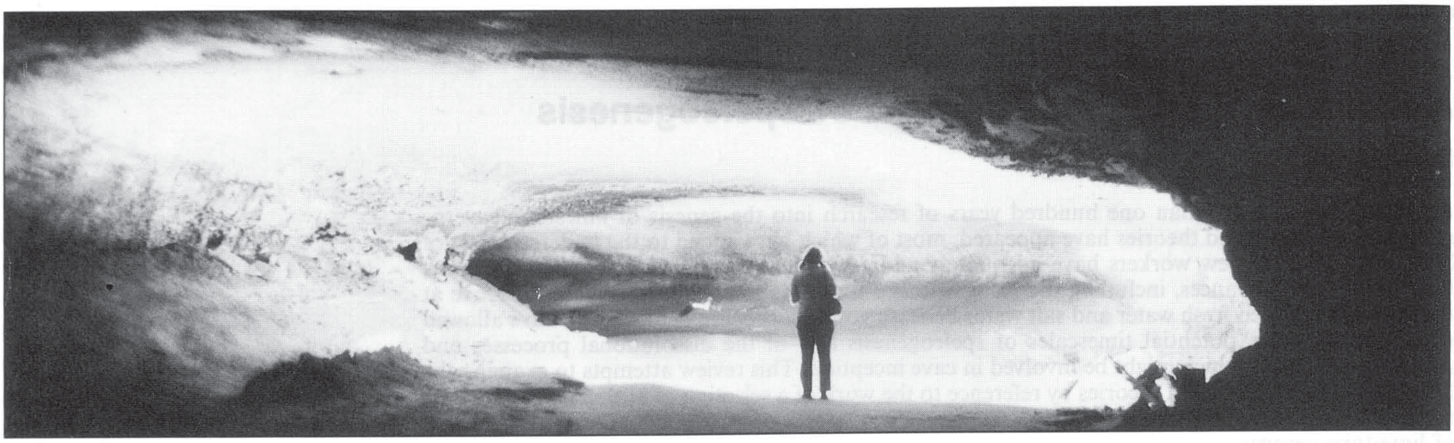
g. The inception skeletons upon which most currently explorable caves developed *were* probably in existence prior to whatever earth movements led to their modern geological setting and *certainly* in existence prior to the superimposed development of modern landscapes.

These and other theoretical conclusions are developed and discussed in detail within the parent thesis and various elements will form themes of papers to be compiled when time permits. In the interim the following review of many of the earlier contributions on the subject of carbonate speleogenesis is presented as a document of potential interest to other workers in this field and is particularly involved in examining the relevance of earlier work to the understanding of cave inception. The review has been slightly modified from its original form (Lowe, 1992, chapter 2), yet inevitably there remain many points in the text where value judgements are presented in the context of the then gestating thesis rather than in a purely objective and unbiassed way. As acknowledged within the thesis, the work could not have been accomplished without the benefit of the ideas of all the previous authors discussed in the review. To belittle the efforts of these workers, some of whom are still active and productive in their research, was not the intention. It is *emphasised* that whereas much of the thesis depends upon my own observations in many areas, most of my lateral interpretations of other authors' findings presented below have no background of personal experience. As the fruits of lateral thought, the alternatives offered are purely hypothetical. These hypothetical re-interpretations may themselves be dismissed by others who do possess the first-hand knowledge that the present author lacks.

Caves have been a part of human life since the earliest times, having been used as convenient, if not necessarily comfortable, places to live, to hide, to store things or to bury the dead. The date when the first thought was given to the questions of why and how caves exist, is unfathomable. Reference to caves, as accepted parts of the environment is common in even the oldest of extant literature, but not until much later is there any recorded philosophy concerning cave formation. Most of the early speculation was fanciful, mystical or mythological and not until the 19th century did anything approaching a scientific evaluation take place. Much of the available early literature has been examined and chronicled by Shaw (1979) and will not be reconsidered here.

THE EARLY SCIENTIFIC APPROACH (pre-1930)

A distinct phase of scientific thought is represented by the views of speleogenesis which emerged in the late 19th and early 20th centuries, prior to the cave work of W M Davis. Around the turn of the century, workers eminent in the field of geomorphology began to show interest in surface karst and inevitably there was an increase in speculation about the nature of underground karst and the movement of underground water. Jovan Cvijić (1893) published a monograph in which he discussed a sequence of surface landform development applicable to the Dinaric karst region. Some years later (Cvijić, 1918) he produced another contribution, in which he related the development of karst landforms to sub-surface hydrology. Hydrographic zones, which by this or other names have played a major role in karst controversy throughout the 20th century, were pointed out and their significance underlined for the first time in this important paper, which was translated into an English digest form by Sanders (1921). Cvijić's zones were termed, in descending order, the dry zone, transitional zone and saturated zone. This zonation was later to be emphasised by Swinnerton (1932). According to Cvijić, most groundwater movement took place at or immediately below the water-table and, hence, the transitional zone would be the major focus of cave formation.



One of the trunk phreatic tubes in the nearly horizontal limestones of the Mammoth Cave System, Kentucky, which so influenced early American thinking on speleogenesis. (Photo: Tony Waltham).

Between the appearance of Cvijić's 1893 and 1918 publications there were several significant events in the context of speleogenesis. One such was the meeting, in the field, of Cvijić's professor, Albrecht Penck, and the American geomorphologist, William Morris Davis. These two, with a group of Penck's students, visited and studied the karst areas of central Yugoslavia, and such was their almost patriarchal influence in the wide field of geomorphology that the conclusions they reached had repercussions that are still being felt today. From the wide range of individuals' contributions, two conflicting models began to appear. Alfred Grund (1903) was of the view that the karst aquifer could be broadly divided into two zones above and below a water-table. The lower zone was saturated by essentially stagnant water whereas in the higher zone there was water circulation within open caves. The water-table which separated the two zones was considered to be of regional extent.

Katzer (1909) considered that there was no zone of saturation and hence no water-table, but that all water within the karst aquifer flowed along open cave passages between surface sinks and risings at a lower topographical level. Further, he considered that the passages were independent of each other since the intervening rock mass was essentially impermeable. Arguably, the latter consideration has stood the test of time to a greater extent than the former, though neither can be said to be totally incorrect. Edouard Martel (1921) was perhaps the most influential supporter of Katzer's viewpoint, on the basis of his own underground experience, which was second to none at that time. His idea of surface water sinking at high level and carving routes through the easiest combination of fissures and bedding planes to valley bottom risings is not without its attractions. However, the stated importance of a process to ongoing cave development does not necessarily imply or prove a relevance to cave inception and the earliest stages of cave formation.

Following the translation of Cvijić's (1918) paper by Sanders (1921) and the appearance of Martel's *"Nouveau traité des eaux souterraines"*, also in 1921, little of notable relevance to general speleogenesis was published until 1930. In the interim a number of short and parochial papers, including early work by Malott and Swinnerton, appeared. Some of those originating in North America were consulted and quoted by the authors of several major publications during the next phase of cave research. If similar studies were under way in eastern Europe, they were overlooked or ignored by western workers and it is not unreasonable to suppose that much of the ensuing controversy regarding cave formation was fuelled by narrow, regionally biassed views. Theories based upon observations within the relatively undeformed carbonate sequences of the eastern USA would, inevitably, conflict with theories deduced from observation of the relatively highly tectonised sequences in Alpine Europe. The only major work to appear in Europe during the next phase, that of Otto Lehmann, expressed ideas quite unlike those in contemporary American publications.

THE 'GOLDEN AGE' (1930-1942)

The term Golden Age is coined here to signify the period between 1930 and 1942, during which major publications by Davis, Swinnerton and Bretz, as well as the often understated contributions of Lehmann, Gardner and Malott, changed the face of speleogenetic research. The same time span was referred to as the *"classic period"* by Watson and White (1985). In the light of more modern research and the inevitable criticism of these major works which has emerged in recent years (eg Watson and White, 1985), it might justifiably be said that the Golden Age has become tarnished. However, whatever their shortfalls, and many are recognisable with hindsight, it is undeniable that each of the researchers discussed below made a major contribution to the knowledge and understanding of speleogenesis. In passing, it is noted that the criticism of Davis's work by Watson and White (1985) appears to have been particularly harsh. As pointed out by Swinnerton (1932, p.664), who emerged intact from Watson and White's scrutiny, *"Davis has examined the tacitly accepted ideas of cave origin and has propounded a counter working hypothesis, hoping that observational investigation might thereby be stimulated"*. Then, as now, this described the very essence of scientific method. If Davis, by his lack of personal field observation, was forced to use his powers of deduction as opposed to induction, his legacy should be seen as no less valuable than that of Dalton, who had never seen an atom, or Mendel, who had never seen a gene. These and many others, have left only partially correct, interim, statements, but they were respected in their time as they should be today.

W M Davis

"... what is now most needed in the investigation of cavern origins is not so much a more detailed deduction of expectable features as a closer and more critical observation of actual features."
(Davis, 1930, p.566)

William Morris Davis was eighty years old when his paper *"The origin of limestone caverns"* was published in 1930. At the height of his academic career he was professor of physical geography at Harvard University and it is for his studies of geomorphology that he is best remembered. The classic essay on caves, with its broad applications to karst geomorphology and hydrology in general, appeared towards the end of his distinguished career and only three years before his death. By his own admission it lacked the detailed personal observation which was attained by his younger, more active, successors. However, it is interesting to note that whereas Davis suffered severe criticism from Watson and White, (1985) for this lack of first hand knowledge, later workers during the Golden Age were also described as lacking such experience: Bretz (1942, p.676), at the close of the Golden Age, stated, of the more modern studies, that *"none* [was] *based on a really extensive acquaintance with caves"*.

Davis was probably the first worker to emphasise the potentially vast timescales involved in cavern formation. His cavern formational ideas followed the broad theme of Grund and he believed that cave development would begin below the water-table as soon as fresh water was able to replace the primitive, formational, salt water contained within a limestone sequence. He visualised slow dissolutional development of groundwater routes, possibly to very great depths, below the water-table, and made a distinction between this process and that of cave formation at shallower depths. He saw the deep dissolutional processes as possibly extending across long periods of geological time, tens or hundreds of millions of years, with the shallower, though still sub-water-table, first stages of cave formation occurring in somewhat shorter time spans. This deep phreatic and subsequent shallow phreatic dissolution process took place during his *"first cycle of cavern development"*.

Though these broad concepts are in very close agreement with views expressed by Lowe (1992), Davis's hypothesis lacked a detailed and convincing model for the actual guidance of his deep drainage routes. Additionally his argument was limited not so much by his ignorance of the chemistry of calcite dissolution (as suggested by Watson and White, 1985) but by his reliance upon this reaction as the sole mechanism of dissolution. He was unaware of the possible role of saltwater *(sensu lato)* in dissolutional processes and was probably equally unaware of the nature and extent of buried palaeokarsts.

In Davis's view, true cave passages would be formed, below the water-table, by modification of once deeper drainage routes when these were raised closer to the surface, and hence closer to the water-table. This apparent rise could be in response to uplift, as a result of the removal of overlying strata, or when the drainage regime was altered by relative downcutting of surface valleys. These caves, conceived at depth and developed below the water-table, were in turn raised above the water-table and drained, with a new erosive cycle being initiated in a lower drainage route. In this way it was possible to explain the presence of caves which, as seen by Davis, must have developed below a water-table, but which were standing at levels clearly above the contemporary water-table. Once above the water-table a second cave development cycle began. During this phase the pre-existing cave passages might capture and carry surface drainage water, might eventually become partially or totally blocked by carbonate deposits, rock fall or fill, and would ultimately be removed by erosion.

Outside the main stream of his thesis, just as with the other workers writing at this time, there exists a plethora of linked or partially linked concepts. It is beyond the scope of this review to examine the full spectrum of these, even to the degree of discussing their relevance to more modern views of speleogenesis. In passing, however, Davis's recognition of the possible links between mineralization and caves (1930, p.619 *et seq*) was far sighted and it is not unreasonable to suggest that much written speculation on the subject of mineralization would have been still-born had this aspect of his work been more widely appreciated.

Of all the cave origin models to emerge during the Golden Age, that of Davis has suffered more adverse criticism, even ridicule, than any other. The thoughts of Harlen Bretz though generally supportive of Davis's thesis, are more widely accepted, this judgement being based by some upon the assessment that Bretz had consciously, or perhaps unconsciously, managed to make the deep dissolutional concept seem less extreme. Davis knowingly left himself and his model open to attack by later workers, particularly those with more practical underground experience. Many of these later views have been valid but, considering that Davis never claimed a formational model which was applicable to all caves, a creditable part of the original hypothesis remains potentially acceptable.

A C Swinnerton

"The whole problem of cave origin awaits systematic study and presentation."
(Swinnerton, 1929, p.82.)

Following close upon Davis's (1930) treatise, came another major and no less impressive consideration of cavern formation by Allyn C Swinnerton (1932). Throughout this publication Swinnerton's respect for Davis and his theories is apparent, though the main trend of his arguments is away from the deep phreatic dissolution favoured by Davis and towards cave formation close to and above the water-table. In many ways Swinnerton's work follows on from the ideas of Cvijić and the threefold division into hydrographic zones is re-emphasised. Though admitting to the possibility of dissolutional processes being active in the deeper parts of the phreatic zone, his opinion was that this was of relatively minor importance and that the major locus of dissolutional cave formation was in a fluctuating shallow phreatic or water-table zone. Beyond this he recognised the importance of ongoing cave development in the vadose zone, following drainage of previously flooded passages, by similar mechanisms to those outlined by Katzer, Martel and other earlier workers.

Swinnerton's paper includes several significant elements in addition to his major hypothesis of shallow phreatic cave origin. He was particularly interested in the observation that many cave passages are sub-horizontal and that they commonly occur at several levels, one above another. These *"tiered caves"* were reported to be present in folded as well as unfolded, gently dipping, sequences - that is, the sub-horizontal trend of the passages appeared to truncate even extreme geological structure. To explain this he postulated that the passages formed within a relatively thin vertical range demarcated by the high and low positions of a fluctuating water-table between wet and dry seasons. What seems to be lacking in this view is a means whereby the water is able to fluctuate within an essentially impermeable rock mass and yet concentrate its dissolution potential into a restricted volume which will ultimately become a cave passage. No illustration nor detailed example of such a shallow phreatic conduit which ignores the constraints of geological structure is provided, so it is not possible to comment on the interpretation presented by Swinnerton, nor even to be certain what he considered to comprise an extreme geological structure (cf the discussion of W E Davies' terminology, below). Possible mechanisms of *"tiered cave"* formation must also be considered. Swinnerton's view of sequential piracy of drainage to progressively lower levels, in response to surface erosion cycles, is superficially acceptable and broadly in line with most observed relationships. The actual means by which the piracy is achieved are vague.

It is pointed out by Watson and White (1985) and White (1988) that many of the earlier workers on cave origins suffered from a basic ignorance of hydrodynamics and solution chemistry - Swinnerton being a notable exception to this rule. Modern views of these two complex subjects are, however, far removed from those accepted by Swinnerton, and are yet evolving. Swinnerton's views were based upon an untested (in this context), incomplete understanding of both. It is thus a moot point whether his more knowledgeable consideration of these two fields represented a step forward which was any more significant than the advances of earlier workers who had grasped and applied new, but only partially developed, concepts in other important disciplines.

Possibly overlooked by some contemporary, and later, workers, was Swinnerton's undiscussed quotation from Haug (1921, p.362), which might have gained greater interest had it been translated: *"Les rivières souterraines présentent alors des biefs étagés, avec des lacs et des cascades et, quelquefois, les galeries formant siphon, les eaux remontent à un niveau plus élevé. Leur cours peut même descendre ainsi audessous du niveau de base de la rivière."* [Subterranean streams have tiered courses, with lakes and waterfalls and, sometimes, siphon passages, the water rising to a higher level. Their course may even drop below the base level of the stream.]

O Lehmann

Sweeting (1972, p.129) describes Otto Lehmann's (1932) *"Die Hydrographie des Karstes"* as a major work. It was based upon a knowledge of the Classical and Dinaric karst areas, the Alps and the French Causses, thus being the only European contribution to the subject of speleogenesis to gain serious recognition during the Golden Age. In contrast to Davis's work, which was essentially deductive, that of Lehmann was inductive and based upon his many field observations in the areas listed above. This being said, Lehmann's work was, if citations are any indication, largely ignored by North American and British researchers. The reasons for this are probably two-fold. Lehmann's text book was written in German, and it appears that no English translation, even of a digest form, nor any English language paper covering his main ideas, appeared. Access to and understanding of his ideas would have been difficult for many British and North American workers, even if they had acknowledged the desirability of looking beyond the apparently definitive views which were concurrently emerging in the west.

Secondly, Lehmann had joined the ongoing eastern European controversy regarding hydrographic zones. His work was strongly supportive of views expressed by Katzer (1909) and Martel (1921), to the effect that cave rivers are the major active agent in cavern formation and that a water-table and associated hydrographic zones either do not exist or are unimportant in this context. Possibly Lehmann's concept of a water-table was not the same as that in contemporary use by the workers he contradicted and, if so, some of his ideas can be viewed sympathetically in the light of later understanding. He repudiated Cvijić's (1918) views, which had upheld Grund's (1903) groundwater theory and which were being mirrored, at least in part, by the theories of Davis (1930) and the findings of Swinnerton (1932). In the west the models devised by Davis and Swinnerton formed two opposing faces of speleogenetic controversy and the European dispute appeared not to impinge upon the problem. In fact, so far as western workers were concerned, particularly those in North America, Lehmann's ideas must have been essentially out of favour as soon as they were published and the controversy that they perpetuated was but a small part of the broader view of speleogenesis which was emerging.

In eastern Europe, however, Lehmann seems to have exercised a

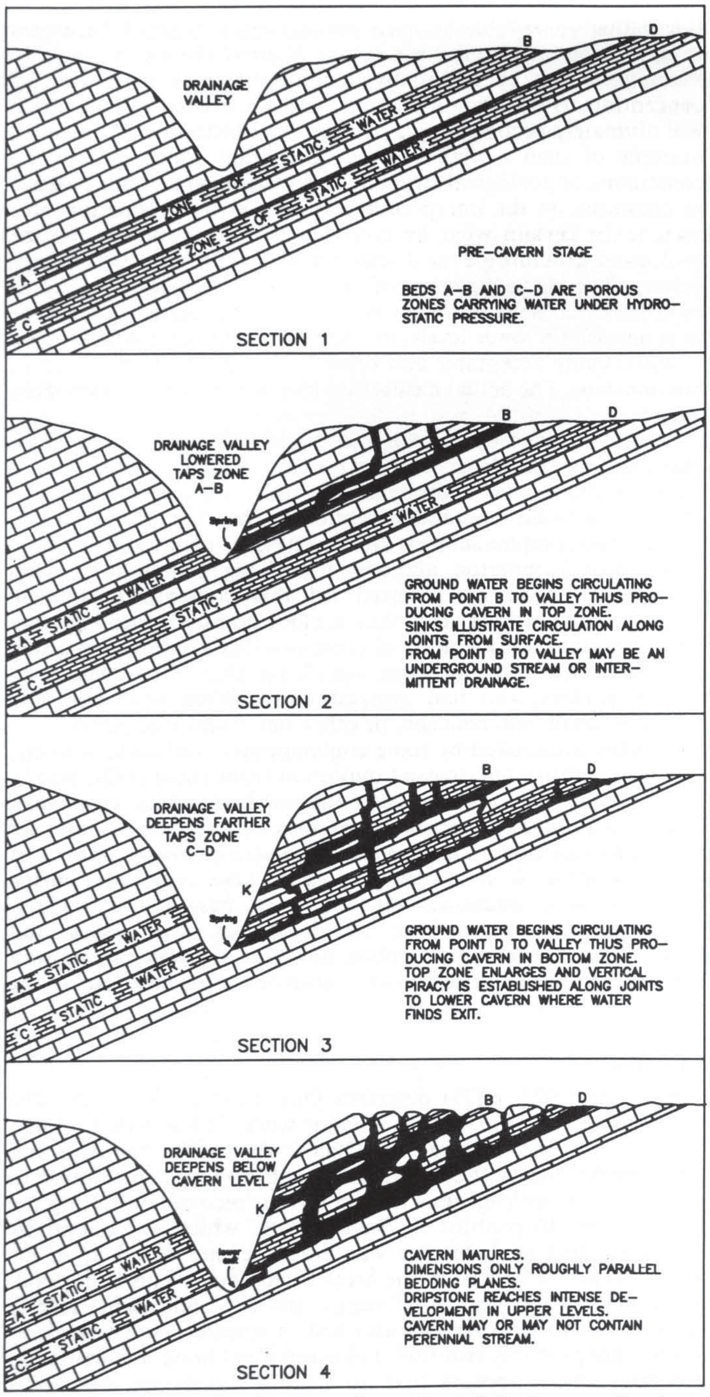


Figure 1. Sketch sections illustrating the origin and development of a large limestone (redrawn from Gardner, 1935, Figure 1).

patriarchal, almost mesmeric, influence upon karst geomorphologists and hydrologists, not unlike that of W M Davis in the western world in earlier years. According to Bögli (1980, p.vi), *".... the hypothesis of O.Lehmann was untenable. Thus the spell was broken which had blocked progress for a long time, but which, however, had also stirred up contradiction and thus encouraged new research."* The breaking of this spell occurred in 1974, some forty two years after its publication, but earlier work had noted supposed weaknesses within Lehmann's hypothesis (eg Zötl, 1961).

Lehmann considered that cave formation could not commence within a rock mass unless *"initial cavities"* (Urhohlräume in German) with a minimum width of 2mm were present. He also stated that for water motion to be initiated a minimal primary perviousness (hydrographische Wegsamkeit in German) was necessary, since both ingress and egress of water to and from the rock mass are required. However, his main hypothesis, which was considered valid by most European workers until finally discounted by Zötl (1974), was that, *"karst cavities are independent and have no relation to one another."* (O Lehmann, 1932, p.15). On the evidence of water tracing experiments in the Dachstein of Austria, Zötl (1961, p.129) postulated *"... a large connected body of karst water"* in the test area. Zötl also pointed out that the number and width of joints decreases downwards and that, *"This observation reveals the basic fact that within massive karst there is in general a dense network in the upper regions with smaller open joints reaching much deeper, even below the base level."* (Zötl, 1961, p.130). These and similar observations, which were reiterated some years later (Zötl, 1974), were those which finally broke the hold of Lehmann's dogma upon eastern European karst studies. However, Bögli (1980, p.iii) points out that he (Bögli) had been aware for some time prior to Zötl's work, that the physical principles upon which Lehmann's hypothesis was founded, and the actual hypothesis, were not borne out by underground observations. In the light of concepts discussed by Lowe (1992) it now appears that the ideas of Lehmann might have been misinterpreted by some later workers and that much of his basic philosophy is at least partially valid in terms of speleo-inception.

Despite the fall from grace of Lehmann's more dogmatic statements regarding the general processes of underground karst formation, his work remains a major milestone of the Golden Age. Among his observations he noted the so-called "karst-hydrological contrast" and commented upon the mechanism whereby swallow holes could become springs, the dual function features being termed "estavelles". The term "karst-hydrological contrast" was used with reference to the apparent discrepancy between the great number of potential seepage points and sinks upon the karst surface into which precipitation can be absorbed, as compared to the relatively few springs from which the karst water re-emerges. This is possibly not a totally valid viewpoint, insofar as it is impossible to prove the presence of a comparable set of seepage output points in parallel with the assumed seepage input points. What it does point out is that, as an underground drainage system matures, many minor conduits are effectively abandoned as an integrated drainage system, targeted upon a major spring, is developed. However, the many minor conduits continue to exist either actively or passively (if their drainage has been pirated) and during abnormally wet conditions many more resurgent seepages will be apparent than during 'average' flow. This drainage fluctuation and the associated return to ordinarily beheaded conduits account for the commonly observed features of high-level flood resurgences and bedding plane or joint seeps. The counter proposal to Lehmann's *"karst-hydrological contrast"* would be that during the early stages of karstification in sub-horizontal limestones, a large number of joint/fault guided input seeps would be balanced by a potentially large number of bedding guided output seeps. In a hypothetical sub-vertical sequence the converse relationship would hold; combinations between the two extremes would occur in gently to steeply dipping sequences.

J H Gardner

Of all the researchers in the 'Golden Age' it was James H Gardner (1935) who came closest to the concepts incorporated in the new appraisal of carbonate speleogenesis which forms the core of Lowe's (1992) thesis. He is generally remembered for ideas regarding cave formation above a water-table, but several seemingly insignificant 'asides' are now seen to be prophetic and important to the overview of speleogenesis. His main concern was in seeking an explanation for the origin of *"large caverns"*, and to this end he became somewhat pre-occupied. The significance of his suggested formational mechanisms to speleogenesis in general is smothered by a great weight of arguments to prove the applicability of his hypothesis to the genesis of Mammoth Cave and Carlsbad Caverns, systems which have little in common beyond large size.

His view was one of essentially down-dip passage formation commencing in *"carrier beds"* or *"aquifers"*, possibly of no great thickness, which form part or parts of a thick limestone or dolomite sequence. These beds were considered to be especially liable to dissolutional cave development but, as pointed out by Warwick (1953, p.48), he did not state the nature of the properties which determined this liability. However, two sentences in particular (Gardner, 1935, p.1260) are important in the context of cave inception - *"In most cases the aquifer is porous limestone or dolomite. It may, however, consist of open spaces along bedding planes of limestone; or a sandstone bed intercalated in limestone strata might serve as the original circulating medium until solution acts on the adjacent calcareous beds."* The second sentence is particularly provoking.

In addition to talking about these carrier beds or aquifers, he refers to stratigraphical control of caverns and quotes the example of The Dry Branch, in Kentucky, where *"the stream and its cavern follow down the bedding planes and remain confined to a zone of limestone strata which probably does not exceed fifty feet in thickness"*. He considered that this cave was typical of the initial stage of cavern growth.

The triggering mechanism of cave formation according to

Gardner's hypothesis requires downcutting by a surface valley to cut through the dipping limestone sequence. Steep dips are said to be particularly suitable, though he points out that lower gradients are potentially susceptible and the dip need only to be *"enough to start the circulation of water by gravity"*. Before valley downcutting the up- and down-dip sides of the carrier bed (relative to the valley) were seen as being saturated by static water. Breaching by the valley would lead to drainage of the up-dip saturated zone and subsequent vadose invasion and cave development within the carrier bed. Continued downcutting might eventually lead to the intersection of a lower carrier bed, at which stage, according to the theory, the upper level would be abandoned as its drainage was pirated by caves formed in the lower bed. All of this is illustrated by a series of cross-sections which comprise his figure 1 (redrawn and reproduced as figure 1). The potential importance and regional significance of the down-dip continuations of the intersected carrier beds were not pursued by Gardner.

In modern terminology Gardner described vadose cave formation within stratigraphically determined zones, wherein inception and gestation were due to the dissolutional effects of essentially static phreatic water. Of great significance is his continued reference to *"sulphur water"*, which is recorded as giving way to salt water at depth. Gardner admits to limited dissolution by groundwater below the water-table, but not to the extent required by Davis's two-cycle theory. Having accepted this limited dissolutional activity he goes on to add his support to Swinnerton's one-cycle theory, seeing most of the cave formational process as being by active streams above the water-table.

As pointed out above, Gardner's view of cave formation is not far removed from the conclusions of Lowe's (1992) study, but one major link was missed. An obvious thought process links the carrier beds or aquifers postulated by Gardner (1935) with the inception horizons described in Lowe's thesis, though the two concepts differ in detail. Possible chemical similarity is suggested by a repeated reference to *"sulphur water"* and it is unfortunate that this term, derived from Gardner's borehole drilling experience, is not amplified (though Egemeier, 1981, believes that it refers to hydrogen sulphide in solution). Also Gardner's description of cavern galleries occurring at several levels - five major levels in Mammoth Cave - is explained in terms of *"cross formational piracy"*, initiated by the surface downcutting described above. This is the major missing link of the hypothesis insofar as if the strata between the carrier beds were impermeable, as required (Gardner, 1935, p.1258), there is no reason why drainage in a high level system should be pirated by a newly activated drainage system in a lower, discrete stratigraphical horizon. Here again, Gardner was tantalisingly close to providing the link, but did not follow the argument to its limit. In describing the sequence of events that occur during the formation of a multi-level system he states (Gardner, 1935, p.1263), *"Open joints are first drained of their water by circulation along the bedding below. In time, the water of the overlying cavern, whether a perennial stream or intermittent, is captured by a lower cavern stream and so on, resulting in complex development."* This somewhat glib statement begs the entire question of *"cross formational piracy"*. Recognition of sub-water-table dissolution along joints or faults between the separate carrier beds, before drainage, would have simply provided the necessary link.

Over all, Gardner's work could have provided a positive step in the understanding of speleogenesis. Concepts such as initiation [*sic*], the pre-cavern stage, multi-level caves, stratigraphical control, ore deposition, sulphur water and salt water all deserved recognition and follow-up work. On the other hand some of the ideas were less inspired. The laborious argument of similarity between Carlsbad Caverns and Mammoth Cave, pronouncements upon time-scales of speleogenesis, the consideration of the effects of earthquakes and the statement of support for Swinnerton's vadose formational model, to the extent (p.1273) that *"Large caverns .... originate and develop in thick terranes of limestone above the water-table."*, might, over the years, have obscured the major importance of the work as a whole.

C A Malott

Clyde A Malott was a professor at Indiana State University. The essence of his *"invasion theory of cavern development"* is encapsulated in abstract form (Malott, 1937) and the formal supportive publication seems never to have appeared. Various corroborative evidence, particularly with reference to his work in southern Indiana, is reported (White, 1988, p.268) to be interwoven within several descriptive papers published between 1929 and 1952. The earlier among these pre-date the work of Gardner and at least one predates the work of Swinnerton and Davis. Davis did acknowledge an indebtedness to Malott for help during the assembly of his (1930) paper, but Malott's not inconsiderable contribution, or at least its earlier manifestations, appears to have been largely overlooked by both Davis and Swinnerton and rates only passing mention from Gardner. In his 1937 abstract, however, Malott refers to the work of Davis, Swinnerton and Gardner. Whereas it is considered above that Gardner overstated the importance of vadose streams in cave formation, Malott was of the opinion that Gardner had not given underground streams a paramount part in cavern development. As with Gardner, Malott had a noticeable preoccupation with *"large caverns"*, but unlike Gardner he provided the missing link to the progression from speleo-inception to cave development. If a formal paper had followed the abstract, it can only be assumed that the ideas would have been further developed.

This being said, the statement in Malott's second paragraph (below) is remarkably close to the views presented by Lowe (1992), though having divined the earliest stages of speleogenesis it is arguable that he made the wrong jump to the next part of the process. Warwick (1953, p.46) reproduced this vital extract from Malott's 1937 abstract: *"Primitive, illy* [*sic;* 'poorly' in Warwick's review] *integrated, and somewhat selective three-dimensional passage ways are developed in limestone regions **below** the water table"*. There is no doubt that Warwick appreciated the significance of this rather bald statement and emphasised the word 'below' by the use of italics.

According to Malott's theory, lowering of the water-table leads to a concentration of groundwater within *"selected parts of the earlier prepared primitive systems"*. Next, surface streams are captured and enter the existing underground routes and these diverted waters will ultimately *"carve out and integrate the large caverns at or near the watertable"*. On the strength of his studies in southern Indiana, Malott believed that this capture of surface drainage was the major factor in cave development.

Malott pointed out, in the same abstract, that when such enlarged channels lose the stream that formed them, they tend to become infilled by fallen rock or calcite deposits, which have the effect of obscuring the evidence of stream carved floors. True as this is, it has no real relevance to the argument of speleogenesis.

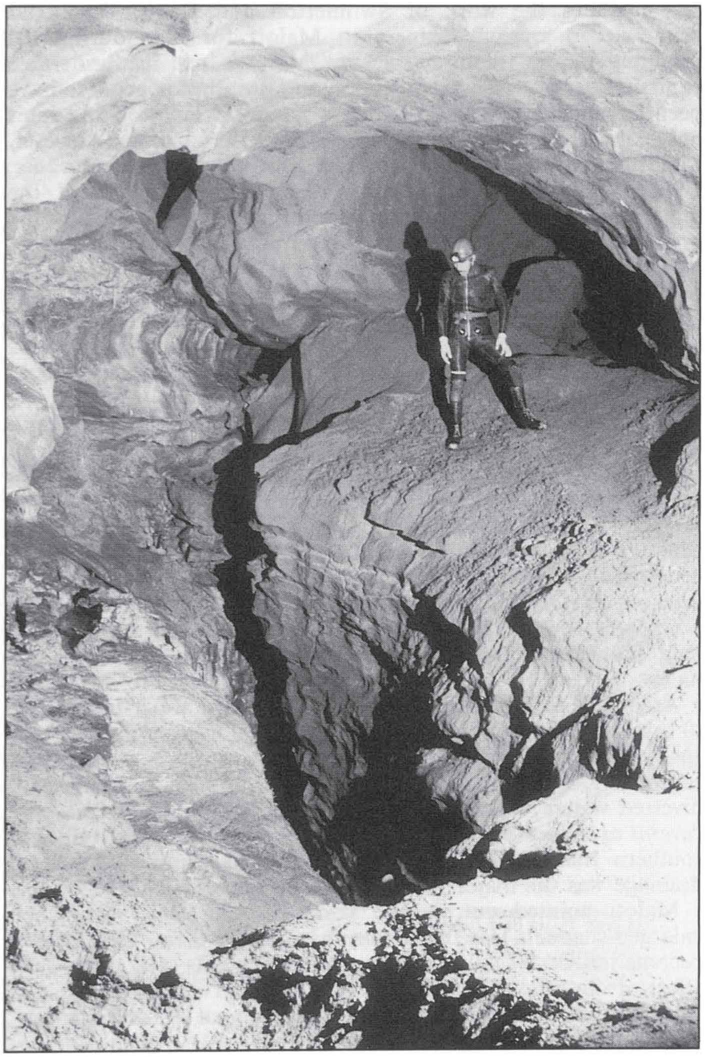
J H Bretz

"Even if Davis' thesis should have been established to the reader's satisfaction, much more fieldwork is needed on the subject of limestone caverns."
(Bretz, 1942, p.809.)

