Groundwater Flow in Low-Permeability Environments

C. E. Neuzil

Water Resources Division, U.S. Geological Survey, Reston, Virginia

Certain geologic media are known to have small permeability; subsurface environments composed of these media and lacking well developed secondary permeability have groundwater flow systems with many distinctive characteristics. Moreover, groundwater flow in these environments appears to influence the evolution of certain hydrologic, geologic, and geochemical systems, may affect the accumulation of petroleum and ores, and probably has a role in the structural evolution of parts of the crust. Such environments are also important in the context of waste disposal. This review attempts to synthesize the diverse contributions of various disciplines to the problem of flow in low-permeability environments. Problems hindering analysis are enumerated together with suggested approaches to overcoming them. A common thread running through the discussion is the significance of size- and time-scale limitations of the ability to directly observe flow behavior and make measurements of parameters. These limitations have resulted in rather distinct small- and large-scale approaches to the problem. The first part of the review considers experimental investigations of low-permeability flow, including in situ testing; these are generally conducted on temporal and spatial scales which are relatively small compared with those of interest. Results from this work have provided increasingly detailed information about many aspects of the flow but leave certain questions unanswered. Recent advances in laboratory and in situ testing techniques have permitted measurements of permeability and storage properties in progressively "tighter" media and investigation of transient flow under these conditions. However, very large hydraulic gradients are still required for the tests; an observational gap exists for typical in situ gradients. The applicability of Darcy's law in this range is therefore untested, although claims of observed non-Darcian behavior appear flawed. Two important nonhydraulic flow phenomena, osmosis and ultrafiltration, are experimentally well established in prepared clays but have been incompletely investigated, particularly in undisturbed geologic media. Small-scale experimental results form much of the basis for analyses of flow in low-permeability environments which occur on scales of time and size too large to permit direct observation. Such large-scale flow behavior is the focus of the second part of this review. Extrapolation of small-scale experimental experience becomes an important and sometimes controversial problem in this context. In large flow systems under steady state conditions the regional permeability can sometimes be determined, but systems with transient flow are more difficult to analyze. The complexity of the problem is enhanced by the sensitivity of large-scale flow to the effects of slow geologic processes. One-dimensional studies have begun to elucidate how simple burial or exhumation can generate transient flow conditions by changing the state of stress and temperature and by burial metamorphism. Investigation of the more complex problem of the interaction of geologic processes and flow in two and three dimensions is just beginning. Because these transient flow analyses have largely been based on flow in experimental scale systems or in relatively permeable systems, deformation in response to effective stress changes is generally treated as linearly elastic; however, this treatment creates difficulties for the long periods of interest because viscoelastic deformation is probably significant. Also, large-scale flow simulations in argillaceous environments generally have neglected osmosis and ultrafiltration, in part because extrapolation of laboratory experience with coupled flow to large scales under in situ conditions is controversial. Nevertheless, the effects are potentially quite important because the coupled flow might cause ultra long lived transient conditions. The difficulties associated with analysis are matched by those of characterizing hydric conditions in tight environments; measurements of hydraulic head and sampling of pore fluids have been done only rarely because of the practical difficulties involved. These problems are also discussed in the second part of this paper.
from drilling in many parts of the world. In still other settings, distributions of ions in groundwater suggest the presence of extensive low-permeability bodies which act as semipermeable membranes.

Low-permeability environments are most commonly associated with fine-grained sedimentary deposits such as shales and clays. These occur in active and ancient sedimentary basins and in accretionary wedges at plate boundaries. Crystalline rocks generally prove to be relatively permeable at large scale because of extensive fracturing. However, at depth they may have regionally low permeability because of healing or filling of fractures [Walder and Nur, 1984]. Even in shallow crystalline rocks, fractures may be sparse, leaving relatively large unfractured areas [Balk, 1937], or they may be numerous but unconnected as Marsily [1982] suggested is the case in the French Massif Central. Certain other terranes may have small permeability, including parts of the oceanic crust at depth [Anderson et al., 1985] and salt and other evaporite deposits [Kreitler et al., 1985]. Groundwater flow in such environments, which are low in permeability at the scale of interest, is the subject of this review. Flow under partially saturated conditions, occurring near the surface and in soils, will not be considered. The term "low permeability" is here applied to media with hydraulic conductivity to water of approximately $10^{-8}$ m/s or smaller.

Fluid flow in permeable rocks has been extensively studied and analyzed in several contexts, notably in the development of groundwater resources and the recovery of petroleum. By comparison, less effort has been spent to analyze the flow behavior in tight, or poorly permeable, rocks. However, the past several years have seen this begin to change. There are several reasons. It has, for example, become clear that aquifer behavior prior to development and the response to development is often mediated by adjoining tight formations [Walton, 1965; Gill, 1969; Neumann and Witherspoon, 1972; Bredehoef et al., 1983]. It has also been recognized that the flow in tight formations has an important role in the evolution of groundwater flow systems over geologic time. Attention has turned recently to the flow in evolving sedimentary basins [Sharp, 1978; Tóth, 1978; Cathles and Smith, 1983; Smith et al., 1983; Garven and Freeze, 1984a, b; Bethke, 1985], which affects the migration of hydrocarbons and the formation of ore bodies. The role of fluids in mechanical processes [Pincus et al., 1982] dictates the importance of groundwater flow in low-permeability portions of the crust for understanding the mechanical behavior of these regions. This has recently been emphasized by Westbrook and Smith [1983] for sedimentary accumulations and Walder and Nur [1984] for crystalline rocks. Knowledge of the flow in tight media is also necessary for understanding the geochemical evolution of their pore fluids and solids. Such regions are a unique environment where fluid and solid phases have been in contact for long periods while little chemical mass has been exchanged with the surroundings. A recent and highly significant motivation for the study of low-permeability environments is geologic disposal of toxic materials so as to minimize their exposure to moving groundwater [Bredehoef et al., 1978].

The study of flow in low-permeability environments has been possible at widely disparate size and time scales. Performance and analysis of controlled experiments in the laboratory and in situ permit the investigation of various flow phenomena and direct measurement of their parameters on relatively small scales of distance and time. At the other extreme are quantitative, theoretically based analyses of large-scale and long-term flow behavior which cannot be observed. This raises the important point that where the permeability is small, flow experiments encompassing significant volumes are not practical. Even areally extensive in situ tests, such as aquifer-confining layer tests, sample the tight confining layer but a short distance in from its boundary. Usually a small fraction of the confining layer is involved. Tests using boreholes in tight rocks likewise involve relatively small volumes of the rock. Thus, compared with the volumes of interest, practicable laboratory and in situ tests are small in scale. Yet, these are often the only source of information on which to base analyses. Only in systems which are at steady state is it sometimes possible to directly measure the hydraulic conductivity of large volumes of tight media. The problem is much more complex if the flow is in a transient state because the changes occur extremely slowly. One is, therefore, unable to observe the large-scale transient response to a known stress which often helps characterize more permeable formations [e.g., Randolf et al., 1985].

Extrapolation of flow behavior in tight rock from small to large scale is fraught with difficulties. As indicated, the size scale dependence of permeability is a problem; similar and less well understood difficulties are associated with the nonhydraulic conductivities governing various types of osmotic flow. A less well recognized but highly significant problem concerns the time scale dependence of storage properties. This results from descriptions developed for relatively short lived transient flow being applied to long-term behavior in low-permeability environments.

The slow transient response of the flow and long periods of time involved add another challenging dimension to the problem. Slow geologic processes acting on the system, which are normally ignored in studies of groundwater development, can be the dominant control of the flow. Hydrologic conditions may reflect cumulative geologic changes extending well back into the past. Indeed, it is often difficult to determine whether present conditions represent steady flow in equilibrium with stable boundary conditions, whether they are slowly transient in response to natural geologic changes, or whether they are due to other effects entirely, such as osmosis.

To the various difficulties mentioned should be added those encountered in practice which are related to determining existing hydraulic conditions in the field. Normally routine procedures such as measuring hydraulic head and sampling pore fluids are much more difficult in tight materials.

