The development of limestone cave systems in the dimensions of length and depth

by

D.C. FORD¹ and R.O. EWERS²

INTRODUCTION

During the 40-50 years since the «classical» papers upon limestone cavern genesis (page 5) were published, the exploration and topographic mapping of solutional cave system throughout the world has greatly increased. Today known systems possess great lengths and depths of galleries in complex three-dimensional array. The Hölloch, Switzerland (Figure 1) is an excellent example, containing 140 + km of galleries and 800 m of relief above further basal, waterfilled passages that are inaccessible. Yet no general solution to the problem of three opposed opinions set forth in the classical papers has been presented in widely circulated geological literature.

As in other fields of study, it is convenient to divide the three dimensional problem into two-dimensional analyses. This paper is concerned with cavern genesis in the dimensions of length and depth i.e. in long section between groundwater sinkpoints and springs. These are the dimensions of principal concern in the classical papers. Principles of cavern genesis in plan view (length and breadth) will be considered in a later paper.

A satisfactory academic definition of a cave is difficult to devise. The Oxford English Dictionary describes it as «an underground hollow usually with an horizontal opening». As such, caves occur in most types of rock and in every geographical region of the world. But it is established that the greatest caves and the greatest regional densities of caves are found in soluble rocks. They are karst features, existing because of the predominant importance of aqueous

¹) Department of Geography, McMaster University
²) Museum of Natural History, West Palm Beach, Florida, U.S.A.
solution processes. The majority of known karst caves, including all save one of the 30 lengthiest systems, occur in limestone. There are extensive karst caves in other soluble rocks such as dolomite, gypsum, calcareous sandstone, etc., and even in quartzites, (Szczerban and Urbani 1974) but these tend only to reproduce some of the morphologic features of limestone caves with few additional significant features peculiar to themselves.

A karst cave is here defined as a solutional conduit or other void that possesses dimensions large enough for turbulent flow of water to occur. From Howard (1964), the minimum dimensions of such a conduit, if of circular cross-section, will vary from 5 to 16 mm diameter under the range of all likely hydrologic gradients. Voids that are too small for turbulent flow to occur are considered to represent a «pre-cave» state. Although caves by this definition will permit turbulent flow, it must not be supposed that the flow is turbulent within them. Many very early caves may evolve entirely within a laminar flow regime. To this extent the definition is unsatisfactory for it utilises an arbitrary dimensional cut-off point.

A karst cave system comprises one or more caves extending continuously between inputs and outputs.

Figure 2 illustrates these definitions. Solutional caves are of two kinds: Figure 2a shows isolated linear or irregularly shaped voids that are not connected to any inputs or outputs by conduits of the minimum dimensions specified above. In scale, these non-integrated caves may range from small vugs to isolated very large rooms encountered in mining operations: Krason and Wojcik (1965) cite cavities of 3,000 m$^3$ capacity. Bauer and Zöttl (1972) describe other recent instances. Such caves often have an origin and development that is distinct from that of the other kind of caves or of cave system, being specific productions of diagenetic, syngenetic processes, etc., or some combination of these processes together with solution by meteoric waters. These caves are not considered further.

Figure 2b shows two simple integrated caves. They are connected continuously to an input and an output respectively. Within the rock, each terminates in fissures of pre-cave dimension into which the groundwater is dispersed or from which it is collected. Figure 2c illustrates a fully integrated system; cave conduits are continuous between input and output.

Like the classical and recent literature cited below, this paper is concerned
with solutional cave systems rather than the simpler caves. Probably most caves that have attained explorable dimensions are portions of fully integrated system, (Figure 2c). But it is important to understand that the most significant characteristics of their three-dimensional skeletons may by determined when they are only partially integrated as in Figure 2b, and inaccessibly small.

As drawn, the caves and system of Figure 2 are of the simple singlephase type: the location of their inputs and outputs has not changed in any of the dimensions during their evolution. The great majority of larger, accessible cave systems appear to be multi-phase: due to relocation of the inputs and/or outputs over time, they possess drained and inactive conduits above others that chan-
nel the modern groundwater flow. An analogy may be made with the organization of drainage paths and channels in a conventional river basin. Overland flow or interflow there occupies the same place in the system as «pre-cave» groundwater circulation does underground. Overland, etc., flow surfaces occupy more than 95% of the area of a river basin except in badlands; similarly, fissure surfaces channelling pre-cave flow are much greater in extent within a limestone mass than are cave systems. The speleological equivalent of badlands where percentage channelled area is great is the maze cave with very closely spaced passage connections, e.g., Anvil Cave, Alabama, (Vernedoe 1964). Like badlands in their realm, the dense maze is comparatively rare underground. Retention hollows and the larger dewponds at the surface, lacking channelled connection with the channel net, are equivalent to the non-integrated vugs and rooms of Figure 2a. Developing or fully developed channel nets are analogous to the simple caves and systems of Figure 2b/c. The equivalent of the three-dimensionally complex, multi-phase cave system such as the Hölloch is the river basin with many abandoned channels on its floodplain and remnants of older floodplains, (suites of terraces), rising above it.

It is a most difficult task to detect and measure true First Order channels in a surface stream net. The most powerful point to be drawn from this analogy is that underground the speleologist rarely detects and cannot measure directly conduits that are of order lower than, probably, Fourth or Fifth. They are too small to be explored; physically they may only be measured in the grossest approximation by means such as pump test and pulse analysis, (Ashton 1966). Therefore, the system information that is available underground is never as complete as that which may be obtained in a surficial basin and the speleologist must proceed more frequently by inference, (Ford 1971).