J Harlen Bretz, a geologist from Chicago, is commonly better remembered for his support of Davis's theories than for his own valuable contribution to the speleogenesis debate. His detailed views, published in 1942, were based on study of more than a hundred caves formed in Cambrian to Permian carbonate sequences and in varied structural environments within North America.

Bretz had the advantages, if such they were, of being able to draw upon the entire wealth of published thought from the Golden Age, of conferring with Swinnerton in the field, and of reading the first contributions of the Scientific Period before publishing his own treatise, which marked the end of the Golden Age. The title of the mammoth work by Bretz, *"Vadose and Phreatic Features of Limestone Caverns"*, is an indication that views of speleogenesis were becoming less restricted. Bretz is described by Watson and White (1985, p.114) as *"a brilliant thinker"* and the clarity of the arguments that he presented leaves no doubt that this view was correct. However, one is tempted to speculate on how Watson and White reconciled their assessment of Bretz with his almost unqualified support for the Davisian theory of speleogenesis, which they describe (1985, p.113) as *"...generally wrong"!*

In many ways Bretz produced a theory which embodied much of the original two-cycle concept of Davis (1930), but included the sort of evidence, gleaned from personal observations, that the earlier work had lacked. The major modification of the two-cycle theory was the recognition of a third aspect of speleogenesis, which was considered to post-date the first stage of deep phreatic dissolution and pre-date the stage of drainage and vadose modification. During this second, intermediate, phase the cave passages, still below the water-table, were filled with stagnant water. Deposits of "unctuous" red clay and silt, derived in part as an insoluble residue of limestone dissolution, slowly settled through the stagnant water, ultimately filling (or almost filling) the



The classic keyhole profile, here in White Scar Cave, Yorkshire, which lay at the core of Bretz's recognition of vadose and phreatic cave morphology (Photo: Tony Waltham).

flooded passage. After relative uplift and drainage of the clay/silt-filled passages, consolidation and desiccation would inevitably lead to a lowering of the upper surface of the fill such that invading surface streams would be able to exploit the drainage route. During the third phase (which is the same as Davis's second cycle), the active streams would remove all or part of the fill, leaving remnant fill-forms, which Bretz recognised as visual proof of his concepts.

Among the other ideas discussed by Bretz was his recognition of a number of features which he considered diagnostic of phreatic origin or vadose modification. The former category includes such structures as bedding plane anastomoses, tubes and half-tubes, passage networks or spongeworks and solutional pockets. The vadose features recognised by Bretz were fewer in number and were generally to be found superimposed upon typically phreatic features. They include rock-mill or pothole type holes in passage floors, gravel deposits and erosional features preserved in cave fill. With the possible exception of gravel deposits, some of which can be explained as having been laid down under phreatic conditions, the Bretz criteria may be considered to remain valid. However, more recent understanding of dissolutional and hydro-dynamic processes might provide somewhat different formational mechanisms to those described by Bretz.

Bretz is less well remembered outside North America for his later studies during the Scientific Period. During this time he explored and studied many more caves and published a number of papers (eg Bretz, 1949, 1956 and 1960). It would be out of place to discuss his later conclusions within this section, except to say that he remained convinced that most caves were initiated and developed in the phreatic zone, where they were subsequently wholly or partially filled by red clay or silt. Removal of fill and passage modification happened in a later cycle, when the passages were (relatively) raised above the water-table. Thus his views, first published in 1942, remained essentially unchanged.

In conclusion, a comment by Watson and White (1985, p.115) which was repeated by White (1988, p.269) merits some consideration - *"Although Bretz was a firm supporter of the Davis hypothesis, one gets the distinct impression that Bretz's deep phreatic was nowhere near as deep as Davis' deep phreatic".* This is in itself a strangely subjective comment from authors who show such a strong bias towards inductive rather than deductive science. However, even putting this aside, the observation seems merely to drag on a controversy of little real meaning. Whatever the viewpoint, depth is relative and limited, in this context, by the bounding surfaces of a particular geological environment. In a hypothetical sub-horizontal carbonate sequence 100m thick, with its base bounded by impermeable strata, dissolutional activity at the basal contact might be considered deep. In a sequence 1000m thick, or in the same 100m sequence following a steep dip to greater depths, cavern development 100m down might be considered shallow. The point which should be discussed is whether inception and development can occur through the full thickness of a carbonate aquifer or whether, in dipping strata, inception and development can proceed beneath great thicknesses of overburden and at considerable distances from outcrop.

"Karst landscapes are the result of the dissolution of bedrock, usually limestone or dolomite, by runoff, by infiltration and by deep *circulation* of groundwaters."
(White, 1984, p.227.) [present author's italics!]

THE 'SCIENTIFIC PERIOD' (1941-1978)

Watson and White (1985) refer to the period between 1942 and 1957 as *"The Hiatus"*, on the basis of a dearth of cave studies in North America during this interval. The subsequent period, from 1957 to the time of their paper, was termed the *"Modern Period"*. The opinion expressed here is that any division of this type should be based upon the first appearance of a scientific contribution in the context of a modern analytical approach. On this basis the term 'Scientific Period' is coined to encompass the contributions, in the 'modern analytical context', including and subsequent to those of Moneymaker (1941) and Rhoades and Sinacori (1941), up until close to the end of the nineteen-seventies. By adopting this approach, the observations of Moneymaker and the views of Rhoades and Sinacori can be considered as the vanguard of the 'Scientific Period' whilst the masterwork of Bretz (1942), though appearing chronologically later, can be viewed as the parting review of all that was best of the empirical and deductive output of the 'Golden Age'. The term itself, Scientific Period, is indicative of the growth of science as a whole, but is particularly meant to reflect a phase during which experimental, analytical and numerical techniques were the godhead of speleogenetic endeavour. Although old fashioned field observation and a modicum of deduction continued to take place, the obvious was periodically or habitually overlooked in the face of massed data, representing 'scientific fact'.

It is emphasised that the dividing points used here are subjective and hinge partly upon sympathy for the lack of recognition of Moneymaker's work and an admiration of the objectivity of Rhoades and Sinacori. Recognition of a general (but not across the board) change in the perspective of speleogenetic research after the work of Ford and Ewers (1978) prompts the placing of the end of the Scientific Period at this date. It is not suggested that the scientific approach has ceased, rather that the minor revolution in the overview of the speleogenesis problem has necessitated (or deserved) a different and less pedantic philosophy. The nomenclature of Watson and White (1985) and its turning points are no less apt in a strictly chronological and parochially North American context.

It is less easy to write about the milestones or benchmarks of speleogenetic research during the Scientific Period and more recent times than those of the Golden Age. The number of contributors has increased almost exponentially, as has been the case in science generally. At the forefront of activity was a dramatic increase in the number of people who ventured underground or generally into the field in karst areas. Added to that was a growth in the number of those who 'dabbled' or specialised in science of all kinds and a concurrent growth of university departments around the world where geomorphology or karst studies became recognised as specialities. Cave studies diversified to include pure chemistry, mathematics, statistics, geology, hydrology, archaeology, zoology and myriad interconnected sub-disciplines. Eventually cave or karst studies *per se* gained a foothold in a small number of universities and polytechnics, becoming, at least at post-graduate level, a legitimate research option.

During this period many papers have dealt with major or minor aspects of cave studies or with parallel and applicable aspects of the contributing sciences. Among these a number of workers have produced significant milestones, beginning with the observations of Moneymaker and Rhoades and Sinacori, passing on to those of the Modern Era, which persists today. Below, a number of major benchmarks are considered as individual sections. Between and subsequent to these sections, other workers whose diligent and painstaking work has added to the modern view of speleogenesis are discussed. Scrutiny of the reference lists in such reviews as Sweeting (1972), White (1988) and Ford and Williams (1989) will confirm not only that there is a disparity between what is seen as significant by different authors, but also that, of necessity, the present work does not deal in detail with more than a small proportion of those contributions which are actually or potentially relevant.

B C Moneymaker

"The lower limit of cavernous rock was, with a few exceptions, controlled by the position of a thin bed of bentonite and shale near the base of the Cannon limestone."
(Moneymaker, 1941, p.80.)

The observations of Berlan C Moneymaker (1941) seem, at first sight, to be less of a herald of the Scientific Period than the theoretical considerations of Rhoades and Sinacori (1941), published some few months later. It seems that the latter might already have been submitted for publication by the time the former appeared. It is unlikely that such tangible evidence, or reference to it, could have been avoided or overlooked by Rhoades and Sinacori, yet no citation appears in their text. Mutual awareness is indicated by Moneymaker's acknowledgement of, and text reference to, Rhoades (though not in the context of the Rhoades and Sinacori paper), but Moneymaker was not acknowledged reciprocally. This was only the beginning of a long, still persisting, period of obscurity for Moneymaker, whose views have rarely merited more than a line or two in reviews of speleogenesis. Examined more closely, however, the rather informal comments which accompany the reported observations can be perceived as being nearly three decades ahead of their time.

Moneymaker's contribution was mainly one of observation rather than of revolutionary theory. During the course of his work for the Tennessee Valley Authority in the nineteen-thirties he had access to the records and/or cores of numerous small and large diameter boreholes which were drilled for purposes of exploration and grouting across a variety of potential dam sites. Additionally he was able to visit and inspect many open-cut excavations made for the same purposes. A common link between many of the boreholes and excavations was the proving of natural dissolutional cavities, often of large size, at various levels, both above and below the water-table, which was taken as approximating to local river level. Moneymaker was careful to point out that cavities proved by small diameter drilling were of limited value as indicators of cavity size since, effectively, they present only a limited, vertical dimension. The full height of the cavity might not be penetrated by the drill and the horizontal extent, be it wide bedding or narrow joint, would be unfathomable. He also presented some rather basic statistics relating number of large diameter holes and deep holes to the abundance of cavities encountered at various depths. Other statistics were presented, in table form, relating the maximum and average vertical dimensions and maximum and average depths below river level of cavities within specific rock formations at 30 localities. The values are of relevance only within individual beds or at individual locations, but the broad limits were of cavities up to 64 feet (c.20m) high, up to 222 feet (c.68m) below river level.

Among the other observations, which were mainly concerned with the factors mentioned above, are descriptions of a number of situations where the *"downward limit"* of cavern formation was marked by beds of bentonite and/or shale. Other non-carbonate or impure carbonate lithologies are also reported within the various cavernous rock formations listed. Moneymaker's view of the function of these beds, the shales and bentonites in particular, was that they acted as aquicludes and had *"protected"* the underlying rock from dissolution. Such a function, plausible as it seems at face value, may not be the only interpretation of the role of these non-carbonate and impure lithologies.

It is also productive to consider Moneymaker's explanation of why cavities were more numerous and of greater average size above than below the water-table. His view was that this indicated initial dissolution below the water-table followed by enlargement above the saturated zone after drainage following downcutting of surface streams. He points out that all dissolutional cavities above the existing water-table were related to earlier water-tables, having occupied positions below, coincident with and at various levels above, saturation level. However, he was unsure as to how much enlargement could be attributed to activity at these various levels. Though not formally stated, the view of speleogenesis he sums up is not far removed from the elegant rationalisation of earlier controversy supplied by Ford and Ewers (1978) at the end of the Scientific Period. Taken in tandem with the implications of his recording and observation of simultaneous dissolutional development at more than one level, in rock formations within the saturated zone, the overall, unformalized view is remarkably close to ideas which are re-emerging today (eg Worthington, 1991; Lowe, 1992).

R Rhoades and M N Sinacori

"The deep underground flow postulated by Davis is strictly compatible with hydraulic theory; there can be no doubt that it occurs very generally."
(Rhoades and Sinacori, 1941, p.786.)

Roger Rhoades and M N Sinacori (the christian name of the second author is one of the great mysteries of speleogenetic literature!) provided a more obvious turning point into the Scientific Period than Berlan Moneymaker. Their paper (Rhoades and Sinacori, 1941) was essentially theoretical and is one of the earliest examples of the mathematical consideration of underground water flow. By making use of a differential, Laplace equation derived from Darcy's law of flow and the equation of continuity, they were able to define two sets of curves. In certain specified conditions these two sets of curves, which intersect at right angles and are referred to as flow lines and equipotential lines, define a flow net. Thus, by drawing lines which represented the two sets of curves they were able to model graphically the solutions to a wide variety of theoretical **two-dimensional** seepage problems.

In broad terms, they were able to demonstrate that not only were the deep and shallow flow paths and dissolutional developments postulated by Davis and Swinnerton likely to occur, but also they could be present at the same locality. According to their model, there could be significant dissolutional development at depth in the early stages of the drainage of a limestone aquifer towards a specific discharge point. With time, the drains closest to the point of discharge would be preferentially enlarged and a master conduit would be cut back, headwards, from the rising, along a line which would follow the local water-table. This back cutting would in turn modify the flow lines in such a way that deep-seated water movement and dissolution would become proportionally less important and shallow water movement and dissolution would become proportionally more important.

Rhoades and Sinacori pointed out that assumptions had to be made to fit the behaviour of water in a limestone terrain into the idealised flow pattern suggested for flow in a more isotropic aquifer. They summed up on the basis that, if suitable openings did exist within a limestone mass, *"the flow will pass from joint to joint in the closest approximation to the ideal arcuate path which the configuration of the joint system will permit."* (p.791). This observation is particularly significant, even more so in that the authors go on to comment that, *"Because the spacing, open character, size, distribution, and intersection of the joints depart from uniformity throughout the rock, the analogy between natural flow through rocks and flow through an ideal homogeneous medium departs from perfection to a variable but significant degree."* This open-ended caveat in no way belittles the overall importance of their statement, but it does allow the type of actual and incipient drains postulated by Lowe (1992) to fall within the limits of theoretical possibility.

Conspicuous by its absence is a consideration of the actual chemistry of the dissolutional development that the authors model. It must be assumed that the development to which they refer was accepted as being due to the effect of carbonic acid dissolution. Some years later Alfred Bögli (1964) put into words the fact (which must have been noted by many, but stoically sidestepped) that the majority of cave formation could not be explained by the one simple mechanism! However, the inadequacy of the solution chemistry tacitly assumed by Rhoades and Sinacori does not diminish their theoretical model. The alternative chemistry suggested by Bögli (1964a,b) and similar reactions deduced by later authors might go all or part of the way towards explaining dissolution in a zone where only non-aggressive water would be expected to flow.

General considerations 1941 - 1960

This long period, partly corresponding to the *"Hiatus"* of Watson and White (1985), was punctuated by the later works of researchers already discussed and the earlier writings of some of the authors discussed below. Within this time there was no acknowledged single contribution which stands as a speleogenetic milestone, though it can be argued that a paper in Russian by Durov (1956), not published in English until 1979, should be elevated to such a position. Somewhat earlier, the first significant paper by T D Ford (1952) represents a different type of milestone, in that it married the skills and observations of the geologist with those of the speleologist, to produce a set of valuable conclusions, still at least partially valid today.

Prodigious quantities of painstaking experimental work and observation were the hallmark of the period. Striking confirmation of this is supplied by the reference sections of the many later texts dealing with karst hydrology, water and solution chemistry or geomorphological analysis and other speleological sub-disciplines. The volume and sophistication of this analytical work did not match that of the following decade when research by geomorphologists of the 'Bristol School' reached its acme. Nonetheless, it appeared that during this time an understanding of cave formation and karst hydrology was beginning to emerge, cemented together by sheer weight of observation and numerical data.

Not all of the conclusions reached on the basis of these diligent studies have stood the test of time. Modification or even abandonment of certain concepts formulated at this time has been necessary in the face of increased data, re-evaluation of the original data or method, or simply in the face of new ideas. A particularly good example is provided by aspects of the work of Corbel. His various pronouncements upon the links between climatic conditions and the broad aspects of karstification not only generated a surge of similar investigations, but formed the basis of considerations in numerous karst geomorphology text books. Discussion of Corbel's approach and deductions by Jakucs (1977, pp.104-108) is not only critical, but scathing. It appears to represent an extremely well-informed contradiction, verging upon the destructive. Such an attack upon the established 'truths' of karst geomorphology (even ones only briefly established) is a rarity; for this kind of criticism to occur in an English language publication is virtually unique.

Despite the seeming shadow of Corbel's work upon analytical speleology/karstology, the period doubtlessly produced a considerable volume of relevant data. Studies in western and, more particularly, eastern Europe during this time included the early work of many authors who will not be considered in detail, such as Gams, Trombe, Zötl, and H Lehmann, as well as Jakucs and Bögli. In America papers from Kaye, Howard and Weyl reported results which would subsequently be of great importance.

What must be questioned is the wisdom whereby so much effort, both underground and in the laboratory, was applied to confirming and refining the accepted views of karst solutional chemistry, rather than to an objective questioning of the overall mechanisms involved. If more attention had been paid to analysis of water samples (particularly those from incipient passages rather than open passages and drips) with regard to their contained ions rather than hardness/aggressiveness, pointers to the potential importance of strong acid dissolution to early and ongoing speleogenetic processes might have been apparent at an earlier date. However, since relatively little work has been done along these lines even today, the outcome of such an objective approach is currently imponderable.

W E Davies

"In the initial stages of solution primitive tubes, pockets and similar solution features are developed at depth in the ground water. In a dense tight limestone initial solution may be dominantly intermolecular; later solution is mainly along joints and fractures. The product of this stage is a series of random, nonintegrated openings."
(Davies, 1960, p.17.)

William E Davies, a geologist with the United States Geological Survey, had produced a number of papers related to caves and underground drainage prior to publication of his more widely known work in 1960. His research during this period is described as a milestone by Watson and White (1985, p.116) and there is no doubt that his studies revealed new and intriguing perspectives of speleogenesis. Much of his reasoning remains fundamentally applicable, even in the light of more recent lateral viewpoints; other aspects benefit from similar lateral consideration.

Davies postulated a four-stage sequence of shallow phreatic cave development which is best summed up in the words of his own abstract (Davies, 1960, p.5):

1. random solution at depth in a zone of saturation to produce nonintegrated solution tubes and pockets.
2. integration of tubes into mature caverns at the top of the zone of saturation during a period when the water-table was uniform in altitude and flow was constant for a long period of time (direction of flow was toward major valleys).
3. deposition of clastic fill under alternating conditions of saturation and aeration.
4. relative uplift of the cave above the zone of saturation with modification of passages by deposition of speleothem, erosion to [?of] fill material and collapse.

The obvious missing factor in the above brief synthesis is that of vadose invasion during the fourth stage. It is perhaps implied by the reference to erosion of fill but, reasonably, erosive modification of the carbonate bedrock would be a process of far greater significance. However, this omission, whether or not intentional, in no way detracts from the remainder of the speleogenetic model.

The title of the paper, *"Origin of caves in folded limestone"*, is, to a degree, misleading. In common with other similar papers, there is a lack of text figures showing cave elevations, rather than simplified plans, in relation to geological structure. Though the elevation view is that which ought, logically, to be tied to cross-sections of fold or fault structures, it is rare to find even the most stylised illustration of this type. Davies' paper is typical of this omission and includes only a number of profiles and related cross-sections of short lengths of cave passage, somewhat loosely tied to dip and strike. There is no attempt, even by the compromise of relating structure to cave plans by means of dip arrows or similar data, to provide a regional view. Thus, on reading closely, it becomes apparent that rather than the sharp structures that might be implied by the title, Davies is mainly concerned with describing cave development in dipping (albeit steeply dipping) sequences as against in flat-lying beds. This in itself is a significant step forward insofar as many earlier workers, W M Davis included, had failed adequately to conceptualise this type of cave formation. Equally, other workers had not considered it a problem.

Paramount among Davies' observations was that passages within folded limestones maintain a uniform level, a situation which he tied directly to dissolutional development at or close to the piezometric surface. The arguments he put forward appear to distil down to the formation of an alternation of bedding controlled passages and linking passages guided by *"joints"* or, rarely, by faults. If this reading of Davies' thesis is correct, then he was close to describing ideas discussed by Lowe (1992), whereby drainage constrained within one susceptible horizon will sidestep along a joint to exploit a different but hydrologically advantageous horizon rather than persist along a less advantageous route in the same bed. If a fault rather than a joint were implicated then the relationship would be even more easily explained insofar as the drainage would flow preferentially up, down or along the fault plane to regain the same horizon, or a suitable alternative, rather than extend along strike in a less appropriate lithology. According to more recent concepts (Lowe, 1992), however, passages developed at progressively lower topographic levels would probably have existed in incipient or immature form whilst the entire sequence was saturated. Only when the piezometric surface began to be relatively lowered would the most advantageous combinations of incipient routes be chosen, these 'pirating' the drainage of adjacent incipient routes down to the same level. Those incipient routes below the water table would continue to function in the same way until, with further relative drops in the water-table, they too were given their chance to pirate the main underground flow.

Insofar as it is possible to draw conclusions from, and comment upon, the illustrations, each of the examples included in Davies' figure 2 (1960, p.8) (redrawn and reproduced as figure 2) is explicable in the context of the new ideas alluded to above. The passage segments shown appear to bear a strong relationship to the bedding of the rock, and hence, presumably, to lithostratigraphy. Passages shown to be strike orientated can be seen in the cross-sections provided to be preferentially developed within the (possibly schematic) limits of one or a small number of beds. Those passages shown as being orientated across strike (ie up and down dip) have a vertical component which might be indicative of rift-type development upon a fault or joint plane. Illustration G (in Grand Caverns, Virginia) is confusing since the profile is broken into two sections of strike passage separated by a section across strike. The relationships of these three sections are obscure.

Another of the points raised by Davies concerns multiple level passages which occur in a relatively small number of the caves he had studied. He goes to some lengths to discuss and justify a linking of these multiple levels to the spacing of river terraces on the surface. This is not a particularly difficult concept to accept in certain circumstances, though there has been much misunderstanding regarding underground passage levels and their relation to the levels of surface features. In the conditions of passage formation described by Davies, where (according to the new model and probably Davies' own) incipient or immature openings are enlarged at or close to the water-table, it would be expected that the drop from one level to another on the underground drainage would be linked to drops in surface level. In a slightly more sophisticated view, it might be that the drop in the underground water-table is triggered by surface streams cutting down to the next suitable horizon which is capable of pirating the cave drainage, as was suggested by Gardner (1935) and Lowe (1992) The most important point made by Davies in this context is that, *"With no significant exceptions these passages maintain uniform slope and vertical spacing in the caves."* (Davies, 1960, p.8). Sadly he does not indicate whether the vertical spacing reflects a simple drop down dip within the same bed or if it reflects a true vertical drop from one susceptible horizon to another.

In the latter context it must be added that Davies also remarks (p.11) that *"...certain beds of limestone are favorable to cave development".* He then goes on to specify (p.11) that *"The most prominent zone of solution is along the junction of the Coeymans and New Scotland Limestones".* This concentration of dissolution along a formational boundary is perhaps the most significant observation thus far in the paper. Davies also describes another distinct dissolution zone within the Lenoir-Mosheim Limestones. It is not specified here that the zone follows the formational boundary, but the wording implies that this might be the case.

Davies also considers the significance of passage shape and makes a strong case for the influence of stress zones upon cross-section at various depths, up to 500 feet (150m) or more, below the (presumably) contemporary surface. This will not be considered in detail here, except to point out that the two broad passage shapes Davies describes might be equally well explained as being a reflection of their guidance mechanism. His broadly elliptical passages (in a minority) could reflect dissolution along a joint or fault or bedding plane. Subrectangular, triangular or trapezoidal forms which constitute the majority in his study area are explicable in terms of dissolution within a susceptible bed with severe constraint to dissolution in the beds above and below, with or without the contribution of breakdown processes. Vadose entrenchment is not mentioned as contributing to the development of any of the passages described in the published account. The lack of any significant reference to underground vadose drainage within Davies' examples is perhaps its most puzzling aspect. It is not clear whether the lack is a true reflection of the situation within Davies' field areas, whether he deliberately omitted vadose effects as being of secondary importance, or whether he overlooked the diagnostic criteria of vadose modification and the importance of vadose processes to his arguments.

Another positive point to consider among Davies' observations is that many of the fill deposits in the caves he studied contain sulphate (sulfate) minerals, particularly gypsum. He describes (1960, p.15) fill beds which contain up to 50% of coarse gypsum crystals. Davies' explanation that the gypsum was surface derived and washed in by the same waters that washed in the fill may be inadequate. If such was the source, it is unreasonable that the gypsum would be preserved as recognisable crystals. Some degree of dissolutional or abrasional rounding, if not the total dissolutional removal, of the crystals would be expected by surface derived water, which would still be undersaturated with regard to sulphates and other salts. It is possible that such gypsum crystals formed *in situ*, within the fill, probably crystallising from supersaturated solutions as the fill desiccated after drainage. The source of the sulphates was perhaps carbonate dissolution by sulphuric acid, the latter derived from oxidation of sulphide impurities present at the requisite levels within the local sequence. A compromise view between the latter explanation and that of Davies would be that supersaturated solutions, from a hypersaline environment on the surface, flooded into the caves and gypsum was precipitated in static or near static conditions underground due to changed physical and/or chemical conditions.

As with several earlier workers, particularly Bretz (1942), Davies showed considerable interest in the cave fill. His view accepted the presence of fine-grained deposits (the unctuous red clay of Bretz) which are residual to the dissolution of limestone. Additionally, however, he notes more typically clastic deposits of silt, sand and gravel grade. An interesting grading sequence is described, initially coarsening upwards, then fining upwards. Davies describes a series

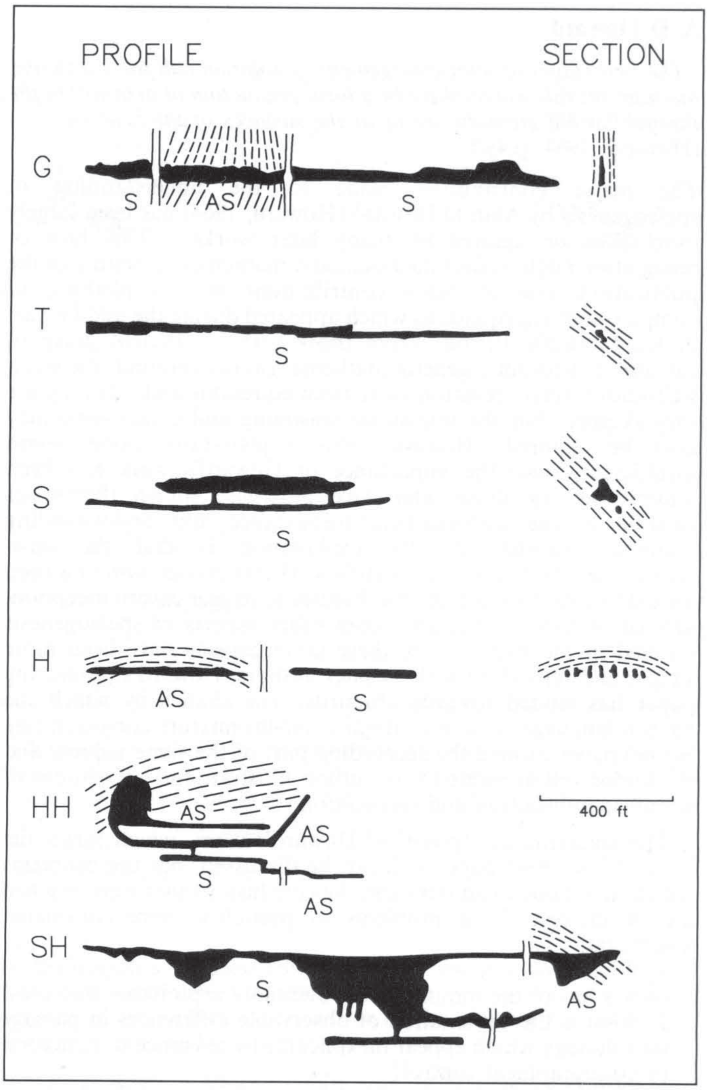


Figure 2. Relationships between cave passages and geology (redrawn from Davies, 1960, Figure 2). Note that "S" = along the strike, and "AS" = across the strike of the succession. Letters on left refer to locations: G = Grand Caverns, Virginia; T = Trout Cave; S = Sinnit Cave; H = Hamilton Cave; HH = Hellhole; SH = Schoolhouse Cave, all in West Virginia.

of depositional events wherein the lowest, floor, deposits are the fine, unctuous clays derived from the host rock by dissolution. These are followed by upward coarsening stream deposits (the silts, sands and gravels) which reflect an increase in the energy of the underground stream as the available passage space is reduced and the stream occupying the smaller space remains of similar volume. It is uncertain whether this mechanism is realistic or whether a simple flow increase might provide a better explanation. Whichever is the case a critical point is reached where the passage is effectively blocked by clastic debris, flow becomes sluggish or static, and finer debris settles to complete the passage fill. As far as it goes, this explanation of the fills observed is reasonable. Still puzzling, however, is the lack of any description of vadose entrenchment into the carbonate bedrock. If there is an explanation for the presence of powerful streams which are only capable of depositional rather than erosional activity, it is not apparent in the text.

One final point pertaining to Davies' speleogenetic views concerns his reference, within his first stage of cave formation, to the dissolutional development of random, non-integrated cavities below the piezometric surface. Such a mechanism is, indeed, implicated within the earliest phases of speleo-inception (eg Worthington, 1991; Lowe, 1992). However, it is also apparent that the seeming randomness of the dissolutional cavities is illusory. According to Lowe (1992) inception probably begins along a limited number of susceptible horizons, and not necessarily at every point on every available horizon. At this stage the rest (the vast majority) of the carbonate sequence is effectively insoluble (due to the lack of access for potential solvents), except along those lines, guided by faults and joints, where the solvents derived from inception horizons are capable of intra- or inter-horizon migration. Such a pattern, though not necessarily of integrated appearance and seemingly random, is predictable with sufficient knowledge of the lithostratigraphy and structure.

A D Howard

"The first stages of joint enlargement by solution call for a different mechanism; this would likely be a local production of acid within the throughflowing groundwater or at the surfaces of the bedrock." (Howard, 1964, p.48.)

The major contribution made to the understanding of speleogenesis by Alan D Howard (Howard, 1964) has been largely overlooked or ignored by many later workers. This lack of recognition might reflect the essentially mathematical nature of the publication, one of many contributions with a plethora of complex-seeming equations which appeared during the middle part of the Scientific Period. From those with insufficient grasp of calculus or without a general mathematical background, there is a well-known recoil reaction away from equations and not only are they skipped, but the interstitial reasoning and conclusions may also be ignored. However, this explanation alone seems insufficient, since the importance of Howard's work has been undervalued by those who have otherwise shown themselves capable of the mathematical forbearance and understanding required. Another possible explanation is that the most revolutionary element of Howard's work, the recognition of a need for a different dissolutional mechanism to trigger cavern inception, seemed of lesser importance than other aspects of speleogenesis covered in the paper. Since these latter aspects have been more simply and fully dealt with by later authors, Howard's pioneering paper has tended towards obscurity. The chance by which the French language version of Bögli's (1964b) mixture corrosion (see below) paper formed the succeeding part of the same volume and the kudos still accruing to its author, illustrate the vicissitudes of scientific publication and recognition.

The mathematical 'proof' of Howard's work, which forms the bulk of his (1964) paper, will not be discussed, but the problems which he set out to address and the conclusions that were reached are of interest. Four problems in particular were considered significant:

1. What processes are responsible for selective enlargement of only a few of the innumerable potentially exploitable fractures?
2. What is the explanation of observable differences in passage morphology which appear inexplicable by reference to structural or stratigraphical control?
3. Why is there an apparent preference for cavern development directly beneath the water-table?
4. What is the mechanism for development along the shortest paths between surface sources and resurgences of groundwater?

Unfortunately, having listed these four problems, Howard did not tackle them separately, but launched into considerations which are partially applicable to them all.

From the point of view of a modern re-examination of speleogenesis, particularly of the factors affecting speleo-inception, his comments regarding the earliest stages of dissolution are of the greatest interest. By assuming initial joint widths of 0.2mm (10% of the value deduced by O Lehmann for his *Urhohlräume* cavities (above), but considerably wider than assumed by S N Davis (below) for his flow calculations) a 10m hydraulic head and an input to output distance of 1km, Howard deduced an initial flow velocity of 30mm per hour. He then examined the traditionally accepted dissolutional mechanism, due to carbonic acid, and by reference to previously published solutional kinetics data (Weyl, 1958) demonstrated that the water flowing into a joint would be fully saturated within the first 10mm. On this basis he commented (p.48) that, *"... it is hard to imagine caves beginning to form ..."*. His conclusion was that a different mechanism was required and that this new mechanism would involve the local production of acid within the rock mass. At this point, rather than following up the detail of the potential acid producing processes he broadly postulated that the acid could be formed by direct or bacterially-assisted oxidation of organic matter in the water or sulphides within the rock, as Kaye (1957) had previously suggested. Rather than developing either concept he then entered into a consideration of the mathematics of the rate of acid creation and rates of passage enlargement.

He derived equations to describe three hypothetically possible speleogenetic situations but involving different interrelations of hydraulic head, length of drainage route from entrance to exit, passage diameter and flow velocity. In the situation described by his first equation no cavern formation would occur, but cases two and three indicated that joint enlargement by this process was a theoretical possibility. The lack of widening under equation 1 conditions was taken as the explanation of the failure of some joints to develop as caves (problem 1, above). However, having demonstrated that, if acid was created, *in situ* cave formation could proceed, he did not demonstrate that acid is created in such a fashion.