The history of hydrogeology-related disciplines appears to have influenced how low-permeability environments have been studied and analyzed. Because relatively permeable environments were studied earlier and more extensively, the study of tight environments has been influenced by experience with more permeable ones. Phenomena associated only with tight media have therefore received less emphasis than they perhaps should. Osmotic flow and related coupled flow phenomena are probably the best example.

Because of the significance of flow in low-permeability environments, the increasing interest in it, and the number of disciplines making contributions toward understanding it, an overview of this emerging field of study seems warranted. The purpose of this review is to provide such a perspective. Specifically, there are four primary aims: (1) to synthesize and place in perspective the diverse research carried out in various disciplines which is important for understanding groundwater flow in saturated low-permeability environments, (2) to consider whether historical bias has influenced the approach to...
the problem, (3) to enumerate questions and issues which, at present, hinder analysis and to discuss appropriate approaches for overcoming them, and (4) to set forth observations on important uncertainties related to the problem.

In practical terms, the problem of flow in low-permeability environments may be summarized as follows: How can knowledge of behavior at small scales be extrapolated to large dimensions and long periods of time? Uncertainties related to small-scale behavior strongly affect the ability to understand and analyze large-scale behavior. Consequently, the first part of the paper considers the basis of our understanding of flow in low-permeability media, namely small-scale experiments and tests. The second part of the paper examines the problem of analyzing low-permeability systems in which flow is occurring at scales that make direct observation of the flow dynamics impractical. It considers how experimental experience has influenced analysis of large-scale flow and when extrapolation of experimental behavior to large scale may be misleading. To conclude, the final part of the paper offers a general, philosophical overview of the problem with observations on inherent difficulties.

The Small Scale: Experimental Characterization of Flow in Tight Rocks

In the present context a useful definition of "small scale" is one at which flow experiments in tight media can be completed in a practical period of time. Thus both laboratory and in situ tests are included.

Experiments have addressed questions related to two rather different perceptions of low-permeability flow. One perception, an outgrowth of traditional groundwater studies, emphasizes the hydraulic character of the flow. From this viewpoint, the primary task is to devise techniques to reliably measure the parameters governing hydraulic flow, namely, the permeability and storage properties. The latter include, in addition to specific storage, the loading efficiency, which describes the response to changes in external mechanical stress. Recent advances have been made which facilitate measurement of hydraulic parameters in the laboratory and in the field.

The other viewpoint emphasizes nonhydraulic flow phenomena which are linked specifically with tight media. Here the primary task is generally viewed as gaining a better understanding of the flow laws. This includes assessing the applicability of Darcy's law in very tight media and investigating osmotic phenomena related to coupled flow. The problems involved in these investigations are generally quite complex, and few, if any, field techniques have been developed for them.

Hydraulic Flow: Measurement of Parameters

Laboratory approaches. Laboratory strategies for measuring hydraulic conductivity of tight geologic materials can be arranged in three classes. These are (1) application of Darcy's law to flow or pressure data from steady or quasi-steady flow, (2) application of a transient groundwater flow description (which implicitly assumes Darcy's law) to transient flow or pressure data, and (3) application of a transient flow description to a mechanical surrogate for flow and pressure data, such as time dependent sample deformation. The last two classes of tests analyze the same phenomenon, namely, transient flow of a fluid in a deformable medium. However, in the second class of tests, which was developed for indurated media, stress change and deformation of the medium are considered only insofar as they affect fluid storage. The third class of tests, which was developed for relatively deformable media, considers strains and stress changes in the medium explicitly.

The prolonged flow transients which occur in low-permeability rocks provide an interesting benefit for laboratory scale testing; the time dependent behavior can be analyzed to compute the specific storage. This usually is not practical for samples of permeable rock in which transient changes in small samples occur too rapidly to be studied. The second and third classes of tests, which analyze time dependent behavior, generally permit determination of both hydraulic conductivity and specific storage.

Steady and quasi-steady tests: When the hydraulic conductivity $K$ is very small, two problems make its measurement in simple steady flow tests difficult. First, the time required to attain steady flow through the sample may be inconveniently long [Olsen et al., 1985]. Second, measuring or generating sufficiently small rates of flow may be difficult [Remy, 1973; Olsen et al., 1985]. The first can be mitigated somewhat by decreasing the sample length in the flow direction. The second problem can be overcome by using large hydraulic gradients. Very small values of $K$ have been measured this way, using hydraulic gradients as large as $10^6$ and more, much greater than those found in nature.

Large-gradient steady flow tests have been described by Ohle [1951a, b], von Englehardt and Tunn [1955], Lutz and Kemper [1959], Gloyna and Reynolds [1961], Young et al. [1964], Pearson and Lister [1973], Coplen and Hanshaw [1973], Kharaka and Berry [1973], Kharaka and Smalley [1976] and Morrow et al. [1981, 1984], among others. Values of $K$ of the order of $10^{-12}$ m/s were obtained in some instances, but large hydraulic gradients were required to obtain them.

An alternative approach is to pump small, predetermined rates of flow through the sample because it can be easier to pump small flow rates than to measure them. Olsen [1966] and Olsen et al. [1985] used a small-bore syringe to pump at rates as small as $10^{-8}$ cm$^3$/s. It permitted measurements of $K$ in the range $10^{-7}$ to $10^{-5}$ m/s with gradients of 0.2 to 40. Olsen et al. [1985] recently have suggested that pumping rates as small as $10^{-7}$ cm$^3$/s are feasible with a differential diameter piston. Summers et al. [1978] used input pumping rates to compute the permeability of Westerly Granite, but with relatively large pumping rates and correspondingly high gradients.

Some advantage is gained by using gas permeants because of their much lower viscosity. However, corrections are necessary [Scheidegger, 1974, p. 172] because of gas slippage. More importantly, the physicochemical interactions between permeant and solids, which may be significant in small pores, will be quite different for gases and water; this makes the relation between gas and liquid permeabilities in many tight media uncertain. However, in certain instances these objectives may not be important. For example, gas and liquid permeability of Westerly Granite measured by Brace et al. [1968] and of Grand Saline salt measured by Gloyna and Reynolds [1961] were in close agreement.

A variant of the steady flow test, the falling head test, has also been used to test tight media. Described by Darcy [1856], the test is conducted by permitting an elevated water level in a standpipe to cause flow through the sample; as the level falls, the gradient across the sample and the flow rate decline. As long as the storage in the standpipe is large compared with that in the sample, the flow remains quasi-steady, that is, the hydraulic gradient within the sample remains approximately linear. Terzaghi [1925] claimed to have measured $K$ smaller
than $10^{-11}$ m/s using falling head tests, and Lambe [1951] formalized the experimental strategy for measuring small $K$ with the technique. More recently, falling head tests were used extensively by Tavernas et al. [1983] for testing tight clays.

The falling head test per se is not particularly advantageous for measuring small $K$. However, a distinct technique specifically suited for tight media is, consciously or not, a variation of it. In this version of the test, a closed reservoir instead of an open standpipe is connected to one end of the sample. At the beginning of the test, the reservoir is abruptly pressurized. The pressure declines with time as water flows through the sample. The great advantage of the closed system lies in the fact that the storage in the reservoir, owing to the equipment and water compressibility, is much smaller than in an open standpipe. An equivalent head decline requires the flow of a much smaller quantity of water into the sample. When $K$ is small, this can sufficiently reduce the time required for the test to make it feasible. This test methodology and analysis was first developed by Bianchi and Haskell [1963], who used a deformable diaphragm with strain gauges to determine flow from the closed reservoir; Nightingale and Bianchi [1970] later used the technique to test clays. Somewhat after Bianchi and Haskell published their paper, Overman et al. [1968] described a similar test in which flow was related to reservoir pressure measured with a pressure transducer. At nearly the same time, Brace et al. [1968] independently developed a similar but somewhat more complex test design and analysis; they used closed reservoirs of arbitrary volume placed at both ends of the sample. After the upstream reservoir is abruptly pressurized, its head declines and the head in the downstream reservoir rises. The analysis permits extraction of $K$ from the pressure changes. Using this method, Brace et al. [1968] measured the conductivity of Westerly Granite for a range of effective pressures, using both gas and water. They obtained values as small as $10^{-14}$ m/s. The mean gradient across the sample ranged between approximately $10^8$ and $10^9$ during the tests.