**Review of the literature**

Substantial reviews have recently been presented by Renault (1967), Jennings (1971) and Sweeting (1972) whilst Bleahu (1974) gives a very full analysis. Only a short summary is needed here. Figure 3 illustrates the conflicting arguments of the classical investigators:

i. (Figure 3a), Martel (1921), Dwerryhouse (1907) and many other early writers supposed that development of explorable caves was limited largely to a vadose drainage zone above a pre-existing or rapidly established watertable. Because this is the zone first encountered by sinking waters, their solvent capacity is optimum in it whilst it is also the zone of greatest velocities of groundwater flow and therefore of greatest potential for cave enlargement by mechanical corrosion.
ii. (Figure 3b). From a wide survey of published materials plus field studies in Missouri, W. M. Davis (1930) and Bretz (1942) concluded that the great majority of large caves or systems had developed at random depth in the phreatic zone, i.e. under permanently waterfilled conditions. Davis proposed a model in which cave-developing ground-water streams followed the deeply curving stream tubes of a Darcy flow net beneath the watertable, i.e. Darcy laws apply until the system is well established and large.

Figure 3. Diagrammatic long sections to illustrate cave system development as proposed by the vadose, deep phreatic and watertable theories.
iii. (Figure 3c). Swinnerton (1932) argued that in a solutionally modifiable medium such as limestone, Darcy’s laws do not apply because even at the outset conditions are anisotropic and inhomogeneous. A given slug of groundwater may disperse into a number of pre-cave routes. That which is shortest, (closest to an already definable watertable in Swinnerton’s geometry), will win in a route competition. The locus of maximum cave development is along or proximate and parallel to, the watertable and the system is propagated from the input end downstream towards the spring.

Rhoades and Sinacori (1941) also advocated a watertable hypothesis but contended that Darcy’s laws apply at inception and that the maximum rate of cave excavation occurs where the Darcy stream tubes converge at the spring. As a result, a gently graded cave conduit that fixes the local watertable is propagated from the spring headwards into the limestone. Most later authors have considered that this paper and that of Bretz (1942) mark the end of the period of formulation of general cave genetic hypotheses. Figure 3 indicates that the hypotheses are mutually contradictory. Although acknowledging the natural variety of cave systems, each of the «classical» writes supposed that his preferred hypothesis explained cave genesis for the general case, reducing any exception to the category of special cases.

Later research by English-speaking speleologists has tended to support the watertable genetic hypotheses (Moore 1966). Sweeting (1950) working in Yorkshire, England, Davies (1959, 1960) in Virginia, Wolfe (1964) in West Virginia, have presented local and regional field evidences and Bedinger (1966) argued from the results of electrical analog simulation of groundwater flow in a fracture environment. Thrailkill (1968) carefully evaluated hydrodinamic and solutional mechanisms and could find no cause for preferred cavern enlargement proximate to a watertable but from observations in Mammoth Cave, Kentucky, argued that the dimension of depth is minor when compared to that of length in a system, that all phreatic flow is therefore proximate to a watertable, and proposed a model that is a compromise of the views of Swinnerton (1932) and Rhoades and Sinacori (1941). His argument begs the question to an extent for if a cave conduit is only a few metres below a watertable but possesses a long profile that is quite discordant to that watertable, it must be allocated to the phreatic hypothesis of Davis and Bretz. In continental Europe, Ek (1961) showed that cave systems in the valley of the Ourthe, Belgium, were indisputably graded to local watertables and Droppa (1964) advanced similar arguments for the extensive, multi-level galleries of the Demanova Caves, Czechoslovakia. But few other authors have offered categoric support for any one hypothesis and it is apparent that there is an abundance of field evidence to support each, (Trimmel 1968, Bleahu 1974).
Contentions of this paper.

The purpose of this paper is to summarize and further develop arguments presented in the speleological literature by Ford (1965, 1968, 1971). They are derived from detailed field studies in caves of Southwest England, the Rocky Mountains of Canada, and more general observations in caves throughout Europe and North America. Independently, Renault (1967, 1968, 1970) has arrived at similar conclusions after detailed work in French caves. Field studies by Waltham (1970, 1971) in Yorkshire, England and by Ewers (1972) in central Kentucky plus hardware simulations by the latter (1972 and in litt.) strongly support many of Ford and Renault’s contentions.

These contentions are that:

1. there is no one general case of cave system genesis that is restricted to the vadose or the phreatic or the watertable locus as the classical hypotheses contend. Rather:

2. there are three common cases and two special cases. From present-day data it cannot be shown that any one of the common cases is in simple majority and it is not relevant to a coherent theory that any one should be. The common cases are:

i. vadose caves, enlarged predominantly by free-surface streams eroding downwards or laterally or both.

ii. phreatic caves, evolved under conditions of total, permanent waterfill until, in the case of fossil system, a comparatively rapid drainage event causes them to be permanently abandoned by their genetic waters.

iii. watertable caves, («shallow phreatic caves», «epi-phreatic caves», etc.) developed along or at shallow depth beneath a piezometric surface that is of extent greater than the cave. The long profiles of the piezometric surface and the conduit are closely accordant in form.

The special cases are:

iv. the non-integrated cave, the vug or larger room, as in Figure 2a.

v. the true artesian cave, which is phreatic in form because the extension of a confining aquiclude below the piezometric surface compels phreastic circulation. Such systems often possess peculiarities, warranting their allocation to a special case (page 25). They are comparatively rare amongst accessible, explored caves.

3. a given cave or system of one phase (Figure 2, type b or c) may be wholly vadose or wholly phreatic or wholly of the watertable type. Or it may be a combination of vadose and phreatic, vadose and watertable, phreatic and watertable. Amongst known caves, the latter combination is rarer than the
others: this may be a reflection of the fact that it is particularly difficult to explore.

4. multi-phase cave systems with substantial relief are usually combinations. It is common to find an active vadose cave feeding an active watertable cave that underlies or intersects a fossil phreatic system that had earlier channelled the genetic waters.

**DEVELOPMENT OF CAVE SYSTEMS**

1. **Early Conditions - structural and lithologic guides and evolution of the phreatic skeleton.**

Cave systems are built of many individual segments, a segment being a portion of passage contained throughout its length in one structural element (e.g. a bedding plane) and possessing a constant orientation in all dimension or nearly so. This section considers the early development of such segments when conditions are phreatic and cross-sectional dimensions are minimal.