Having thus established a possible mechanism of passage inception, Howard went on to consider the ongoing mechanism of cave enlargement. Broadly he stated that once a significant through-flowing discharge is achieved, the emphasis will shift from dissolution by *in situ*-generated acid to dissolution by undersaturated groundwater. He noted that limestone is soluble in groundwater both due to the solubility of calcite in pure water and because of carbonic acid which is contributed by the atmosphere and sources within the soil. Once the rate of flow becomes sufficiently rapid that groundwater passing through the enlarging joints is not completely saturated with respect to calcite, a rapid increase in the rate of dissolution should occur. Only the joints which had previously been widened beyond the size threshold which would allow this increased flow are further enlarged, adding to the explanation of selective joint development (problem 1 above). Only a *"selected few"* (p.51) of numerous original fractures will become cave passages and these are deduced to be those which initially have a high hydraulic gradient acting across them and/or are initially of larger diameter. On the local scale the size factor will be of greater importance, since the gradients of adjacent and similar fractures will also be similar, but on a larger scale, the influence of hydraulic gradient will predominate so that favoured routes will attain the laminar-turbulent threshold before it is attained by potential 'rival' routes.

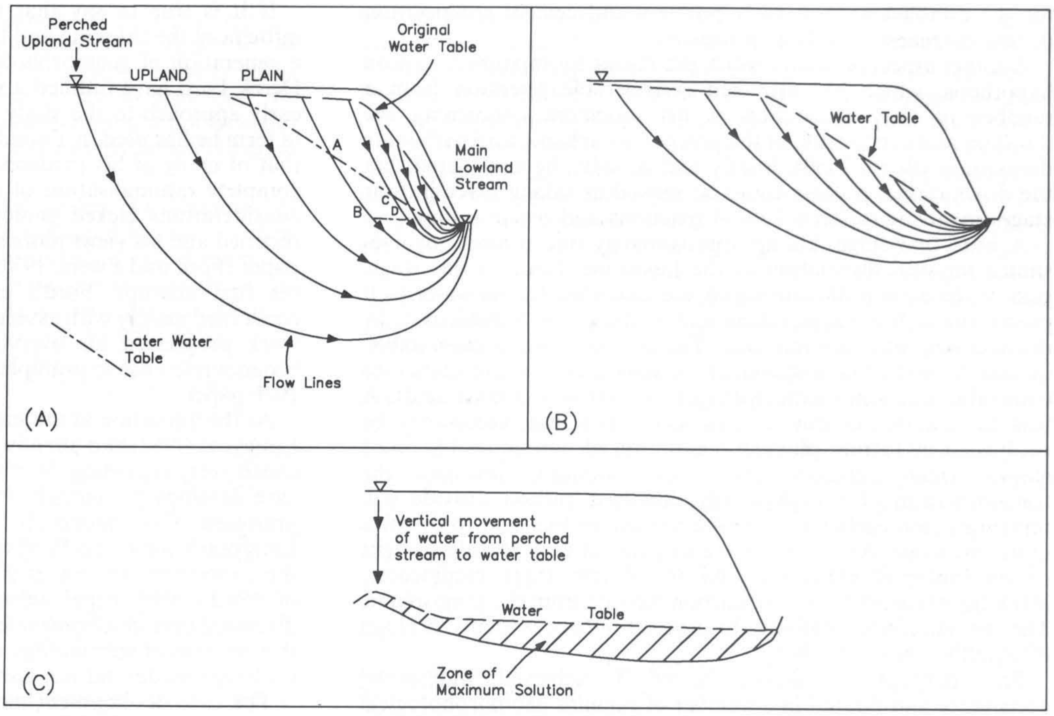
Having discussed the interrelationships which lead to the favoured development of a limited number of initial fractures, Howard examined the development of water-table caverns and caves formed by artesian flow. In discussing factors affecting the development of water-table caverns he assumed that it proceeds without accompanying change of surface topography. This is a gross assumption which is certainly unrealistic in terms of the timescales of underground and surface landform development and probably also produces an unrealistic view of the interrelation of processes and landforms involved. Allowing for this lack of reality, his consideration of the processes involved remains potentially informative. Following his arguments from earlier in the paper he deduced that the groundwater level at the earliest stage of underground drainage is at or just below the surface and that the flow pattern produced will be comparable to that in porous rocks. A deeply arcuate flow net (similar to those figured by Rhoades and Sinacori, 1941, and resembling the model suggested by Davis, 1930) was deduced and the main dissolutional effect, due to locally produced acid, was assumed to be in the area of highest hydraulic gradients. In the specific model illustrated (Howard's plate 13, reproduced as figure 3) this lies close to the edge of a supposed upland area which drains to a lowland stream.

With increased joint widths in the upland edge area there is a tendency for the water-table to drop (figure 3A) and, due to considerations of hydraulic gradients, the flow lines become shallower where the water-table slopes down near the upland edge (figure 3B). Even after conditions of constant discharge have been established, surface streams continue to contribute drainage with a constant hydraulic head. Eventually the enlargement of the favoured passages will be such that the hydraulic head will be negligible over most of the area when compared to that below the stream sinks. Thus the water-table will become graded to these larger entrances (as shown in figure 3C). In Howard's view this would lead to a restriction of nearly all the groundwater movement to a zone directly below a nearly flat water-table. In addition, dissolution will be restricted to just a few passages, due to the transition to the influences of undersaturated water and turbulent flow, as discussed above. This model is considered to provide at least a partial answer to Howard's problem 3, above.

Additional discussion regarding the nature of the drainage in the essentially vadose regime between the surface sinks and the water-table brings a partial answer to the questions of passage morphology posed by Howard's problem 2. However, the view of erosion by turbulent, free-surface flow followed by the aggradation of sediment which infills irregularities, is far from being the complete story, which is perhaps much more to do with variations in structure and stratigraphy than Howard perceived.

Overall, the model of water-table cavern development formulated by Howard is attractive; elements of the model are still acceptable, despite the appearance of supposedly superior models. The major criticisms in the context of this review are that his initial conditions of topography are unrealistic and his model takes only the very broadest cognizance of geological factors. His view

Figure 3. Theoretical view of flow lines, hydraulic gradients, water table development and zone of maximum dissolution (redrawn from Howard, 1964, Plate 13). Howard describes this as a physiographic situation which promotes cavern development, and makes the assumptions that no previous dissolution of limestone has taken place and that cavern development occurs without an accompanying change of surface topography. Neither assumption can be accepted in the context of the inception horizon hypothesis (Lowe, 1992).



assumes that significant cave formation - his earliest, local acid dissolution phase - only commences after a significant surface relief has been produced. This is not necessarily so and certainly does not need to be the case if some other mechanism than hydraulic head drives the initial groundwater circulation (cf S N Davis, below). Howard claims (p.58) that the water-table model *"...should be commonly applicable to all cavern forming situations free from unusual structural, stratigraphic or topographic controls"*, but realistically these conditions will only ever be approximated. The lithology varies within a single bed of limestone, more so across a cavernous sequence, no pattern of jointing is regular and predictable, the dip of a sequence is undulatory when examined in detail, and so on. Thus, though possessing attractive elements and possibly answering several important questions which were current at the time of Howard's study, the main shortfall of Howard's model is that it tries to be too wide-ranging.

In considering the nature of caves formed by artesian flow, Howard stated (p.58) that they are forced to follow steep paths by, *"stratigraphic and structural constraints"*. He considered the situation of water sinking and flowing down-dip within a constrained limestone bed on the shallow limb of an asymmetrical syncline towards a resurgence which lies at a lower topographical level at the upper limit of the steep limb of the syncline (or as Howard viewed it, the steep limb of an adjacent anticline). Consideration of the flow lines produces a convincing picture of the concentration of drainage along the shortest route between sink and rising (providing a partial answer to problem 4, above). Once the idealised flow pattern is deduced, short additional steps produce views of water-table drawdown, water-table passage development and the potential for maze-like passage configurations, all of which are elements of most subsequent cave formational models. As described, the Howard model seems to be flawless, but again it must be appreciated that it could only apply in its fullest sense in the idealised conditions figured in his paper. If the geological structure and its relationship to topography were to be examined in three dimensions rather than two, it is unlikely that such a simple picture of the flow situation would be apparent; what of the fold plunge component, what are the effects of joints superimposed upon the constraint of stratigraphy and lithology, and so on?

Howard's (1964) work is of fundamental importance to the concepts investigated and re-assessed during more recent studies. His ideas of the earliest stages of dissolution mesh with the ideas of the earliest stages of flow suggested by S N Davis (1966) to point towards the need for a speleo-inception mechanism which lies outside the views of most earlier and later workers. Though it is possible to see flaws in Howard's thinking, which are due to his need to approximate or assume various parameters involved in his arguments, it is difficult to fault the mathematics and the reasoned deductions which are built from the approximations. Though his deduced mechanisms might be appropriate within the assumed limits of his data and in the uncomplicated situations devised to illustrate the model, his tacit assumption of more universal applicability for the full spectrum of the model might be seen as being too optimistic. However, it seems reasonable to accept that aspects of speleogenesis described by Howard, particularly his view of the initial phases of dissolution, might be widely applicable on the microcosmic scale, as elements of local activity contributing towards a more regional whole.

A Bögli

"In other words, the formation of the majority of caves cannot be explained by the present conception of the solution of carbonates." (Bögli, 1964b, p.62.)

Alfred Bögli, a Swiss chemist and speleologist of considerable reputation, is best known for his work in the Hölloch cave system and for the recognition and first published description of the mechanism of carbonate mixture corrosion (*Mischungskorrosion* or *mélange des eaux*; 1964a, 1964b). Bögli pointed out quite simply that dissolution would not be expected to take place within the phreatic zone by means of the traditionally accepted carbonic acid reaction. He reasoned that within this zone the water would be saturated with respect to calcite and incapable of dissolving further rock. Since, in order to become aggressive, the water would have to take in additional carbon dioxide and since carbon dioxide is unavailable in a water-filled environment, the long-accepted reaction could not take place. Other researchers, prior to Bögli, had noted the same apparent paradox, though the case was perhaps more clearly stated by Bögli. Recognition of the impossibility of carbonic acid dissolution, even at relatively shallow depth beneath the surface, had led to the argument that corrasion must be the dominant cave forming process - an argument which was contradictive of the great majority of observational data.

Study of the solution chemistry and saturation curves indicated to Bögli that the mixing together of two saturated solutions could produce an undersaturated solution, capable of dissolving additional calcite. For example, at a point where a major phreatic flow derived predominantly from surface sinks is intersected by flow (from a joint or similar fissure) which is entirely of percolation origin, mixing takes place. The water derived from surface sinks would be saturated with respect to calcite dissolved under conditions of relatively low carbon dioxide partial pressure. The percolation water would also be saturated, but this saturation would have been achieved at a high carbon dioxide partial pressure. The resultant mixed water is, in theory, undersaturated and capable of dissolving additional calcite. Bögli provided illustrative curves and examples which were at least broadly understandable to non-chemists, and his description of mixture corrosion stimulated later research into this and associated mechanisms, such as that by Wolfgang Dreybrodt (discussed below). It is interesting to speculate whether mixing processes such as those described by Bögli may be equally applicable at pore level within carbonate successions and be at least partially responsible for the establishment of early porosity and cement redeposition during diagenesis and later processes.

Another aspect of Bögli's work, pre-dating his mixture corrosion hypothesis, which has attracted considerable attention from a number of later researchers is his statement concerning the duration and subdivision of the process of carbonic acid/carbonate dissolution (Bögli, 1960). Briefly and broadly, he considered that the dissolutional process could be viewed as taking place in four stages involving different sets of reactions and reaction times.

A brief first stage, taking approximately one second, involves simple physical dissolution of the limestone. Prior to this stage, part of the carbon dioxide which was absorbed by the water as it passed through the atmosphere and/or the soil is transformed, by dissociation, into carbonic acid. The second phase is inseparable, in time, from the first and consists of association of carbonate ions (from the limestone) with hydrogen ions (from carbonic acid). A loss of equilibrium due to this association will necessarily be readjusted by further physical dissolution of limestone. The third stage, which requires about one minute, involves the transformation of the physically dissolved carbon dioxide still remaining into carbonic acid, and the use of this acid to dissolve more limestone. At this point the original carbon dioxide content of the water is exhausted and the fourth stage commences, involving absorption of more carbon dioxide from the atmosphere. The sequence now continues to proceed until, eventually, a chain of equilibria is established.

This concept, or derivatives of it, achieved widespread acceptance and figured in a number of popular geomorphological texts (eg Sweeting, 1972). Though not in total disagreement with Bögli's rationalisation, Jakucs concluded (1977, p.48) that, *"...even if we retain Bögli's four stages for its illustrative and didactic merits, we must keep in mind that the individual stages in nature never occur in succession, but invariably simultaneously and in mutual interplay: that is, the stages as such do not in fact exist".* By the time his text book was published (Bögli, 1978), Bögli's description of the process had changed in detail (though remaining essentially the same) to include six steps with the added rider that if dolomite (providing Mg^{2+} ions) is involved in the dissolution, then the interface conditions become considerably more complex.

Subsequent to his publication on mixture corrosion, Bögli's karst research has continued through the Scientific Period and into the Modern Era, though his views on many aspects of karst studies have remained essentially unchanged. In 1980 many of his ideas appeared in English translation in his *"Karst hydrology and physical speleology"*, which provides a highly readable digest of European, and, to a lesser degree, American, karstological concepts. Though it can be argued that Bögli's expertise is biassed by his cave studies in the Alps or by European conditions in general, it is his early appreciation of the failure of contemporary dissolution chemistry which forms his most significant contribution. His statement of this failure and his suggested alternative mechanism are valid in all parts of the world; only some of his subsequent ideas that deal with cave formation and karst development in general suffer from a degree of parochialism, reflecting a preoccupation with Alpine geological conditions.

A vadose canyon incised beneath a phreatic lift, in Swildon's Hole, Somerset, site of Derek Ford's early work. (Photo: Tony Waltham).

D C Ford

"The multitudinous theories .... are neither correct nor incorrect in the general case, they are irrelevant...."
(White and Longyear, 1962, p.167; quoted by D C Ford, 1965, p.110)

If it is true to say that the work and ideas of W M Davis influenced the thinking and the direction of research of more than a generation of geomorphologists, it is equally true that those of Derek Ford have fulfilled a similar role in more recent times. His early approach to the study of caves, particularly of cave origins (a term he has used in a number of papers), was less dogmatic than that of many of his predecessors, leading eventually to an almost complete rationalisation of cave development theories. His early considerations lacked geological insight, but this omission was rectified and his views more clearly stated and illustrated in a joint paper (Ford and Ewers, 1978; discussed below) some 13 years after his first attempt. Ford's early publications (1965, 1968) were concerned mainly with cavern development in the Mendip Hills. As work progressed his overview became more cosmopolitan and broader reference to examples from wider areas was apparent in his 1971 paper.

As the quotation at the start of this section indicates, White and Longyear (1962) had already pointed out the futility of the ongoing controversy regarding deep phreatic, shallow phreatic and vadose cave development models. It is interesting that even though Ford produced this quotation and other aspects of White and Longyear's work, it is Ford who is generally considered to have lain the controversy to rest. It is also interesting to note that the title of Ford's 1965 paper referred to cave *origin,* his later papers discussed cave *development* and *cave genesis.* It is uncertain whether this revision of terminology mirrors a realisation that his new (and evolving) model did not solve the problem of cavern inception.

The cave development model described and modified by Ford, which grew into the Ford and Ewers model (below) is well known and in its various forms has figured in most speleological and karst geomorphological texts published in the past 20 years. It will not be discussed in detail here insofar as many of the criticisms which can be levelled at Ford's early model reflect his lack of geological consideration, an omission which was rectified by Ford and Ewers.

"Therefore we are dealing with a relatively thin limestone mass in which there is potential for an unusual amount of lithological perching of passageways."
(D C Ford, 1971, p.92.)

S N Davis

"Several possibilities exist that might explain the puzzling fact that solution cavities develop rapidly along joints that are not significantly more permeable than the enclosing rocks."
(Davis, 1966, p.114)

The importance of Stanley N Davis's contribution to the discussion of speleogenesis (Davis, 1966) lies not necessarily in his conclusions but more certainly in the problem that he recognised and raised. Broadly, he pointed out that limestone, even jointed limestone, under normal conditions, is almost completely impervious to groundwater. This view is supported by the limited data available but appears contrary to the widely accepted view of limestone as a potential aquifer rock. The lack of primary, intergranular **porosity,** in the majority of (but not all) carbonate lithologies is readily demonstrated, as is the widespread occurrence of varied degrees of secondary **permeability.** The latter property, frequently attributed to the presence and movement of water along fractures and bedding planes, might, in many cases, be better seen as being a function of speleogenesis. To demonstrate that the rock mass is effectively **impervious** until joints are widened, as stated by Davis, is virtually impossible, though Davis's observational results present an approximation to values expected in pristine rock.

Davis's preamble to the problem includes a number of important observations, not only to his own work, but applicable across the entire field of speleogenetic study. He considered that it is obvious that dissolution cannot proceed unless water moves freely through the rock, and this *"obvious"* statement is in itself a step forward. Most earlier authors had either ignored this need for movement or tacitly assumed a movement as they debated how dissolution could be achieved by theoretically non-aggressive solutions. Davis necessarily took an opposed simplistic view and by assuming the presence of groundwater with a constant capacity for carbonate dissolution he examined the interrelationships between effective porosity, permeability and hydraulic gradient. By the use of Darcy's Law he was able to demonstrate that *"reasonable combinations"* of these variables yielded average flow velocities of 3mm per year in dense limestone and as much as 3m per year in slightly porous limestone. Even at the greater value, he calculated that it would take 100,000 years to increase the effective porosity by 1% (assuming an initial value of 10%). He considered the chosen parameters were liberal and on this basis he concluded (Davis, 1966, p.113) that, *"...large solutional openings in limestone are not caused by the percolation of water through the solid matrix of the limestone".* Though it might be argued that, on the basis of ideas concerning the dissolution of limestone by strong acid, that the theoretical basis and hence the conclusion itself might not be valid, it leads to a consideration of fracture permeability which is highly relevant.

By similar reasoning to that applied to his consideration of porosity, Davis concluded that in areas fractured by faulting flow velocities of many km/year might be expected, if initial openings large enough to permit turbulent flow were present. On this basis he considered that little difficulty exists in developing conditions prerequisite for the extensive dissolution of faulted limestone. A possible flaw in Davis's reasoning is that the visible presence of an open, joint/fault-guided, dissolution-enlarged permeability in a faulted area is not proof that the joint/fault fractures were open at the commencement of dissolution. Supposed provings in wells, tunnels and open excavations described by Davis (p.113) might merely indicate the presence of a developed and/or developing dissolutional permeability. The same rock mass could be expected to contain healed or incipient fractures and there is no possible way to provide conclusive proof that open fractures existed when dissolution commenced. Whether such fault zones contain primary open fissures or not, Davis goes on to consider the situation in (apparently) unfaulted regions where, he assumes, open joints would **not** exist.

He points out that, even in untectonised areas, many dissolutional features have a linear form and that this observation has led to an almost universal conclusion that dissolution preferentially follows joints. This section of his work (pp.113-114) is of great importance. Field tests showed that joint width in the freshest available exposures of several different rock types was certainly less than 20μ and probably less than 10μ. From this, and by referring to evidence from borehole samples, it was possible to state that joint widths in buried rock sequences, where confining pressures are much higher than at the surface, must be smaller and values as low or lower than 0.1μ might be expected. By applying figures within the ranges outlined above to a model of water flow along a joint between smooth planes, Davis was able to demonstrate (p.114) that flow along cracks appreciably narrower than 10μ will scarcely add to the bulk permeability of dense limestone. [NB. the units used above are as used by Davis, in modern terms $1\mu = 1 \times 10^{-6}$m.]

In the light of all the previous considerations Davis was faced (p.114) by *"...the puzzling fact that solution cavities develop rapidly along joints which are not significantly more permeable than the enclosing rocks".* Thus his paper proceeds to explore potential mechanisms whereby this apparent paradox might be explained, and his work presents an interesting parallel to Howard's (1964, discussed above) work on the complementary paradox of the chemistry of joint-guided dissolution. Before embarking upon his explanation of how such dissolution might come about he points out that the data already considered could be misleading. Although virtually closed joints had been observed, these could be linked to open joints which would be amenable to dissolution. Also, the permeability values he had used within his model may have been biassed in such a way that the true significance of the contrast between the permeabilities of joints and limestones was masked. Additionally he notes that dissolution along joints may be confined to shallow depths where joints have opened due to the release of stress, quoting an example of an underground quarry in Kansas where fissure flow is common near the portals but totally absent at depth. He notes that most of this mine is below the local water-table, proved by exploratory drilling; these linked but contradictory observations ought to have prompted an examination of the water-table concept in limestone (cf Lowe, 1992, chapter 9). These possible mitigations of the apparent puzzle lead to the possibility (p.114), which he subsequently develops, that forces other than those supplied by the regional hydraulic gradient may be implicated in the initial stages of water movement through joints.

Davis attributed the initial development of solutional openings in jointed limestones to the effects of *"groundwater pumping"* (p.115), a process caused by strains established within rock sequences by various effects, but notably by earth tides. He presented the hypothesis in a simplified form, accepting that the interplay of chemical and other factors is highly complex, and added the important rider that pumping is probably only important during the early stages of dissolution.

In the 'primitive' situation previously described, incipient or tight joints are present, but the overall permeability is extremely low. Although there will be a natural tendency for water to move due to the regional hydraulic gradient the actual flow quantity is so small as to be negligible (and the amount of dissolution would likewise be infinitesimal). However, water is assumed to gain contact with joints, either from [?surface] streams, fault zones [assumed to be open] or existing cave passages, and is driven along the incipient microscopic pathways by differential movements of the rock on either side. Having suggested a mechanism for pumping the water [grossly simplified in the present account] Davis examined the limitations of initial joint width, hydrostatic heads and the influence of the lithostatic pressures applied by the rock pile above the potential dissolution level. On this basis he showed that groundwater pumping, whether driven by earth tides, earthquakes or other mechanisms, is more reasonably implicated in the initial stages of dissolution than the traditionally accepted processes of flow due to a hydraulic gradient. The reasons for the subsequent decrease in the importance of groundwater pumping in the dissolutional modification of an individual joint/joint system was also demonstrated, being, briefly, the establishment of 'normal' hydraulic flow as the joint (system) is opened to 100μ or more.

A number of minor inconsistencies are apparent in the latter part of Davis's arguments. Prominent amongst these is the element, described above, that the groundwater pumped along joints might derive (among other possibilities) from pre-existing water-filled caves. In view of the other arguments he presents which indicate that development will occur preferentially along one joint once it attains a critical size and that development of adjacent joints will cease, it seems unlikely that water would be forced along essentially closed joints if an open cave passage was available as an alternative route. The mechanism by which pre-existing passages came to develop would also have to be examined, since the very essence of the paper depends upon passages not being able to form without the contribution of groundwater pumping. In all fairness, Davis himself considers that his groundwater pumping hypothesis is speculative, but he is able to back up the theoretical treatment by reference to differential movements measured across a joint in Wool Hollow Cave, California (1966, p.115) and to earthquake- or tide-induced water level fluctuations in wells (p.116). In the latter context he quotes several publications which provide convincing figures for the extent and magnitude of the effects under consideration.

Davis's contribution to the understanding of the very earliest stages of cave formation is possibly of very great importance, though as with several other publications dealing with the beginnings of speleogenesis this significance is often understated or overlooked. Davis receives brief reference and citation in most of the relatively recent western karst texts, White (1988) being a notable exception. Davis employed considerable restraint and reticence in the claims he made for the groundwater pumping hypothesis. Lowe (1992) argued that Davis's ideas are not only applicable to the dissolutional development of joint guided drainage, but also to the earliest stages of dissolution along bedding planes *(sensu lato)*. These two elements may be considered as effectively indistinguishable end members of a continuum of discontinuities, or may be viewed as separate and distinct entities, each embodied with particular distinguishing facets, some of which may be seen to merge, in the context of speleogenesis, within the bounds of the continuum indicated above.

"Well, the thing that Stan Davis is saying about these movements is their importance in getting the water initially into the tiny cracks. The water has to go in or there would be no solution at all......"
(Back; in discussion of Back, Cherry and Hanshaw, 1966, p.125)

Developments in the nineteen-sixties and early seventies

There were, in addition to the individual contributions described in detail above and below, many other publications deriving from serious scientific study of carbonate terrains during this part of the Scientific Period. At the University of Bristol a strong school of karst geomorphologists and hydrologists grew, under the inspiration and guidance of E K Tratman and D I Smith, the latter within the Geography Department. Much valuable, commonly innovative, research was carried out, mainly in the Mendip Hills and in County Clare (Eire), by such workers as Atkinson, Drew, Newson, Smart and Trudgill. To these names must be added that of Derek Ford, though his doctoral research was carried out under the auspices of Oxford University. The list of papers by the Bristol researchers during this period is impressive in terms of both quantity and quality and includes much which was crucial to the later development of many current views of speleogenesis.

The main streamway in Ogof Ffynnon Ddu, whose exploration prompted some re-assessment of the origins of the caves in South Wales. (Photo: Clive Westdale).

A small part of the output of the Bristol workers was concerned with chemistry of cave waters. Throughout the Sixties and into the Seventies other workers pursued the problems of dissolutional mechanisms, and enormous volumes of observational and analytical data, were amassed. The contributions of Bögli and Howard have already been discussed, but later (and some earlier) work by British researchers including Bray, Picknett, Pitty and Stenner seemed to provide confirmation of newly evolving ideas. In North America some of the work of Jacobson, Langmuir and Wigley, among others, was equally significant. A paper by White and Longyear (1962) was noteworthy in that it was concerned mainly with the limiting effects of hydraulic mechanisms upon speleogenetic processes. Among the points raised, it was argued that a critical velocity of flow must exist at which a transition from laminar to turbulent flow occurs, and that there is a corresponding critical conduit diameter, between 5 - 10mm, at which turbulent flow can begin. This concept was later followed up and developed by Atkinson (1968).

It is not intended to discuss these works in detail, insofar as the greatest part, by far, of the effort was targeted upon refining the understanding of carbonate dissolution by carbonic acid, with little energy applied to study or discussion of alternative processes applicable to cave inception. Much of the chemical research carried out during the 'post-Bögli' part of the Scientific Period only added weight to Howard's (1964, p.48) observation that, *"...it is hard to imagine caves beginning to form..."*, though adding appreciably to the understanding of how they might continue to develop. Surprisingly little work was carried out, within the field of chemistry, to fully investigate Howard's mathematically-based arguments, in marked contrast to the numerous follow-up analyses in pursuance of Bögli's theories. If it was seriously noted that most caves **ought not to exist,** the logical viewpoint must have been taken that they demonstrably **did exist,** and the investigation of development processes carried on regardless of the paradox.

It is also significant, certainly in the British context, possibly on a wider scale, that the post-World War II cave exploration 'boom' reached an acme, if not its eventual high point, during the nineteen-sixties. Just four examples of discoveries and studies which led to quantum leaps towards an understanding of cave genesis are provided by the explorations of the West Kingsdale System in the Yorkshire Dales (Brook and Crabtree, 1969), the Ogof Ffynnon Ddu (O'Reilly and others, 1969) and Dan yr Ogof (Coase and Judson, 1977) systems in South Wales and of the far reaches of Swildons Hole in the Mendip Hills. In each of these systems exploration and investigation still continue, but without doubt it was the success of new exploratory methods and initiatives during this time which pointed the way onward and indicated that during earlier exploration and study the obvious had frequently been overlooked.

During this same period there was a significant increase in the number of expeditions, mounted by cavers and cave scientists from the more 'developed' cave exploring countries, to visit and study limestone outcrops in remote or poorly documented areas. Whereas many of the earlier expeditions were of a purely exploratory nature, such as the British expeditions to Provatina and the Epos Chasm in Greece, in 1968 and 1969, many of their findings possessed great scientific interest. Others of the early ventures were designed to include a research component, such as the British expeditions to the Omalos Polje in Crete (1967) and to Ghar Parau (Iran) in 1971 and 1972. In the wake of these, and similar, steps into expeditionary cave study, have followed an ever increasing number of small and large expeditions from Britain, Europe in general, North America and Australia, during which many of the world's more inhospitable and some relatively pleasant limestone areas, have been subjected to rapid or intensive exploration and scientific study.

J Thrailkill

In 1968 John Thrailkill, a member of the Geology Department at the University of Kentucky, published an important paper, based upon the results of doctoral research, in which he considered hydrological and chemical factors which he believed to be pertinent to the excavation of limestone caves. By reference to a brief review of pre-existing theories of cave development [*"origin"*], mainly as understood in North America, he was able to pinpoint shortfalls of the then current understanding of speleogenesis.

A major element of the paper is an analysis of fluid flow through a simple pipe network. By considering the flow of water entering one corner of a network of equal sized pipes and discharging from a distant corner, he showed that rates of flow are similar in all pipe segments. In addition his results indicate that there is essentially no difference between laminar flow and turbulent flow within the limitations of the theoretical model.

A second element of his theoretical discussion (Thrailkill, 1968, p.29 et seq) involved the consideration of various chemical factors relevant to cavern excavation. *"Vadose seepage", "vadose flows", "temperature effect", "mixing effect"* and *"flowrate effect"* were examined in some detail, as well as *"other effects", which include the influence of geothermal gradient, solubility* variation due to hydrostatic pressure, and involvement of complex ions and/or oxidation-reduction reactions with or without organic influences. Generally, his conclusions on each of these topics are sound and several points of great importance were discussed. Many of these conclusions were based upon assessment of Thrailkill's own results and reassessment of those of earlier workers. It is not intended to examine the reasoning further here. However, the actual conclusions reached merit brief consideration and some of Thrailkill's terms, though possibly not strictly incorrect, deserve discussion.

The term *"vadose seepage"* was coined by Thrailkill (1968, p.29) to describe meteoric water which seeps slowly through the vadose zone, beneath the soil zone but above the water-table. His term is described as including water held by surface tension, that undergoing slow seepage and that forming discrete flows. The limits between these types, or indeed their precise natures, are not described. Thrailkill's examination of such seepage water was essentially on the basis of chemical changes (mainly within the calcium carbonate - water - carbon dioxide system). What he was examining, in effect, were the types of flow previously examined by Weyl (1958), whereby the water in small channels rapidly becomes saturated with respect to calcite and is thus theoretically incapable of further cave development activity. Additional consideration is given to the question of the behaviour of such water on encountering either an unflooded cave passage or the water-table.

There appears to be nothing to contradict in the reasoning and conclusions reached at this point, though the points discussed by Thrailkill in subsequent sections might be thought to bear upon the views presented. Additionally it is arguable that the term *"vadose seepage"* is unfortunate and misleading in this context. Although the flows penetrate the supposed vadose zone, their nature is more typical of water movement within the phreatic zone. Individual seepage routes are effectively flooded. Such seeps could be considered to be upward and lateral ramifications of the water-table or, conversely, the water-table cannot be envisaged as a uniform, planar surface. This distinction does not alter the conclusions of Thrailkill's arguments, but it does illustrate a weakness of the hydrographic zone concept which is made clearer still by his distinction of *"vadose flows"*.

Thrailkill's *"vadose flows"* (1968, p.31) are described as, *"Any continuous flow ranging from a small trickle to a large underground stream (above the water table)..."*. The lower size limit, the threshold between a seep and a flow, is thus not clearly defined, nor is it apparent whether the smallest *"flows"* (at the threshold) fill their conduit or whether they are envisaged as having an air surface. The former possibility seems more likely in view of the comment (1968, p.31) that, *"...some* [vadose flows] *are merely the accumulation of vadose seepage..."*. He goes on to describe the larger element of his *"vadose flows"* as either captured surface streams or concentrated surface/soil drainage, such as from the bottom of a sinkhole. The former type is perhaps representative of the general view of a vadose flow - a cave stream flowing in a passage which has air space (for the most part). He points out that the chemistry of such flows had been little studied but that they are generally undersaturated with respect to calcite. He then contrasts this undersaturation with the saturation or supersaturation of vadose seeps, making the obvious observation that this is probably a reflection of their contrasting origin, volume and speed of flow.

Thrailkill's consideration of the *"temperature effect"* (1968, pp 31-32) is purely theoretical and considers the effects of a drop in temperature upon saturated or supersaturated (with respect to calcite) solutions within seeps and flows as they encounter the water-table or vadose zone caves. No firm conclusions are reached, though he points out the theoretical possibility that under favourable conditions more calcite might be dissolved during short periods of slight undersaturation than would be deposited during long periods of considerable supersaturation. This being so, it is not demonstrated that vadose seeps, which can only be considered as ramifications of the phreatic zone, undergo any significant temperature depression at the level of the apparent water-table. More logically a temperature continuum ought to exist between the (supposed) reservoir within the phreatic zone and the points of input of the various seeps. His comment (p.32) that, *"...cooling at the water table is even more likely in the case of vadose flows."*, is a truism.

By considering *"mixing effects"* (1968, pp 32-34) Thrailkill is following the theoretical line of thought begun by Bögli (1964a,b). His conclusions are that although an increase in aggressiveness due to mixing of vadose and phreatic waters is theoretically possible, its potential would be reduced or nullified if one component (especially if from a vadose seep) was supersaturated. Only under optimum conditions would significant undersaturation result. He then argues (p.34) that, *"If the groundwater were not derived from vadose seepage, the mixing effect might be a more significant mechanism..."*. In these views Thrailkill might have become somewhat tangled in his own nomenclature insofar as his *"vadose seeps"* would generally be expected to be **more** saturated, if not actually supersaturated, than the sub-water-table water in the phreatic zone, the latter being derived more from vadose streams than from seepage input. An alternative mixing environment, where vadose seeps meet a vadose streamway with an air surface, is more realistically viewed as phreatic water (the seep) meeting the vadose water of the cave stream.

In discussing *"flowrate effects"* (1968, pp34-36) Thrailkill draws upon earlier published data (U S Geological Survey, 1960, 1962) for the Green River, which flows across an essentially limestone surface, near Mammoth Cave in Kentucky. Briefly, after allowing for the fact that the published data were based upon laboratory rather than field measurements and making the assumption that river water is in equilibrium with the carbon dioxide partial pressure of the normal atmosphere, he showed that the river would be supersaturated with respect to calcite at low flowrates and undersaturated at high flowrates. The conclusion, supported by a mass of chemical data and discussion, is what would be expected intuitively upon consideration of the changing parameters of water volume and velocity between the two flow states.