Shortly after the methodology was presented by Brace et al. [1968], Sanayal et al. [1971, 1972] and Remy [1973] reported using a test and analysis, apparently developed independently, which was essentially the same. In the tests by Sanayal et al. on Precambrian chert, hydraulic oil was used as a permeant, and the upstream reservoir was pressurized thermally rather than mechanically. Conductivities as small as $10^{-10}$ m/s were computed with hydraulic gradients ranging up to $10^4$. Zoback and Byerlee [1975] also tested Westerly Granite, Kranz et al. [1979] tested Barre Granite, and Reda and Hadley [1985] tested welded tuff, using essentially the same technique outlined by Brace et al. [1968].

The advantage of closed reservoir tests, i.e., the smallness of the storage in the reservoirs, renders the assumption of quasi-steady flow inappropriate in many instances. When shut in, the reservoir storage is not always much greater than that of the sample, and fully transient flow in the sample is possible. While the quasi-steady assumption is probably justified for rocks such as granite, which have small storage, Overman et al. [1968], Miller et al. [1969], and Nightingale and Bianchi [1970] used the quasi-steady analysis for clays, which have significant compressive storage. Indeed, it appears that storage effects led Miller et al. [1969] to incorrectly conclude that their sample exhibited non-Darcian flow behavior. Neuzil et al. [1981] analyzed the error associated with the assumption of quasi-steady flow and showed that it could produce the behavior Miller et al. [1969] observed, resulting in underestimates of $K$.

**Hydraulic transient flow tests:** The effect of storage in the sample can be accounted for if the test is analyzed with the transient groundwater flow equation in one dimension, which may be stated as

$$K \frac{\partial^2 h}{\partial z^2} = \frac{\partial h}{\partial t}$$

(1)

Compressive storage, which permits transient flow, is accounted for by the one-dimensional specific storage $S_s [L^{-1}]$. As discussed later in this paper, (1) is strictly valid when no lateral strain occurs in the specimen. With the appropriate boundary conditions to describe the two-reservoir test methodology of Brace et al. [1968], the solution of (1) can be used to analyze fully transient flow. Lin [1978] and Trimmer et al. [1980] adopted this approach, using numerical solutions of (1). Lin tested argillite of the Eleanna Formation and measured conductivities between $10^{-13}$ and $10^{-16}$ m/s with initial gradients across the sample ranging from $5 \times 10^1$ to $5 \times 10^3$. Trimmer et al. tested crystalline rocks, obtaining conductivities as small as $5 \times 10^{-17}$ m/s.

The procedure used by Lin and Trimmer et al. requires additional information about the storage properties of the sample. In essence, $S_s$ must be independently determined from more fundamental properties. Both Lin and Trimmer et al. determined these properties separately to analyze their tests. Hsieh et al. [1981] and Neuzil et al. [1981] developed an analysis and test methodology utilizing an analytical solution of the equation for the test boundary conditions. Their method permits both $K$ and $S_s$ of the sample to be determined from the test data. In a sample of the Pierre Shale, Neuzil et al. [1981] measured $K$ as $2 \times 10^{-12}$ m/s and $S_s$ as $8 \times 10^{-7}$ m$^{-1}$ with an initial gradient of $2 \times 10^3$.

Variations of the test procedure for fully transient flow have been used. Al-Dhahir and Tan [1968] obtained a solution for a somewhat different test procedure. Their procedure requires measurements of flow, negating much of the advantage of shut-in transient tests. Gondouin and Scala [1958] used a similar analysis for testing shales but apparently based it on an incorrect analogy with diffusion. R. H. Morin and A. W. Olsen (unpublished manuscript, 1986) have described a technique wherein a predetermined flowrate from a pump is abruptly applied at the sample boundary. Analysis of the time dependent pressure change at the boundary permits computation of $K$ and $S_s$. Either positive or negative flowrate (injection or withdrawal) may be applied, allowing determination of both rebound and compression values of $S_s$.

Most if not all experimenters who have used the fully transient techniques discussed have used a constant lateral and longitudinal load on the specimen. Some error, probably small, may be expected as a result because of the no lateral strain condition implicit in (1). A more appropriate experimental design would fully constrain the specimen laterally while applying a constant stress longitudinally.

**Mechanical transient flow tests:** The tests described so far require one to measure (or pump) small rates of flow or to measure reservoir pressure changes, which are a surrogate for flowrate measurements. One can also perform tests which permit determination of $K$ and $S_s$ from the transient mechanical behavior of the sample. Such tests have been mainly the denominate of soil mechanics and have been developed for relatively compressible media which can be described by standard consolidation theory. Thus the available methodologies are appropriate for media with high loading efficiencies (see...
The study by Terzaghi [1923] of consolidation of low-permeability media beneath foundations provided the derivation of (1) well before it was developed independently by Theis [1935] and Jacob [1940] for water supply applications. Terzaghi’s derivation recognized that expulsion of pore fluid, at a rate governed by the hydraulic conductivity, was required for consolidation to proceed.

In a step load consolidation test [Terzaghi, 1927; Lambe, 1951; Scott, 1963], the specimen is laterally constrained and subjected to an abrupt loading which remains constant. Pore water is permitted to drain at one or both ends. Solution of (1) for the time dependent strain under appropriate boundary conditions [Suklie, 1969] permits computation of the ratio K/S, by comparing observed and computed strain. Specific storage is computed from the porosity and observed compressibility of the specimen. Hamilton [1964] and Bryant et al. [1975] obtained hydraulic conductivities for deep sea sediments, and Wolff [1970] obtained conductivities for a clay confining bed by analyzing consolidation data. Bredehoeft et al. [1983] used consolidation tests to compute hydraulic conductivity of the Pierre Shale. They obtained conductivities between $10^{-12}$ and $10^{-14}$ m/s with hydraulic gradients [see Mitchell, 1976, Figure 15.7] ranging as high as $10^4$ to $10^5$.

Variations in testing methodology and analyses based on less restrictive assumptions have been advanced as a variety of mechanical test techniques. For example, the case of large strain consolidation was analyzed by Gibson et al. [1967] as a generalization of Terzaghi’s [1923] theory. The application of a varying load to maintain a constant strain rate was analyzed by Smith and Waals [1969]. Lowe et al. [1969] suggested a test in which the applied stress is controlled to maintain a constant hydraulic gradient, and Aboshi et al. [1970] describe a test in which loading is changed at a constant rate. Znidarčić [1982] provides a summary of the various analyses and tests which have been proposed.

Comparability: There is growing confirmation of the ability of nonsteady flow techniques to reliably measure the permeability of many types of tight media. Mitchell and Younger [1967], Wolff [1970], Mesri and Olson [1971], and Silva et al. [1981] found reasonably good agreement between results of steady state tests and values computed from step load consolidation tests. Some systematic discrepancies have been reported by others, but many are relatively minor or appear to be attributable to special circumstances. Seaber and Vecchioli [1966] called attention to permeabilities computed from step load consolidation tests which were substantially smaller than those from quasi-steady tests, but the differences may have been due to the lack of any confining load during the latter. Olson and Daniel [1981] also reported underestimates of K obtained from step load tests. Using theoretical arguments, Znidarčić [1982] suggested that inherent error of a few tens of percent is likely in K determined in constant rate of strain tests. In a rather thorough study of tight clays, Tavenas et al. [1983] compared K computed from step load, constant strain rate, and constant hydraulic gradient consolidation tests with values form steady and open standpipe quasi-steady flow tests. Constant strain rate and constant gradient tests gave values which agreed, within a few tens of percent, with directly measured values. Values computed from step load tests were usually too small by as much as a factor of 10.

Most instances where discrepancies arise between consolidation tests and direct determinations of K involve highly deformable media such as clays. It is likely that because of the high deformability, assumptions in standard consolidation analysis are violated sufficiently to cause errors. Precise measurement of K in highly deformable media is difficult with any technique. As discussed by Smiles and Rosenthal [1968], Smiles [1968], and Kharaka and Smalley [1976], large hydraulic gradients can result in varying degrees of consolidation in the sample; during transient flow the hydraulic properties being measured vary in time and space. Also, the larger deformations during consolidation of soft materials may need to be accounted for [e.g., Gibson et al. 1967]. Despite these potential problems, consolidation techniques seem to be quite useful. At stresses of interest in most hydrogeologic problems, media are sufficiently stiff that these difficulties are not serious.