Cave systems develop within one of three gross structural situations: either the scale of any significant rock folding is greater than that of the system or it is the same or smaller. The first situation, in which the system develops within a flatlyning or monoclinally tilted block, is the more common and applies to the following discussion. The other situations are briefly considered on page 25.

It is accepted that cavern-bearing limestones function initially as fracture aquifers rather than granular aquifers. Inter-granular permeability is rarely significant until a system is aged and well enlarged when it may support route re-organisation of a gradational kind («bypass tubes», page 17). Effectively porosity is very low at the onset of karstification and the watertable is at ground level in most instances (Sawicki 1909, Cvijic 1918, Röglic 1965, Boccher 1977). Nevertheless, there may be significant losses of meteoric water into the rock immediately upon exposure and preferential solutional enlargement of pre-cave voids in fractures commences.

The views of Katzer 1909) upon the zonation of groundwater in karstified rock have been widely supported by later research (e.g. Gvozdetsky 1954, Burdon and Papakis 1963, Bauer and Zötl 1972). Where the limestone formation is comparatively thin (i.e. underlain by an impervious base at shallow depth) there is an uppermost vadose zone underlain by a phreatic zone extending to the base. Where the limestone mass is deep (e.g. extending to hundreds of meters below sea level) the phreatic zone becomes much thicker than in the shallow case but it is underlain by a zone of «tight» rock almost devoid of groundwater circulation. Unit effective permeability of the phreatic zone is greater in
Cave segments may be guided by planes, joints, faults or intercepts of such fissures. Faults are surprisingly unimportant in many systems although certain great ones such as Gouffre Padirac, France, are largely fault-located, as are many unexplorable karst springs. Whether or not a fault is exploited by karst groundwater often appears to be a function of the packing of secondary minerals in it (Ford 1965, Tratman and Ollier 1969). Where exploited, faults may be considered to function in the same manner as the more abundant joints.

Whether bedding planes or joints plus faults are quantitatively the more important guides of cave segments cannot be said. \textit{A priori}, bedding planes should be more important because they are continuous to the boundaries of the limestone mass or more nearly so. Joints are discrete features. In a great many cave systems there is a complex alternation of bedding plane-guided and joint-guided segments and it is apparent that the system could not have been

![Diagram](image)

\textbf{Figure 4.} Families of «dip tubes» (oriented proximate to true dip) in successive penetrable bedding planes are the basic cave units from which many cave systems of explorable dimensions are built. They are connected commonly by «joint chimneys» up which groundwater flows from a lower plane (B) to a higher plane (A).
created were not both available for the simultaneous transmission of groundwater. In Swildon's Hole, England (Ford 1965) approximately 50% of the aggregate known passage length of 7,000 m is guided by bedding plane-joint intercepts. The 10,000 m of passage in Castleguard Cave, Banff National Park, display a similar distribution (Ford 1971c). At the extremes, Deal (1962) reports that all 12,000 m of passage investigated in Jewel Cave, South Dakota, are joint-guided but the system is an example of the artesian special case (page 25). Ewers (1972) investigated 20,000 m of conduits in Pulaski County, Kentucky, and found that 93% were guided by bedding planes alone. Joints and faults are insignificant in the 250 km of passages in the Mammoth Cave-Flint Ridge complex (Deike 1968) and are of only minor importance in the 140 km of the Höllloch.

In any bedded limestone formation there are a great many bedding planes but only a small proportion is utilized during cave formation. Very often these display some clear feature that explains their preferential selection, such as a shale parting, a discontinuous chert filling or slickensiding and minor brecciation indicative of differential slipping along the plane (Ford 1965, Renault 1967, Waltham 1969). Rauch and White (1973) have shown that the different solubility of individual beds or groups of beds in a formation may determine which bedding planes are penetrated significantly.

The earliest cave segments to develop in bedding planes are «dip tubes» (Ford 1971). These are straight or slightly sinuous, phreatic features propagated down the true dip of the strata or within approximately 15° of it. Where the dip is greater than 5° or so the tube is commonly a simple, single conduit. Where it is gentler the single tube may be replaced by a band of anastomosing tubes, the band being a few metres in width and oriented down dip. In the centre of the band being is a principal tube larger than the others which may expand ultimately to consume them (Ewers 1972).

Dip tubes are spaced in sub-parallel array across a given bedding plane during earliest cave development and many families of such tubes are stacked one above another in successive penetrable bedding planes. Ewers (in litt.) has shown that their spacing in a plane is governed by Darcy's laws. At the upstream ends, the tubes terminate at surficial inputs or (later) within the evolving vadose zone. Downstream they distribute water into «pre-cave» cavities in the plane or link to cave segments in joints.

Early development in joint planes has less closely studied. Phreatic flow may proceed up, down or along them. Development up and down forms tubes that are very similar to dip tubes. Development along the joint requires that there be a «target» (an earlier cave other low pressure point) within the joint or at its terminus. The segment that develops will probably follow the shortest route
between input and target: thus, in some cases it will climb or descend aslant the dip and strike of the plane. In joint planes of comparatively high initial transmissivity, parallel and partly ellided tubes may develop. S.N. Davis (1965) has discussed the role that pumping by earth tides plays in the initial penetration of joints.

In many caves the segments are paired in the distinctive manner shown in Figure 4. A joint tube is expanded upwards from a dip tube in a lower plane to intersect a «target» dip tube in a higher plane. The bulk of the discharge of the lower tube then follows the new routing although the old route may be further extended downdip. Development of the joint tube represents a first step in the integration of the phreatic cave skeleton. Lower dip tube plus joint chimney constitute a simple phreatic loop.

In the earliest stages of system development there is in the rock a complex, three-dimensional array of independent dip tubes, loops and lateral joint connections. Greater cave system are constructed by the progressive integration and enlargement of small portions of this array. Our field researches in many karsts indicate very strongly that during the growth of the early array, reservoir capacity of the rock comes to exceed available water input. Vadose and phreatic zones are defined, with the highest parts of the array being drained.