The significance of this undersaturation at high flowrates is, according to Thrailkill, twofold. Firstly, in the event of a stream sinking during flood conditions, the water would be able to dissolve large amounts of calcite in the vadose zone and the upper part of the phreatic zone. Possibly, depending upon the equilibrium conditions of the incoming vadose and *in-situ* water-table phreatic waters with respect to carbon dioxide, mixing effects might allow additional dissolution. Temperature effects at surface, in the vadose zone and at the vadose/phreatic interface, might also contribute to an increased capacity for dissolution during flood conditions. Secondly, Thrailkill points out that during flood activity surface water may be forced back into the aquifer from resurgence level, influencing the composition of the groundwater body. In this situation the chemistry for direct dissolution by the aggressive surface water would be slight, but dissolution will occur due to the effects of temperature and mixing.

In a short section entitled *"other effects"* (1968, p.36), Thrailkill discusses additional ways in which limestone groundwater might theoretically become undersaturated. Among his other statements is the somewhat surprising conclusion that large and rapid flows within limestone aquifers *"swamp out"* the effects of regional geothermal gradients, to the extent that there is no temperature increase with depth and no concomitant decrease in calcite solubility. Whether he is correct is uncertain in view of limited available data, but it would appear that there is room for doubt, particularly within deep aquifer systems.

The effects of hydrostatic pressure upon calcite solubility are examined briefly with the conclusion that theoretically possible pressure rises will cause negligible additional dissolution when compared to those of effects discussed above. His consideration of other ions (notably Mg^{+2} and SO_4^{-2}) and the potential effects of living organisms is still briefer. Both are seen as potential contributors to increased undersaturation but evidence for their quantitative significance was (in 1968) lacking.

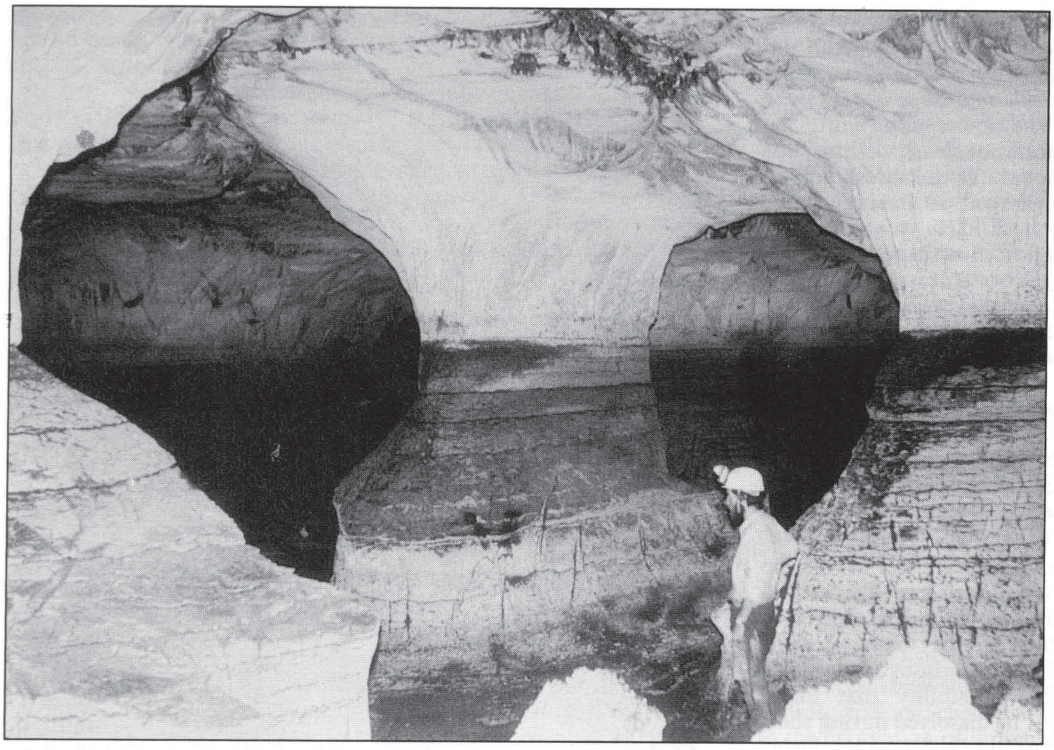
Within the remainder of his paper, Thrailkill concentrates mainly upon a discussion of the patterns of water flow in the phreatic zone. Examples of aquifer systems with uniform surface input, single discrete input and multiple discrete sources are discussed and flow path diagrams derived. On the basis of the theoretical considerations and calculations outlined above he provided a degree of confirmation that in the case of a moderately uniform input of water to a carbonate aquifer deeply penetrative curving flow lines would be generated. These deep flow lines are of great importance to the concepts of speleogenesis outlined by Worthington (1991) and Lowe (1992). Flow lines at shallower depths which are predicted and described by Thrailkill (and by many earlier and later workers) are also considered later in the paper. What must be examined both within and outside the context of Thrailkill's work is the relationship between geological factors and flow lines. This must pay particular attention to whether flow patterns such as those predicted by Thrailkill (and hence a spectrum of associated underground conduits) will develop regardless of geological constraints or whether the geological factors will regiment the geometry of the flow lines (and associated caves). In this context it is suggested here that at this level of resolution the simplistic two-dimensional view of flow lines, as used in most models, is inadequate to describe fully the apparent relationships.

A N Palmer

"Passages may form a two-dimensional pattern that is confined to favourable partings or geologic horizons, or a three-dimensional pattern that follows no single geologic structure."
(Palmer, 1975, p.57)

The important contributions made by Arthur Palmer were not examined within the review section of Lowe's (1992) work, since they were considered in some detail elsewhere in the thesis. His work concerning the origins of maze caves (Palmer, 1975) is probably most widely known and contains ideas which have received a widespread acceptance in the speleological world. Taking a broad view of Palmer's ideas there is little which can be criticised; his theories of maze cave development grew out of and added to the established 'wisdom' of cave formation. Inevitably, elements of this earlier knowledge have since been shown to fall short of reality. The need for and the nature of inception processes and inception guidance, together with a radical view of the timescales of speleogenesis suggested by Lowe (1992), necessitate re-examination of Palmer's ideas. Lateral consideration allows for different emphases to be placed upon some of Palmer's interpretations. Concurrently, some of the relationships reported by Palmer add qualitative support to Lowe's concepts of speleo-inception.

Palmer discussed three broad types of maze cave morphology: anastomotic, network and spongework. The caves were deduced as forming due to carbonic acid dissolution in a number of specific geological/hydrological situations, with strong structural *"control"*, which is equivalent to the structural 'guidance' discussed by Lowe (1992). *"Control"* by joints and (less commonly) faults produces a system of inter-connected loops in two dimensions which are generally constrained in the third dimension by stratigraphical *"controls"*. Palmer's view that three-dimensional maze systems follow no specific geological *"control"* (see quotation which begins



In Sof Omar Cave, Ethiopia, the splendid maze of passages is formed at one horizon in the limestone sequence (Photo: Steve Worthington)

this section) is not necessarily valid if the offsetting of inception horizons by faults, and or simultaneous inception horizon development are realities.

Much of Palmer's reasoning was based upon a quantitative, mathematical approach, which supports the development mechanisms suggested within his model by disallowing potential alternatives. Two over-riding considerations which are stressed and discussed within Lowe's (1992) thesis at least partially negate the essentially mathematical support of Palmer's views, though mathematical counter arguments cannot yet be presented. Palmer's models assumed that the maze caves under discussion have formed totally within the context of modern (currently visible) topography and its interaction with the geology. The models also assumed that all dissolution has been by means of the carbonic acid reaction under conditions of 'normal' hydraulic flow, paying no heed to the requirements of speleo-inception and its specialised driving mechanisms and chemistry. Since a short paper examining the possible origins of maze caves in the context of inception horizon theory and extended speleogenetic timescales is planned for the future, Palmer's maze cave theories are not discussed in detail here.

More recently Palmer has contributed to or produced other important papers, most notably Bakalowicz and others (1987) and Palmer (1991). The former paper presents a theory of maze cave formation due to dissolution by thermal waters within the Black Hills area of South Dakota. Lateral views of Palmer's earlier maze cave work which were presented by Lowe (1992) appear to be equally applicable to this more specific case, which has much in common with Lowe's (1992, chapter 9) overview of cave inception and regional hydraulic gradients within buried carbonate sequences prior to the superimposition of modern topography.

"It is reasonable to assume that most sandstones have greater potential as a path for recharge to the caves than as a source of confinement."
(Palmer, 1975, p.60)

Palmer's 1991 paper reviews the origins and morphology of limestone caves and draws heavily upon his own earlier work and the then currently accepted views of carbonate speleogenesis. More than in his earlier work and to a greater extent than most other workers to that time Palmer shows an awareness of the requirements of speleo-inception, though the review is more concerned with cave development than *"initiation"* (Palmer's term). Throughout the text it appears that he is aware of the presence of inception routes, though he does not refer to them by this name. There is a tacit assumption of mechanisms which will allow water movement and dissolution within buried sequences, but the concept is developed more in the context of maze cave formation than as a general inception process. Other elements of the paper are discussed by Lowe (1992. section 4.5.6) and need not be reiterated here, with the exception of Palmer's contention that:

"Sulfuric acid is also produced by oxidation of pyrite, but the process is usually too slow and the pyrite too dispersed to create more than local pores."
(Palmer, 1991, p.18)

This statement conflicts with the general philosophy of Lowe's inception horizon thesis, but possibly only in degree. It may reflect a failure to recognise that whereas pyrite is generally dispersed, it is also locally concentrated. It is this concentration, in bedding-related zones (inception horizons), linked to the infinitessimally slow progress attributed to the earliest, diffuse flow, phases of speleogenesis which cannot be ignored. These considerations magnify the importance of pyrite (or other sulphides and/or sulphates) beyond that to be expected on the basis of general pyrite levels and normally accepted, post-breakthrough, speleogenetic timescales.

A C Waltham

"Cavernous limestone karst ... is the terrain type whose geomorphological understanding becomes an art, matured only through experience, and then applied sensitively and individually at each new site assessment."
(Waltham, 1989, p.1)

A C (Tony) Waltham has made a broad and significant impression upon cave research during the past twenty years. He has originated and developed a number of ideas which are applicable not only to the mainstream of speleogenetic thinking but also to several areas of applied karst study, and he has managed to bring geological factors to the forefront of cave formational theory.

Perhaps his most significant contribution to the general understanding of speleogenesis was his recognition and acknowledgement of the crucial role played by geological influences in guiding cave development (Waltham, 1970, 1971a, 1971b), an awareness which has been maintained in more recent publications. Waltham's 1970 paper, though appearing to be his first significant writing in this context, is shown by its reference list to postdate the completion, if not the publication, of two other important contributions (Waltham, 1971a, 1974). Thus, similar ideas are expressed within these three publications and another short symposium contribution (Waltham, 1971b). Several of the concepts raised or discussed in these publications were innovative; some have stood the test of time with little need for modification, others now appear dated or somewhat naïve.

Much of the criticism which can be levelled at these earlier publications is of a semantic rather than factual nature - for instance a questioning of the repeated use of the terms *"shale"* and *"control"*. In the case of the former term its use might be defended in that it was readily understood by a broad spectrum of readers, but it can equally be criticised in that it masked the true nature of

In the Lost Johns Master Cave, beneath Leck Fell in the Yorkshire Dales Karst, where the roof tube is guided by a single bedding plane for its entire length. (Photo: Tony Waltham)

certain inter-carbonate clastic/organic horizons that were being discussed. The term **guidance** is preferable to **control** when discussing geological (or other) influences upon speleogenesis. It is also noteworthy that Waltham uses the term *"initiation"* (eg 1971b, p.78) in a context which is probably close in meaning to **inception** as described by Lowe (1992).

A number of elements of Waltham's early work are now less readily acceptable. For instance the idea that, *"Detailed morphological studies ... show that various cave systems include examples of both phreatic and vadose origin."* (Waltham, 1970, p.574) would need to be couched in somewhat different terms in a more modern context. The caveat that, *"...initial phreatic phase ... is tolerated within the definition of a vadose cave."* (Waltham, 1970, p.580) is all too easily overlooked and, in any case, obscures a highly relevant set of relationships. Likewise his figure 2 (1970, p.576) which purports to show passage cross-sections which distinguish between phreatic and vadose *"origins"* can be seen to indicate phreatic origins without or with vadose modification - and possibly much more besides. In another paper (Waltham 1971b, p.75) he refers to, *"...downward flowing vadose waters forming initial cave openings immediately above impermeable horizons."*, a concept which is in some ways contrary to views of inception and of the role of such horizons presented by Lowe (1992).

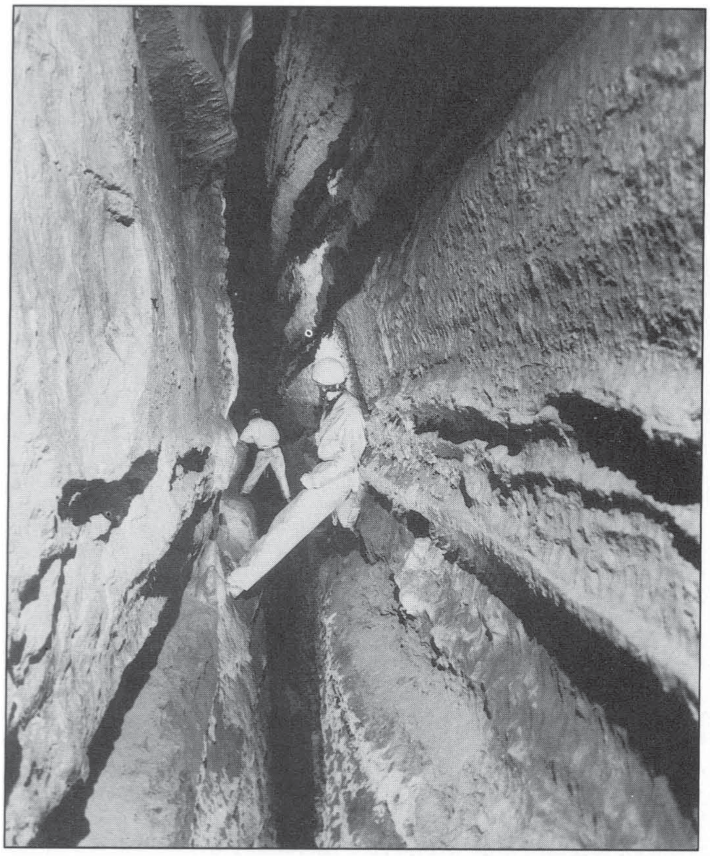
Waltham's emphasis upon the importance of geological factors was (and remains) highly significant, coming at a time when such relationships were overlooked or ignored in the light of apparently concrete analytical data of one sort or another. His specific identification (if not fully correct terminology) of the importance of clastic and other non-carbonate beds, joint or fault discontinuities and synclinal axes as guides to cave development is as pertinent now as it was then. Of similar importance was his emphasis (1970, p.574), following on from the work of Drew (1966), that there is no water-table, in the classical sense, within a cavernous limestone sequence. The lack of a water-table as classically understood or the presence of an apparent water-table within rocks of this type is of vital importance to some more modern theories of cave development.

It is of interest to note that his consideration of *"shale"* partings within the Carboniferous limestone sequence of the Yorkshire Dales (Waltham, 1971a), a paper of fundamental importance to any argument involving cave inception, is one of very few 'cave generated' publications to be accepted by a major British geological journal. Most of the outstanding North American contributions to cave development theory have appeared within reputable geological journals, but in Britain there has been a reticence by editorial panels even to consider such contributions. Of necessity British cave scientists have published valuable geological results overseas, in limited-circulation, cave-orientated journals or within a variety of geographical or geomorphological reviews. Thus the British geological establishment remains largely ignorant not only concerning the actual mechanisms of cave development and underground water movement, but also of the wealth of structural and stratigraphical data which might be gathered by serious and well directed hypogean study.

In his consideration of the caves of Leck Fell (Waltham, 1974), Waltham showed beyond reasonable doubt that, even in seemingly complex cave systems, detailed and careful study will reveal the basic guiding influences of stratigraphy and geological structure. His development model for the several caves which form the Lost Johns System broke new ground in British cave studies, matched at the time only by the apparent sophistication of D C Ford and T C Atkinson's models for the caves of the Mendip Hills. Ford's and Atkinson's models can now be seen to have lacked a degree of realism that could have been, at least partially, added by closer examination of geological detail. Waltham's study, published some years later, included many aspects of the 'Bristol School' philosophy, but achieved a closer approximation to reality, as a result of its view of geological parameters.

The model of cave system development on Leck Fell represented a state of the art view of development chronology within the semi-rigid framework set by the geological constraints. Almost certainly its overall validity could still be argued in relative terms, though it now seems likely that the phases of development might be viewed differently in the context of absolute chronology. The several phases of formation identified and discussed by Waltham could be considered as a quasi-continuous series of development events during which pre-existing, geologically pre-determined, incipient drains were sequentially modified. Following dissolutional enlargement, which became more significant as each particular incipient route approached the top of the contemporary phreatic zone, morphological modification would continue by both dissolution and corrasion under vadose conditions. Complex patterns of short circuiting and re-invasion, superimposed upon a more predictable development sequence, add to the problems of interpretation.

Thus, although it is reasonable to accept the Waltham development phase model in terms of the relative ages of major passage growth and the changing roles played by individual drains in the overall development picture, it is far from certain that this relative timescale has been adequately pegged to points on an absolute timescale of cave formation. The inception of passage formation at several horizons within the speleogenic sequence could have been essentially coeval, the levels at which inception occurred being defined by stratigraphically guided chemical factors



The Traverses in Ogof Ffynnon Ddu, South Wales, where the ledges reflect subtle variations in the limestone lithology. (Photo: Jerry Wooldridge).

and linked by structures such as joints and faults. Later processes, which only began to occur as relative uplift took place, would involve the superimposition of secondary factors, such as local hydraulic gradient, upon the inception horizons and the beginnings of the establishment of true conduit flow utilising the most efficient combination of inception routes and their linking fractures. Thus, although it is possibly still acceptable that development phases within this, or other, cave systems might be tied to absolute periods of time (such as interglacial stages), it must also be accepted that this linking is probably of periods of modification, not of the actual origination **(inception)** of the particular passage system.

H W Rauch

"Pyrite is a potentially important lithological variable which is often overlooked by karst researchers."
(Rauch, 1972, p.21.)

The research thesis submitted to Pennsylvania State University by Henry William Rauch in 1972 arguably represents one of the major benchmarks in the analytical approach to speleogenesis which typified the Scientific Period. The project set out to identify the links, if any, between lithology and cave formation. His analytical results, and the presentation of them, are outstanding examples of scientific method and objectivity, but his experimental parameters, preliminary assumptions and some of his conclusions must now be questioned in the light of more recent ideas. Had his assumptions been less rigid and his experiments the same or, conversely, if his experiments had more broadly interrogated the rigid basic assumptions, somewhat different conclusions might have been suggested. However, in Rauch's favour, it must be said that, even today, only a limited amount of experimental work has been targeted along the lines he might have followed, to test the parameters that he took for granted. It is impossible to reconsider the total philosophy and stated conclusions of Rauch's research within the limited confines of one small part of the present review. Generalized comments concerning the programme of research, the way the results were examined and interpreted and the conclusions that were reached are, however, of relevance to this overview.

It is not intended here to examine in detail Rauch's laboratory method or the statistical basis upon which many of his conclusions depend. The detailed descriptions of his experimental technique and his considerations of various potential error sources indicate a punctilious care and attention to detail. However, as previously stated, it might be argued that, both the experimental and statistical approaches are based upon insecure assumptions.

Briefly, the experiments described are designed to compare the solubilities of individual carbonate samples of various discrete lithologies affected by a steady flow of solvent saturated with carbon dioxide and flowing through drilled holes of sub-conduit diameter. There can be no doubt that the meticulous work produced results valid for a system of this type, but it is doubtful that such a system could appear naturally as a result of the mechanism being tested. Thus the elegant experiment falls down, just as earlier dissolutional modelling experiments using hydrochloric acid and carbonate or water-soluble materials and water fell down, by not adequately reproducing conditions which might realistically be expected in nature. The experiment ignored the problems of inception, the potentially infinitesimal flow in an inception system and the inevitable saturation and concomitant lack of aggressiveness of the supposed solvent with respect to carbonate. Further, having made the gross assumption that tubes of sub-conduit diameter come into existence by an effectively unexplained mechanism and that these tubes may be enlarged by solvent water made aggressive by a limitless supply of carbon dioxide, the additional fatal assumption is made that in nature the tubes would be bored through massive carbonate rock of uniform lithology, as if by a worm digging tunnels through the soil.

In fairness, Rauch's ancillary discussions of field observations indicate an awareness that such is not the reality of cavern development, but these points are not made clear in the experimental descriptions and discussions. It is argued here that to be realistic and applicable to real situations the experiments might have been carried out in a similar manner but examining the effect of passing an essentially non-aggressive 'solvent' along micro-bores at or adjacent to junctions between contrasting 'pure' and 'impure' lithologies. The many practical difficulties of putting such a programme into effect are acknowledged.

The fallacious nature of the statistical consideration hinges upon the obvious impossibility of quantifying 'true' lengths and volumes of cave passage within a rock mass; to do so within individual carbonate formations and members is still less practicable. There can be only a consideration of those passages which have been located and penetrated; no allowance can be made for undiscovered or impassable passages, nor for passages yet beneath the water-table. An argument that these limitations apply equally to all formations and may thus be ignored is also fallacious. Simplistically, one formation may support a limited number of passages all of which are explorable and able to be measured; an adjacent formation might hold no explorable passages but include a greater length and volume of unenterable voids. Elsewhere, blocks of various rock units may appear barren of caves due to a lack of connections to the surface producing recognisable and passable entrances. The entire view of cave statistics which is exemplified by Rauch's treatment and which had been championed by Curl (eg 1958, 1963, 1964 and 1965) and has recently been revitalised by White (1988, p.61 *et seq*) is unrealistic.

Having dwelt on the point that Rauch's approach may not have been totally valid to a realistic consideration of speleogenesis it must be restated that his experimental and observational results are prodigious. These results, linked with an apparent understanding of chemistry which is far beyond the grasp of the present author, combine to produce a number of important conclusions. Some of these are capable of lateral interpretation and are considered below.

Rauch's finding that dissolution rates decrease with increasing impurity content within a carbonate rock must be accepted on the basis of the weight of experimental results, but only in the context of the unrealistic experimental parameters outlined above. It must be asked whether the same relationship would be manifest if the 'solvent' was essentially non-aggressive rather than constantly replenished with carbon dioxide.

Dissolution rate and cave volume are reported to increase as micrite grainsize decreases. A complementary statement indicates that a decrease in cave volume was linked to an increase in sparite content. The chosen explanation for the relationships above is that a greater length of inter-crystalline boundary and hence exposed surface area of carbonate is *probably* present in the more finely crystalline rock types. Although this conclusion might stand up to examination it appears to imply a great importance for inter-crystalline porosity, which is not necessarily encountered in reality. Instead, the different susceptibilities of micrite- and sparite-dominated lithologies to dissolution could be a reflection of their depositional environment and/or history of recrystallization. From another viewpoint it is also possible that differences in strength or rigidity, developed subsequent to diagenesis, and a consequently different pattern and intensity of fracture development, may be involved in the relationships being considered.

A possibly linked conclusion deals with the effect of dolomite content upon dissolutional rates. Broadly, values of 2 to 3% MgO (equivalent to 9 to 14% dolomite) were linked to the maximum rates of dissolution. Rauch equates this initial rise (with increased dolomite content) in dissolution to an increased abundance of *"silty streaks"* and hence an increase of bedding partings and exposed surface area for dissolution. This may be partially correct (though *"silty streaks"* cannot be assumed necessarily to provide access for solvents to exposed surfaces). It again seems possible that this value for optimum dolomite content reflects a depositional environment for rocks with primary dolomite. Such rocks might be considered to exhibit some aspect of their chemistry, not necessarily the dolomite itself, which is not represented in purer, calcitic, lithologies nor in secondarily produced dolostones or dolomitic limestones of greater than 14% $MgCO_3$.

Rauch's conclusions regarding the role of *"silty streaks"* merge into his views dealing with the importance of *"major and minor bedding partings"* (p.367). He reports that *"silty streaks"* in carbonate rocks may be preferentially attacked during dissolution experiments and that in nature both dissolution rate and cave volume generally increase in parallel with an increase in thickness and number of *"silty streaks"*. This direct relationship is attributed to content of clay, dolomite rhombs and/or high-magnesium calcite, together with their associated bedding partings and high intergranular porosity/permeability.

He states (p.366) that *"initiation of cave development is retarded by the thick layers of impure material (either shale bands or silty streaks over about 5mm thick) which usually surround ... major partings."* This statement in isolation is highly contentious and contradicts the results of such workers as Waltham (eg 1971a, b). A directly contrary view is expressed by Lowe (1992), who suggests that many inception routes are guided by such thick, impure layers. Broadly, however, some of Rauch's views on *"silty streaks"* and linked phenomena may be valid or partially valid. Contributions made by various chemical alterations of these argillaceous layers must also be considered.

In his final lithological conclusions Rauch states (p.367) that, *"Pyrite ..... does not significantly affect cave development."*. To this he adds the caveat that, *"Pyrite may be responsible for initial vug development in dolomites and some limestones...."* and that, *"Pyrite could act to increase the carbonate solution rate by generation of sulfuric acid...."*. Views expressed by Lowe (1992) contradict Rauch's first statement (above), by emphasising the very strong indication that many caves owe their inception to the presence of pyrite. Rather than pyrite not affecting cave development, in these cases there could be *no* development without its intervention. The validity of Rauch's caveat is, thus, also clear.

L Jakucs

Professor László Jakucs had been director of the Institute of Physical Geography at the József Attila University at Szeged, Hungary, since the early nineteen-sixties when, in 1977, a revised and enlarged English translation of his text, *"A karsztok morfogenetikája"* Morphogenetics of karst regions] was published. Having been an active speleologist since his youth and having made a considerable reputation as a karst investigator by the time he reached his mid-twenties, his book is based upon a wealth of personal expertise. Of more importance, the contents of the book, though they draw heavily upon published material, show an approach which is sufficiently mature to reconsider not only his own early ideas, but also some of the ingrained 'truths' of traditional karst geomorphology. Much of the previously published work that he cites is of eastern European origin and many of the papers considered do not appear within the reference lists of similar texts originating in the West; indeed most such works have probably not appeared in English translation.

There is much of interest to the karst hydrologist and geomorphologist in the book, but much of the content is of no relevance to this review. The general interest and quality of the book being accepted, it is for Jakucs' broader views of and observations upon cave development that he is included here. These views, already broader than those in most comparable works, encourage further lateral thought and are illustrated by reference to quotations from his book.

"This short chapter is insufficient to exhaust all the potential factors that may affect non-karstic corrosion of limestone in nature......... some of the fundamental axioms of karst evolution will have to be scrapped."
(Jakucs, 1977, p.53)

Within a chapter entitled *"The concept of karst corrosion"*, Jakucs examines what might be termed the traditional views of carbonate dissolution. The effects of pure water in physically dissolving calcium carbonate are briefly discussed, but it is a discussion of the role of carbon dioxide in the process which Jakucs terms *"hydrocarbonate dissolution"* (Jakucs, 1977, p.28 *et seq*) which forms the bulk of the chapter. The term *"hydrocarbonate dissolution"* is taken here to be a quirk of translation and to be synonymous with the process of dissolution by carbonic acid. Within this section, Jakucs considers the dissolution of calcium carbonate in terms of two (gas+liquid) and three (gas + liquid + solid) phases and is particularly at pains to emphasise that most of the carbon dioxide within natural precipitation derives from the soil rather than from the atmosphere. In the same section he considers the effects of temperature, atmospheric pressure and hydrostatic pressure upon the processes of dissolution, both in terms of direct influences and the perturbations due to mixing and/or changes of physical conditions. Following on from this is a consideration of the concept of mixture corrosion (Bögli, 1964a,b) and the supposed time requirements for the steps involved in *"hydrocarbonate dissolution"*. The latter discussion hinges upon Bögli's (1960) research in which he identified four stages within the dissolutional process. This need not be re-examined here, except to say that Jakucs concludes (p.48) that, *"...even if we retain Bögli's four stages for its illustrative and didactic merits, we must keep in mind that the individual stages in nature never occur in succession, but invariably simultaneously and in mutual interplay: that is, the stages as such do not in fact exist"*.

Of greater importance is the latter part of Jakucs' dissolution chapter, which is subtitled *"Non-karstic corrosion of limestone"*. Within this short section (Jakucs, 1977, pp.48-53) he describes numerous alternative mechanisms by which the chemical decomposition of limestone can be achieved. The organic products of humification are considered briefly, but the main emphasis is placed upon the role of inorganic acids, such as sulphuric and nitric acids, which are produced by a wide variety of oxidation reactions, with or without the intervention of bacteria. Additional non-karstic corrosional processes, possibly of lesser importance, rely upon complex exchange reactions between calcium carbonate and other salts or upon the solubility of phosphates and ammonium salts.

"..limestone corrosion is wrought not only by the carbonic acid content of water, but just as effectively by other organic and inorganic acids and other compounds..."
(Jakucs, 1977, pp.105-106)

Jakucs considers the influence of limestone composition upon karst development in a chapter entitled *"Petrovariance as a factor of karst corrosion"*. Having made the potentially crucial comment (1977, p.55) that, *"Accessories and contaminants exert a considerable influence on the behaviour of limestone under corrosion."*, he goes into the detail of how and why impurities are *inimical* to cave formation. Much of this discussion concerns the effects of magnesium carbonate upon the solubility of calcium carbonate. From a review of existing European literature on the subject he concludes that the solubility of calcium carbonate, or limestone, is substantially reduced by even a minimal amount of magnesium carbonate (p.60). This conclusion is at variance with that reached by Rauch (discussed above). Within the same section Jakucs repeats a commonly held view (p.55) that, *"The various deposits and fillings of caverns are also composed largely of these insoluble residues"*. This concept, most commonly associated with the work of Bretz (1942) is reported by Jakucs as being supported by a number of European workers, including Bögli. The present author is doubtful whether such a process can contribute *"largely"* to cavern deposits, at least in the way implied by Jakucs (and the earlier workers). Notable by its absence in this section of Jakucs' text is any consideration of how certain impurities might be **advantageous** to carbonate dissolution, an omission which is surprising in view of his discussion of *"non-karstic corrosion"* in the previous chapter.

"Permeability in the limestone is one of the most fundamental prerequisites of karstification."
(Jakucs, 1977, p.70)

A subsequent section of this same chapter (pp.68-71) deals with *"Influence of the lithology limestone structure upon karst corrosion".* Following on from the significant quotation reproduced directly above, Jakucs points out that Quaternary limestones are *"most favoured"* (p.70) due to having a high degree of textural porosity rather than any excess of structural fissuration. This statement is probably true, though Jakucs appears to have been unaware of earlier work on cave formational processes affecting Quaternary carbonates in the coastal environment and he makes no mention of the dissolutional implications of fresh/salt water mixing.

Putting aside the Quaternary rocks as a special case in view of their primary porosity, Jakucs states that the highest values of open fissure development are exhibited by Mesozoic limestones and that the most conspicuous karst phenomena occur in Triassic, Jurassic and Cretaceous rocks. This conclusion must be re-examined. He links the state of development to an assumed timescale of karstification which can no longer be supported. Later work has shown that speleogenesis and general karstification can reach a significant level of development within tens or hundreds of thousands (less in porous coastal zone deposits) rather than millions of years, so that this factor cannot reasonably be seen as relevant to a comparison of karstification upon rocks of different age. His conclusion is based upon data which are biassed towards the European situation, where Mesozoic rocks which have been relatively recently tectonised are the predominant host of surface and underground karst features. A parallel situation is apparent in the younger mountain chains, formed subsequently to those of Europe, where the karstified rocks are mainly of Tertiary age. Within the Palaeozoic rocks of the region best known to Jakucs there is a relative dearth of potentially suitable rock types, tectonised or otherwise, at outcrop or in suitable situations for sensible determination of their fissure density. On the other extreme, his Quaternary data refer only to tufaceous deposits.

"And since the bedding of limestones is, in a general way, connected with the fluctuations in deposition (primary fluctuations in pelite content), bedding-plane clefts let water pass precisely along those surfaces where the rock is most contaminated."
(Jakucs, 1977, p.71)

The final point to be considered in detail from Jakucs' text derives from the quotation above. As with other examples already discussed it is possible that Jakucs has intuitively grasped a point of great importance to the mechanisms of speleogenesis but has narrowly missed making a potentially vital lateral step. Having pointed out that water will inevitably be constrained by such bedding surfaces he deduces that in this situation dissolution residues will choke fissures (and presumably the bedding itself, though this is not stated) and thus be inimical to speleogenesis. From this he further deduces that well-bedded limestones are less prone to karstification than are thickly bedded, homogeneous limestones. Here again it might be said that he is correct in having identified and brought attention to two extremes of a continuum, but he has missed the potential importance of lithological interrelationships within that continuum. Additionally his consideration of two such extreme rock groups does not take into account any alternative constraints upon karstification which may be inherent in their litho-tectonic aspect. The whole question of the importance of bedding planes with their associated 'impurities' to speleo-inception within a generally massive rock sequence is overlooked, as are the potential links of these horizons with the agents of non-karstic erosion and general contaminants discussed above.