Few instances have arisen where it is possible to compare K obtained from both hydraulic and mechanical transient tests with steady flow measurements. Moreover, little attention has been given to the comparability of specific storage, and therefore hydraulic diffusivity, derived from the two types of transient tests. One reason is that many tight media cannot be tested with both types of transient techniques. Mechanical tests are inappropriate for stiff media, such as argillites and crystalline rocks, because only a small fraction of the measured strain is caused by fluid flow and resulting pore volume change. Deformable media present the problems mentioned above when tested with hydraulic transient techniques.

Comparative data are, however, available from tests on the Pierre Shale. Because of its high loading efficiency but moderate compressibility, it is amenable to testing by both consolidation and transient hydraulic techniques. Bredehoeft et al. [1983] presented preliminary data on the Pierre Shale. More extensive data are now available and are presented in Figure 1. The hydraulic and mechanical transient flow tests (Figure 1a) are in reasonable agreement. Despite a spread in the data, they characterize K of the shale sufficiently well to be useful in analysis of the flow system. Mechanical test data points at the same stress level each represent a different sample. The spread in Figure 1a may largely reflect a subtle lithologic diversity; the variation in K with effective stress in a single sample is generally quite regular. The hydraulic transient test data also represent tests on several samples, some at multiple stress levels. In these the variation in K with effective stress in a sample was less regular.

The ability of the two different transient test techniques to characterize S, and, by extension, hydraulic diffusivity of the Pierre Shale (Figure 1b), is less satisfactory. In Figure 1b, diffusivities computed from hydraulic transient tests tend to be larger than those from mechanical transient tests (and thus S, tends to be smaller) by as much as 2 to 3 orders of magnitude. It has been recognized [e.g., Cooper et al., 1967] that the response during a transient test is generally more weakly dependent on S, than K; relatively large errors in S, are therefore possible. Many of the assumptions implicit in the description of transient flow provided by (1) are only approximately satisfied, among them the assumptions of no lateral strain and a perfectly elastic porous skeleton. While insufficient to strongly affect computed K, the associated errors may strongly affect computed S,.

Another shortcoming of currently available techniques is their inability to measure small values of S,. Small S, is found in stiff rocks which cannot be tested mechanically. Hydraulic transient tests are also unable to measure small S, because the response curves become so similarly shaped that they cannot...
be distinguished and only $K$ can be computed [Neuzil et al., 1981]. Computation of hydraulic diffusivity under these conditions is possible at present only with independent determinations of rock and grain compressibility and porosity.

A distinction between hydraulic transient and mechanical transient tests can be made on purely practical grounds. Of the two, mechanical transient tests, and step load consolidation tests in particular, are performed more routinely and use more readily available equipment. Equipment requirements for the relatively newer hydraulic transient techniques can be severe [e.g., Neuzil et al., 1981; Trimmer, 1982] and relatively few sets of test equipment have been built.

**Indirect estimates of permeability:** It is sometimes possible to estimate the permeability of tight materials using the relation between permeability and other properties. Brace et al. [1968] and Brace [1977] investigated the relation between electrical resistivity of saturated granite, its pore sizes, and permeability. Brace et al. [1968] found a consistent relation between resistivity and permeability in the Westerly granite which they extrapolated to estimate the permeability at large confining loads. Brace [1977] examined the theoretical relation between the formation factor, pore hydraulic radius, and permeability and showed that it held for a variety of porous media, including tight crystalline rocks. Heard and Page [1982] estimated permeability of the Westerly granite and Stripa granite at elevated temperatures and pressures by relating it to changes in porosity. Because microcracks transmitted the water, they related permeability to the cube of porosity.

The drawback of these techniques is that the information needed to make the estimates is often difficult to obtain. For example, establishing the relation between resistivity and permeability for the Westerly granite required Brace et al. [1968] to make several permeability measurements. The relation discussed by Brace [1977] is more general but requires the pore hydraulic radius, which is difficult to measure. The measurements of porosity change used by Heard and Page [1982] are

---

**Fig. 1.** Hydraulic properties of the Pierre Shale derived from a steady flow test and transient hydraulic and mechanical tests. The transient hydraulic test methodology used was that described by Hsieh et al. [1981] and Neuzil et al. [1981]; the transient mechanical tests were step load consolidation [e.g., Scott, 1963]. The data are plotted against effective stress and an equivalent depth in meters using the convention that effective stress increases at approximately $1.3 \times 10^6$ Pa/m: (a) hydraulic conductivity and (b) hydraulic diffusivity.
also inconvenient to make. These techniques seem to be of greatest value when the required data are already available or when estimating permeabilities over a wider range of conditions than that for which measurements are available.

**Measurement of loading efficiency:** One- and three-dimensional loading efficiencies represent the ratio of pore fluid pressure change to external stress change for conditions of one- and three-dimensional strain, respectively. In principle, the one-dimensional efficiency, \( \zeta \), can be measured in a step load consolidation test by measuring the internal pore pressure increase at the moment of load application. In practice, however, the measuring system disturbs the sample response by acting as a fluid sink because of its own compressive storage. The result is a somewhat muted pressure response with a time lag. Gibson [1963] has discussed these effects, and Whitman et al. [1961] and Perloff et al. [1965] have presented transient flow analyses which account for them. Measuring systems constructed to minimize storage have been used by Hasbo [1960] and Mitchell and Younger [1967]. Mersi et al. [1976] described tests to measure three-dimensional efficiency, \( \beta \). They derived an expression for \( \beta \) for use in undrained tests which accounts for the apparatus storage.

**In situ measurement of hydraulic parameters.** In situ testing has a different significance in low-permeability environments than in permeable ones. The volumes of tight rock tested using in situ methods are larger than in laboratory tests but are still relatively small, particularly compared with those involved in tests in permeable media. Perhaps more importantly, they are restricted to the region close to a borehole or the boundary with a permeable formation; for practical test periods a hydrodynamic disturbance penetrates only a small distance into tight rock. The tests described in this section involve the rock essentially at a point or along a line (borehole tests), or over a surface at the boundary between a low-permeability formation and one of significantly higher permeability (aquifer-confining layer tests). The primary advantage of in situ procedures is thus the ability to test large one- or three-dimensional samples of tight formations. Tests of large three-dimensional volumes cannot be made except, to a small degree, by relatively indirect methods which are also described in this section.

As in the case of laboratory testing, standard testing techniques are applicable, in principle, even when the permeability is small. In practice, however, the problems of implementing such tests can be severe. For example, the use of observation wells is generally impractical because of the long time required for the effects of injection or withdrawal to reach them. Similarly, long periods are required for water levels in a well to recover from pumping.

**Single borehole and periodic tests:** The difficulty of using observation wells in tight media dictates an important role for single-borehole tests. Significant contributions to the theory of single-borehole tests have come from both the geotechnical and groundwater disciplines. Such tests may be categorized as either slug (or recovery) tests or injection/withdrawal tests.

Slug tests are single-borehole tests carried out by abruptly increasing or decreasing the fluid head in the well and observing the transient response as it returns to equilibrium. The test was introduced by Ferris and Knowles [1954] using a line source or sink to represent the borehole. Gibson [1963] was the first to analyze the recovery from an abrupt head change in a piezometer with storage; he obtained a solution for spherically symmetric flow about a spherical piezometer tip. However, the usefulness of the solution is limited because tests are often carried out through the walls of a cylindrical borehole and because anisotropy may be expected to prevent the flow from being spherically symmetrical in many instances. An important advance was made when Cooper et al. [1967] generalized the earlier solution of Ferris and Knowles to account for borehole or standpipe storage.

The slug test procedure of Cooper et al. [1967] is not itself particularly suited to measuring small permeability. The rate of response to a slug depends on the rate at which water flows between borehole and formation and the storage in the borehole or standpipe. When the transmissivity is small, the response time can be inordinately long. This can be overcome to some degree by decreasing the standpipe diameter. However, practical limitations to the smallness of the standpipe render the test impractically long if the transmissivity is very small.