**DIFFERENTIATION OF PHREATIC AND WATERTABLE CAVE SYSTEMS THE «FOUR STATE MODEL»**

The factors that determine whether a phreatic or watertable system will develop below the vadose zone are indicated in Figures 5 and 6. The frequency of penetrable fractures of all kinds that can be utilised during development of the early array (i.e. the spacing of segments of that array) is the principal determinant. This frequency is referred to hereafter as «fissure frequency». It will be appreciated that fissure frequency in limestones is very variable indeed and that the variation is, in reality, a continuum. In Figures 5 and 6, which survey the flat lying and monoclinally dipping cases respectively, the real continuum is reduced to four different states of frequency. State 1 is the minimum required for any significant groundwater circulation in given conditions of topographic relief and climate. If the frequency is less (State 0) no cave systems will develop; the rock may be dissected and removed entirely by surface streams, etc. State 4 is the minimum frequency required for the generation of a conduit along, or proximate and parallel to, the stable watertable of a phase. Higher fissure frequencies than this will merely yield a cave conduit that mirrors the watertable gradient ever more closely. States 2 and 3 are specified because each supports an intermediate class of cave that is geometrically di-
Distinct, as illustrated in Figure 7.
In Figures 5 (1) and 6 (1) fissure frequency is State 1. Piezometric gradients are steep and the watertable is high in the rock because reservoir capacity is low and the minimum flow path to the spring is long. The cave system developed through integration of the early array takes a deep course through the phreatic zone because there is no efficient alternative at lesser depth. In Figures 5 (4) and 6 (4) fissure frequency is State 4 and conditions are the converse. The vadose zone is deeper than in State 1, the piezometric gradient is low; if the volume of the trunk system is so enlarged that it can discharge all available water, the piezometric surface is lowered into it and it becomes an ideal watertable cave.

It will be seen that this model is consonant with some principles of karst groundwater circulation advocated by Swinnerton (1932) and Rhoades and Sinacori (1941). Where fissure frequency is high enough, their arguments are valid in part. The error of these authors was to presume that frequencies of State 4 or greater were the expected norm in limestones. They do not appear to be. Further points are made in the Discussion.

Classification of phreatic and watertable systems.
A classification of systems is presented in Figure 7. It is based upon the topographic relationship of system and piezometric surface when that surface has stabilised after the early stage of lowering through the expanding array. Such stabilisation has usually occurred before the system attains explorable dimensions. The classification is a minimum one, ignoring intervening subtypes that exist in reality.

1. (State 1), the bathyphreatic cave; type example - caves of the eastern front, Sierra de El Abra, Mexico. Bathyphreatic caves may be composed of many segments, as illustrated, but together these constitute a single loop beneath the piezometric surface.

2. (State 2), multiple loop phreatic cave; type example - the Hölloch, Switzerland. With higher fissure frequency the phreatic system exploits a path that brings it to the zero-pressure surface at many points between input and output. But fissure geometry and frequency will not permit it to follow the surface. Location of the stable piezometric surface is fixed by the apices of loops.

3. (State 3), mixed phreatic-watertable cave; type example - Swildon's Hole II-XII, England. Segments or multiples of segments developed proximate and parallel to the piezometric surface alternate with phreatic loops.

4. (State 4), ideal watertable cave; type example - Domica Caves, Czechoslo-
Figure 5. The four states of fissure frequency that differentiate types of phreatic and watertable caves, here drawn for the case of flatlyng strata. Developed systems for states 1 and 4 (the end members) are also shown.
Figure 6. The four states of fissure frequency (as in Figure 5) drawn for the case of steeply dipping strata where the system outlets are in the general direction of true dip. Developed systems for states 1 and 4 are shown.
1. BATHYPHREATIC CAVE

2. DEEP PHREATIC CAVE WITH MULTIPLE LOOPS (AND SHOWING GRADATIONAL FEATURES)

3. CAVE WITH MIXTURE OF PHREATIC & WATER TABLE-LEVELLED COMPONENTS

4. IDEAL WATER TABLE CAVE.

Figure 7. Hypothetical examples of the four types (states 1-4) of phreatic and watertable caves.
Figure 8. Subsequent gradational features common in phreatic caves: A. «isolated vadose entrenchment» of the upward apex of phreatic loop. B. development of a «bypass tube» above the downward apex of a phreatic loop. C. development of a «paragenetic» gallery upwards to the piezometric surface.
vakia. The system parallels the piezometric surface which, as a consequence of passage enlargement, may have been lowered into it.

The type examples are drawn from our personal experience. Equally good instances exist elsewhere. Many bathyphreatic systems that are short are known (up to some hundreds of metres in length). Complete examples of long system are very rare because of the evident difficulties of exploration where they remain waterfilled; where drained, loop bases tend to be sealed by clastic debris, limiting access. The El Abra examples are developed in Cretaceous reef and massive backreef facies with sparse but large joints. (Fish 1973 and in litt.). Vertical phreatic shafts carried groundwater down and up at least 200 m below a palaeo-watertable. A modern spring is fed from a phreatic shaft plumbed to -30 m. Above the spring, the same shaft extends upwards 80 m until terminated by hillside erosion and contains older spring outlets (Fish 1973).

The Hölloch (Bögli 1970) is a multi-phase system in a nappe of Cretaceous limestones. It is a multiplicity of loops below three or more successive watertables. There are no watertable-graded segments. Vertical amplitude of looping into the modern phreatic zone is known to exceed 170 m at high stages of flow. Swildon’s Hole II-XII, developed in steeply dipping, massive Lower Carboniferous limestones, is a sequence of modern watertable segments created by gradational processes (below) that are interspersed with short phreatic loops of 15 m vertical amplitude. This active system underlies a palaeo-system composed of longer phreatic loops with 20 m vertical amplitude and fewer, shorter graded segments (Ford 1965). The Domica Caves have developed in massive, steeply dipping strata that were intensely fractured by Carpathian tectonism. Many large, flat-roofed passages exist side by side at a uniform level. Apparently, as older galleries became partly silted new ones were created alongside them (Jakucs 1976).