Although much of what is said above appears to be critical of certain aspects of Jakucs work, the implied criticism is slight and only to the degree that having recognised certain crucial elements in the arguments covering the full spectrum of karstification, he has commonly been swayed towards following the lines of a published *status quo* rather than considering possible lateral views which are suggested by his observations. This lack of major lateral adventure appears uncharacteristic of his otherwise outspoken and forthright approach to his studies, as typified by his consideration of Corbel's work (discussed above). The examples of his views examined briefly here derive from the theoretical, preliminary part of his text. Other aspects of this theoretical section and much of the descriptive, second part of the book are equally interesting and this text, together with those by Bögli (1980) and Dreybrodt (1988), provides a fascinating counterpoint to similar works by English-speaking authors.

D C Ford and R O Ewers

Many of Derek Ford's earlier ideas (mentioned briefly above) can be recognised in the paper that he produced with geologist Ralph Ewers in 1978. The major factor that differentiates this from Ford's earlier papers and qualifies it for a pivotal position between the Scientific Period and what was to follow, is the increased consideration of geological structure and stratigraphy, presumably reflecting Ewers' expertise. Though Ewers is best known for his contribution to the 'Ford-Ewers' speleogenetical model he has produced many other papers dealing with geologically linked aspects of cave formation and related topics.

There was more to the Ford and Ewers model than a rehash of earlier work with geological embellishment, though it might be said that the model's basic principle was advanced by Ford himself as early as 1965. According to Watson and White (1985, p.117), Ford and Ewers provided a seemingly ultimate answer to the hoary controversy of deep phreatic, shallow phreatic or vadose cave origins. Perhaps cynically, it might be said that a controversy is not actually answered by removing it, but that is exactly what Ford and Ewers accomplished. They produced the closest approach (to that date) to a 'general theory of cave formation', not by producing a single hypothesis which acknowledged all observations and fulfilled all theoretical requirements, but by effectively 'moving the traditional goalposts'. In short, they reiterated what Derek Ford had already said - that there was no single hypothesis and each of the three arms of the longstanding (deep phreatic, shallow phreatic or vadose cave development) controversy had equal validity. As pointed out by White (1988, p.294), *"... 'Do caves form above, at, or below the water table?' To all of these possibilities, the Ford-Ewers model answers, 'yes!'"*

Since its original appearance in 1978 the Ford-Ewers model has been examined, re-examined and quoted in detail by numerous workers, including the original authors. It is generally agreed to provide a comprehensive view of carbonate speleogenesis and yet, in reality, elements of the model, such as the fissure-frequency concept, remain relatively poorly tested hypotheses. The model will not be discussed in detail within this review; there is, indeed, little that can be criticised in terms of the broad picture of carbonate cave development which is provided. However, examination of the Ford-Ewers model, particularly its more recent restatements (White, 1988; Dreybrodt, 1988; Ford and Williams, 1989), indicates that scant regard is paid to the problems of inception of carbonate dissolution and gestation of sub-conduit underground drainage systems. Additionally, a proportion of the relationships and states of development figured by Ford and Ewers which are intended to illustrate close links between stratigraphy, structure and speleogenesis may be viewed as partially contradicting this supposed interdependence.

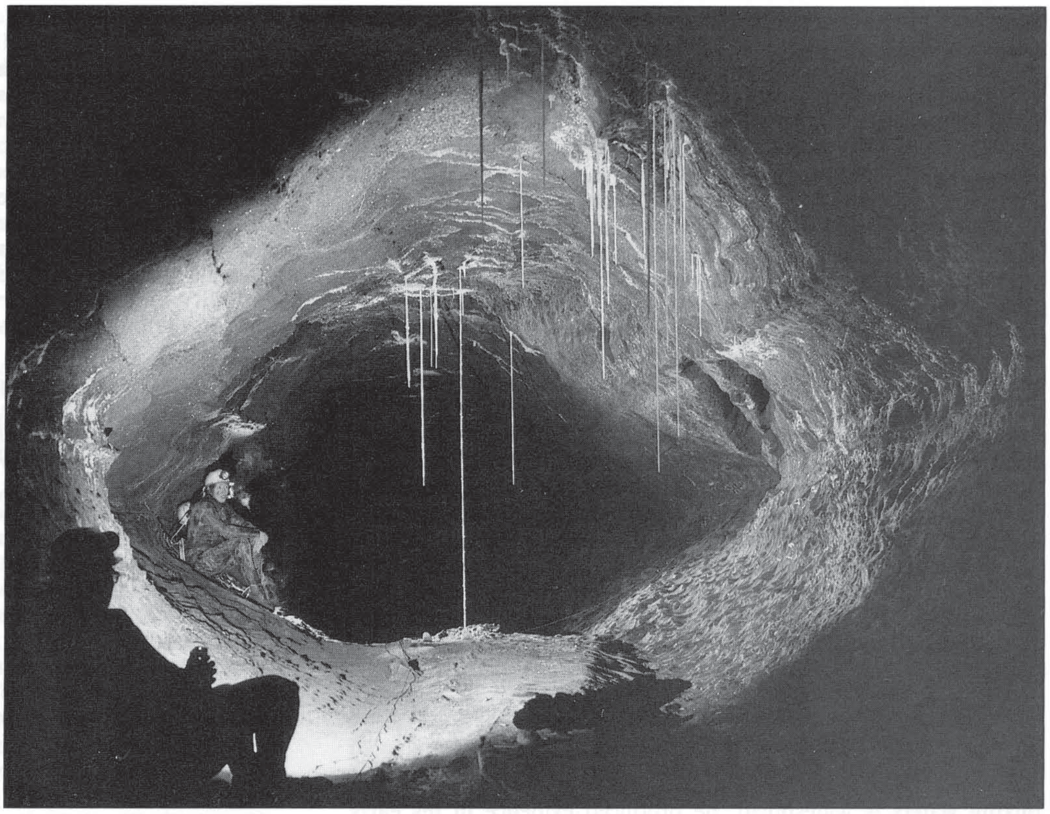
Examination of the requirements of inception and early gestation processes provides a possible prologue to the stages outlined by Ford and Ewers, without presenting any significant conflict with the main thrust of their arguments. However, consideration of their figures (and supporting text) in the light of lateral concepts (eg Worthington, 1991; Lowe, 1992) not only removes many of the potentially contradictive elements, but also supplies a degree of confirmation of the links which are essential to the model.

Other contributors to the Scientific Period

As pointed out above, enormous effort was expended during the Scientific Period in an attempt to refine pre-existing conceptions of cavern genesis and to produce an ultimate answer to the questions of how and why caves and related karst features develop. For many workers the perceived routes towards these ends lay in the amassing of analytical data (of several distinct types) or in mathematical modelling and 'complex' equations. Some studies were misdirected, based as they were upon concepts which were incomplete, if not partially incorrect. Much painstaking and diligent work was devalued and its importance obscured by attempts to use the results in support of the aged concept of carbonic acid dissolution. After Bögli's (1964a, b) acknowledgement that this concept was insupportable, emphasis shifted from the constrained view of carbonic acid dissolution to consideration of various processes of mixture corrosion. Earlier inconsistencies were partially explained, well enough that there was still no incentive to postulate any generally applicable alternative dissolutional mechanism.

However, it was during the Scientific Period that publications began to appear which dealt with supposedly *aberrant* forms of dissolutional cave development, which involved strong inorganic acids, particularly sulphuric acid. Comments concerning the

A splendid phreatic tube is Dan-yr-Ogof, South Wales, where the inception horizon is clearly recognisable from the enlarged profile (Photo: Paul Deakin).



possible role played by sulphuric acid were made, *en passant,* by authors such as T D Ford (1963), but an account written in Russian by Durov (1956), which was not published in English until 1979, provides the first real departure from traditional ideas about dissolution since karst studies began. Earliest among the English language workers to examine 'special' cases of cave formation by strong acids was Morehouse (1968), whose work was followed by the early research of Egemeier (1973).

An unusual contribution during this phase of speleogenetic investigation was made by Bedinger (1966). His work involved the development and study of an *"electric-analog"* model of cave formation. The significance of his findings is uncertain, any potentially valid model being applicable only on the most local of levels in 'real' conditions or more widely in idealised situations which probably do not exist in nature.

Palaeokarstic features, the identification and effects of which are potentially of great significance to concepts of speleogenesis, began to attract serious attention during the later part of the Scientific Period, though earlier work by Dixon (1910, 1921) and T D Ford (1952) anticipated much of what was to follow. Papers by Walkden (eg 1972a,b, 1974, 1977) provided descriptions of palaeokarsts, with particular reference to those developed within Dinantian carbonate sequences in Derbyshire. The close connections between palaeokarst horizons and volcanic activity, in this area at least, were pointed out, but at this stage there was little attention paid to any possible links with speleogenesis. Less widely known, and tenuously linked to the same subject, were short papers by Shelley (1967) and Turner (1968), which discussed thin coal seams within the higher part of the Dinantian carbonate succession of the Yorkshire Dales, including a seam exposed in Notts Pot and other caves in the neighbourhood of Leck Fell.

During the latter part of this same period groundwater chemistry and processes of dissolution in coastal environments began to be re-examined in greater detail and from a somewhat different viewpoint than had previously been applied by workers such as W M Davies, Swinnerton and Bretz. Early work by Back and Hanshaw, which dealt with this coastal situation as well as other controversial aspects, such as dolomitization and dissolution in general, was indicative of much which was to be written later. Insofar as most of these early ideas were later refined, the work of Back and Hanshaw (and others) is considered below as part of the output of the Modern Era. A paper by Ollier (1975) is possibly the first significant publication, other than the work of Back and Hanshaw, to deal with processes of cavern development in the 'freshwater lens' context. Though Ollier's treatment of the subject appears to be targeted very much upon compromising his own observations with the traditional cave formational models, much of the modern understanding of this specialised aspect of speleogenesis is included. Also of importance is a paper by Jennings (1968) which deals with supposed syngenetic cavern development in young, in some cases only partially consolidated, aeolian calcarenites. The timescales of speleogenesis implied by this supposed syngenetic development were, at the time of publication, considerably shorter than any described or hypothesised by other workers. Similar work on speleogenesis in aeolian calcarenites has since contributed to the conclusions reached by Back, Hanshaw and their co-workers, mentioned above.

THE MODERN ERA

As pointed out above, Ford and Ewers (1978) put an end to the fruitless argument which had recurred throughout the Golden Age and the Scientific Period concerning the relationships of cave formation to hydrographic zones. Not only did they propose a compromise which satisfied most of the earlier discussion, but they also began to pay (Ford and Ewers, 1978, p.1785) the smallest degree of *en passant* attention to other chemical reactions (possibly) of local importance to speleogenesis. As the nineteen-eighties progressed a number of new initiatives or avenues of understanding began to develop. Concepts which had previously been touched upon as relevant to speleogenesis, but had subsequently been ignored or forgotten, began to reappear in literature; publications which originated outside the perceived boundaries of cave research gave new insight into karst processes and timescales.

During the Modern Era, as during the preceding Scientific Period, the amount of published and, it is assumed, unpublished work relevant to cavern genesis and related processes has increased dramatically. To choose the work of any individual or group of workers as being of outstanding value in the context of the era is difficult, more so since many of those involved remain active in their research field and undoubtedly they have much still to contribute. There have been considerable and valuable overviews of karst science in general, such as those by Bögli (1980), Jennings (1985), Trudgill (1985), White (1988), Dreybrodt (1988) and Ford and Williams (1989). All of these authors have drawn upon personal expertise reaching back into or beyond the Scientific Period, as well as utilising a wealth of published data reflected in their significantly personal reference lists. These texts succeed in presenting the 'state of the art' view of karst science and, as might be expected, different viewpoints are apparent between early and later works or between contemporary works by authors who began with geological, geomorphological or chemical backgrounds.

Within the same period there have been a number of books published which are less obviously centralised upon karst studies *per se,* but which include either chapters or contributions on cave related topics or contain data of vital importance to a modern understanding of speleogenesis. Those by Scoffin (1987) and James and Choquette (editors, 1988) are of particular interest. About half of a symposium volume edited by LaFleur (1984) is concerned with the activity of groundwater in carbonate terrains and several of the contributions provide early insights into concepts which would be developed later in the decade. Similarly, a number of chapters within *"Solute Processes"*, edited by Trudgill (1986), are of great value for their collation, amplification and explanation of a mass of extant and, in some cases, contradictory data.

The importance to speleogenesis of such aspects as palaeokarsts and palaeo-caverns, strong acid dissolution, stratigraphical guidance and freshwater/saltwater interface dissolution has become more apparent during the nineteen-eighties, though the full or correct significance has not been immediately obvious to all of the workers involved. An attempt is made below to consider, briefly, the significant steps away from pre-existing concepts made by authors such as W Back, B B Hanshaw, J S Herman, C A Hill, E H Kastning, J E Mylroie, R A L Osborne, P L Smart and their several co-workers. Much of their research is on-going and other aspects are yet to be investigated as new clues to the significance of earlier observations are noted. Additionally a number of less obvious contributions, previously considered of only peripheral importance, will be mentioned and their possibly wider implications noted.

W Dreybrodt

Wolfgang Dreybrodt has successfully taken up the obvious challenge presented by the broad postulates of Bögli's mixture corrosion thesis and has produced convincing evidence of the theoretical importance of groundwater mixing to carbonate dissolution. Not only has he shown that corrosion as a result of mixing waters is widespread, he produced evidence in his early work (1981) which appeared to demonstrate that no alternative mechanism nor chemistry is required to account for speleo-inception. The validity of this standpoint in the light of some of Dreybrodt's own later work and more recent views such as those of Worthington (1991) and Lowe (1992) must be re-assessed, however, and Dreybrodt himself points out:

"Little is known of the initial phase of karstification since direct observation is not possible. It is inferred, however, that a system of penetrable primary fissures must exist....."
(Dreybrodt, 1988, p.6)

The brief quotation above forms part of the Introduction to *Processes in Karst Systems,* a section of text which includes a number of points of great importance to the modern view of speleogenesis. Dreybrodt is one of the few modern authors who points out (1988, p.5) that surface karst development is closely related to the state of underground *"karstification"*. Also, perhaps unintentionally, or as a linguistic quirk, the quotation appears to equate the initial phase of karstification (*sensu lato*) with the inception phase of speleogenesis. It is equally of interest that, as quoted above, Dreybrodt is only able to say that a system of penetrable primary fissures is *"inferred"* to exist. This point, of seemingly small significance, is to some degree at variance with his earlier (1981) views that no special mechanism nor chemistry is required to account for the inception of underground drainage. If mixture dissolution processes can provide even the earliest stages of cave development, as he concluded in 1981, the point is not re-emphasised in this part of his 1988 text. Nor is any explanation offered for the *"system of penetrable primary fissures* [which] *must exist....."* (1988, p.6). The problem of the potential presence of a primitive permeability, possibly related to regional tectonism, is as yet little studied, but was addressed on a purely theoretical level by Lowe (1992, chapter 8).

These points do not detract from Dreybrodt's overwhelming but yet theoretical confirmation and amplification of Bögli's mixture corrosion hypothesis. The weight of chemical evidence, consideration of plotted curves and laboratory models and the 'need' for an explanation of deep phreatic and shallower forms of dissolution lead to an intuitive leap to grasp and accept the mixture corrosion hypothesis. However, realistically, empirical evidence is lacking and the hypothesis remains just that. Observations which might add a degree of confirmation might equally be explained by other dissolutional effects. However, from the standpoint of 'devil's advocate', there seems to be little doubt that processes of mixture corrosion, as originally postulated by Bögli and refined by Dreybrodt and other workers, are at least partially responsible for a proportion of dissolutional speleogenesis in the post-inception development phases.

More recently another important paper was published (Dreybrodt, 1990) in which Dreybrodt returns to a fundamentally chemical and mathematical modelling of the early stages of karst evolution. Although paying an apparent lip-service to the constraints and limitations imposed by geology, in this paper Dreybrodt has reverted to his earlier visions of speleogenesis. The calculations which support his eventual model appear daunting, but the conclusions reached are quite clear cut and, against the background of ideas developed by Worthington (1991) and Lowe (1992), they must be considered reactionary rather than revolutionary. It is not possible for a non-mathematician to comment sensibly upon the equations employed by Dreybrodt. A possibility remains that, even if the mathematical reasoning is faultless, the basic concepts and framework of his discussion and model are, at least in part, slightly removed from reality.

To put the main objection to Dreybrodt's seemingly powerful arguments into perspective, the following statement (Dreybrodt, 1990, p.639) is helpful: *"There is general agreement among karst geologists that the initial phase of karstification in limestone rock takes place under phreatic conditions by gradual enlargement of primary fissure porosity due to solutional attack by CO_2-enriched groundwater"*. As in his earlier writings (1981, 1988) described above, Dreybrodt ignores the one geological consideration upon which all the processes he describes depend. He assumes the presence of primary fissures of sufficient width to carry groundwater flow under the influence of a hydraulic head, without considering how such fissures attain these dimensions either with regard to chemistry or a driving mechanism. This discrepancy appears not to be noted by Dreybrodt and in the context of this lack of recognition it is not surprising to find that neither the work of Howard (1964) nor that of S N Davis (1966), discussed above, is cited.

It is this tacit assumption that traditionally accepted chemistry and flow mechanisms can account for the full spectrum of speleogenetic activity which detracts from what would otherwise be an excellent model for the development stages of cave formation. Inception processes are ignored by the Dreybrodt model. It seems that Dreybrodt's (1981) belief that processes of groundwater mixing can account for all observed speleogenetic activity is maintained, but now, as then, the mechanisms for the quantum leap from joints of (possibly) infinitesimal width to those 0.2mm wide are ignored. Although Dreybrodt acknowledges the utilisation of joints narrower than 0.2mm (from 0.01 - 0.1mm) as flow paths, his mathematical consideration and the various curves generated show no reference to such narrow structures.

Thus the major criticism of the latest (1990) Dreybrodt model of speleogenesis lies not only in its lack of an explanation or mechanism for speleo-inception, but also in his apparent unawareness of the lack. As a model for development beyond inception, there is little that can be fully contradicted, but several points which might bear minor reconsideration and amendment. One point of particular interest, which reappears throughout the paper, is the attainment of a *"breakthrough time"*, a concept noted earlier by Atkinson (1968). The dissolutional widening of a fissure begins with the operation of slow, fourth-order kinetics along its entire length. As the fissure is widened at the input end of the system, first-order kinetics begin to operate and their region of activity slowly propagates along the fissure towards the outlet. At the moment of breakthrough flow rate increases dramatically and subsequent development is likewise accelerated. Dreybrodt links the *"breakthrough time"* transition to the transition from laminar to turbulent flow conditions reported by many other authors. He also states (p.644) that, *"..it marks the end of the* initial *phase of karst evolution"*.

The statement reproduced above is possibly crucial to the application of a partial compromise to the disparity between the Dreybrodt model and the model suggested by Lowe, (1992). It might be argued that Dreybrodt's considerations and calculations add to the understanding of aspects of speleogenesis raised by earlier workers and supply valuable theoretical data applicable to problems such as speleogenetic timescales. On these counts there is little that cannot be reconciled to Lowe's theoretical model, the differences generally being of degree rather than of disagreement. If, in line with the statement above, it is accepted that the *"breakthrough time"* marks the end of the *initial* phase of karstification (or speleogenesis) it should be underscored that this *initial* phase is preceded by an **inception** phase. This presents a seeming contradiction of terms which must be accepted as a legacy of earlier ideas, the past and ongoing misuse of the term *"initiation"* by many workers, and the general lack of awareness or consideration of the very earliest parts of the speleogenetic process. It was recommended by Lowe (1992, chapter 4) that although some workers had used the term *"initiation"* in a sensible and clearly defined manner, others had not, and the term would be better abandoned. In recognition of the former duality of initiation's meaning, **inception** and **gestation** were suggested as replacements. The former covers the processes, commonly ignored, which bridge the gap between rock with no caves (in the broadest sense) and rock with caves. The inception phase passes transitionally into the gestation phase (which compares closely with most workers' idea of *"initiation"*), covering the time span and processes between the conception of solvent motion in dissolutional voids (caves *sensu lato*) and breakthrough to turbulent flow conditions.

W Back, B B Hanshaw, J S Herman (and others)

"Solution effects of the cavern type are limited to the compact Walshingham Limestone. The less compact limestones - the slightly cemented dune sands - show certain solutional effects but apparently they are too porous to permit sufficient concentration of migrating surface water to form caverns."
(Swinnerton, 1929, p.82 [concerning Bermuda])

Many of William Back's better known publications, some of which date back into the Scientific Period, were written in co-operation with Bruce Hanshaw as either the main or second author. Janet Herman's joint contributions with Back are also notable, but several additional authors have co-operated with Hanshaw, Back and Herman on diverse aspects of their research. As with Mylroie (see below) much of the output has dealt with processes of dissolution within young carbonate rock sequences in coastal environments, and it is this element of their work which is of greatest relevance to the overview of speleogenesis.

Studies undertaken by Back and his several co-workers, centred upon the Caribbean area, began during the Scientific Period. The ideas have developed progressively as work has continued and the combined, evolved view is briefly considered here. In 1966, Back and others contributed a paper dealing with water/mineral equilibria in an artesian carbonate aquifer. The aquifer studied, mainly Tertiary age limestones with subordinate dolostones, lies in central Florida. Relationships of calcite, aragonite and dolomite were examined by application of thermodynamic concepts to well water samples derived from depths of 100 to 1500 feet. Water level data from the wells allowed the piezometric surface to be contoured, thus defining a high in the recharge area and implying an outward flow from the high, perpendicular to the contours. Broadly, analytical results were as might be expected; saturation with respect to calcite, aragonite and dolomite increasing outwards from the recharge area. Anomalous calcium/magnesium ratios were explained as resulting from preferential dissolution of high-magnesium calcite, which is more readily soluble than dolomite.

Other ancillary aspects of the study are of great relevance to the entire speleogenesis debate. Primarily the research confirmed that dissolution of limestone and dolostone occurs at depths several hundred feet (considerably deeper than 50m) below the supposed water-table, hence allowing cave and channel formation. In discussing the paper (Back and others, 1966, p.126), Back opined that more dissolution occurs in areas of salt water than fresh water. This concept was not followed up, but perhaps reflects an early and almost accidental encounter with interface mixture dissolution. Finally, the research included limited age determination of groundwater. Results indicate groundwater flow rates of about 8m per year. This velocity value is in itself of little worth, without being able to consider the detailed nature of the hydraulic head and gross permeability. On the most basic level of speleogenetic argument, however, it is an absolute value which indicates that groundwater movements as small as 2cm per day may be seen to be involved in dissolutional activity within the phreatic zone. Mechanisms active in deep phreatic situations and in conditions of such slow flow cannot be those traditionally associated with carbonate dissolution.

A more recent paper (Back and others, 1986) examines mixing dissolution in the coastal areas of Yucatan. The processes described are not unlike those reported by other workers in similar areas, and they will not be examined in detail here. Of particular interest, however, is a short section of the abstract:

"Such dissolution has probably been a significant control on permeability and porosity distribution in carbonate rocks in the geologic record."
(Back and others, 1986, p.137)

This view is highly significant in the broad context of speleogenetic timescales and the potential links between caves visible today and those formed in more remote times.

Janet Herman's most significant contributions in the context of this review take the form of joint papers with Back and Luis Pomar. Ideas incorporated in these papers have much in common with ideas included in earlier work by Back, Hanshaw and their other collaborators, but they are discussed separately since a different field area to that covered by most of Back's studies is involved. In two closely linked papers (Herman, Back and Pomar, 1986, 1987) the groundwater geochemistry and aspects of speleogenesis in the Mediterranean islands of Menorca and Mallorca are discussed. Water samples from caves and wells were analyzed and a broad range of salinities recorded. It was deduced that the observed compositions reflected mixing of end members of fresh groundwater and Mediterranean saline water and consequent dissolution or precipitation of carbonates. A series of interrelated chemical reactions was predicted and subsequently tested/simulated by a computer model.

Discussions of the observed results and computer predictions are complex and not considered in detail here. However, the results support conclusions similar to those reached elsewhere in coastal environments. The fresh/salt groundwater mixing zone is a highly active focus for water+rock interaction, leading to the increased development of porosity and secondary permeability. Coastal limestones exhibit extensive cavernization on all scales. Additionally it is acknowledged that since the mixing zone will have fluctuated in response to sea-level and climatic change and/or tectonism, dissolutional (and various associated) activity will have occurred at different levels at different times. The latter observation has the unstated corollary that dissolutional cave passages associated with individual sea-level stands might be expected to be preserved at various levels, at, above or below the contemporary sea-level.

E H Kastning

"Most passages of large caves in the area ... are conformable with the Segovia beds; the largest conduits are excavated in part along highly burrowed, massive beds of marl ..."
(Kastning, 1984, p.369.)

The essential elements of Ernst Kastning Jr's contribution to the development of the modern understanding of speleogenesis are encapsulated in a symposium contribution (Kastning, 1984) drawn mainly from ideas presented in several earlier papers. Much of the collated information can be construed as supportive of more recent ideas, though most of Kastning's interpretations stopped short of the conclusions reached by Worthington (1991) and Lowe (1992). If his approach had been more objective and if the data had been considered from a lateral viewpoint Kastning might have deduced a similar inception mechanism. As with a number of other works discussed elsewhere in this review, there are indications that the major thrust of Kastning's reasoning was directed towards refining the 'established geomorphological truth' rather than considering any possible fallacies contained within that 'truth'. This being said, Kastning's view, constrained as it was by pre-existing thought, provides a clear and concise distillation of the post-Ford and Ewers concept of cave development, to which are added several crucial elements deriving from the author's field work and theoretical considerations.

Following a concise review of the history of cave formational controversy from the North American point of view Kastning lists the parameters which determine the final dissolutional conduit configuration as, *"... geologic structure, lithic character of the bedrock, topographic conditions and hydraulic gradient within the aquifer."* (Kastning, 1984, p.351). Having made the point that these affect the final form and that any individual factor or combination of factors can be dominant locally, he makes the valid, perhaps vital observation that it is the structural framework which determines initial groundwater circulation routes. Following on, however, his view that the dip/strike of strata in combination with the position of the water-table will determine whether passages will form horizontally (in the shallow phreatic zone) or steeply down dip (in the vadose zone) is suspect. Firstly it can be argued that there is no such interplay of rock attitude with the water-table in this context since the water-table would not exist **in the form implied** without the presence of primary conduits. Also the concept of vadose zone cave development can now, as a general rule, only be viewed as a secondary development process, superimposed upon and guided by parts of a pre-formed pattern of dissolutional (phreatic zone) drainage routes. Formation of the latter pattern is probably guided primarily by the interaction of stratigraphy (and hence dip/strike) and structure within a rock mass which is effectively saturated. That is, any pore or void space within the rock is water-filled, though the water is not necessarily mobile, and if a water-table were deemed to exist it would have to be postulated either at the surface (in an exposed sequence) or (in buried rocks) at the base of any overlying impermeable bed. Only at a later stage of speleogenesis is an apparent water-table generated, as a consequence of uplift and/or surface downcutting and the change in function of pre-existing dissolutional drains.

The argument above concerning Kastning's correct recognition of the importance of stratigraphy/structure but misapprehension of the role of the water-table may seem of limited relevance to speleogenesis as a whole, but is quite crucial. Kastning (1984, p.352) cites the works of Palmer (1972, 1977), D C Ford (1971) and Ford and Ewers (1978), the essence of which contributions is still to be found in the latest reviews of speleogenetic mechanisms. It is vital to the consideration of speleogenesis presented by Lowe (1992) that the ageless water-table controversy and the misconceptions concerning the nature of the apparent water-table in (potentially) cavernous carbonate sequences have no bearing upon the general processes of speleo-inception. The inverse relationship is, however, of great moment: the locally determined level of an apparent water-table in a cavernous carbonate sequence is largely a reflection of the proving, or lack of proving, of active conduits and associated sub-conduit feeders within the rock and the relationship of these conduits to the resurgence(s) to which they **now** drain.

More positively, Kastning's symposium contribution (and his papers from which it derived) provides a useful indication not only of stratigraphical guidance of inception routes, but also the nature of the inception levels and how their presence and importance may be obscured by the effects of later, particularly vadose, processes. Of the areas that are discussed in detail, the data for the Sonora district (Texas) and the Helderberg Plateau (New York) are particularly important. Kastning's conclusion (1984, p.377) that, *"Cave passages are usually excavated along bedding-plane partings or fractures that are the most open initially, or are favourably oriented along prevailing hydraulic gradients. These openings will enlarge at the greatest rates and become the master conduits of well-integrated cave systems."*, is perhaps only partially correct if the ideas put forward by Worthington (1991) and Lowe (1992) are valid.

J E Mylroie (and others)

"The study of many of the complexities of karst landforms often leads to a classic 'can't see the forest for the trees' situation." (Mylroie, 1984, p.158)

Much of John E Mylroie's early work, culminating in a significant symposium contribution (Mylroie, 1984) might be considered to fit broadly within the constraints of contemporary karst and speleogenetic thinking and to add only local insight rather than general revision to the overview. More recently, however, his work, including that published jointly with James L Carew and Peter N Vogel, has provided important lateral views of several aspects of cave formation. Foremost among these are descriptions of processes and timescales of speleogenesis within late Pleistocene aeolian calcarenites and associated rocks in the Bahamas (eg Mylroie and Carew, 1986 and 1990; Vogel, Mylroie and Carew, 1990). Perhaps the most important aspect of Mylroie's work is his contribution to the understanding of speleogenetic timescales and the processes responsible for cave development in young rocks at or close to sea level. Much of his output on these subjects has been in the form of joint papers with Carew and Vogel.

By a consideration of cave passages in the Bahamas and in glaciated areas, Mylroie and Carew (1986) were able to show that a conduit of 1m diameter *could* form in 10,000 years. They make the important point (p.249) that, *"Caves that are in apparent equilibrium with the surficial environment may have completed their macroscopic enlargement relatively recently, as their geomorphic setting suggests, but the initiation of their development could be more proportional to the age of the enclosing rock"*. This is more readily understandable in the full context of the enclosing text. Expressed somewhat differently the point being made is that caves which appear related to the modern landscape, for instance a post-glacial landscape, could have achieved most of their growth within this modern episode, but their inception and early stages of development could date back towards the formation of the host rock.

More recently Mylroie and Carew (1990) and Vogel, Mylroie and Carew (1990) have updated and added to ideas concerning cave development in aeolian calcarenite. Details of studies centred on San Salvador Island in the Bahamas, but including work elsewhere in the Caribbean region, have been described in several papers during the nineteen-eighties (eg Mylroie, 1984; Carew and Mylroie, 1987). The results of the San Salvador studies show close agreement with those reported from Yucatan (Back and others, 1986) and other parts of the Bahamas (eg Smart and others, 1988). They are not examined in detail here.

Recognition and demonstration of the intense dissolutional activity in freshwater lens marginal mixing zones is, however, of paramount importance to an overview of speleogenesis. The relation of the observable extent of rock dissolution to absolute timescales provides irrefutable proof of rapid cave development under suitable conditions. It is a small step to accept such rapid speleogenesis under the conditions of climate, porosity and hydrology met in the ideal lens environment beneath a modern tropical or sub-tropical island. However, the further implications of the processes and timescales involved, in the context of *"the present is the key to the past"*, are less easily assimilated. Likewise it is less straightforward to evaluate the potential for and the possible extent of the operation of similar processes within older, less permeable and better consolidated carbonate sequences in cooler climatic belts. Lateral extensions of 'freshwater lens theory' appear possible, if not probable, but await examination.

C A Hill

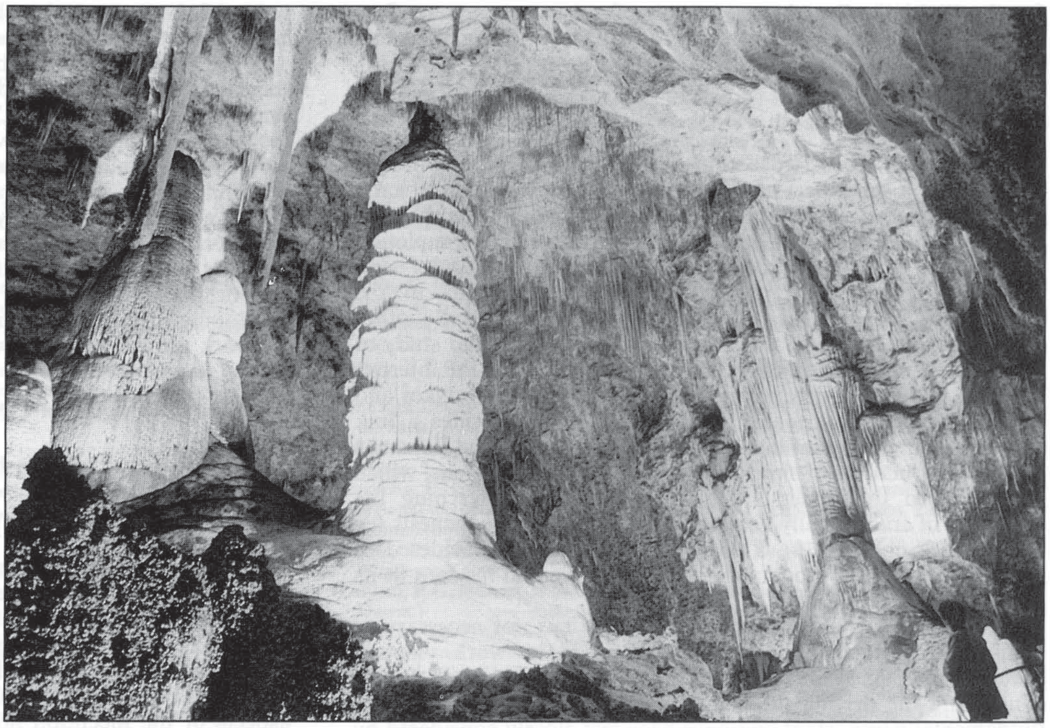
Carol Hill has produced an impressive list of publications dealing with a variety of cave minerals, especially saltpetre (potassium nitrate) and other, less common, nitrates. From the point of view of this review, however, her most significant and crucial publications have been those dealing with the caves of the Guadalupe Mountains, particularly her considerations of the speleogenesis of Carlsbad Caverns (Hill, 1981, 1985, 1986, 1987).