To overcome this problem, Bredehoefi and Papadopulos [1980] modified the analysis of Cooper et al. [1967] to accommodate a shut-in test procedure specifically designed for low-permeability formations. This innovation closely parallels the development of laboratory hydraulic transient techniques which were discussed earlier. Interestingly, an analysis for a shut-in slug test using a spherical piezometer tip had been outlined earlier by Gibson [1963], but its significance seems to have been overlooked. In a shut-in test, the storage in the well is due to water and system compressibility, which is much smaller than that due to water level changes in a standpipe. The response time is proportionally reduced.

Shut-in slug tests of the type described by Bredehoefi and Papadopulos [1980] were compared with conventional slug tests and recovery techniques in moderately low permeability rocks in New Mexico by Dennehy and Davis [1981]. In this case the shut-in test gave a value for transmissivity approximately 10 times larger than the value from the open slug test. This error may have resulted from the potentially erroneous methodology outlined by Bredehoefi and Papadopulos; Neuzil [1982] showed that the test procedure they described did not ensure the necessary equilibrium condition at the beginning of the test. Bredehoefi et al. [1983] used shut-in slug tests in the Pierre Shale to measure hydraulic conductivities between \( 4 \times 10^{-12} \) and \( 10^{-11} \) m/s. The smallest transmissivity they measured was \( 3 \times 10^{-10} \) m²/s in a 70 m test section.

Injection/withdrawal tests were first discussed by Gibson [1963] in a paper describing a constant head test. At the initiation of such a test, fluid is injected or withdrawn, changing the hydraulic head in the borehole or piezometer by an arbitrary amount. Thereafter the flow is controlled to maintain the arbitrary head constant. Analysis of the flow with time permits computation of \( K \) and \( S_r \). As in the case of his slug test analysis, Gibson's [1963] solution applies to the case of spherically symmetrical flow. Wilkinson [1968] later considered three dimensional and purely radial flow to a cylindrical borehole. He analyzed constant head tests in clays to obtain in situ values of \( K \) ranging from \( 10^{-7} \) to \( 10^{-9} \) m/s. Mieussens and Ducasse [1977] used a similar test procedure to measure in situ \( K \) ranging from \( 2 \times 10^{-8} \) to \( 4 \times 10^{-9} \) m/s with corresponding hydraulic diffusivity between \( 10^{-6} \) and \( 8 \times 10^{-6} \) m²/s. Mieussens and Ducasse also presented an analysis for determining the anisotropy of \( K \) when a short test section is located at the bottom of the borehole.

Periodic and related tests involve repeated stressing of the well and observation of the outwardly propagating head changes in one or more observation wells. An immediately apparent disadvantage is the necessity of an observation well; if the hydraulic diffusivity of the rock is small, the observation...
well may have to be impractically close to the excitation well in order to obtain a measurable response.

One approach which is amenable to analysis is sinusoidal pressure variations in the excitation well. The fluctuations produced in the observation well are damped in amplitude and shifted in phase, both of which are characteristics of the formation diffusivity. An analysis of the problem was presented by Black and Kipp [1981] wherein they demonstrated the practicability of the method in rocks with diffusivities as small as \(10^{-4}\) m²/s. Although rocks of reasonably small permeability may have diffusivity this large if the specific storage is quite small, much smaller diffusivities characterize many tight rocks. The technique therefore does not appear generally applicable in low-permeability environments.

A similar procedure is the repeated application of pulses or slugs in the excitation well. Walter and Thompson [1982] investigated this technique with the intention of applying it to tight formations. In spite of the limitation of utilizing an observation well they were able to measure transmissivities as small as \(3 \times 10^{-7}\) m²/s. It appears that the method would not be usable for significantly smaller transmissivities.

Aquifer-confining layer tests: Among the earliest motivations for hydrologists to study the behavior of tight confining layers was their influence on adjoining aquifers. The standard (Theis) analysis of aquifer drawdown during pumping was often seen to be inaccurate as a result of leakage from confining layers. This problem was first addressed by incorporating quasi-steady state [Jacob, 1946; Hantush and Jacob, 1955; Hantush, 1956] and then a more realistic transient [Hantush, 1960] confining layer leakage into the analytical solution for drawdown in the aquifer. Although not intended for identifying confining layer properties, Hantush's [1960] analysis permits computation of the product \(K_S\) for the confining bed when sufficient data are obtained to distinguish among his family of type curves. Bredehoef et al. [1983] took advantage of this ability to estimate properties of Cretaceous shales adjoining the Dakota aquifer in South Dakota.

A similar tack was taken by Witherspoon et al. [1962], with the difference that their techniques was motivated specifically by the desire to characterize the confining layers. Witherspoon et al. argued that the measurement of aquifer drawdown by itself was insufficient to provide good indications of confining bed properties in many instances, an argument supported by Bredehoef et al. [1983, p. 42]. Instead, they suggested that head measurements in the confining layer itself are desirable and presented a technique for analyzing such a test. Over the next several years this method was refined by Witherspoon and Neuman [1967] and Neuman and Witherspoon [1968, 1969a, 1969b, 1972]. Neuman and Witherspoon [1972] applied the technique to an aquifer-confining layer system in California; they determined diffusivities for the confining layers there to be of the order of \(10^{-5}\) m²/s.

Wolf [1970] presented a slightly different approach to the analysis of head data in a confining layer. He approximated aquifer drawdown by a step function. Head changes in the confining layer were matched to analytical type curves to obtain the hydraulic diffusivity. Wolf obtained a value of \(10^{-6}\) m²/s for confining beds in Maryland. Wolf and Papadopulos [1972] presented a similar analysis for the case where the confining bed is vertically inhomogeneous.

Indirect in situ determinations: As noted, the slow rate of propagation of pressure disturbances in low-diffusivity rocks means that only relatively small volumes of rock can be tested in reasonable periods of time. This limitation can be bypassed to some extent where a well-documented disturbance has occurred in the past.

Old excavations are analogous to load decrements in consolidation tests with the advantage that large volumes of rock and periods of years or decades are involved. Bromwell and Lambe [1968] instrumented a clay layer beneath a building excavation in order to measure the fluid heads. Comparing their head data with solutions of (1) from consolidation theory, they estimated the hydraulic diffusivity of the clay as \(3.5 \times 10^{-6}\) m²/s, a value 6 times larger than that obtained by applying laboratory consolidation tests. Lutton and Banks [1970] found subnormal heads in the clay shales near the Panama Canal. The heads were apparently still responding to excavation and resulting unloading which occurred 70 years before the measurements. Like Bromwell and Lambe, they compared theoretical behavior with observed to estimated in situ diffusivities of a clay shale unit a few tens of feet thick. They obtained values between \(10^{-6}\) and \(10^{-8}\) m²/s. Interestingly, these values compared well with laboratory-determined values, even though well-developed joints and slickensides were present in the shales. Vaughan and Walbanke [1973] used the same approach to compute the diffusivity of the London clay beneath a 7-year-old highway cut. They computed a diffusivity of the order of \(10^{-8}\) m²/s for the uppermost few meters of the clay.

Riley [1970] estimated the hydraulic properties of low-permeability lenses in an aquifer by analyzing their strain under changes in effective stress. Vertical strain in a 123-m section of well bore was measured using an extensometer. Effective stress changes resulted from cyclically varied fluid head in the aquifer, caused by agricultural pumping, and the subsequent groundwater flow between the lenses and the aquifer. Riley's analysis recognized the analogy between the field behavior and that in laboratory mechanical transient tests; in essence, the observations were analyzed as large-scale consolidation test data. The actual analysis was considerably complicated by the complex nature of the stress changes and the variability of the lens thickness. Riley estimated \(K\) of the lenses as being \(2.9 \times 10^{-11}\) m/s. He distinguished between \(S_i\) under conditions of virgin consolidation, which he estimated at \(7.5 \times 10^{-4}\) m⁻¹ and \(S_i\) in the elastic range of deformation, which he estimated at \(9.3 \times 10^{-6}\) m⁻¹. Numerical simulations by Helm [1975] indicated that Riley's estimates were quite reasonable. Helm [1976] also found that somewhat improved agreement between computed and observed strain could be gotten if stress dependent properties were used. Specifically, his results suggested that \(K\) of the lenses could vary between \(3.3 \times 10^{-11}\) m/s and \(2.9 \times 10^{-12}\) m/s with the observed changes in stress.