Subsequent gradational features in phreatic systems.

After phreatic systems of States 2 and 3 have been stabilised and enlarged for substantial periods of time, they may be subject to gradational modification which change their character towards that of a higher state (e.g. ideal watertable caves). These modifications are entrenchment, bypassing and paragenesis (Figure 8).

Isolated vadose trenches are features of mature State 2 systems that convert portions of them to State 3 morphology. Consequent upon system volume expansion or at low stages of flow, the piezometric surface is lowered into the upper apices of phreatic loops. Groundwater flows across tops of loops as free surface streams which entrench vadose canyons downwards. Entrench-
ments of 20 m or more are known (Ford 1965). The vadose trenches are termed «isolated» because their water inputs and outputs are phreatic. In vadose cave system at least the inputs to a given vadose segment will also be vadose. «Bypass tubes» (Ford 1965) or «tubes en raccord» (Renault 1968) are groundwater shortcuts developed above the downward apices of phreatic loops. They are common features of State 3 systems where, together with isolated vadose trenches, they may compose a large part of the watertable-graded component. Individual bypass tubes may attain lengths as great as 50 m. Tubes develop across many loops by exploiting minor groundwater penetrations of intervening bedding planes and joints that were impenetrable when the early phreatic skeleton of the system evolved but which have developed some transmissivity subsequently as a result of proximity to the large, waterfilled conduits. However, in some instances no penetrated fissures are apparent in bypass tubes: they have developed by linking intergranular pores, a process of solutional mining.

Bypass tubes develop most commonly at the head of a phreatic system or at other points along it where connected vadose caves are able to supply quantities of coarse clastic bedload. Downward phreatic loops are natural sediment traps, the coarser material tending to infill the apices during low stage. The consequent reduction of phreatic channel cross-section favours a temporary but substantial increase of hydrostatic head on the upstream side when there is rapid flooding, and as a result the hydraulic gradient across the loop may be extreme. This abets bypass formation. Once bypassed, a phreatic loop becomes wholly aggraded so that later supplies of coarse bedload are moved directly through the flat bypass tube above it and into the next phreatic loop downstream where they contribute to the development of a second bypass. Bypassing is a gradational process that propagates downstream.

Isolated vadose trenches and bypass tubes are more common in systems in steeply dipping strata than flatlying because phreatic loops in the former have a simpler, two-segment geometry. Paragenesis (Renault 1968, Bretz 1942, White and White 1969) is a process that operates most powerfully where strata are flatlying. Renault has also emphasised that it is more common in downstream portions of a system.

Paragenetic passages commonly originate as much enlarged dip tubes in the phreatic zone. As a consequence of enlargement, hydrostatic head and head and groundwater velocity are reduced, permitting the permanent deposition of a portion of any insoluble suspended load. This armours the bed and lower walls prohibiting solution there. The passage enlarges upwards above a deepening column of fine-grained fill. The vertical amplitude of paragenesis exceeds 50 m in some instances (Renault 1968). The roofs of many paragenetic passa-
ges are remarkably flat. Paragenetic passages appear to be common in drained portions of the Mammoth Cave - Flint Ridge System of Kentucky, where the paragenetic roffs have often been interpreted as original watertable galleries.

**Actual measures of fissure frequency**

From Figures 5, 6 and 7 it is interesting to determine what are the actual measures of segment separation in an early array that will yield the four different states of enlarged stable systems. There is entirely insufficient data upon fissure frequencies in cave-bearing limestones for a general solution to this question. The most valid measurements are from very detailed cave mapping. Fissure frequencies measured on natural limestone surfaces (e.g. limestone pavements, Williams 1966) are maxima; there is drastic reduction at a depth of a very few metres. Fissure frequencies measured in quarries tend also to be considerable overestimates because of unloading fracture, as inspection of any nearby cave will show.

However, the distinction between States 3 and 4 may be estimated for a given limestone from the dimensions of bypass tubes these are «ideal watertable cave» segments. Twenty examples have been measured in caves of the Mendip Hills, England, and the Rocky Mountains of Canada. For bypassing to occur, bedding planes penetrated by dip tubes can be spaced no more than 10 m apart in the lithologic column and penetrated joints intersecting them, no more than 15-18 m apart. These are the minimum frequencies for State 4 conditions at those two sites. Strata there are massive, crystalline, well bedded Palaeozoic rocks that dip steeply. Where strata are flatlylyng, groundwater perching or the sandwiching of a few thin, well-jointed beds within a generally massive formation may achieve State 4 results where average fissure frequency is lower than this. Similarly, young limestones (Tertiary and Quaternary) often possess much higher primary permeability than older rocks. As this measure increases so also may the maximum spacing of fissures that is required for the State 4 result.

In the central Mendip Hills the England it is possible to link «fissure frequency» with the hydrologic measure, «effective porosity». From the nature of the caverns, the fissure state at and below the modern watertable is intermediate between State 3 and 4 (Ford 1971). From analysis of flood hydrograph recessions, Atkinson (1971).

Has shown that the effective porosity of the rock is approximately 1%. Probably, one thirtieth of this amount is represented by large cave systems which pass water comparatively rapidly (Atkinson 1975). The remaining porosity is composed of the skeletal array of conduits, «pre-cave» voids and non-
Figure 9. To illustrate the increase of fissure frequency in a limestone mass with the passage of time after onset of karstification and the effect of increase upon the geometry of successive caves developed in a multi-phase system. The figure is hypothetical but based upon the situation in the central Mendip Hills, England.
Integrated caves. This finding suggests that the four significant states of fissure frequency in cave system genesis are differentiated at \( \leq 1\% \) effective porosity.