Carlsbad Caverns and other lesser systems in the Guadalupe Mountains of New Mexico have been the focus of much scientific interest since the early days of cave study. As will be apparent from the outline of Carol Hill's conclusions presented below, much of the early speculation was along the traditional lines of cave formational thinking and most of the theories of genesis produced were significantly removed from reality. In his otherwise excellent paper of 1935, Gardner considered the formation of the Carlsbad system at great length in the context of his own cave development hypothesis. His arguments were intended to prove a common formational mechanism for Carlsbad Caverns, Mammoth Cave and other *"large caverns"*. With hindsight it can be stated that he was attempting the impossible; if he succeeded in convincing himself of a supposed similarity of origin, it must now be accepted that his interpretation of the evidence was awry and his conclusions were illusory. Other workers of no less repute have published their thoughts on Carlsbad Caverns, most notably Davis (1930) and Bretz (1942 and, more specifically, 1949). These, and other contributions prior to Hill's work, as well as some later publications (eg Bullington, 1985) have been based upon the false assumption that Carlsbad Caverns and adjacent systems were conceived and developed by the traditionally accepted process of dissolution by dilute carbonic acid.

A review of various theories of formation for caves in the Guadalupe Mountains by D G Davis (1980) pre-dates much of Hill's work and although various unorthodox development mechanisms are considered and evaluated, the then prevailing confusion was not definitively clarified. Davis's discussion of models presented by Jagnow (1977), Palmer and others (1977) and Egemeier (1973), which are not considered separately here, is nonetheless an interesting preface to Hill's work. Hill's earliest publication on the Guadalupe Mountains (Hill, 1981) found inadequacies in each of the earlier proposed mechanisms. Though accepting the superiority of Hill's overview and developed model, the present author envisages that elements of the earlier ideas might still be applicable to the history of cave development in the Carlsbad Cavern area,.

Broadly, Hill's formational mechanism relies upon the *in-situ* generation of sulphuric acid by the oxidation of hydrogen sulphide, a process which had been suggested previously by, for example, Morehouse (1968) and Egemeier (1973). Solutions rich in hydrogen sulphide are postulated to have risen along joints (presumably open, though this is not stated explicitly and nor is their origin explained) having been generated at depth within hydrocarbon-rich deposits beneath the Delaware Basin east of the Guadalupe Mountains. As the solutions rose they came into contact with a zone of relatively oxygen rich groundwater and within the zone of mixing the hydrogen sulphide underwent oxidation to sulphuric acid. This highly aggressive mineral acid dissolved large volumes of carbonate bedrock in and adjacent to the mixing zone, producing

The Big Room of Carlsbad Caverns, New Mexico, where solution by sulphuric acid was a major process in the cave development. (Photo: Andy Eavis).



much of the complex ramification of Carlsbad Caverns. Very strong arguments are presented in support of this mode of dissolutional cavern development, which relies mainly upon a complex reaction chain first suggested by Egemeier (1981), which can be simplistically considered (ignoring the ionised nature of some components) as:

$$H_2S + 2O_2 \Longleftrightarrow H_2SO_4$$
$$H_2SO_4 + CaCO_3 \Longleftrightarrow CaSO_4.2H_2O + CO_2$$

Both reactions are to some degree reversible. However, the process or processes involved in the development of the Guadalupe caves are far more complex than is implied by these simple equations. There must have been some proto-cave development mechanism operating to open the primary pathways which allowed access of both the upward moving sulphurous brines and the downward moving oxygen rich groundwater (termed Solution I caves by Hill); a similar mechanism must also have produced a secondary permeability which allowed the lateral transmigration of the major proportion of the reaction products. The presence of massive beds of gypsum (hydrated calcium sulphate) within Carlsbad Caverns is generally taken as an additional indication of the chemical reactions involved in the speleogenesis of the area; though providing fairly convincing evidence of reactions it does not necessarily confirm the origin of *all of* the reactants. Native sulphur is also recorded from many parts of the Guadalupe caves. Though this might be assumed to represent material precipitated by alternative reactions during oxidation of sulphur-bearing compounds, the potential involvement of biological agents, particularly bacterial activity, cannot be discounted. Several processes whereby sulphur may be precipitated by bacteria have been described (Ehrlich, 1981) and it is apparent that other possible mechanisms of mineral acid generation involving biological intermediaries are relatively commonplace in the broad geological context and not unknown in the context of speleogenesis, as was tentatively suggested by T D Ford (1965) and more recently argued by Ball and Jones (1990). Thus, although the processes and reactant sources proposed by Hill provide a reasonably sound basis for her views of cave formation in the Guadalupe Mountains there seems to be room for further study into the processes (if any) that were active prior to, in tandem with, or after the proposed influx of sulphurous brine from beneath the Delaware Basin.

The immense importance of Hill's work from the standpoint of a modern reappraisal of speleogenesis is that here the principle of dissolutional cavern development by reactions other than carbonic acid dissolution has been suggested, demonstrated and widely accepted. Whereas at present acceptance by many workers is on the level of *"a special case"* it can be argued that rather than being a *special case* the development of the Guadalupe caves may be an *'extreme case'* and that similar processes have been, and continue to be, active in many other, but less easily demonstrable, environments. In addition, in the context of the requirement for some kind of proto-void system to allow access of sulphurous brines from adjacent areas, it may also be productive to consider the potential palaeogeography of the Capitan Reef Belt and examine the possibility that syngenetic dissolutional processes analogous to those studied in modern freshwater/saltwater mixing zones might have been involved. Preservation of such syngenetic voids, which may take the form of palaeo-caverns or simply an extended vughy porosity, must be considered a strong possibility in the light of studies in the Tongan Islands (Lowe and Gunn, 1986; Lowe, 1989a) and other areas (see R A L Osborne, below).

R A L Osborne

"The term multiple karstification is used to describe karst processes affecting the same body of rock on a number of distinct occasions during its history."
(Osborne, 1986, p.57)

Though the contributions made by Robert A L Osborne to an understanding of speleogenesis may not be immediately apparent, they are of considerable importance to various elements of some modern cave development hypotheses. Osborne has carried out detailed studies, particularly of stratigraphy and lithology, in numerous caves in New South Wales, Australia. His findings appear to present conclusive evidence for *"multiple karstification"*, wherein pre-existing karst landforms, including caves, may provide guidance for, or be incorporated within, later features. His results from Wellington Caves (1983) and Timor Caves (1986), New South Wales, are particularly significant.

In the former caves, careful study of sediments indicated multistage infilling of existent cave passages followed by renewed cave formation, partly utilizing the earlier features. Though no date for the original cavernisation within carbonates of Devonian age is suggested, several later phases which can be tied to approximate datum points by fossil evidence indicate that the earliest dateable events are at least of Tertiary age and possibly older. Though by no means a great age geologically, this indication of Tertiary speleogenesis is significant since it proves that the initial cave development in the area must have predated these events. Even the idea that caves were forming in Tertiary times would have been considered unreasonable, until relatively recently, by those karst workers who saw cave formation as a Pleistocene or, less commonly, post-Pleistocene phenomenon.

In Timor Caves, Osborne carried out similar detailed stratigraphical studies. By dint of the presence of dateable basaltic rocks he was able to show that caves existed in the Devonian bedrock prior to Late Cretaceous times (at least 73 million years ago).

Another aspect of Osborne's sedimentological studies is of crucial importance . Among the lithified Tertiary sediments visible in the Wellington Caves are beds of sandstone and conglomerate with characteristics indicative of turbiditic deposition. Such unequivocal evidence that turbidites may extend deeply into cave systems and, with associated deposits, can totally infill pre-existing passages, is of great significance. Osborne suggests a mechanism for the "initiation" and propagation of sub-aqueous mass-flow which is unlike that postulated in Tonga (Lowe, 1989a), but this divergence detracts from neither possibility.

P L Smart (and others)

Much of the recent work of Peter Smart and his co-workers, which is but a small part of his considerable contribution to karst and speleogenesis studies over the years, is ongoing. Publication has been only in relatively brief 'review' form pending completion of projects and more complete processing of results and completion of ongoing projects by research students in the Geography Department of Bristol University and other collaborators. Central to the work is the exploration and study of Blue Holes, underwater caves, in the Bahamas. Also of interest, however, is a fine review paper dealing with neptunian dykes and related features (Smart and others, 1988). This provides a link between sedimentary infillings in active formation in the Bahamas and relict structures of similar type in this and other areas, which is helpful to the understanding of arguments concerning the extension of traditionally accepted timescales of speleogenesis (eg Lowe, 1992, chapter 10). This review will not be considered in detail, being seen as moderately comprehensive and modern in its approach and conclusions.

As with the work of Mylroie and others it is not necessary to analyze the results of the Bahamian Blue Holes work. Aspects covered particularly in the publications of Whitaker and Smart concerning the driving mechanism of sub-surface fluid flow are fundamentally important and have been less closely considered by other workers in comparable fields. How far, if at all, their ideas impinge upon speleogenesis in other environments must also be considered. Smart and Whitaker (1988) briefly explore the evolutionary process which spans the change from intergranular flow [primary porosity] to fissure and/or cavernous flow [secondary permeability]. Their time and environmental framework is relatively limited and it is possible that the ideas they introduce are capable of appreciable extrapolation. Other, possibly crucial, observations dealing with the role of microbially derived strong acids in the freshwater lens environment are included in a paper by Bottrell and others (1990). The results described in this short and relatively obscure paper may be seen as providing one more link between processes of syngenetic cave growth and the wider field of speleogenesis in mature carbonate sequences.

Other significant contributions

The volume of published material bearing directly or indirectly upon karst and speleo-genesis to emerge during the Modern Era is almost certainly far greater than the sum of all that produced in the previous 100 years. The contributions discussed above are those which are either generally accepted as being pivotal to modern understanding of the subject or those which offer insights applicable to recent research findings such as those of Worthington (1991) and Lowe (1992). There are, in addition, innumerable less widely known papers with aspects applicable to this current reasoning, amongst which are the earlier papers by the present author upon which many of his conclusions depend (Lowe, 1982, 1983, 1985, 1986, 1989a, 1989b, 1989c; Lowe and Gunn, 1986). Examination of the reference lists and bibliographies which form important parts of modern karst and karst-related text books, some of which are mentioned above, will emphasise that only a small number of recent workers listed are cited and discussed here. In parallel with the papers listed are the results of many studies which, though not necessarily of lesser stature, are less obviously of direct relevance to the theme of this review. Work such as that by Cooper (1986), Gillott (1978), Helz and others (1987), Pye and Miller (1990), Vear and Curtis (1981) and Ueda and Sakai (1983) is outside the mainstream of speleo-genetic study, but provides insight potentially applicable to speleogenesis in its fullest sense.

On the broader scale it is impracticable to discuss the many and valuable publications on speleogenesis and karst studies in general by researchers who have not been considered individually within the above review. In the United Kingdom, workers such as Atkinson, T D Ford, Gunn, Sweeting, Trudgill, Warwick and many others have produced important and often substantial contributions, only elements of which appeared relevant to the central theme of this review and its parent thesis. Likewise contributions by authors such as Jennings, Lauritzen, Miotke, Quinlan, Trombe, White and Williams, from Europe, North America and Australasia, contain observations which are relevant to many aspects of speleogenesis but which were mainly peripheral to the theme of Lowe's (1992) studies. Doubtlessly very much of potential importance by these workers and others, writing in languages other than English, has been overlooked.

A PhD thesis submitted by Worthington (1991) during the final stages of Lowe's research appears to represent the most significant step forward from 'classical' carbonate speleogenesis dogma since the revolution in thinking which was heralded by the Ford-Ewers model. The data presented and the conclusions drawn by Worthington are in some ways complementary to those presented by Lowe (1992). In other ways, and particularly in his mathematical treatment of some elements of speleogenesis, Worthington's methods and conclusions differ from Lowe's. It is neither practical nor desirable to attempt a dissection of Worthington's still-developing ideas in the current context, but limited references to his work are included above where appropriate.

CONCLUSIONS

During more than one hundred years of cave research ideas concerning the origins and development of underground drainage systems have changed. Few workers have adequately addressed the questions presented by the earliest, inception, stages of speleogenesis, and some ideas which were potentially applicable to the problem have been largely overlooked or ignored. This review has attempted to point out elements of earlier work which are either directly applicable or applicable following lateral interpretation of published results, to the overview of speleo-inception. Papers which explain and develop the arguments behind aspects of the inception horizon hypothesis of limestone cavern origin (Lowe, 1992) are currently in preparation.

Acknowledgements and epilogue

Most of this review is extracted from my unpublished PhD thesis, with only minor modifications. All those whose help and support allowed completion and acceptance of that document are thanked again, particularly John Gunn, Tony Waltham and Trevor Ford. Mostly, however, my appreciation must pass to all those earlier and contemporary workers whose published ideas have provided the fuse to spark off my own thoughts. If, in the pages above, I have criticised their work, I hope the criticism has been fair and in the spirit of scientific progress. When I was about 13 or 14 years old a school chemistry teacher told me that a good hypothesis is one which leads to further research, providing either confirmation or new hypotheses. Since then I have believed that the mere act of publishing, no matter what the weight of evidence or reputation of the writer, is no proof that an interpretation is valid. The scientist's rule of thumb should be, "That sounds fine, but what if....?" Soon after beginning my research I was faced by two obvious questions - "What if the Ford-Ewers Model isn't the ultimate answer?" and "What if cave formation by carbonic acid dissolution is not inevitable in *pure* carbonates?". Subsequently many more "what if?" questions became apparent. No doubt my own hypothesis will generate still more.

REFERENCES

A number of the reference works listed below, marked [*], originally appeared in languages other than English. Where English translations or digests were not available the substance of the contribution has been accepted on the basis of later appraisals by reliable authors.

Atkinson, T C, 1968. The earliest stages of underground drainage in limestones - a speculative discussion. Proceedings of the British Speleological Association, Vol.6, 53-70.

Back, W, Cherry, R N and Hanshaw, B B, 1966. Chemical equilibrium between the water and minerals of a carbonate aquifer. National Speleological Society Bulletin, Vol.28, No.3, 119-126.

Back, W, Hanshaw, B, Herman, J S and Van Driel, J N, 1986. Differential dissolution of a Pleistocene reef in the ground-water mixing zone of coastal Yucatan, Mexico. Geology, Vol.14, No.2, 137-140.

Bakalowicz, M J, Ford, D C, Miller, T E, Palmer A N and Palmer, M V, 1987. Thermal genesis of dissolution caves in the Black Hills, South Dakota. Geological Society of America Bulletin, Vol.99, 729-738.

Ball, T K and Jones, J C, 1990. Speleogenesis in the limestone outcrop north of the South Wales Coalfield; the role of micro-organisms in the oxidation of sulphides and hydrocarbons. Cave Science, Vol.17, No.1, 3-8.

Bedinger, M S, 1966. Electric-analog study of cave formation. National Speleological Society Bulletin, Vol. 28, No.3, 127-132.

Bögli, A, 1960. Kalklösung und Karrenbildung. Zeitschrift für Geomorphologie, Supplement 2, 4-21. [*]

Bögli, A, 1964a. Mischungskorrosion - ein Beiträg zur Verkarstungsproblem. Erdkunde, Vol.18, 83-92. [*]

Bögli, A, 1964b. Érosion par melange des eaux. International Journal of Speleology, Vol.1, 61-70. [*, read in translation.]

Bögli, A, 1978. Karsthydrographie und physische Speläologie. (Berlin, Heidelberg: Springer Verlag.) [*, see next entry]

Bögli, A. 1980. Karst Hydrology and Physical Speleology. (Berlin, Heidelberg, New York: Springer-Verlag.)

Bottrell, S H, Smart, P L, Whitaker, F and Raiswell, R, 1990. Geochemistry and isotope systematics of sulphur in the mixing zone of Bahamian blue holes. Applied Geochemistry, Vol.6, 97-103.
Bretz, J H, 1942. Vadose and phreatic features of limestone caves. Journal of Geology, Vol.50, 675-811.
Bretz, J H, 1949. Carlsbad Caverns and other caves of the Guadalupe Block, New Mexico. Journal of Geology, Vol.57, 447-463.
Bretz, J H, 1956, Caves of Missouri. Missouri Geological Survey and Water Resources, Series 2, 39.
Bretz, J H, 1960. Bermuda, a partially drowned late mature, Pleistocene karst. Bulletin of the Geological Society of America, Vol.71, 1729-1754.
Brook, D and Crabtree, H, 1969. Exploration Journal. (Leeds: University of Leeds Speleological Association.)
Bullington, N R, 1985. Geology of the Carlsbad Caverns. 149-151 in Permian carbonate/clastic sedimentology, Guadalupe Mountains: Analogs for Shelf and Basin Reservoirs. Cunningham, B K and Hedrick, C L (editors). PBS - SEPM Annual Field Trip, April 1985, Publication 85-24.
Carew, J L and Mylroie, J E, 1987. A refined geochronology of San Salvador Island, Bahamas. 33-44 in Proceedings of the third symposium on the geology of the Bahamas: San Salvador, Bahamas, College Center of the Finger Lakes Bahamian Field Station, Curran, H A (editor).
Coase, A C and Judson, D M, 1977. Dan yr Ogof and its associated caves. Transactions of the British Cave Research Association, Vol.4, Nos 1 and 2.
Cooper, A H 1986. Subsidence and foundering of strata caused by the dissolution of Permian gypsum in the Ripon and Bedale areas, North Yorkshire. 127-139 in The English Zechstein and Related Topics. Harwood, G M and Smith, D B, (editors). Geological Society Special Publication No.22.
Curl, R L, 1958. A statistical theory of cave evolution. National Speleological Society Bulletin, No.20, 9-22.
Curl, R L, 1963. On the definition of a cave. Proceedings of the Third International Speleological Congress, Vienna, 43-47.
Curl, R L, 1964. On the definition of a cave. National Speleological Society Bulletin, Vol.26, 1-6.
Curl, R L, 1965. Caves as a measure of karst. Journal of Geology, Vol.74, 798-830.
Cvijić, J, 1893. Das Karstphänomen. Geographische Abhandlungen herausgegeben von A. Penck, Vol.5, Part 3, 218-329. [*, part read in translation]
Cvijić , J, 1918. Hydrographie souterraine et évolution morphologique du Karst. Recueil des Travaux de l'Institute de Géographie Alpine, Vol.4, No.4, 375-426. [*, read in the translated version of Sanders, 1921.]
Davies, W E, 1960. Origin of caves in folded limestone. National Speleological Society Bulletin, Vol.22, Part 1, 5-18.
Davis, D G, 1980. Cave development in the Guadalupe Mountains: a critical review of recent hypotheses. National Speleological Society Bulletin, Vol.42, 42-48.
Davis, S N, 1966. Initiation of groundwater flow in jointed limestone. National Speleological Society Bulletin, Vol.28, No.3, 111-118.
Davis, W M, 1930. Origin of limestone caverns. Geological Society of America Bulletin, Vol.41, 475-628.
Dixon, E E L, 1910. Unconformities on limestone and their contemporaneous pipes and swallow holes. Report of the seventy-ninth meeting of the British Association for the Advancement of Science, Winnipeg: 1909, 477-479.
Dixon, E E L, 1921. The unconformity between the Millstone Grit and the Carboniferous Limestone at Ifton, Mon[mouthshire]. Geological Magazine, Vol.21, 157-164.
Dreybrodt, W, 1981. Mixing corrosion in $CaCO_3$-CO_2-H_2O systems and its role in the karstification of limestone areas. Chemical Geology, Vol.32, 221-236.
Dreybrodt, W, 1988. Processes in Karst Systems. (Berlin, Heidelberg, New York, London, Paris, Tokyo: Springer-Verlag.)
Dreybrodt, W, 1990. The role of dissolution kinetics in the development of karst aquifers in limestone: a model simulation of karst evolution. Journal of Geology, Vol.98, No.5, 639-655.
Durov, S A, 1956. On the question about the origin of the salt composition of karst water. Ukrainian Chemical Journal, Vol.22, 106-111. [*English translation by Olaf Muller, 1979, in Cave Geology, Vol.1, 185-190.]
Egemeier, S J, 1973. Cavern development by thermal waters with a possible bearing on ore deposition. Stanford University Thesis. 88p.
Egemeier, S J, 1981. Cavern development by thermal waters. National Speleological Society Bulletin, Vol.43, 31-51.
Ehrlich, H L, 1981. Geomicrobiology. (New York: Marcel Dekker Inc.)
Ford, D C, 1965. The origin of limestone caverns: a model from the central Mendip Hills, England. Bulletin of the National Speleological Society of America, Vol.27, 109-132.
Ford, D C, 1968. Features of cavern development in Central Mendip. Transactions of the Cave Research Group of Great Britain, Vol.10, Part 1, 11-25.
Ford, D C, 1971. Geologic structure and a new explanation for limestone cavern genesis. Transactions of the Cave Research Group of Great Britain, Vol.13, 81-94.
Ford, D C, and Ewers, R O, 1978. The development of limestone cave systems in the dimensions of length and depth. Canadian Journal of Earth Sciences, Vol.15, 1783-1798.
Ford, D C and Williams, P W, 1989. Karst Geomorphology and Hydrology. (London: Unwin Hyman.)
Ford, T D, 1952. New evidence for the correlation of the Lower Carboniferous reefs at Castleton, Derbyshire. Geological Magazine, Vol.89, 346-356.
Ford, T D, 1963. The Goyden Pot drainage system, Nidderdale, Yorkshire. Transactions of the Cave Research Group of Great Britain, Vol.6, No.3, 81-90.
Ford, T D, 1965. An ultimate food source for cave life? British Hypogean Fauna and Biological Records of the Cave Research Group of Great Britain, No.1, 24-25.
Gardner, J H 1935. Origin and development of limestone caverns. Bulletin of the Geological Society of America, Vol.46, 1255-1274.
Gillott, J E, 1978. Effect of deicing agents and sulphate solutions on concrete aggregate. Quarterly Journal of Engineering Geology. Vol.11, 177-192.
Grund, A, 1903. Die Karsthydrographie. Studien aus Westbosnien. Geographische Abhandlungen herausgegeben von A. Penck, Vol.7, 103-200. [*]
Haug, E, 1921. Traité de Geologie. Vol.1. (Paris.) [*]
Helz, G R, Dai, J H, Kijak, P J, Fendinger, N J, and Radway, J C, 1987. Processes controlling the composition of acid sulfate solutions evolved from coal. Applied Geochemistry, Vol.2, 427-436.
Herman, J S, Back, W and Pomar, L, 1986. Speleogenesis in the groundwater mixing zone: The coastal carbonate aquifers of Mallorca and Menorca, Spain. International Speleological Union, Spain 1986, Vol.1, 13-15.
Herman, J S, Back, W and Pomar, L, 1987. Geochemistry of groundwater in the mixing zone along the east coast of Mallorca, Spain. Karst Water Resources (Proceedings of the Ankara - Antalya Symposium), 467-479.
Hill, C A, 1981. Speleogenesis of Carlsbad Caverns and other caves of the Guadalupe Mountains. Proceedings of the 8th International Speleological Congress, Bowling Green, Kentucky, 143-144.
Hill, C A, 1985. Speleogenesis of Carlsbad Cavern and other caves of the Guadalupe Mountains, New Mexico. Newsletter of the National Speleological Society Section on cave geology and geography, Vol.12, No.2, 30-31.
Hill, C A, 1986. Carlsbad Cavern and other caves in the Guadalupe Mountains, New Mexico: A sulphuric acid genesis related to the oil and gas fields of the Delaware basin. Proceedings of 9th International Speleological Congress, Barcelona, Vol.1, 267-269.
Hill, C A, 1987. Geology of Carlsbad Cavern and other caves in the Guadalupe Mountains. New Mexico Bureau of Mines and Mineral Resources Bulletin 117.
Howard, A D, 1964. Processes of limestone cave development. International Journal of Speleology, Vol.1, Parts 1 and 2, 47-60.
Jakucs, L, 1977. Morphogenetics of Karst Regions. (Bristol: Adam Hilger.)
Jagnow, D H, 1977. Geologic factors influencing speleogenesis in the Capitan Reef Complex, New Mexico and Texas. Unpublished MSc Thesis, University of New Mexico.
James, N P and Choquette, P W, (editors), 1988. Paleokarst. (New York: Springer Verlag.)
Jennings, J N, 1968. Syngenetic karst in Australia. 41-110 in Contributions to the study of karst. Williams, P W and Jennings, J N, (editors). Australian National University: Research School of Pacific Studies, Publication G5.
Jennings, J N, 1985. Karst Geomorphology. (Oxford and New York: Blackwell.)
Kastning, E H, 1984. Hydrogeomorphic evolution of karsted plateaus in response to regional tectonism. 351-382 in Groundwater as a geomorphic agent. LaFleur, R G, (editor).(Boston: Allen and Unwin.)
Katzer, F, 1909. Karst und Karsthydrographie. Zur Kunde der Balkanhalbinsel. (Sarajevo: Kajon.) [*]
LaFleur, R G, (editor), 1984 Groundwater as a geomorphic agent. (Boston: Allen and Unwin.)
Lehmann, O, 1932. Die Hydrographie des Karstes. Enzyklopädie der Erdkunde. (Leipzig and Vienna: Deuticke.) [*]
Lowe, D J, 1982. The geology of central Levka Ori, Crete. [including photogeological maps]. Journal of Sheffield University Speleological Society, Vol.3, No.2, 41-44.
Lowe, D J, 1983. The Anglo-Canadian Rocky Mountains Speleological Expedition - 1983. Transactions of the British Cave Research Association, Vol.10, No.4, 213-244.
Lowe, D J, 1985. Karst development and cave formation in the Bocock Peak area, B.C., Canada. Transactions of the British Cave Research Association, Vol.12, No.2, 33-44.
Lowe, D J, 1986. [geological contributions in] Canada. [The Anglo Canadian Rocky Mountains Speleological Expeditions, 1983 and 1984]. ACRMSE. (Private publication.)
Lowe, D J, 1989a. Tonga '87 - the Report of the 1987 speleological expedition to 'Eua Island, Kingdom of Tonga. 28+ii pp, 18 figures, photographs. (Private publication.)
Lowe, D J, 1989b. The geology of the Carboniferous Limestone of South Wales. 3-19 in Limestones and Caves of Wales. Ford, T D, (editor). (Cambridge: Cambridge University Press.)
Lowe, D J, 1989c. Limestones and caves of the Forest of Dean. 106-116 in Limestones and Caves of Wales. Ford, T D, (Editor). (Cambridge: Cambridge University Press.)
Lowe, D J, 1992. The origin of limestone caverns: an inception horizon hypothesis. Unpublished PhD Thesis, Manchester Polytechnic/Council for National Academic Awards.
Lowe, D J and Gunn, J, 1986. Caves and limestones of the islands of Tongatapu and 'Eua, Kingdom of Tonga. Transactions of the British Cave Research Association, Vol.13, No.3, 105-130.
Malott, C A, 1937. Invasion theory of cavern development [abstract]. Proceedings of the Geological Society of America for 1937, p.323.
Martel, E A, 1921. Nouveau traité des eaux souterraines. (Paris: Delagrave.)[*]
Moneymaker, B C, 1941. Subriver solution cavities in the Tennessee Valley. Journal of Geology, Vol.49, 74-86.
Morehouse, D F, 1968. Cave development via the sulfuric acid reaction. National Speleological Society Bulletin, Vol.30, No.1, 1-10.
Mylroie, J E, 1984. Hydrologic classification of caves and karst. 157-172 in Groundwater as a geomorphic agent. LaFleur, R G (editor). (Boston: Allen and Unwin.)
Mylroie, J E and Carew, J L, 1986. Minimum duration for speleogenesis. Proceedings of the 9th International Speleological Congress, Barcelona 1986. Vol.1, 249-251.
Mylroie, J E and Carew, J L, 1990. The flank margin model for dissolution cave development in carbonate platforms. Earth Surface Processes and Landforms, Vol.15, No.5, 413-424.
Ollier, C D, 1975. Coral island geomorphology - the Trobriand Islands. Zeitschrift für Geomorphologie, Vol.19, No.2, 164-190.
O'Reilly, P M, O'Reilly, S M and Fairbairn, C M, 1969. Ogof Ffynnon Ddu, Penwyllt, Breconshire. (Swansea: South Wales Caving Club.)
Osborne, R A L, 1983. Cainozoic stratigraphy at Wellington Caves, New South Wales. Proceedings of the Linnaean Society of New South Wales, Vol.107, Part 2, 131-147.
Osborne, R A L, 1986. Cave and landscape chronology at Timor Caves, New South Wales. Journal and Proceedings of the Royal Society of New South Wales, Vol.119, 55-75.
Palmer, A N, 1972. Dynamics of sinking stream systems: Onesquethaw Cave, New York. National Speleological Society Bulletin, Vol.34, 89-110.
Palmer, A N, 1975. The origin of maze caves. National Speleological Society Bulletin, Vol.37, Part 3, 56-76.
Palmer, A N, 1977. Influence of geologic structure on groundwater flow and cave development in Mammoth Cave National Park, U S A. 405-414 in Karst Hydrogeology: proceedings of the 12th Congress of the International Association of Hydrogeologists, Huntsville, Alabama. Tolson, J S and Doyle F L (editors). International Association of Hydrogeologists Memoir, Vol.12.
Palmer, A N, Palmer, M V and Queen, J M, 1977. Speleogenesis in the Guadalupe Mountains, New Mexico: gypsum replacement of carbonate by brine mixing. Proceedings of the 7th International Speleological Congress, Sheffield, England, 336-339.
Pye, K and Miller, J A, 1990. Chemical and biochemical weathering of pyritic mudrocks in a shale embankment. Quarterly Journal of Engineering Geology, London, Vol.23, 365-381.
Rauch, H W, 1972. The effects of lithology and other hydrogeologic factors on the development of solution porosity in the Middle Ordovician carbonates of central Pennsylvania. Unpublished PhD Thesis, Pennsylvania State University.
Rhoades, R and Sinacori, N M, 1941. Patterns of groundwater flow and solution. Journal of Geology, Vol.49, 785-794.
Sanders, E M, 1921. The cycle of erosion in a karst region (after Cvijić). Geographical Review, Vol.11, 593-604. [see Cvijić, 1918.]

Scoffin, T P, 1987. An introduction to carbonate sediments and rocks. (Glasgow and London: Blackie.)
Shaw, T R, 1979. History of cave science. (Crymych: Anne Oldham.)
Shelley, A E, 1967. Analyses of two coals from the Great Scar Limestone near Ingleton, Yorkshire. Proceedings of the Yorkshire Geological Society, Vol.36, Part 1, No.3, 51-56.
Smart, P L, Palmer, R J, Whitaker, F and Wright, V P, 1988. Neptunian dikes and fissure fills: an overview and account of some modern examples. 149-163 in Paleokarst. James, N P and Choquette, P W (editors). (New York: Springer Verlag.)
Smart, P L and Whitaker, F, 1988. Controls on the rate and distribution of carbonate bedrock dissolution in the Bahamas. Proceedings of the fourth symposium on the geology of the Bahamas, 313-321.
Sweeting, M M, 1972. Karst Landforms. (London: Macmillan.)
Swinnerton, A C, 1929. The caves of Bermuda. Geological Magazine, Vol.66, 79-84.
Swinnerton, A C, 1932. Origin of limestone caverns. Geological Society of America Bulletin, Vol.43, 662-693.
Thrailkill, J V, 1968. Chemical and hydrologic factors in the excavation of limestone caves. Bulletin of the Geological Society of America, Vol.79, 19-45.
Trudgill, S T, 1985. Limestone geomorphology. (London: Longman.)
Turner, J S, 1968, A note on the Meal Bank coal horizon around Ingleborough. Transactions of the Leeds Geological Association, Vol.7, 265-268.
Ueda, A and Sakai, H, 1983. Simultaneous determinations of the concentration and isotope ratio of sulfate- and sulfide-sulfur and carbonate-carbon in geological samples. Geochemical Journal, Vol.17, 185-196.
U S Geological Survey, 1960. Quality of surface waters in the United States, parts 1 - 4. U S Geological Survey Water Supply Paper 1520, 1-641.
U S Geological Survey, 1962. Quality of surface waters in the United States, parts 1 - 4. U S Geological Survey Water Supply Paper 1571, 1-773.
Vear, A and Curtis, C, 1981. A quantitative evaluation of pyrite weathering. Earth Surface Processes and Landforms, Vol.6, 191-198.
Vogel, P N, Mylroie, J E and Carew, J L, 1990. Limestone petrology and cave morphology on San Salvador Island, Bahamas. Cave Science, Vol.17, No.1, 19-30.
Walkden, G, 1972a. The mineralogy and origin of interbedded clay wayboards in the Lower Carboniferous of the Derbyshire Dome. Geological Journal, Vol.8, Pt 1, 143-159.
Walkden, G, 1972b. Karstic features in the geological record. Transactions of the Cave Research Group Great Britain, Vol.14, No.2, 180-183.
Walkden, G, 1974. Palaeokarstic surfaces in Upper Viséan (Carboniferous) limestones of the Derbyshire Block, England. Journal of Sedimentary Petrology, Vol.44, No.4, 1232-1247.
Walkden, G, 1977. Volcanic and erosive events on an Upper Viséan carbonate platform, north Derbyshire. Proceedings of the Yorkshire Geological Society, Vol.41, Part 3, No.28, 347-366.
Waltham, A C, 1970. Cave development in the limestone of the Ingleborough district. Geographical Journal. No.136, 574-585.
Waltham, A C, 1971a. Shale units in the Great Scar Limestone of the southern Askrigg Block. Proceedings of the Yorkshire Geological Society, Vol.38, Part 2, No.3, 285-292.
Waltham, A C, 1971b. Controlling factors in the development of caves. Transactions of the Cave Research Group of Great Britain, Vol.13, No.2, 73-80.
Waltham, A C, (editor), 1974. The Limestones and Caves of north-west England. (Newton Abbot: David and Charles.)
Waltham, A C, 1989. Keynote address: Karst, caves and engineering - the British experience. 1-8 in Engineering and environmental impacts of sinkholes and karst. Beck, B F, (Editor). Proceedings of the third multidisciplinary conference on sinkholes and the engineering and environmental impacts of karst. St. Petersburg Beach, Florida, 1989. (Rotterdam: Balkema.)
Warwick, G T, 1953. The origin of limestone caves. 41-61 in British Caving. Cullingford, C H D, (editor). (London: Routledge and Kegan Paul.)
Watson, R A and White, W B, 1985. The history of American theories of cave origin. Geological Society of America, Centennial Special Volume 1, 109-123.
Whitaker, F F and Smart, P L, 1990. Active circulation of saline ground waters in carbonate platforms: Evidence from the Great Bahama Bank. Geology, Vol.18, 200-203.
White, W B, 1984. Rate processes: chemical kinetics and karst landform development. 227-248 in Groundwater as a geomorphic agent. LaFleur, R G, (editor). (Boston: Allen and Unwin.)
White, W B, 1988. Geomorphology and Hydrology of Karst Terrains. (Oxford: Oxford University Press.)
White, W B and Longyear, J. 1962. Some limitations on speleo-genetic speculation imposed by the hydraulics of groundwater flow in limestone. National Speleological Society, Nittany Grotto Newsletter, Vol.10, 155-167.
Worthington, S R H, 1991. Karst hydrogeology of the Canadian Rocky Mountains. Unpublished PhD Thesis, McMaster University.
Zötl, J G, 1961. Die Hydrographie des nordostalpinen Karstes. Steierische Beiträg zur Hydrogeologie (Graz), Vol.2, 54-183.[*]
Zötl, J G, 1974. Karsthydrogeologie. (Wien: Springer.)[*]

D J Lowe
23 Cliff Way
Radcliffe on Trent
Nottingham NG12 1AQ

CAVE SCIENCE Vol. 19, No. 3, December 1992
Transactions of the British Cave Research Association

Cave Detection using Electrical Resistivity Tomography

Mark NOEL and Biwen XU

Abstract: Electrical resistivity tomography (ERT) is a new geophysical imaging technique which models geological structure from potential measurements made on an array of surface electrodes. Using a pole-pole configuration and N electrodes, it is possible to record N(N-1)/2 independent values of apparent resistivity which are used to recreate a vertical image section by a backprojection technique. Recent field experiments above Long Churn Cave in Yorkshire employed a 20 electrode array with 5m spacing multiplexed to an Abem Terrameter. When inverted, the data produced a vertical resistivity section in which the cave passage can be clearly distinguished as a high resistance, elliptical anomaly.