Analysis of purely hydrodynamic stress and response is also possible in aquifer-confining layer systems. In instances where well documented hydraulic stresses on an aquifer have occurred for an extended period, numerical simulation of the system will permit estimation of the confining layer properties. If a relatively long period of aquifer-confining layer interaction is simulated, the estimated properties will reflect a proportionally thick zone of the confining layer. Efficient numerical algorithms for simulating transient leakage from confining layers are now becoming available [Herrera and Yates, 1977; Premchitt, 1981; L. Torak and R. L. Cooley, unpublished manuscript, 1986] and should permit wider application of this technique.

In situ testing: Discussion: The small volume of rock involved in most in situ tests in low diffusivity rocks is a
stronger disadvantage than may be first be realized [Faust and Mercer, 1984; Moench and Hsieh, 1985]. The rock influencing borehole tests, for example, is a narrow annulus surrounding the borehole. This is the same region most subject to disturbance by drilling the hole; in tight media, permeability enhancement by mechanical disturbance is perhaps the most likely effect of drilling, although permeability decreases caused by clogging or by "smearing" soft formations is also possible. This problem was first considered by Wilkinson [1967, 1968], who suggested using an analytical solution by Gibson [1966] to account for the effects of "smear" around a piezometer during constant head injection tests. Faust and Mercer [1984] and Moench and Hsieh [1985] have shown that slug tests will be strongly influenced by the properties of such borehole "skins" when the shut-in technique for tight formations is used.

Tests involving confining layers and aquifers are subject to a similar disadvantage. To reduce the time for the test, the piezometer in the confining layer must be placed close to the aquifer. However, the zone near the aquifer may not be representative of the bulk of the confining bed, particularly if the contact is gradational. There are also practical problems associated with these tests. It is difficult to precisely locate and effectively isolate piezometers in a confining layer at depth.

Finally, it is worth noting that like laboratory tests in tight media, and for analogous reasons, in situ tests have utilized unrealistically large hydraulic gradients. In spite of these problems, the tests do offer the advantage of sampling large vertical sections in the case of borehole tests and large areas in the case of aquifer-confining layer tests. Inhomogeneities such as those due to fractures, and which otherwise would be missed, may be found with these tests.

**Transient flow under experimental conditions.** Published results of transient tests are remarkably consistent in one aspect; in all instances, including media with some of the smallest permeabilities yet measured, solutions of (1) appear to provide good descriptions of test behavior. Results of consolidation tests (except in very deformable media) have long indicated this was the case in soil-like media. The results of more recently developed hydraulic tests, which depend more directly on the flow, now provide stronger evidence that this is true, and for a wide variety of lithologies. Laboratory tests in crystalline rocks [Trimmer et al., 1980; Trimmer, 1982] and argillite [Lin, 1978] and both laboratory and in situ tests in shale [Neuzil et al., 1981; Bredehoeft et al., 1983] all were well described by the appropriate solutions of (1). Thus, under experimental conditions, (1) appears to be an adequate model of transient flow even in the tightest rocks.

**Nonhydraulic Flow Phenomena**

A distinguishing feature of all tight media is the smallness of the pores through which flow occurs. With a large fraction of the pore fluid in close proximity to the solid surfaces, the physicochemical phenomena which occur at the surfaces may produce significant flow effects. These are not accounted for in the Navier-Stokes flow model (on which the theoretical basis of Darcy's law rests [Hubbert, 1956]) and are called nonhydraulic flow phenomena in this paper. Reported nonhydraulic effects in tight media include those which indicate other than direct proportionality between the hydraulic gradient and flow (non-Darcian behavior) and those which indicate pore fluid flow in response to driving forces other than the hydraulic gradient (osmotic or coupled flow). Geologic media in which significant coupled flow occurs are called geologic membranes in this paper.

**Applicability of Darcy's law.** Over the past several decades many experimentalists, including Deryagin and Krylov [1944], von Engelhardt and Tunn [1955], Lutz and Kemper [1959], Hansbo [1960], Miller and Low [1963], Mitchell and Younger [1967], and others (see compilations by Kutilek [1972] and Elnagger et al. [1974]) have reported apparently non-Darcian behavior in low-permeability media. The effects reported include changes in apparent conductivity with hydraulic gradient and so-called "threshold gradients," below which little or no flow occurs.

Mechanisms suggested to explain non-Darcian behavior are generally changes in the properties of water, such as viscosity, near solid surfaces. Study of the physics of water lends some support for this view. Several discussions of water properties [e.g., Low, 1961; Mitchell, 1976, chap. 6; Forslund and Jacobson, 1975; Clifford, 1975] note the large volume of evidence for altered water properties near interfaces. However, the nature of the changes and the distance they extend from the interface are disputed. As Clifford [1975] notes, the nature and behavior of water near boundaries is one of the most controversial aspects of the study of water.

While reported deviations from Darcy's law are numerous, it is difficult to defend them from the suspicion that they are due to subtle experimental errors. Many measurements were made at or near the limit of resolution of the technique used. Some investigators, most notably Olsen [1965], have been able to demonstrate that some anomalies are almost certainly erroneous. Possible sources of error are numerous and include error in measuring gradients [Olsen, 1965], small leaks [Neuzil et al., 1981], bacterial and particulate clogging [e.g., Gupta and Swartzendruber, 1962], transient flow during a steady flow test [Olsen, 1962; Smiles and Rosenthal, 1968; Smiles, 1969; Pane et al., 1983], changes in the solid matrix [e.g., Miller and Low, 1963; Olsen, 1966; Hansbo, 1960], and gas generation and dissolution [Fettke and Copeland, 1931; Olsen, 1962]. Discussions of many of the various possible errors are presented by Mitchell and Younger [1967] and Olsen and Daniel [1981]. Interestingly, another type of nonhydraulic flow phenomenon, osmosis, is often credited with producing effects mistaken for non-Darcian behavior [Wentworth, 1944; Kemper, 1960; Low, 1961; Bolt and Groenevelt, 1969; Jackson, 1967; Miller et al., 1969]. In particular, Olsen [1969, 1985] presents persuasive evidence that chemical or electrical potential gradients are generated internally in argilaceous specimens during hydraulic flow tests, causing osmotic in addition to hydraulic flow. His experiments in kaolinite display a linear relation between flow and hydraulic gradient. However, under a hydraulic gradient only, the trend of the data does not pass through the origin. By applying external chemical or electrical potential differences, the trend could be adjusted to intersect the origin. Olsen [1985] concluded that the applied external osmotic forces cancelled those generated internally, eliminating the osmotic flow. This seems to offer a rational explanation for many reports of a threshold gradient without appealing to non-Darcian flow.

Other negative evidence for the existence of true non-Darcian flow in experiments to date include (1) carefully conducted experiments, such as those of Olsen [1965] and Smiles and Rosenthal [1968] which exhibited Darcian behavior and (2) the fact that there is little consistency in the types and magnitudes of reported anomalies [see Kutilek, 1972]. A rea-
A reasonable conclusion at present is that the case for observed non-Darcian behavior in experimental data is weak, an opinion also voiced by Mitchell [1976, p. 350]. Some deviations are always likely in any difficult measurements. It is probable that the bias of the experimenter influences whether deviations from predicted behavior are viewed as anomalies or as experimental error and noise.

Nonetheless, it would be premature to dismiss the possibility of unanticipated flow behavior in low-permeability media. This is simply because we lack experimental data for flow in tight media under realistically small hydraulic gradients. As we have seen, extremely large hydraulic gradients (up to the order of $10^6$) have been utilized in order to produce measurable flow rates. Various refinements such as shut-in transient flow methods have been introduced, but with the object, it seems, of reducing the time necessary for testing [Hsieh et al., 1981] and permitting computation of specific storage rather than reducing the applied head gradient. As useful as these advancements have proved to be, the hydraulic gradients they require are still large.