**Increase of fissure frequency with time.**

At the onset of karstification, initial fissure frequency varies within and between formations. With passage of time after that, it increases so that later cave system in a multi-phase complex tend to higher state. This is illustrated in Figure 9 which models the general situation in the Mendip Hills over, probably, most of Pleistocene time (Ford and Stanton, 1968). Throughout the six steps of the model (Figures 9a-f) frequency diminishes with depth. In the earliest situation (9a) it was low even close to the surface and State 2 phreatic systems composed of few loops with great vertical amplitude developed, (9b). Whilst these aged (9c), unincorporated portions of the phreatic skeletal array were extended and their frequency increased. When spring positions were lowered by allogenic processes (9d) the higher frequency yielded State 2 systems composed of many loops of lesser amplitude. Repetition (9e, f) yielded State 2 systems with so many small loops that trenching and bypassing produced State 3 geometry whilst systems were still expanding to explorable dimensions.

An important feature of the model is that, although the bulk of groundwaters circulating at a given time are quickly integrated to the developed cave system, significant quantities are able to circulate elsewhere throughout the array and so extend it and increase its frequency. The proof of this is to be found where a spring position is lowered, as in 9c to 9d, etc. In principal, system response may take the form either of vadose entrenchment of the existing large galleries or of development of a new, lower phreatic system. In practice, the latter response is almost invariably the one that occurs in karst regions known to us. This indicates the presence of an accessible array beneath the older cave system. This is true even in the limestone Rockies of Canada where, as a consequence of valley glaciations, some potential spring positions were lowered by large amounts very rapidly.

**Contrast of cavern genesis in flatlying and steeply dipping strata.**

In the speleogenetic context, the range 2-5° of dip separates strata which may be considered flatlying from those steeply dipping. There are important probabilistic differences between the two cases.

Where strata dip steeply, bedding planes (which are the continuous fissures in
the rock) tend to entrain the groundwater to great depth. Their gradients are greater than initial hydrologic gradients. There is a quasi-artesian trapping effect by the superincumbent rock beds, as Glennie (1954) has noted. The inclined attitude of the limestone formation ensures that it will penetrate to considerable depth beneath the land surface even if the formation is very thin. Thus, Katzer's deep phreatic zone of comparatively low fissure frequency will tend to develop. The results of these various effects are that the deeper phreatic types of cave systems are particularly common in steeply dipping strata.

Per contra, watertable cave systems are particularly common in flatlying or gently dipping strata. Penetrable bedding systems planes, the efficient continuous entities, extend to the perimeter of the limestone mass where they outcrop to offer many potential spring points. It is only joints or faults (discrete features) that may entrain groundwater to great depths. Because the rocks are flatlying, conditions are favourable for the perching of groundwater streams upon dolomitised beds, shale bands, etc., again prohibiting deep penetration. In contrast, such aquitards play a quasi-artesian role in steeply dipping strata. Where rocks are flatlying, Katzer's two-zone model (vadose and phreatic zones) will apply unless there is exceptional vertical permeability, such as exists in Yorkshire, England (Waltham 1969). The effective porosity of the phreatic zone is comparatively high as a result, favouring the development of watertable cave systems. Many of the world's lengthiest examples of State 3 and State 4 systems are found in Kentucky and Indiana, where the thickness of the soluble, flatlying limestone formation is no more than 80-90 m and frequently less. Finally, flatlying limestones are often capped over wide areas by impermeable strata. In the initial stages of groundwater penetration of the aquifer, the cap inhibits overhead infiltration down joints so that potential of the joints to entrain groundwater to substantial depth is not realised. Such capping is ineffective where the strata dip steeply.

It may be noted in passing that the American literature which more than any other has championed watertable hypotheses of cavern genesis, has drawn heavily for field data upon the eastern U.S.A and mid-continent where limestone formations are predominantly thin and flatlying. Regions of steep dip are better sites to investigate fundamentals of speleogenesis because 'there the «noise» introduced by perching, paragenesis, etc., is much reduced.

**Fissure frequency and topography.**

In his discussion of variations of fissure frequency, Renault (1967, 1970) stresses the increase that occurs close to valley walls (20-100 m) as a result of the lateral release of load there. Thus, cave systems draining from the interior of a
Figure 10. The differentiation of the two principal types of vadose caves: A. the early condition of no significant vadose zone. Groundwater drainage is channelled through an early condition of no significant vadose zone. Groundwater drainage is channelled through an early array of phreatic tubes. B. consequent upon volumetric expansion of the array, a vadose zone is created where in streams greatly enlarge portions of the phreatic routes = «Drawdown Vadose Cave». C. streams are introduced to an established vadose zone at new input position and carve new routes that may utilise very little of the early phreatic array = «Invasion Vadose Caves».
plateau may suddenly ramify as they approach a spring in a valley side. Many systems develop entirely within the unloaded side zone, following courses sub-parallel to the longitudinal trend of sideslopes. The extensive, multiphase Demanova Caves of Czechoslovakia are an excellent instance (Droppa 1964). Short cave systems that have developed entirely to serve as shortcuts through spurs in meandering valleys, at valley confluences, etc., (a category with innumerable examples, e.g., Malott 1931, Ek 1961, Jennigs and Sweeting 1963, 1968) or to link vorfluter floors or glades in tropical karsts (Gradzinski and Radomski 1963, Sweeting 1972, Brown and Ford 1973) are other instances. Fissure frequencies are higher in these unloaded zones, favouring State 3 or 4 systems over States 1 and 2.

Renault has extended the scale of his argument to contend that alpine mountain masses are similarly unloaded when compared to plateau massifs. This point receives less support because deep phreatic caves (States 1 and 2) are common, perhaps predominant, in alpine regions such as the Rocky Mountains of Canada, the Kalkstein of Austria and the Hölloch massif of Switzerland. As a consequence of their structure, steep dip situations dominate in such areas.

**Differentiation in tightly folded strata.**

The foregoing discussion referred to differentiation of phreatic and watertable caves where the host rock may be considered as a flatlying or monoclinally-tilted massif. Where there is much tight folding (i.e. the scale of cave systems is greater than the scale of the folding) fissure frequency is generally high as a consequence of the high stressing of the rock. As a result, watertable systems are more common than in other structural situations as Davies (1960) has shown.