The development of dependable methods for cavity location is of paramount importance for civil engineering in areas of soluble rock (Waltham *et al.* 1986; Waltham & Smart, 1988) but interest also extends to applications in karst geomorphology, hydrology and speleological discovery. Among the earliest geophysical attempts at cavity location in limestone were experiments carried out by Palmer (1954) who recorded variations in electrical resistivity. This approach was explored by later workers who tested various configurations of current and potential electrodes (eg. Day, 1964; Dutta *et al,* 1969; Myers, 1975) or numerical or graphical approaches for interpreting the data (eg. Habberjam, 1969; Bristow, 1966). The application of classical geophysical techniques to the detection of caves in limestone has been reviewed by Wigley and Brown (1978) and by the British Geological Survey who describe a theoretical and experimental evaluation of several methods (McCann *et al.,* 1987; Suddaby & Hallam, 1982). Butler and Murphy (1980) also describe a comparative study at a specially constructed test site of a wide range of geophysical methods for cavity detection. Historically, the subject also records research involving gravity measurement (Colley, 1963), geomagnetic surveying (Lange, 1965) and seismic reflection or refraction methods (Franklin, 1977; Watkins, 1967). Novel proposals which await rigorous assessment include passive acoustic recording (Lange, 1972), ground temperature measurement (Noel, 1985), airborne thermography (Rinker 1974), radar (Borkowski, 1990), electromagnetic methods (Scollar, 1990), sonar (Unterberger, 1977) and dowsing (Wilcock, 1990).

In this paper we present the results of field trials aimed at locating a cave using the new geophysical technique of electrical resistivity tomography, or ERT, (Noel & Walker, 1991; Noel & Xu, 1991). The method differs from the conventional approach to electrical resistivity survey in using a large array of electrodes and by recording the maximum number of independent measurements that are possible on the array. Electrical measurements are made using a resistivity meter multiplexed to the array and these are then inverted to yield a vertical resistivity section through the underlying structure by applying a backprojection algorithm. The term 'tomography' refers to the mathematical procedure whereby an *internal* structure (ie. the underlying geology) is being reconstructed from data collected on the external boundary (ie. the earth's surface). The concepts have been derived from a novel medical imaging technique in which anatomical conductivity structure is determined from electrical data obtained from a set of 16 skin electrodes (Brown, 1986). Our preliminary results certainly encourage further evaluation of resistivity tomography for locating cavities within karst and other terrains.

THEORY

The electrical resistivity of rocks, ρ, is principally a measure of water content since in these materials current is conveyed by ionic, rather than by electronic, conduction. Sedimentary lithologies display a wide range of electrical properties, viz., $10<\rho<10^4$ Ohm.m (McNeill, 1980). In comparison to the enclosing limestone the air within a karst cavity, although possibly saturated with water vapour, can be considered an insulator and it therefore appears as a target with infinite contrast to the resistivity prospection method. Geophysical resistivity measurement is normally made using a source of alternating current injected through a pair of soil electrodes C1 and C2, while potential is recorded between a second pair of electrodes, P1 and P2. This 4-electrode arrangement overcomes errors arising from probe contact potentials or contact resistance and the unpredictable variation in telluric potentials. If V is the potential difference recorded between P1 and P2 when current I flows through the ground from C1 to C2 the 'apparent resistivity', ρa, is defined as

$$\rho a = G \,.\, \Delta V/I$$

Where G is a factor which depends on the geometry of the array. If the electrodes are equally spaced along a line with the potential electrode pair innermost (Wenner configuration) then

$$G = 2\pi s$$

where s is the interelectrode spacing. When conducting a survey with this configuration, it is generally assumed that the measurement relates to a region situated beneath the midpoint of the array and at a depth equal to half the current electrode spacing (eg. Griffiths & King, 1981). By expanding the electrode array about this midpoint, the proportion of current flowing at depth is increased and the apparent resistivity data then provide an approximate vertical log of resistivity change. Moreover, by repeatedly translating and expanding the electrode array, a series of resistivity soundings can be obtained and then merged to yield a classical electrical 'pseudosection' (eg. Griffiths *et al.,* 1990).

Resistivity Tomography

Although instruments are now available which automate the collection of pseudosection data from electrode arrays, a key limitation remains the low spatial resolution achievable by this empirical technique. In the new method outlined here, the first priority is to collect *all possible* independent measurements of apparent resistivity on the array and to use these as the basis for an improved inversion to the structure based on a tomographic backprojection model.

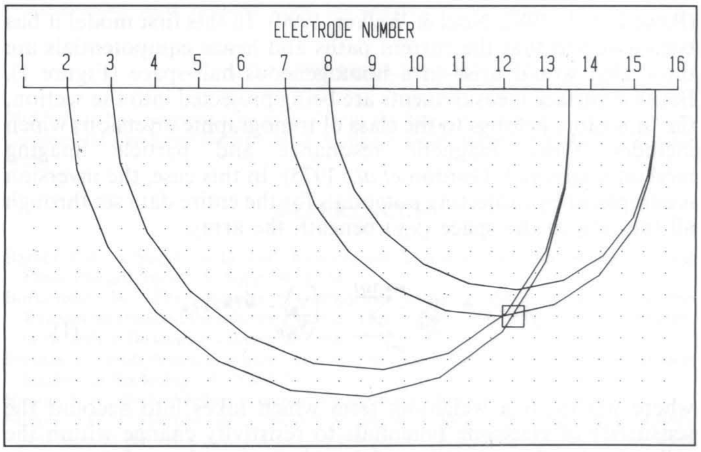


Figure 1. The method of simple backprojection used to generate resistivity sections. Two stages in the image reconstruction are shown for a single cell (small square) beneath an array of 16 electrodes, as an example. The potential measured between electrodes 2 and 3 when current is passed between 13 and 14 is compared to the potential expected for a uniform region. The cell resistivity is then adjusted according to this ratio and the procedure repeated as a sum (see equation 1) for the measurements 7, 8 15 and 16 and all other projections in the data set which intersect the cell.

With an array of N electrodes it is possible to measure a total of N(N-3)/2 independent values of apparent resistivity with the four-probe method or alternatively, the greater number, N(N-1)/2, if a pole-pole configuration is used (Barber & Seager 1987; Noel & Xu, 1991). Thus for an array of 20 electrodes we can collect a 'Data Set' of 170 (or 190) measurements which exceeds the 57 values attainable with a Wenner pseudosection scheme. Geophysical reconstruction of the underlying structure uses a two-dimensional computer model which begins with a uniform vertical section divided into a mesh of rectangular cells. The electrode potentials V_m measured at each step are compared to the potentials V_u expected for a homogeneous subsurface and used to adjust the resistivities of cells which lie between the equipotentials extending

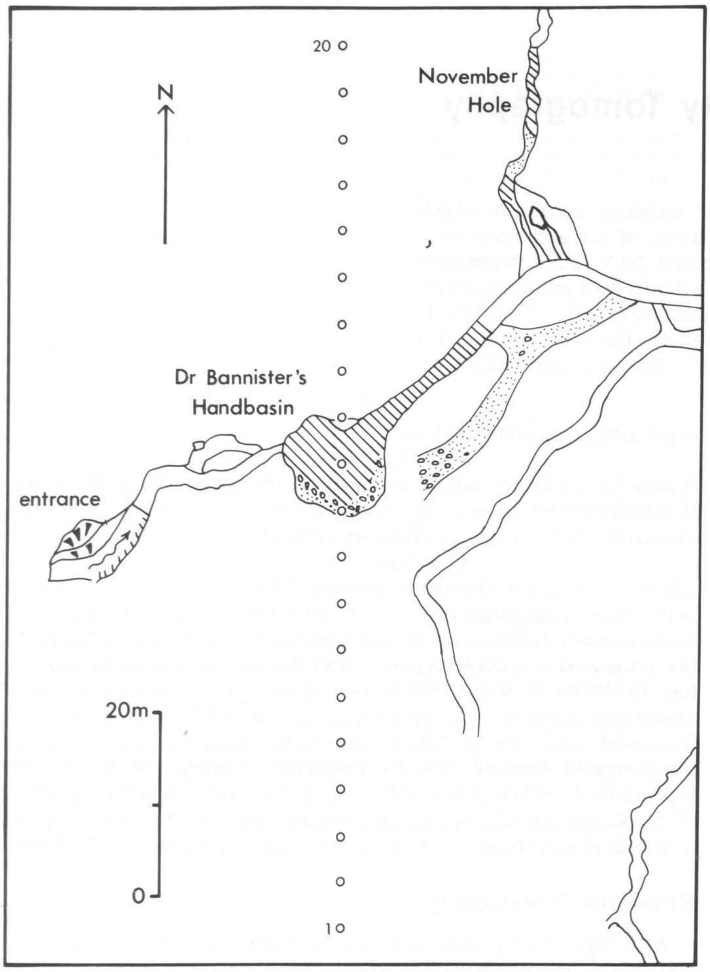


Figure 2. The position of the 20 electrode array (circles) in relation to Dr. Bannister's Hand Basin in the Upper Long Churn Cave system. The electrodes comprised 0.8cm diameter mild steel rods, inserted into the soil to a depth of about 15cm, with a spacing of 5m.

from the measurement electrodes according to the equation

$$\rho_m/\rho_u = V_m/V_u$$

(Powell *et al.,* 1987; Noel & Walker, 1990). In this first model it has been assumed that the current paths and hence equipotentials are those that would arise in a homogeneous half-space (Figure 1). Because surface measurements are being projected into the section, the procedure belongs to the class of tomographic inversions which includes X-ray, magnetic resonance and particle imaging methodologies (c.f. Gordon *et al.,* 1975). In this case, the inversion is completed by projecting potentials for the entire data set through all the cells in the space (x,y) beneath the array:

$$\rho_{x,y} = \frac{\rho_u}{\Sigma w} \cdot \sum_{p=1}^{N(N-3)/2} \left(\frac{V_m}{V_u}\right)_p \cdot w(x,y)_p \qquad (1)$$

where $w(x,y)_p$ is a weighting term which takes into account the sensitivity of electrode potentials to resistivity change within the cell at (x,y) and is a function of the relative positions of the current and potential electrodes and of the underlying cell. Derivation of the weighting function w is beyond the scope of this paper but is described fully in Noel & Xu, (1991). For a given array, this set of weights can be precomputed and stored as a look-up table for use later by the inversion algorithm.

This first model bears obvious similarities to the geometric reconstruction technique proposed and tested by Bristow at Higher Kiln Quarry in Devon (Bristow 1966) although a less heuristic strategy is taken here in the 'projection' of measured anomalies toward the subsurface structure. Our second model takes a more formal approach to the inversion by considering each potential measurement, P_n, within the data set to comprise the convolution of the resistivity structure [R] with the corresponding sensitivities, [S] in (x,y) space,

$$P_n = k\,[S]^*[R]$$

where k is a constant. The resistivity structure is then found from the N measurements in the Data Set thus:

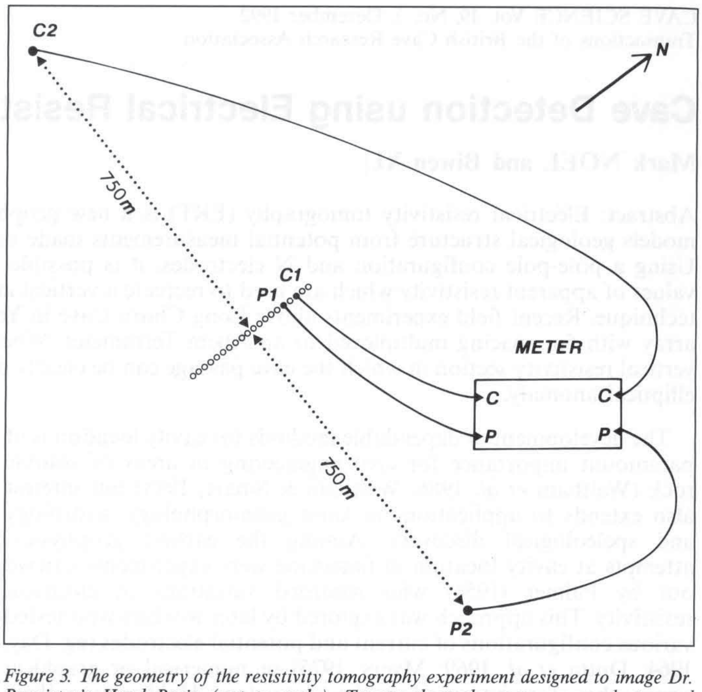


Figure 3. The geometry of the resistivity tomography experiment designed to image Dr. Bannister's Hand Basin (not to scale). Twenty electrodes were arranged at equal spacings along a 95m line over the cave while two more electrodes (C2, P2) were placed about 750m away on opposite sites of the array, i.e. at 'pseudoinfinite' locations. Current was injected between one array electrode, C1, and C2. The potential difference was then measured between P2 and one other array electrode, P1, using an Abem Terrameter. By changing the positions of C1 and P1 on the array it was thus possible to record 190 unique values of potential difference.

$$[R] = \frac{1}{k.N} \sum_{n=1}^{N(N-1)/2} P_n \cdot [W]_n \qquad (2)$$

where $[W]_n = [S]_n^{-1}$

The set of $[W]_n$ can be found from $[S]_n$ by singular value decomposition (since the equations are underdetermined) and these coefficients are again precalculated and stored in order to speed later computation.

These two inversion schemes have been evaluated against numerical model data (Noel & Xu, 1991) while field trials of the simple backprojection model of eqn. (1) have already been conducted over archaeological structures (Noel & Walker, 1990).

FIELD TRIALS AT LONG CHURN CAVE

In designing a field experiment the first requirement was to select a level site underlain by a cave passage with simple, known geometry. A second criterion was to ensure that the electrode array could be accurately located with respect to the target. Finally, a cave passage at moderate depth was sought in order to maximise the likelihood of detection using this preliminary technique. It was decided that a segment of the Long Churn Cave System, near Ingleborough, Yorkshire would fulfil these criteria and location of the test was made by reference to the survey compiled by Milner & Milner (1977).

An array of 20 mild steel electrodes were placed in a line crossing at right angles the projected position of Dr Bannister's Handbasin, a large chamber approximately 20m from the upstream entrance to Upper Long Churn Cave (Figure 2). The inter-electrode spacing was 5m. Installing the array at this site presented some problems because of the restricted choice of soil pockets among the extensive limestone pavement. A pole-pole electrode configuration was employed in which one current and one potential electrode (C1 & P1) were selected on the array while electrodes C2 and P2 are placed sufficiently distant that, from a geophysical standpoint, they appeared to be at infinity. To achieve this objective, it was necessary to insert these remote electrodes some 750m NW and 750m SE of the main array (Figure 3). 190 measurements of apparent resistivity corresponding to the complete data set were then gathered using an Abem Terrameter SAS300B which was connected to the array via a switch box or multiplexer.

The field data were inverted using the model given in eqn. (2) and the results are shown in Figure 4. The passage is clearly visible as a diffuse, high resistance anomaly directly beneath the midpoint of the electrode array. Other, high-resistance features which appear to extend almost to the surface may represent open joints in the

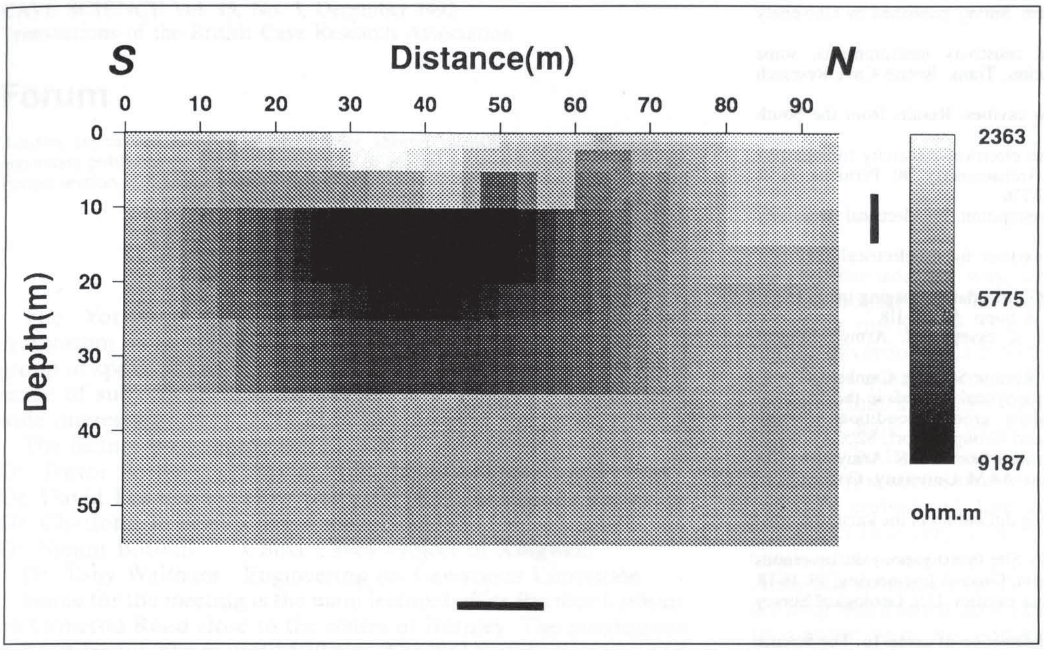


Figure 4. Electrical resistivity tomogram representing a vertical section through Dr. Banniseter's Hand Basin. Darker shades correspond to higher resistivity. Numbers on the grey scale relate to values of resistivity in ohm.metres. Horizontal and vertical scales are equal. The horizontal and vertical bars mark the surveyed position of the chamber (see Figure 2).

roof of the chamber. The geophysical section, however, exaggerates the scale of the cave implying that the passage is ~40m in diameter as opposed to the ~11m estimated from the published survey. This error almost certainly arises from the assumption, implicit in the model, that current paths are undisturbed by inhomogeneities in the subsurface. Nevertheless, although the conductivity contrast is clearly very high and current paths must be distorted, the inversion method has clearly succeeded in resolving the cave passage in vertical section. Improvement to the inversion could probably be achieved through a process of iterative refinement involving current path raytracing or finite difference methods (eg. Griffiths *et al.*, 1990) and these options are being explored.

For the sake of comparison, data corresponding to a Wenner pseudosection survey were computed from the field data by superposition and the contoured result is shown in Figure 5. In this graphic the cave is very difficult to resolve, underlining the improved spatial and resistivity resolution attainable by the tomographic method.

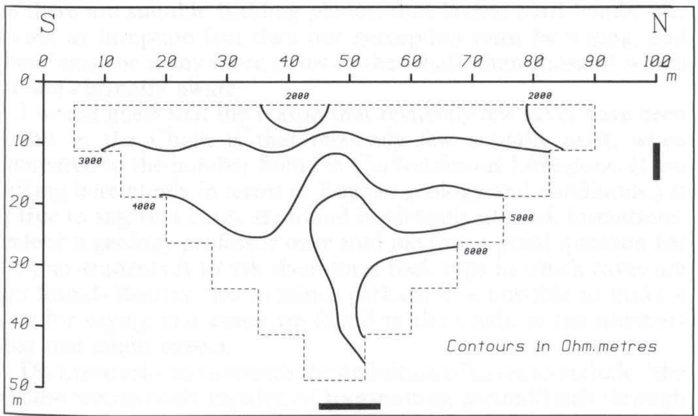


Figure 5. Contoured electrical resistivity pseudosection through Dr. Bannister's Hand Basin based on a Wenner configuration and a total of 20 surface electrodes. Data for this section were obtained by arithmetic superposition from the tomographic data set. The horizontal and vertical bars marks the surveyed position of the chamber (see Figure 2).

CONCLUSIONS

Geophysical detection of natural and artificial cavities is now an important element of site assessments made for environmental, engineering, hydrological and archaeological purposes. Theoretical studies indicate the potential that geoelectric methods have for locating cavities owing to the very large conductivity contrast that occurs between most rocks and the enclosed air. This conjecture is supported by the results of field trials made by the British Geological Survey using resistivity soundings and electromagnetic surveys over natural cavities and mineshafts (McCann *et al.*, 1987; Suddaby, 1982). However, a conventional pseudosection provides a tapering geophysical section with limited spatial resolution at depth, in contrast to an ERT image which is also better constrained by the larger dataset on which the inversion is based.

Results obtained from our preliminary tests indicate that resistivity tomography is able to resolve a cavity at modest depth within limestone terrain whereas, on the other hand, a conventional pseudosection yielded ambiguous results. Further development of the ERT method should be primarily directed at improving both the spatial and conductivity resolution of the inversion model. Here a primary requirement is to increase the size of the input dataset. This can be accomplished by deploying a larger number of electrodes but accurate placement will then be essential if spacing errors and hence image artifacts, are to be avoided (Noel & Xu, 1991). Research in these topics is currently underway and further field experiments with ERT are planned over both deeper and smaller cave passages.

ACKNOWLEDGEMENTS

This research is supported by grants from the Science and Engineering Research Council and The Royal Society. Mr J. Hodgson kindly helped to build the electrode array.

REFERENCES

Barber, D.C. & Seager, A.D, 1987. Fast reconstruction of resistance images. Clinical Phys. Physiol. meas., 8, Supp. A, 47-54.

Borkowski, W., 1990. Results of subsurface radar geophysical studies of the Krzemionki banded flint mines, Poland. Archaeometry '90, Pernicka, E. & Wagner, G.A. (eds.), Birkhauser, Basel, 687-696.

Bristow, C., 1966. A new graphical resistivity technique for detecting air-filled cavities. Studies in Speleology, 1, 204-227.

Brown, B.H., 1986. Applied potential tomography. Phys. Bull., 37, 109-112.

Butler, D.K. & Murphy, W.L., 1980. Evaluation of geophysical methods for cavity detection at the WES cavity detection test facility. U.S. Army Engineer Waterways Experiment Station, Vicksburg, Technical Report GL-80-4.

Colley, G.C., 1963. The detection of caves by gravity measurement. Geophysical Prospecting, 11, 1-9.

Day, J.G., 1964. Cave detection by geoelectrical methods, Part 1, Resistivity. Cave Notes, 6(6), 41-45.

Dutta, N.P., Bose, R.N. and Saika, B.C., 1969. Detection of solution channels in limestone by electrical resistivity method. Geophysical Prospecting, 18, 405-414.

Franklin, A.G., 1977. Effects of subsurface cavities on wavefront diagrams. U.S. Army Engineering Waterways Experiment Station, Vicksburg.

Gordon, R., Herman, G.T. & Johnson, S.A., 1975. Image reconstruction from projections. Sci. Amer., 223, 56-68.

Griffiths, D.H. & King, R.F., 1981. Applied Geophysics for Geologists and Engineers, Pergamon, London.

Griffiths, D.D., Turnbull, J. & Olayinka, A.I., 1990. Two-dimensional resistivity mapping with a computer-controlled array. First Break, 8, 121-129.

Habberjam, G.M., 1969. The location of spherical cavities using a tri-potential resistivity technique. Geophysics, 34, 780-784.

Lange, A.L., 1965. Cave detection by magnetic surveys. Cave Notes 7(6), 41-54.

Lange, A.L., 1972. Mapping underground streams using discrete natural noise signals: a proposed method. Caves and Karst, 14, 41-44.

McCann, D.M., Jackson, P.D. & Culshaw, M.G., 1987. The use of geophysical surveying methods in the detection of natural cavities and mineshafts. Q. J. Eng. Geol., 20, 59-72.

McNeill, J.D., 1980. Electrical conductivity of soils and rocks. Geonics Ltd, Mississauga, Canada, Technical Note TN-5.

Milner, A. & Milner, B., 1977. The Alum Pot System. Survey published by University of Leeds Speleological Association.
Myers, J.O., 1975. Cave location by electrical resistivity measurements, some misconceptions and the practical limit of detection. Trans. British Cave Research Assoc., 2, 167-172.
Noel, M., 1985. Geothermal detection of shallow cavities. Results from the South Wales karst. Geophys. J. R. astr. Soc., 81, 312.
Noel, M. & Walker, R., 1990. Development of an electrical resistivity tomography system for imaging archaeological structures. Archaeometry '90, Pernicka, E. & Wagner, G.A. (eds.), Birkhauser, Basel, pp. 767-776.
Noel, M. & Xu, B., 1991. Archaeological investigation by electrical resistivity tomography. Geophys. J. Int., 107, 95-102.
Palmer, L.S., 1954. Location of subterranean cavities by geoelectrical methods. Mining Mag. Lond., 91, 137-141.
Powell, H.M., Barber, D.C. & Freeston, I.L., 1987. Impedance imaging using linear electrode arrays. Clinical Phys. Physiol. meas., 8, Supp. A, 109-118.
Rinker, J.N., 1974. Infrared thermal detection of caves. U.S. Army Engineer Topographic Laboratories report 74-226.
Scollar, I., 1990. Archaeological Prospecting and Remote Sensing, Cambridge, 347.
Suddaby, D.L. & Hallam, J.R., 1982. The use of geophysical methods in the detection of natural cavities, mineshafts and anomalous ground conditions. British Geological Survey, Engineering Geology Research Group Report, 82/5.
Unterberger, 1977. Radar and sonar probing of subsurface at U.S. Army Corps of Engineers Waterways Experiment Station. Texas A&M University, College Stn., Texas, Special Report.
Waltham, A.C. & Smart, P.L., 1988. Civil engineering difficulties in the karst of China. Q. J. Eng. Geol. Lond., 21, 2-6.
Waltham, A.C., Vandenven, G. & Ek, C.M., 1986. Site investigations on cavernous limestone for the Remouchamps Viaduct, Belgium. Ground Engineering, 19, 16-18.
Watkins, J.S., 1967. Seismic detection of near-surface cavities. U.S. Geological Survey Prof. Paper 599-A.
Wigley, T.M.L. & Brown, M.C., 1978. Geophysical detection of caves. In: The Science of Speleology, Ford, T.D. & Cullingford, C.H.D., (eds.), Academic Press, London.
Wilcock, J.D., 1990. Dowsing at Higher Kiln Quarry. William Pengelly Cave Studies Trust Ltd. Newsletter, 58, 8-11.

Dr M. Noel and Biwen Xu
Department of Geological Sciences
University of Durham
Science Laboratories
South Road
Durham DH1 3LE

CAVE SCIENCE Vol. 19, No. 3, December 1992
Transactions of the British Cave Research Association

Forum

Readers are invited to offer review articles, shorter scientific notes, comments on previously published papers and discussions of general interest for publication in the Forum section of Cave Science.

CAVES SYMPOSIUM

The Yorkshire Geological Society is holding a half-day symposium on Saturday 13th March 1993, with the title 'Caves'. A group of speakers has been invited to give short lectures on a wide range of subjects so that the meeting will appeal to people with wide interests.

The lecture programme is:

Dr. Trevor Ford	Evolution of the Castleton Cave System.
Dr. David Lowe	How Caves Develop: everybody knows.....
Dr. Charlotte Roberts	Cave Archaeology.
Dr. Simon Bottrell	China Caves Project in Xingwen.
Dr. Tony Waltham	Engineering on Cavernous Limestone.

Venue for the meeting is the main lecture hall at Burnley College, on Ormerod Road close to the centre of Burnley. The programme will last from 1.30 p.m. until 5.30 p.m. The Y.G.S. welcomes any and all non-members of the society who wish to join the meeting.

No booking is necessary.

Further enquiries to Paul Kabrna, 0282-813772.

THE "PROBLEM" OF CHALK CAVES

G.J. Mullan

David Lowe (1992) appears to have come up with an interesting theory, that of Inception Horizons, which adds a further factor to those which have previously been deemed necessary for cavern genesis to occur in a given rock formation. These are: low primary permeability, high secondary permeability and a degree of solubility in water. To these we must now add suitable inception horizons, although I rather suspect that these will prove more significant in determining where caves form rather than whether.

When he attempts to apply this theory to cavern genesis in the Chalk, however, he appears to be trying to hold opposing views simultaneously. On the one hand he seems to be saying that the reason for the perceived lack of caves in Chalk is that it is too pure, and too lacking in suitable fractures and bedding planes, to permit of many, if any, inception horizons; but on the other he argues that as there are suitable bedding planes, flint layers, marl bands, etc, to act as inception foci then our perception must be wrong, and there must be many more caves in the Chalk than those of which we are currently aware!

I would guess that the reason that relatively few caves have been found in the Chalk is that relatively few actually exist, when compared to the number found in Carboniferous Limestone. (I am talking here purely in terms of English geology and conditions.) It is true to say, that caves are found in virtually all rock formations. Indeed a geology professor once told me that a good question for 1st year students is to ask them for a rock type in which caves are not found. Bearing this in mind, perhaps it is possible to make a case for saying that caves are found in the Chalk in the numbers that one might expect.

Dr. Lowe asks us to stretch the definition of caves to include "the earliest micro-voids capable of transmitting groundfluids through a previously impermeable rock mass" (Lowe, 1992). I can accept this providing that it is acknowledged that it is the action of the groundfluids that has made the difference between the non-cavernous and the cavernous state, rather than any tectonic process. On this definition caves can be found in such unlikely rocks as quartzitic sandstones (see Mullan, 1989 and Wilson, 1991 for some English examples), where, of the required factors outlined above (ignoring inception foci for the moment) the only one present is high secondary permeability.

If caves can form in such an unlikely material, it therefore follows that a statistically larger number will be found in a chalk formation where two of the required factors, high secondary permeability and suitable solubility, are to be found. Again, it follows naturally from this argument that the greatest number of caves will be found where the greatest number of the preconditions are realised, in the Carboniferous Limestone. This is in accordance with perceived reality.

To put this another way, one can say that in both the theoretical and the actual case, the number of caves to be found in the Chalk stands in the middle of a continuum whose end members are the highly cavernous Carboniferous Limestone and the infrequently cavernous quartzite rocks. It may be that that Dr. Lowe would wish to argue that a smaller number of caves is presently known in the English Chalk than would be deduced from the above discussion. This may be true, I have not carried out a strict statistical analysis, but does not necessarily alter my conclusion as a further factor has yet to be considered. One very good reason for this apparent lack is the fact that very little cave exploration has been carried out in chalk areas, as compared with Carboniferous Limestone areas. After all, until the mid 1980s, when cavers went looking for them, nothing was known about the sandstone caves of Northumberland. In similar vein, his comment that the type of cave that he is describing need never be obvious at the surface would be greeted with wry amusement by any Mendip Digger who has devoted himself to uncovering such hidden caves — in the Carboniferous Limestone.

Dr Lowe states that the primary porosity of chalk (14.4 to 46.0%) is only high when compared with limestones in general (0.67-2.55%) but then goes on to state that sandstones fall into the range 3.32-39.8%. Whilst this supports his statement that even the most compact sandstone is more porous than unfractured crystalline limestone, surely it also indicates that chalks, especially English Chalk at the top end of the range, is even more porous and therefore much less likely to develop the concentrated fissure flows that can lead to cavern genesis.

A further statement of Dr. Lowe's that must be challenged comes in his paragraph about well-sinking and dowsing. My views on dowsing have been aired before, but Barrington and Stanton (1977) give a good account of the problems of dowsing "success stories" that shows the dangers of taking folk lore at face value. My present concern is with the way that consciously or not he gives the impression that he accepts a potential use of this "art" without any real facts to back this up. He says that an unknown proportion of wells have encountered dissolutional voids. This, surely, must follow: if such voids exist in any proportion in the Chalk, which they do, then of course a proportion of successful wells (the ones about which one hears) will encounter them. This is simple probability and requires no arcane arts to fathom it. If it could be shown that dowsed wells were disproportionately successful at striking these voids then it would be worth saying. It follows therefore that there is no evidence at all for the final part of this statement, as to whether such cavities are bedding or fracture guided.