Several workers [Olsen, 1966; Mitchell and Younger, 1967; Smiles and Rosenthal, 1968; Gairon and Swartzendruber, 1975; Olson and Daniel, 1981; Pane et al., 1983] have recognized that the artificial conditions created by high gradients may affect flow behavior. Without experimental observations, the applicability of Darcy's law at small gradients can only be inferred. The gap in observational data is clearly illustrated by Figure 2, which is a plot of measured $K$ versus the hydraulic gradient, or range of gradients, used in tests described in the literature. Referring to Figure 2, we see that by taking 1 as an upper limit for representative in situ hydraulic gradients, our observation of flow behavior under these conditions extends to $K$ no smaller than approximately $10^{-9}$ m/s. Flow over nearly the entire range of low permeability, as defined in this paper, has not been observed under realistic gradients. At the lowest permeabilities, flow has been measured only with gradients several orders of magnitude larger than any in nature.

Because of this observational gap, the use of Darcy's law to describe flow under these conditions represents a hypothesis; to be on firm analytical ground, it should be tested. It is therefore desirable to extend experimental observation of flow behavior to gradients representative of in situ conditions. There are two possible approaches: (1) devising methods for measuring smaller flow rates between reservoir and sample, and (2) developing techniques for monitoring transient fluid pressures within a sample during transient flow testing. At present, the smallest flow rate measurements have apparently been achieved using a hydraulic transient flow test (Figure 2, data set 3), while steady flow tests have measured flow rates approximately 10 times larger. However, steady flow tests may offer the best chance for significant improvement, particularly with small flow rate pumps, as suggested by Olsen et al. [1985]. Steady flow tests also have the advantage of greatest simplicity and generality. Accurate measurement of transient pore pressures within a sample has been limited by storage in the measuring systems and the technical difficulty of the measurements [Ortiz, 1986].

**Coupled flow in geologic membranes.** The problem of applicability of Darcy's law, as addressed in the previous section, is distinct from that posed by flow in geologic membranes or
media in which coupled flow is significant. In the theoretical framework of coupled flow, developed through irreversible thermodynamics [e.g., Katchalsky and Curran, 1967] Darcy’s law is generally held to apply as a special case to describe the flow in response to a hydraulic gradient only.

Experimental experience indicates that certain tight geologic materials, notably argillaceous media, behave as membranes; they exhibit coupled flow. This refers to the fact that flows of one kind are caused by driving gradients of another kind. For example, fluid flow in these media is found to be driven not only by hydraulic head gradients (hydraulic flow) but also, through coupling, by gradients of chemical and electrical potential and of temperature [Mitchell, 1976]. This nonhydraulic fluid flow as a result of coupling is called osmosis. Coupling also exists between all of the driving gradients mentioned and flows of solute, electrical charge, and heat. The flows, \( J_k \), are usually assumed to be linearly related to the gradients \( X_k \), as expressed by \[ J_k = \sum_{k=1}^{n} L_{ik} X_i \quad i = 1, 2, \ldots, n \] (2)

where the subscripts \( i \) and \( k \) indicate the various types of flows and driving gradients, respectively, and their coupling coefficients. The driving gradients \( X_k \) and flows \( J_i \) which are important in geologic media are vector quantities, while the coupling coefficients \( L_{ik} \) may be scalar or tensor quantities (one such quantity is hydraulic conductivity). Onsager’s relations [Katchalsky and Curran, 1967; Bear, 1972] suggest that the coefficients \( L_{ik} \) are related as

\[ L_{ik} = L_{ki} \] (3)

By analogy with hydraulic conductivity, the coefficients for osmotic flow may be thought of as conductivity coefficients of a special sort. Thus, for example, the coefficient linking fluid flow to the gradient in chemical potential may be considered a chemical osmotic conductivity.

Of most direct concern in this paper is the flow of groundwater described by (2) (hydraulic flow and osmosis). However, the other flows described by (2) are also central to the problem. For example, flows of solute and electrical charge alter the chemical and electrical potential fields, in turn affecting chemical osmosis and electro-osmosis.

Flow of chemical solute is of particular interest in geologic membranes. The appropriate form of (2) will show that the flow of solute relative to the pore fluid is caused by several driving forces, including the hydraulic gradient. The latter coupling with the hydraulic gradient is manifested as the relative retardation of solutes when solutions flow through a membrane, a phenomenon known as ultrafiltration, hypofiltration, or reverse osmosis. The coefficients governing osmosis and ultrafiltration are related by (3) (see the discussion by Katchalsky and Curran [1967, pp. 119–126]).

The ratio of chemical osmotic conductivity and hydraulic conductivity can be shown to be a measure of the membrane’s ability to exclude, or filter, solute [Katchalsky and Curran, 1967]. Called the reflection coefficient or osmotic efficiency, this ratio is one for perfect membranes (complete exclusion of solute) and less than one for “leaky” membranes. A heuristic explanation of ultrafiltration can also be given in terms of the exclusion of ions from small pores by electrical properties of the clay [e.g., Kemper and Evans, 1963; White, 1965; Greenberg, 1971; Mitchell, 1976].

Laboratory experiments by many workers have shown that many clays and argillaceous rocks exhibit the properties described above. As early as 1948, Wyllie [1948] showed that shale samples exhibited the electrical properties of membranes. Ultrafiltration was demonstrated in bentonite by Kemper [1960] and in bentonite and other clays by McKelvey and Milne [1962]. Hanshaw [1962] demonstrated the ultrafiltration and electrode properties of montmorillonite and illite. Electro-osmotic and thermo-osmotic flow of the pore fluid were observed in a saturated soil by Taylor and Cary [1960]. Osmotically developed pressures were observed in bentonite and reconstituted shale by Kemper [1961], in bentonite by Kemper and Rollins [1966], and in undisturbed argillaceous rocks by Young and Low [1965]. The latter experiments are important because they used natural rather than prepared or reconstituted samples. These and other of the earlier experimental investigations are discussed by Berry [1969].


The general applicability of (2) and (3) in geologic media is not established. While these phenomena are demonstrable in the laboratory, certain clays and argillaceous rocks exhibit the properties described above. As early as 1948, Wyllie [1948] showed that shale samples exhibited the electrical properties of membranes. Ultrafiltration was demonstrated in bentonite by Kemper [1960] and in bentonite and other clays by McKelvey and Milne [1962]. Hanshaw [1962] demonstrated the ultrafiltration and electrode properties of montmorillonite and illite. Electro-osmotic and thermo-osmotic flow of the pore fluid were observed in a saturated soil by Taylor and Cary [1960]. Osmotically developed pressures were observed in bentonite and reconstituted shale by Kemper [1961], in bentonite by Kemper and Rollins [1966], and in undisturbed argillaceous rocks by Young and Low [1965]. The latter experiments are important because they used natural rather than prepared or reconstituted samples. These and other of the earlier experimental investigations are discussed by Berry [1969].

The flows, \( X_k \), are usually assumed to be linearly related to the gradients \( X_k \), as expressed by [Katchalsky and Curran, 1967; Bear, 1972]

\[ J_k = \sum_{k=1}^{n} L_{ik} X_i \quad i = 1, 2, \ldots, n \] (2)

where the subscripts \( i \) and \( k \) indicate the various types of flows and driving gradients, respectively, and their coupling coefficients. The driving gradients \( X_k \) and flows \( J_i \) which are important in geologic media are vector quantities, while the coupling coefficients \( L_{ik} \) may be scalar or tensor quantities (one such quantity is hydraulic conductivity). Onsager’s relations [Katchalsky and Curran, 1967; Bear, 1972] suggest that the coefficients \( L_{ik} \) are related as

\[ L_{ik} = L_{ki} \] (3)

By analogy with hydraulic conductivity, the coefficients for osmotic flow may be thought of as conductivity coefficients of a special sort. Thus, for example, the coefficient linking fluid flow to the gradient in chemical potential may be considered a chemical osmotic conductivity.

Of most direct concern in this paper is the flow of groundwater described by (2) (hydraulic flow and osmosis). However, the other flows described by (2) are also central to the problem. For example, flows of solute and electrical charge alter the chemical and electrical potential fields, in turn affecting chemical osmosis and electro-osmosis.