There are a few instances where the scale of the system and of the rock folding are essentially the same, because the cave distribution is an effect of the fold pattern. Instances are Grotte St. Andre-de-Cruzieres, Grotte En-Gorner in France (Renault 1967), Butler-Sinking Creek System in Virginia (Deike 1960) and Greenbrier Caverns, West Virginia (Rutherford 1971). The latter system contains more than 35 km of mapped passages distributed between three syn-
clines and two intervening anticlines. In such cases, trunk conduits tend to follow the axes of synclines and may be of phreatic or watertable type. On the synclinal flanks and across the intervening anticlinal crests there are dip tube systems if fissure frequency is low: joint mazes of artesian type are formed if joint frequency is high.

*The special case of artesian phreatic cave systems.*

Cave systems that are regular and dense mazes of passages that developed essentially within one erosional phase are rare. But they are spectacular features when drawn on a plan and clearly exerted a strong effect upon W.M. Davis’ thinking (1930). In most instances, they have developed in strata with State 4 or higher fissure frequency that were held in a bathyphreatic (State 1) circulation during cavern genesis by the presence of an impermeable overburden. By their nature such trapping situations tend to be maintained for longer spans of time than single phases in most common (non-artesian) caves. Long duration and very low hydraulic gradients that are a consequence of the large reservoir capacity of the rock, induce a maze of roughly equidimensional cave segments. Wind Cave and Jewel Cave, South Dakota, (Deal 1962) are the type examples.

**DIFFERENTIATION AND CLASSIFICATION OF VADOSE CAVE SYSTEMS**

The differentiation of vadose systems is shown in Figure 10. Initially, there is no vadose zone (Figure 10a) but as the phreatic array is expanded and phreatic system integration commences, the piezometric surface is drawn down. Vadose stream enlargement of portions of the phreatic skeleton commences (Figure 10b). Later in the phase or during ensuing phases, new surface streams may be introduced which carve new, generally steeper systems through the drained zone. Often, they utilise very little of the early phreatic skeleton. The two classes of vadose systems may be termed:

1. **Drawdown vadose caves**: type example, G.B. Cave, England. In such systems, 90% of the cave volume may have been created under vadose conditions but the segment skeleton is largely or entirely phreatic in origin. The initial phreatic morphology (commonly preserved in passage roofs where it may be difficult to inspect) is mature, comprised of smooth and well-rounded erosional forms. G. B. Cave is analysed in detail in Ford 1964: see also Smart and Stanton 1974.
2. Invasion vadose caves: type example, Spluga della Preta, Italy. The concept of «invasion waters» in cavernous vadose zones has been used widely since Malott (1937) termed it to describe small streams that chanced to intersect drained phreatic galleries and perhaps follow them for some distance. As illustrated, invasion vadose systems are subsequent developments in time. Often, their erosional morphology is wholly vadose in character (especially where development is dominantly vertical) but an initial, rough and poorly formed phreatic component may be preserved. Tratman (1957) has termed this «pa-raphreatic» and attributed it to intermittent flooding. The type example, Spluga della Preta, is a system of vertical shafts in limestone aggregating 800 m in depth and terminating in a siphon perched on underlying dolomites. It is very rare for caves of this class to constitute a complete system (i.e. to extend from sinkpoint to spring). Generally, they feed waters to one of the other classes of caves. (*)

Both drawdown and invasion caves are common. Often, the portions of a system that are of explorable dimensions prove to be a mixture of the two, for example, many of the domepit-and-shaft-drain combinations of the central Kentucky karst (White et al. 1972).

Factors affecting the extent and magnitude of vadose cave development.

The extent of vadose caves of either type is a function of the depth of the vadose zone and of any lithologic, etc. effects tending to divert groundwaters from a simple, vertical descent. The magnitude (passage size) of vadose caves is a function of the magnitude of their streams and the duration of erosion; whilst this is true of all other classes as well it has particular significance for the vadose cases.

Depth of the vadose zone is a function of the relief of a limestone mass above its springs and of the gradient of the basal watertable. The importance of relief is evident; the deepest vadose caves occur in montain regions. Where relief is lower, variations of watertable gradient may cause significant variations in vadose depth. Ceteris paribus, gradient is inversely proportional to fissure frequency. Where frequency is State 1 or 2, chance effects in the selection of phreatic routes to which vadose systems drain may create substantial local variations (many tens of meters) in the depth of the vadose zone (Ford 1965).

(*) In earlier discussion, (Ford 1968, 1971) invasion systems were termed «primary vadose caves» being primary in the sense that no earlier, non-vadose cave had existed along their specific route. Drawdown systems were termed «secondary vadose caves» because they develop from earlier phreatic conduits. This terminology is now rejected because it may confuse the conception of sequence. In a given vadose rock mass, any drawdown systems (secondary) will be contemporary with or older than any invasion systems (primary).
In general groundwater hydrologic models, water in the vadose zone is presumed to drain vertically downwards. In many karst regions, this will be true only in the sense that it does not drain upwards. Hundreds, perhaps thousands, of kilometers of lateral vadose passages are now mapped. They are particularly common where strata are flat-lying and perching upon shale bands, etc., occurs. In Kentucky and Indiana many vadose galleries traverse several kilometers to gain a net descent of 10 meters.

Magnitude of streams is very important because the vadose zone is the first to be encountered by water flowing from sink to spring. Suppose in given circumstances that a system may not attain explorable dimensions unless the stream is of Fourth or higher order. If sinking streams do not exceed First Order in magnitude, much or all of a vadose zone may be utilised to combine to Fourth Order. Development of explorable vadose systems is retarded or prohibited. This effect is most marked in holokarst areas (limestone regions that are not supplied by allogenic streams). It particularly inhibits enlargement of drawdown systems to explorable size. Invasion systems tend to have a greater vertical component and, when they fall vertically, even very small streams may create large shafts by splash and film solution processes.