The conclusion to be drawn from the above is straightforward. Dr. Lowe considers that apparent number of caves to be found in Chalk is less than he would predict. However another look at the factors involved leads to the possibility that exploration in the Chalk has discovered about as many caves as might be expected.

REFERENCES

Barrington, N. and Stanton, W.I. 1977. The Complete Caves of Mendip and a View of the Hills. Cheddar Valley Press, Cheddar pp233-4

Lowe, D.J. 1992 Chalk Caves Revisited, Cave Science 19.2 pp55-58.

Mullan, G.J. 1989 Caves in the Fell Sandstone of Northumberland. Proceedings of the University of Bristol Spelaeological Society 18.3 pp430-7.

Wilson, L.J. 1991 The Caves of North Northumberland Vol IV. Newsletter of the University of Bristol Spelaeological Society 7.1 pp6-7.

G. J. Mullan
38, Delvin Road,
Westbury on Trym,
Bristol BS10 5EJ.

THE "PROBLEM" OF CHALK CAVES — A REPLY

D. J. Lowe

The response to my short paper (Lowe, 1992b) has been more immediate and far greater than I expected. Several correspondents wished to know more about the "Inception Horizon Hypothesis" (Lowe, 1992a), others provided further information about caves (some previously unknown to me) in the Chalk and other Mesozoic carbonates of the United Kingdom. In responding directly to these enquirers I have attempted to pass on the gist of my still unproven hypothesis and, when requested, the benefit of my geological expertise, whilst emphasising the limitations of my first-hand experience of Chalk caves. Graham Mullan's response, presented as a contribution to the *Cave Science* "Forum" rather than directly to me, requires some comment, mainly to clarify points made and opinions expressed in the original article. I appreciate the opportunity offered by the Editor of *Cave Science* to provide this response. The opportunity also allows me to emphasise the nature and backgound of my PhD thesis and the several 'papers' and talks drawn directly from it which have already been presented or which may appear in the future.

The 'Chalk paper' was, as stated, taken from a larger thesis and was not designed to present all the arguments behind the Inception Horizon Hypothesis, which is a tentative and largely qualitative lateral reappraisal of earlier discussions of cave formation. This small and more or less self-contained part was exhumed to present (I thought) a manageable introduction to the lateral view, which will be covered more fully by subsequent papers. It is stressed that the parent thesis is a philosophical consideration of much that has been written before, made in the light of wide practical experience during 25 years. It is equally important to note that the term *hypothesis* is used consistently in its literal sense (cf the Concise Oxford Dictionary) of: *"Supposition, made as basis for reasoning, without assumption of the truth, or as starting point for investigation..."*. Although I personally see truth within the Inception Horizon Hypothesis and believe that it explains many of the conflicts and shortfalls of current wisdom, follow-up investigations, which will prove or disprove its principles, are yet to be commenced.

Several points raised by Graham Mullan are subjective, just as some of my own original statements were, and these need not be contradicted in detail here. It is worth illustrating, however, that viewpoint and view may alter if the implications of the Inception Horizon Hypothesis are accepted. For instance, part of the reason that few caves are known within the Chalk might be that no one has bothered to look for them, as Graham suggests. Having spent much of my caving career digging for, rather than exploring, caves, I can sympathise with the comment regarding those who dig in the Mendip Hills. The difference between this situation and that on the Chalk outcrops is that the diggers on Mendip have many indications that caves do exist - only the way into them is obscure. The Chalk caves to which I allude may be too small to explore (sub-anthropocentric), they may be non-integrated (Ford and Ewers, 1978), or any potentially explorable parts may be as yet uncut by the effects of surface erosion. How many of the caves found by digging in the Mendip Hills are non-integrated, and how many "obvious" caves in the Carboniferous Limestone would be visible or explorable if the presently exposed Dinantian outcrops had not been revealed and shaped by removal of overlying beds and the incision of fluvial or glacial valleys? Most of the caves that we do know in the Chalk have been revealed by *recent* marine erosion.

A major element of the Inception Horizon Hypothesis, not laboured in the Chalk context, is a potentially *great* extension of speleogenetic timescales. Inception processes, and hence later phases of cave growth, may thus be deduced to take place in buried carbonate sequences. It becomes a moot point whether the localised cave systems we see today are a function of the present landscape, as most earlier workers have supposed, or represent intersected and modified artifacts of pre-existing and basinally extensive inception systems. My view is that these "modern", explorable, systems have evolved from and are superimposed upon fragments of an original inception skeleton. The existence of this skeleton is, from my viewpoint, made clear by the cave passage growth which was made possible and rendered visible by the imposition of the current landscape. However, even though I accept that basin-wide inception systems still exist, carrying underflow or beheaded underground drainage beyond the truncated fragments we may explore, I do not propose to start digging through several hundred feet of Namurian strata on Ingleborough to locate the hidden and flooded inception routes below. Nor will I be attempting to locate the inception routes which (I believe) extend from the South Wales North Crop, beyond modern truncators such as the Clydach Valley, into the centre of the "South Wales Coalfield Synclinorium", or the waterfilled conduits deduced to exist in the limestones deep beneath the Bath-Bristol Basin.

Some of Graham's comments are probably misunderstandings due to my use of terms and concepts long familiar to me, without adequate explanation. The term "tectonic maturity" for instance was used, without explanation, to describe the different aspects presented by generally young rocks which have undergone little tectonism (in a broad sense) since consolidation, and older rocks, or those younger rocks in active tectonic belts, which have suffered severe and/or repeated disturbance. I certainly did not intend to imply that tectonic processes triggered the step from non-cavernous to cavernous rock. On the other hand, although the Inception Horizon Hypothesis labours the importance of stratigraphical (loosely 'bedding plane') guidance of cave inception, I find it impossible to refute the local involvement of tectonic features, notably faults and joints, as links in the inception skeleton, allowing groundfluids into the rock mass and linking potential inception horizons.

The Inception Horizon Hypothesis and my research thesis in general deal specifically with carbonate speleogenesis and I have not read the papers Graham mentions which describe caves within quartzitic sandstones of the Fell Sandstone Group. In principle, however, I see no reason why caves may not be conceived and developed in any "soluble" rock. I do not intend to discuss what we mean by soluble, but the siliceous cement of a true quartzite is potentially soluble, as is carbonate cement in other types of sandstone or protolithic carbonates in limestones. The processes and timescales involved are different and a suitable mechanism must exist to allow the earliest groundfluid motion (see for example Davis, 1966; Pye and Miller, 1990; Lowe, 1992a). The "quartzite karst" developed on and within the Proterozoic rocks of the Roraima Formation of Venezuela (eg White and others, 1966) is well-known and spectacular. To what extent does the scale of this dissolutional landscape reflect the potentially very great length of time that these rocks have been exposed, and to what extent does it reflect the high local rainfall, or any other factors? Without more detailed study I don't know the answer to these questions. As Graham (and a geology professor) point out, there are few rocks which will not support cave development, and specifically development by dissolution, but the dissolutional processes involved, the magnitude of these effects in the context of each rock type, and the timescales required for the effects to become visible, may be widely different. My considerations have been limited to dissolutional effects within limestones, but in passing I would hazard that whatever rock type is indicated as supporting dissolutional cave development, some horizons within a given 'formation' will be more susceptible to void inception than others.

I don't think that I am trying to hold opposing views simultaneously, as Graham suggests in his second paragraph, but again I fear I may be unable to see the paradox he suggests due to my closeness to the problem. In the context of the Inception Horizon Hypothesis much of the Chalk may be deduced to be too pure to support cave *inception,* but the atypical horizons within the Chalk may also be deduced to present potential inception foci. Thus it is suggested that the lack of caves within much of the unfractured Chalk mass is not so much a reflection of its high primary permeability as of a lack of a dissolutional mechanism to conceive voids in the purer lithologies. This consideration is equally applicable in the Carboniferous Limestone, where inception *within* pure beds is rare (even if bedding planes exist) but inception related to impure beds is common. My arguments make no attempt to deny that subsequent speleogenetic processes allow caves to ramify upwards or downwards into carbonates (or non-carbonates!) of any purity.

In my discussion of porosity and permeability, the important point being made was that the total permeability of Chalk is several orders of magnitude higher than its primary permeability, figures which I consider significant in that they indicate that fissure flow dominates over diffuse flow. This is important in context, since it suggests that the traditional view that there are few caves in the Chalk due to dominance of diffuse flow is erroneous. On the basis of the Inception Horizon Hypothesis I argue that *if* there is "fissure" flow, as permeability values and well provings indicate, there must be "fissures" (though some of these might be "caves"!). There is, however, no dissolutional mechanism within the existing wisdom to allow enlargement of the "fissures" inside commonly accepted speleogenetic timescales. If mixture dissolution *(sensu*

Bögli, 1964) is brandished as the answer, one is faced with the question of why the full thickness of the Chalk "sponge" is not split by dissolutionally enlarged rifts carrying "fissure" flow? Perhaps it is, but we don't perceive them in this way. On the other hand the "fissure" flow in the Chalk might be linked not only to the presence of 'potential' fissures but, just as in other rock types, it might be related to:

a. the intersection of 'potential' fissures and inception horizons, or,
b. in situations where an active inception horizon is truncated by a 'potential' fissure (generally a fault), the flow might utilise the fissure as a bridge to adjacent inception horizons.

In both situations the presence of atypical rock chemistry within or adjacent to the inception horizons could provide the dissolutional environment necessary for void inception and gestation.

On the questions of whether the number of caves known in the Chalk is that which would be expected and whether the number lies in the middle of a continuum, I must also stand by my original statements. This immovability reflects my basic refusal to be led by statistics, the results of a 'science' I freely admit is beyond my comprehension. Whilst I accept that the number of caves that we know about in the Chalk lies somewhere between the number of caves we know about in some other rocks, I would not accept that the end-members are Carboniferous Limestone and quartzite rocks. I know some extensive areas and thicknesses of Carboniferous Limestone which contain fewer "known caves" than does the Chalk, and some with no known caves at all. Since there are "known caves" in quartzite rocks, these rocks must be at the other end of the continuum from the Carboniferous Limestone with no "known caves". In my ignorance I do not see how we can apply statistics to a question with so many variables or unknowns. Surely to be truly valid, statistics of this type would have to include considerations of present and past outcrop area, unit thickness (if you could adequately recognise separate units to consider), rock age, how long the rock had been exposed to speleogenetic processes, the present and palaeo-climates, how these had affected the originally consolidated unit to produce the sequence we see preserved today, how many people had spent how many hours looking for caves in each rock type, etc - not to mention deciding what constitutes a cave in the first place. As I said, I don't claim to understand statistics, so maybe I am simply wrong, but all I said was that the Inception Horizon Hypothesis allows the prediction of many more caves in the Chalk than we currently know, and that these caves may not be obvious to those who merely examine the land surface.

Before closing this response with a brief discussion of my reference to "dowsing" I will try to defend one point which was challenged by Graham on the basis of his views of "dowsing" and his faith in statistics. Graham states that there is no evidence to support my view that some of the dissolutional cavities located by well sinkers, *"...must have been guided by bedding partings rather than fractures."* (Lowe, 1992b, p.56). Firstly, although I accept Graham's comment that the sample information which has survived passage into folklore is not statistically valid, I must point out that I did not claim this validity; nor does its invalidity affect my conclusion. I state that an *unknown proportion* of [all the] wells sunk have encountered dissolutional cavities, and it is this striking of "caves" which is important, not the failure to strike them in other holes or the relative proportions of hits and misses. Earlier in the article I had discussed the potential importance of "fissure" flow, responsible for the formation of these water-filled cavities and, by implication, the non-integrated (natural entrance-less) caves described in France. Having convinced myself that "fissure" flow is common within the Chalk and that some of the "fissure" flow is in reality flow within dissolutional conduits (as illustrated by those cavities which have been explored) I concluded that some of the voids were guided by bedding related rather than simply structural features. This judgement was, indeed, subjective and possibly suspect, being based upon my own (uncited) experience of the relative importance of "fracture-guided" and "bedding-guided" passage segments in rocks of many lithologies and many ages. Though I have explored caves (particularly in Tonga) which are *almost* totally fracture-guided, even here *"at least some"* (to use the phrase I used in the Chalk paper) passage segments are bedding-guided. Later in the paper I discussed examples of bedding-guidance within the Chalk which have been seen and described, though not by me. Although I might have put the cart before the horse in the paper, I still consider that my statement is valid.

My thesis includes several non-committal references to the (black) art of dowsing. I have not studied the supposed evidence for the validity of this "technique", nor listened to the sometimes hysterical cries of those who condemn its practitioners. Whether or not dowsing is unscientific and unsubstantiated quackery, as some supposedly objective observers would insist, a great many individuals practice the "technique". [In passing I can reminisce about how, some 25 years ago, my 6th Form cronies and I laughed and laughed at a pamphlet describing acupuncture, which sounded like Oriental torture but is now becoming more and more accepted by the medical establishment]. Much of what is written and reiterated about dowsing can readily and conveniently be dismissed as "folklore", yet elements of "real" science appear to interfinger with the mythology (eg Lange and Kilty, 1992). Any scientist proposing a new hypothesis or describing his own, lateral interpretations of other peoples' hard-won data must be to some degree egocentric - I am not *sufficiently* egocentric to dismiss (some elements of) dowsing out of hand and I prefer to remain agnostic. If my references to dowsing, which stressed the folklore element, have upset other scientists I would caution them to examine their own objectivity. Whilst having neither need nor desire to support the cases for or against "dowsing" I will make two points. Whether it is by coincidence or as a consequence of the "infinite number of monkeys" theory, as some other function of the laws of probability or for any other reason, some published "results" present a remarkably close approximation to predictions of *regional* underground flow routes made on the basis of geological study and the Inception Horizon Hypothesis. On a more general level, I find it difficult to rationalise the conflict between the concept that a "creature of little brain" such as a pigeon, can navigate whilst blindfolded, and the possibility that similar but generally unexploited abilities are hidden away somewhere within the complex and highly convoluted human brain.

All of that being said, I have accepted the criticism that my treatment of dowsing in the Chalk article is not (and cannot be) statistically valid, based as it is upon folklore. At the same time and at the risk of seeming flippant, I must add that the science (or ?art) of statistics appears far more arcane and frightening to the uninitiated than dowsing, and has demonstrably been responsible for far more human misery. Attempts to apply statistics to some elements of cave science (eg Curl, 1965; Rauch, 1972) can be criticised in much the same way as my own treatment, in that they apply the science of statistics to a subset of "all caves", which can be described as "known caves", with no allowance for the possible size or content of the undiscussed (and patently unfathomable) subsets of "unknown caves" and "unexplorable, sub-anthropocentric caves".

REFERENCES

Bögli, A, 1964. Mischungskorrosion - ein Beiträg zur Verkarstungsproblem. *Erdkunde,* Vol.18, 83-92.

Curl, R L, 1965. Caves as a measure of karst. *Journal of Geology,* Vol.74, 798-830.

Davis, S N, 1966. Initiation of groundwater flow in jointed limestones. *National Speleological Society Bulletin,* Vol.28, No.3, 111-118.

Ford, D C and Ewers, R O, 1978. The development of limestone cave systems in the dimensions of length and depth. Canadian Journal of Earth Sciences, Vol.15, 1783-1798.

Lange, A L and Kilty, K T, 1992. Natural-Potential responses of karst systems at the ground surface. *Proceedings of the Third Conference on hydrogeology, ecology, monitoring and management of ground water in karst terranes, Nashville, 1991.* [Dublin, Ohio: National Ground Water Association.]

Lowe, D J, 1992a. The origin of limestone caverns: an inception horizon hypothesis. Unpublished PhD thesis, Manchester Polytechnic.

Lowe, D J, 1992b. Chalk caves revisited. *Cave Science,* Vol.19, No.2, 55-58.

Pye, K and Miller, J A, 1990. Chemical and biochemical weathering of pyritic mudrocks in a shale embankment. Quarterly Journal of Engineering Geology, Vol.23, 365-381.

Rauch, H W, 1972. The effects of lithology and other hydrogeologic factors in the development of solution porosity in the Middle Ordovician carbonates of central Pennsylvania. Unpublished PhD thesis, Pennsylvania State University.

White, W B, Jefferson, G L and Haman, J F, 1966. Quartzite karst in south-eastern Venezuela. International Journal of Speleology, Vol.2, 309-314.

D. J. Lowe
23 Cliff Way
Radcliffe on Trent
Nottingham NG12 1AQ

UNDERGROUND DATA LOGGING FOR CAVE SCIENTISTS

Mike Bedford and David Gibson, of the BCRA's *Cave Radio & Electronics Group* are drawing up a specification for some underground data logging equipment. We envisage the construction of some very small, low power, "intelligent sensors" using specialised integrated circuits not yet available on the "amateur electronics" scene, and software we have already developed for related applications. The sensors will be the size of a 35mm film container, or a box of matches, and will be fully cave-proof. The sensor will collect perhaps 40,000 16 bit data values, for example one reading per minute for 28 days. The sensors will be programmable via an infrared serial link from a hand-held terminal; and this will also be the route for the dumping of collected data.

Hand-held "ruggedised" personal computers are becoming available, and often feature an infrared link as standard. Our Group has field tested two such computers. Programming the sensor will, in the more advanced "models", allow the user to write short control programs in Basic, or to specify the sample frequency, and so on.

At this stage, we need guidance on the worth of a project such as this. We can imagine all sorts of data that could be collected, but we do not know whether it would be of use to a researcher. Examples are air temperature, pressure, humidity and flow rate; water temperature, flow rate, depth, turbidity, pH, conductivity, fluorescence. We would be grateful for any comments on the worth of monitoring these, or any other, parameters; on the required frequency of measurement and so on. If anyone has a specific project in mind we will try to incorporate it into our specification. However, I should stress that these sensors are not going to be available for a while. Much depends on the precious commodities of time and money.

David Gibson
Cave Radio & Electronics Group
12 Well House Drive, LEEDS, LS8 4BX

BCRA CAVE RADIO & ELECTRONICS GROUP

Index to Newsletters 6-9

The CREG is a "special interest group" of the BCRA. It now produces a quarterly journal of some 25 pages, together with a quarterly newssheet. Its aim is to encourage the development and use of radio communications and other electronic and computer equipment in caving and related activities. Apart from cave radio the Group also covers surveying, lighting, photography and cave detection.

Prior to September 1992 the journal and newssheet were combined into a single newsletter. Newsletters 6-9 were published quarterly between December 1991 and September 1992. This index includes abstracts or introductions to each article where these were provided, otherwise just a brief description of the contents. An index to issues 1-5 is in preparation.

Reprints of all publications are available. For information about the Group please write to the secretary, David Gibson, 12 Well House Drive, Leeds LS8 4BX.

Cave Radio

Bibliography of Underground Communications, Nick Williams
This 30 page publication, listing over 350 references to cave radio and related work, was issued as a supplement to newsletter 7, and is now available separately.

Field Pattern of a small magnetic dipole, David Gibson, No.6, pp1-2
Gibson explains how the the field lines behave and how interaction between the induction and radiation fields causes circular polarisation which upsets radio-location. Non-mathematical. Gives a list of further reading.

A guide-wire communications system - part 1, Dick Glover, No.6, pp6-7
Glover recalls his involvement in early work on guide wire radio.

How big should a loop antenna be?, David Gibson, No.6, pp12-14
"Antenna sizes for cave radio receivers are governed by the required signal to noise ratio as well as transmission distance and available power. When a receiving loop exceeds a certain size (and is matched for noise) then the range depends only on the transmitting antenna size and power. There is therefore a case for using small portable loops for receiving, and larger antenna for transmitter stations." *Very mathematical*

Optimal design of a cave radio - summary, David Gibson, No.7, p5
Summary of the concept of Optimal Design to be developed in future articles

Footnote on David Gibson, No.8, p7
Introduces the first of a series of articles on optimal design

UK VLF licence regulations, David Gibson, No.7, p5
Gibson summarises the regulations, giving references, and querying what the regulations are supposed to mean.

Are Q-multipliers any use for cave radio? David Gibson, No.7, pp8-10
"Q-multipliers are often used in cave radio designs. Their purpose is to increase the selectivity of the tuned receiver loop. A common design relies not on explicit circuit action but on parasitic effects. The operation of this design, and a common r.f. Q-multiplier are explained. It is argued, though, that the use of Q-multipliers is not necessary, and that they do not provide any useful addition to a low-frequency receiver.

Modulation methods and signal/noise ratio, David Gibson, No.8, pp1,8-11
"Baseband audio signals are modulated onto a carrier for radio transmission. The performance of a communications link can be expressed in terms of the range achieved for a given transmitter power drain. If the link has been designed to that receiver noise and antenna thermal noise are insignificant then the range is limited by atmospheric noise. It can be shown that the significance of this noise depends not only on the modulation method but also on the type of circuit used for demodulation. Types of modulation and demodulation are summarised (some reader knowledge of the methods is assumed, and their merits (signal/noise ratio, circuit complexity) are compared.

SSB modulation review, David Gibson, No.8, pp12-16
"This article gives a brief comparison of the three common methods of single sideband modulation, and introduces a fourth method based on digital filtering. SSB modulation is commonly achieved by either i) DSB modulation and r.f. filtering, ii) the 'out-phasing' or 'Hilbert Transform' method of generating quadrature audio signals using a wideband phase-shift network, and iii) the so-called 'third method' of generating quadrature intermediate frequency signals. The fourth method, proposed here, is a digital filter implementation of a wideband phase-shift network.

Expedition Radio, Wookey, No.8, pp20-22.
"In 1990 Cambridge University Caving Club (CUCC) was lent some 150MHz radios by Philips and found them extremely useful. The year after we had to make do with significantly less expensive 27Mhz CBs. This report should give you some idea of exactly what can be achieved with such equipment, and of the pros and cons of taking radios [for underground guide-wire use] on your next expedition.

A digital DSB×AM modulator - introduction, David Gibson, No.8, p24
Introduction, outlining a PWM digital modulator for DSB and AM transmission which was subsequently described in No. 10.

Review of the Watsonlinc, David Gibson, No.9, p8
Review of a device to aid telephone communications.

An aluminium tape antenna, David Gibson, No.9, pp21-26
"Aerials for cave radios are usually constructed by winding a loop from as many turns as possible of copper wire. A loop made of a few turns of Aluminium tape is a novel alternative. This design arose by considering the concept of the 'performance' of a cave radio in terms of range and time of transmission related to 'cost-factors' such as the mass of the antenna. With this approach the number of turns is not a factor; and aluminium performs twice as well as copper.

Cave Radio, David Gibson, Cambridge Underground 1990, pp28-33
A nontechnical account of how induction radio works. References to others' work are made and discussed.

Number of turns has no effect on loop antenna performance, David Gibson, reprint in C.U.90, pp34-37
"The performance of a small, very low frequency, loop antenna is derived in terms of the transmitted field strength obtained in return for cost factors such as mass and power. It is shown that the number of turns is not a contributory factor. It is also shown that aluminium performs better than copper in this respect. For a receiving antenna the number of turns does not affect the signal to thermal noise ratio." *Very mathematical*

Smaller battery can improve loop antenna performance, David Gibson, reprint in C.U.90, pp38-41
"The performance of a small low frequency transmitting loop antenna is related to the signal strength and transmission time obtained for a given antenna and battery mass. It is shown that for a given total mass, the best performance is obtained when the battery mass is equal to the antenna mass. If the system consists of a heavy battery and light antenna there is an advantage in reducing the mass of the battery and building a heavier antenna. The subjects of matching the amplifier to the antenna, and choice of battery type are mentioned briefly." *Very mathematical*

Radio amateurs' VLF band, No.6, p4

Miscellany, Peter Eggleston, No.8, p4

Guidewire system uses CB radio, Peter Eggleston, No.8, p4-5

Q-multipliers, Frank Reid, No.8, p5

Emitter tracks cave cracks, [Molefone application], No.9, p7

Software Reviews

Surveyor 88 (QMC), Phil Ingham, No.6, p16
A computer application system for the processing of cave survey data (D.A.Crowl), Phil Ingham, No.6, p18
Digital representations of Karst (J.A. Ganter), Phil Ingham, No.6, p19

Surveying Hardware

Notes on a laser gyrocompass, David Gibson, No.6, p8-9
Principles behind a laser gyrocompass using the Sagnac effect. Gives references.

Accuracy problems in ultrasonic and light-beam rangefinders, David Gibson, No.6, p9
Humidity and temperature affect the speed of both sound and light. This can be significant.

Improving directivity in ultrasonic rangefinders, David Gibson, No.6, pp9-10
A reflector is suggested.

A laser rangefinder, circuit sketch, David Gibson, No.6, pp10-11
"Frank Reid and I both commented in issue 3 on a method of making a laser rangefinder. These notes attempt to explain the principle in more detail.

Phased-array sonar for rangefinding, David Gibson, No.7, p1,2
A suggestion (only semi-serious) for using an array of small transducers to produce a directed beam.

Navigational Aids for cave surveyors, Dick Glover/David Gibson, No.8, pp22
Glover resurrects an idea from 1965. Gibson explains why it would be difficult to implement, being based on time-of-flight measurements through rock.

GIS, KIS & Fractals - the Xanadu effect, Dick Glover, No.8, pp23-24
Glover pokes fun at geographers

Global Positioning Underground, David Gibson, No.9, pp18-20
"The Global Positioning System (GPS) uses a hand-held receiver to pick up signals from four of the 24 orbiting GPS satellites and uses the information to calculate the location of the receiver to a high degree of accuracy, of the order of meters and less. GPS satellite signals have a frequency in the GHz range and will obviously not penetrate underground. Is it possible to design a GPS system which will work underground?. The answer is probably "no", for reasons which *will* be explained. However, other forms of global positioning are possible. One such system uses accelerometers and gyrocompasses to work out distance and direction traversed. Solid state gyros are now available for under £200.

Surveying Software

Uniform survey data - letter, Stuart France, No.6, p15
France proposes a uniform method for the presentation of cave survey data.

SMAPS surveying software, No.8, p2

Surveyors' Group, No.9, p2

Survey Software, Wookey, No.9, p4-5

Lighting

Digestive biscuits, FX2s and abseiling, David Gibson, No.6, p7
An idea for charging a battery whilst abseiling, and more . . .

The aven explorer, David Gibson, No.7, p7
A very bright lamp can use small batteries if it is only required to last a short time. NiCds can provide the high current needed.

A very big flashgun, David Gibson, No.9, p13-14
How to connect lots of little flashguns together

LED lighting (Speleonics) No.9, p2,7

Lighting, Martin Black, No.8, p4

Alkaline cells for lighting, David Gibson, No.9, p5

Cave Detection

Cave detection with $\rho,\sigma,\mu,\varepsilon$ - part 1, David Gibson, No.7, pp7,10

Cave detection with $\rho,\sigma,\mu,\varepsilon$ - part 2, David Gibson, No.8, pp6-7
Gibson outlines the problems in detecting changes in these physical parameters, and makes some suggestions about the electronics.

Project Greensites, review by David Gibson, No.8, p16
A review of the SWCC project to investigate cave detection methods.

Geophysical Techniques for cave detection - part 1, Phil Ingham, No.8, pp17-19
Ingham explains some techniques and how relevant they may be to cave detection.

Subsurface interface radar, p24
Extract from a book which describes how this method was used in an archaeological study.

Cave Detection, Mark Noel, No.8, p4

Dowsing, John Wilcock, No.9, p4

Batteries

Tips for using NiCd cells, David Gibson, No.6, p7
Brief notes on potting and charging

No-nonsense NiCd charging, David Gibson, No.9, pp15-17
How to deal with the latest type of NiCd cells. Article includes a circuit for a "cell zapper" designed to fuse the whiskers which can grow across cell plates.

A battery Discharger, David Gibson, No.9, p17
A circuit which will discharge a battery to a measured end-point then switch off.

Drills

Bosch Drill Battery Duration Tests, Mark McLean, No.9, p9

Battery Powered Drills in Caves, Nick Williams, No.9, pp9-,127
How to lose the guarantee on your Bosch drill and related horror stories.

News and letters

Guidelines for contributors, No.6,p5

Radio amateurs' VLF band, No.6, p4

SMAPS surveying software, No.8, p2

Surveyors' Group, No.9, p2

LED lighting (Speleonics) No.9, p2,7

Bat Detector, No.9, p7

Cave Detection, Mark Noel, No.8, p4

Lighting, Martin Black, No.8, p4

Miscellany, Peter Eggleston, No.8, p4

Guidewire system uses CB radio, Peter Eggleston, No.8, p4-5

Q-multipliers, Frank Reid, No.8, p5

Dowsing, John Wilcock, No.9, p4

National Association of Mining History Organisations, Peter Eggleston, No.9, p4

Survey Software, Wookey, No.9, p4-5

Alkaline cells for lighting, David Gibson, No.9, p5

Emitter tracks cave cracks, [Molefone application], p7

Spring [CREG] Symposium report, Peter Eggleston, No.9, p6-7

B.C.R.A. Research Funds and Grants

THE JEFF JEFFERSON RESEARCH FUND

The British Cave Research Association has established the Jeff Jefferson Research Fund to promote research into all aspects of speleology in Britain and abroad. Initially, a total of £500 per year will be made available. The aims of the scheme are primarily:

a) To assist in the purchase of consumable items such as water-tracing dyes, sample holders or chemical reagents without which it would be impossible to carry out or complete a research project.
b) To provide funds for travel in association with fieldwork or to visit laboratories which could provide essential facilities.
c) To provide financial support for the preparation of scientific reports. This could cover, for example, the costs of photographic processing, cartographic materials or computing time.
d) To stimulate new research which the BCRA Research Committee considers could contribute significantly to emerging areas of speleology.

The award scheme will not support the salaries of the research worker(s) or assistants, attendance at conferences in Britain or abroad, nor the purchase of personal caving clothing, equipment or vehicles. The applicant(s) must be the principal investigator(s), and must be members of the BCRA in order to qualify. Grants may be made to individuals or small groups, who need not be employed in universities, polytechnics or research establishments. Information and applications for Research Awards should be made on a form available from S. A. Moore, 27 Parc Gwelfor, Dyserth, Clwyd LL18 6LN.

GHAR PARAU FOUNDATION EXPEDITION AWARDS

An award, or awards, with a maximum of around £1000 available annually, to overseas caving expeditions originating from within the United Kingdom. Grants are normally given to those expeditions with an emphasis on a scientific approach and/or exploration in remote or little known areas. Application forms are available from the GPF Secretary, David Judson, Rowlands House, Summerseat, Bury, Lancs. BL9 5NF. Closing date 1st February.

SPORTS COUNCIL GRANT-AID IN SUPPORT OF CAVING EXPEDITIONS ABROAD

Grants are given annually to all types of caving expeditions going overseas from the U.K. (including cave diving), for the purpose of furthering cave exploration, survey, photography and training. Application forms and advice sheets are obtainable from the GPF Secretary, David Judson, Rowlands House, Summerseat, Bury, Lancs. BL9 5NF and must be returned to him for both GPF and Sports Council Awards not later than 1st February each year for the succeeding period, April to March.

Expedition organisers living in Wales, Scotland or Northern Ireland, or from caving clubs based in these regions should contact their own regional Sports Council directly in the first instance (N.B. the closing date for Sports Council for Wales Awards applications is 31st December).

THE E. K. TRATMAN AWARD

An annual award, currently £25, made for the most stimulating contribution towards speleological literature published within the United Kingdom during the past 12 months. Suggestions are always welcome to members of the GPF Awards Committee, or its Secretary, David Judson, not later than 1st February each year.

BRITISH CAVE RESEARCH ASSOCIATION PUBLICATIONS

CAVE SCIENCE — published three times annually, a scientific journal comprising original research papers, reviews and discussion forum, on all aspects of speleological investigation, geology and geomorphology related to karst and caves, archaeology, biospeleology, exploration and expedition reports.
Editor: Dr. Trevor D. Ford, 21 Elizabeth Drive, Oadby, Leicester LE2 4RD. (0533-715265).

CAVES & CAVING — quarterly news magazine of current events in caving, with brief reports of latest explorations and expeditions, news of new techniques and equipment, Association personalia etc.
Editor: A. Hall, 342 The Green, Eccleston, Chorley, Lancashire PR7 5TP. (0257-452763).

CAVE STUDIES SERIES — occasional series of booklets on various speleological or karst subjects.
Editor: Tony Waltham, Civil Engineering Department, Trent Polytechnic, Nottingham NG1 4BU. (0602-418418, ext. 2133).

No. 1 Caves & Karst of the Yorkshire Dales; by Tony Waltham & Martin Davies, 1987.

No. 2 An Introduction to Cave Surveying; by Bryan Ellis, 1988.

No. 3 Caves & Karst of the Peak District; by Trevor Ford & John Gunn, 1990.

CURRENT TITLES IN SPELEOLOGY — annual listings of international publications.
Editor: Ray Mansfield, Downhead Cottage, Downhead, Shepton Mallet, Somerset BA4 4LG.

CAVING PRACTICE AND EQUIPMENT, edited by David Judson, 1984.

LIMESTONES AND CAVES OF NORTHWEST ENGLAND, edited by A. C. Waltham, 1974. (out of print)

LIMESTONES AND CAVES OF THE MENDIP HILLS, edited by D. I. Smith, 1975. (out of print)

LIMESTONES AND CAVES OF THE PEAK DISTRICT, edited by T. D. Ford, 1977. (out of print)

LIMESTONES AND CAVES OF WALES, edited by T. D. Ford, 1989.

Obtainable from B.C.R.A. Sales
B. M. Ellis, 20 Woodland Avenue, Westonzoyland, Bridgwater, Somerset TA7 0LQ.