Drawdown systems, therefore, are best developed where streams can collect upon adjacent or superjacent non-karst rocks and attain at least Second or Third Order before sinking into the limestone. Particularly extensive and complex systems develop where many such streams are closely spaced when they cross a contact with limestone. This is the situation in Yorkshire, England, and the Burren, Eire where steep-sided mountain caps of shale yield high drainage densities to underlying limestone platforms (Waltham 1974, Tratman and Ollier 1969). Many of the most spectacular instances of invasion systems are to be found in regions where surface drainage is much deranged, so that large streams are diverted from original drawdown sinks and create invasion caves. Glaciation is most effective in this respect (as indicated in Figure 10c) and the combination of high relief and derangement creates such alpine caverns as Spluga della Preta or Yorkshire Pot, Crowsnest Pass (Thompson 1976).

DISCUSSION
A set of principles of cavern genesis and differentiation in the dimension of length and depth has been presented that indicates that explorable vadose, phreatic and watertable systems will all be common occurrences at the global scale, although within any small region a particular class may be predominant. When all systems have been located and explored in their entirety, it
may transpire that one class is predominant globally - though this has certain-
ly not been the trend of field findings during the past twenty years. Even if this
should be the case, it will not negate the necessity for a general theory that in-
corporates all classes because known examples of all classes are already of
considerable magnitude and play integral roles in the genesis of many complex
systems.

Attention has been directed to general principles, factors and derivations.
There are many sub-types of systems that are not classified and details of com-
mon aberrations (such as a few wholly phreatic segments high up in a vadose
cave) are not considered.

The principles derive largely from empirical observation, the most important
contribution stemming from detailed reconstructions of genetic sequence in
systems that are mixtures of fossil and active components. Our field sample
together with that of Renault aggregates more than 1,500 systems or portions
thereof. It includes most types of limestone as well as gypsum and delomite,
and most structural and topographic situations. There is a deficiency in exam-
ples of Tertiary and Quaternary Limestones in tropical settings: if these devia-
te it is predicted that it will be in the direction of greater frequency of State 3
and State 4 systems vice States 1 and 2, because fissure frequency and primary
permeability tend to be higher than in older rocks. Simulations of cave initia-
tion using salt blocks, gypsum locks, electrolyte solutions, sand models and
computer models (Ewers 1966, 1972 and in litt.) have contributed to our un-
derstanding but are more relevant to the problem of genesis in length and
breadth.

The principles resolve much of the contradiction of the classical papers. The
proponents of the vadose theory and of the phreatic theory also based their ar-
arguments upon field evidences and interpreted these correctly in most instan-
ces. Their samples were insufficiently varied or contained error. Proponents
of watertable theories deduced their models from sets of first principles of
groundwater flow and these, also, were correct in part. But they are not as wi-
dely applicable as was supposed because fissure frequency is insufficiently
high.

The leading vadose theorists such as Martel and Dwerryhouse were pioneer
cave explorers. Such was the exploration technology of their age that the acti-
vely enlarging systems that they knew were vadose systems; early exploration
of phreatic and watertable caves was restricted largely to fossil examples.
They interpreted the kinds of aqueous erosion that they observed (vadose) to
apply in most situations.

W.M. Davis based his phreatic theory largely upon written accounts and, par-
ticularly, maps. His interpretations of what he saw there were correct but the
early cave maps contained many gross errors, often being sketches from memory. The map of Mammoth Cave that Davis used has been described as «fiction» by cave historians (J.F. Quinlan, pers. comm.). His supporter, J.H. Bretz, saw more than 500 caves in Missouri and correctly determined that they were not vadose in character. But most were too short to permit their unequivocal assignment to the phreatic case rather than the watertable alternative. Helwig (1965) and Reams (1968) have shown that Bretz often misinterpreted the nature and significance of the clastic infillings that terminated his caves a short distance underground.

W.M. Davis’ phreatic model fits phreatic cave systems to Darcy flowlines (and thus to Darcy’s Laws) throughout their enlargement. Rhoades’ and Sinacori’s formulation (1941) applies Darcy’s laws only to the initiation of systems, as these are defined in the present text. Swinnerton’s alternative watertable formulation (1932) rejects the laws even at initiation.

Our laboratory model experiments indicate that primitive cave systems distribute according to the laws within the plane of a fissure. But with integration between different fissures, Swinnerton’s condition of route competition applies. Because integration appears early in the development of the phreatic skeleton, Swinnerton’s views are more nearly correct than those of Rhoades and Sinacori. However, the latter authors made a most valuable contribution when they introduced the concept that systems develop from the spring headwards into the limestone mass (see also Mandel 1966). Caves propagate from the head (or sinkpoint) but a great many, perhaps the majority of systems integrate from the spring in a progression of headward steps. This explains much of their complexity on the plan, which will be enlarged upon in a subsequent paper.

ABSTRACT
Karst caves are defined as solutional cavities 5-16 mm in diameter and discussion is limited to cases where such continuously extend to a surficial input or output or both. Three opposed sets of general genetic hypotheses («the classical hypotheses») have been presented for such caves, arguing that the majority develop 1) in the vadose zone 2) in the phreatic zone 3) proximate and parallel to a watertable. It is contended here that vadose, phreatic and watertable caves are all of common occurrence and may be linked in one genetic theory. A four state model is proposed in which ideal phreatic and watertable caverns are end members: in a given massif of soluble rock the state (cave type) that develops is a function of the frequency of fissures penetrable by groundwater. The water-table type is the high frequency end member. Fissure frequency increases with passage of time after onset of karstification and gradational features may also develop to modify phreatic types. Vadose caves may be of «drawdown» type (following an initial phreatic path) or «invasion» type (developing a new path through rock drained by earlier caves). Extensive cave systems may comprise vadose, phreatic and/or watertable developed contemporaneously.

REFERENCES

DEVELOPMENT OF LIMESTONE CAVE SYSTEMS


KRASON J. and WOJCICZ Z., 1965 - «The deep karst of the Boleslawiec Syn-