Flow of chemical solute is of particular interest in geologic membranes. The appropriate form of (2) will show that the flow of solute relative to the pore fluid is caused by several driving forces, including the hydraulic gradient. The latter coupling with the hydraulic gradient is manifested as the relative retardation of solutes when solutions flow through a membrane, a phenomenon known as ultrafiltration, hypofiltration, or reverse osmosis. The coefficients governing osmosis and ultrafiltration are related by (3) (see the discussion by Katchalsky and Curran [1967, pp. 119–126]).

The ratio of chemical osmotic conductivity and hydraulic conductivity can be shown to be a measure of the membrane’s ability to exclude, or filter, solute [Katchalsky and Curran, 1967]. Called the reflection coefficient or osmotic efficiency, this ratio is one for perfect membranes (complete exclusion of solute) and less than one for “leaky” membranes. A heuristic explanation of ultrafiltration can also be given in terms of the exclusion of ions from small pores by electrical properties of the clay [e.g., Kemper and Evans, 1963; White, 1965; Greenberg, 1971; Mitchell, 1976].
their mechanisms are complex; various theoretical explanations of observed behavior have been suggested by Kemper [1960, 1961], Kemper and Evans [1963], Hanshaw [1962], Kemper and Rollins [1966], Kemper et al. [1972], Kemper and Quirk [1972], Kharaka and Berry [1973], Groenevelt and Elrick [1976], Groenevelt et al. [1978], and Marine and Fritz [1981], among others. Although this work is important for understanding the mechanisms of coupled flow phenomena, it has limited predictive applicability in groundwater systems because of its complexity. A phenomenological approach, as used by Letey and Kemper [1969] and Olsen [1969, 1972], seems to offer more promise. Instead of attempting to elucidate the chemophysical phenomena involved, the latter experimentalists sought to test the applicability of (2) and (3) and to measure the coupling coefficients in specific systems. If the phenomenologic laws described by (2) and (3) apply in geologic membranes, then, in principle, the associated groundwater flow can be understood if the appropriate values of the coupling coefficients can be measured, their variability determined, and the driving gradients mapped.

Measurement of coupling coefficients: Like hydraulic conductivity, measurements of coupling coefficients can be made using a variety of techniques, including direct measurement with (2). Tests using (2) analyze steady state flow and are analogous to steady state hydraulic tests to determine K. This approach has been used almost exclusively; little development of transient testing techniques has been done. The equation for water flow may be written in the form of (2) for hydraulic, chemical, and electrical driving forces as [Olsen, 1972; Greenberg, 1971]

\[ q = -K \alpha h - K_e C_{\alpha} - K_c V E \]  

(4)

where \( C_{\alpha} \) is the concentration of dissolved solids \([M/L^2]\), E is electrical potential \([L^2M/QT^2]\), and \( K_e[LT^2] \) and \( K_c[QT/M] \) are coupling coefficients (Q denotes units of electrical charge). If dissolved concentrations are high, it may be necessary to cast (4) in terms of activities rather than concentrations. Equation (4) describes fluid flow within the membrane. Experimentally, however, it is difficult to measure the quantities of interest within a tight sample. As in permeability determinations, the conditions external to the sample are most readily monitored. For this purpose (4) may be rendered in difference form as

\[ q = -\frac{K}{L} \Delta h - K_e \frac{C_{\alpha}}{C_B} - K_c \frac{E}{L} \]  

(5)

where \( C_A \) and \( C_B \) are the solute concentrations on either side of a membrane of thickness L. Linear (steady state) gradients in solute concentration generally are not attained during testing if the sample is of appreciable thickness. However, the relation between fluid flow and external concentration differences implied by (5) applies even for nonlinear concentration gradients within the membrane when the net fluid flux is zero, which is the usual experimental condition.

Equation (5) and similar equations may be used in a variety of experimental configurations to measure the coefficients. This approach was used by Olsen [1969, 1972] to measure K, K_e, and K_c in kaolinite cakes. Olsen [1972] found that under increasing effective stress, \( K_e \) and \( K_c \) all decreased, but the decrease in K was greatest. At 0.1-MPa load, K was \( 1.3 \times 10^{-6} \) m/s, \( K_e \) was \( 7.0 \times 10^{-9} \) m^2/V/s, and K_c was \( 9.8 \times 10^{-11} \) m^2/s; at 70 MPa, K was \( 1.6 \times 10^{-11} \) m/s, \( K_e \) was \( 2.5 \times 10^{-10} \) m^2/V/s, and K_c was \( 4.3 \times 10^{-12} \) m^2/s. Letey and Kemper [1969] made similar determinations in a bentonite-water-salt system. Their results at unspecified load suggest that K_c was \( 1.1 \times 10^{-11} \) m^2/s for their system. Mitchell [1976, p. 355] has tabulated several values of K_c for geologic media, nearly all of which fall in the range \( 10^{-9} \) to \( 10^{-8} \) m^2/s.

Equation (4) and the analogous equation for solute flow may be combined with statements of mass conservation for solution and solute to obtain descriptions of transient behavior associated with coupled flow (see the discussion of large-scale membrane effects). One such transient phenomenon is osmotic consolidation of a membrane, which occurs when osmosis causes water to flow out, producing a mechanical strain. Both chemical and electrical osmosis have been observed to produce this effect [Greenberg, 1971; Mitchell, 1976]. Greenberg [1971] analyzed chemicoosmotic consolidation to compute coupling coefficients. Results of tests using bentonite cakes at 0.1-MPa loading suggest a K_c of \( 6.5 \times 10^{-11} \) m/s with a K of \( 2.2 \times 10^{-11} \) m/s.

The results cited, particularly Olsen’s [1972], are intriguing because they indicate that coupling is significant at moderate depths in clays with only moderately low K and that its relative importance may be even greater in tighter materials and at greater depths. However, despite this extensive body of experimental evidence, the significance of coupled flow in situ remains controversial (see the discussion of large-scale membrane effects). A question arises as to whether further experimental work is likely to play a significant role in resolving the issue of large-scale behavior in situ. This may be related to waning activity in this line of investigation during the last decade.

While further experimental work alone is not likely to resolve questions surrounding larger-scale problems, certain extensions of the experimental groundwork appear to be not only desirable, but necessary. The lack of data for coupled flow in undisturbed geologic media under in situ conditions is particularly conspicuous. Other important questions requiring experimental investigation include (1) under what conditions the phenomenologic relations (2) and (3) are adequate, (2) how coupling coefficients vary on what they depend, (3) how the clay mineralogy (e.g., kaolinite versus montmorillonite) affects the values of the coupling coefficients, (4) whether coupling coefficients exhibit significant anisotropy, and (5) in systems with several different ions, the nature of their interaction.

Important questions also surround coupled flow behavior within the membrane medium itself, particularly when this behavior is transient. Experimental techniques utilized so far for the study of geologic membranes have been limited in this respect. Specifically, they have utilized only observations of conditions external to the membrane, and, with the exception of osmotic consolidation studies by Greenberg [1971], they have considered only steady flow. Thus there is a dearth of direct information on transient behavior within the membrane with which to compare theoretical predictions.

If coupled flow proves to be a significant component of certain flow systems, an understanding of these systems will require answers to the questions posed above. Problem specific testing, like that routinely done for hydraulic properties, will also be required to determine the values of the coupling coefficients. This is because of the multiplicity of factors affecting the coefficients, which include effective stress, absolute dissolved concentrations, the particular ions involved, both absorbed and in solution, the lithology of the membrane, and, presumably, spatial changes in the type as well as the concentration of ions.
LARGE LOW-PERMEABILITY FLOW SYSTEMS

... and man, who finds himself constrained by the want of time, or of space in almost all of his undertakings, forgets, that in these, if in any thing, the riches of nature reject all limitation.

Playfair [1802, p. 137]

As suggested in the introduction, large-scale groundwater movement in regions known to have low permeability on a small scale is problematical. This is particularly true when the flow is transient, because the flow dynamics are difficult, if not impossible, to observe. Indeed, it may be difficult to determine whether or not the groundwater flow is transient and which are the relevant driving processes. “Large scale” is applied here to size and time scales significantly larger than those involved in direct testing, and ranging upward to encompass regional flow and geologically significant periods of time.

The motivation for attempting to analyze large-scale flow in low-permeability environments can be conveniently considered within the framework of prediction problems, inverse problems, and detection problems. Prediction has obvious significance for certain purposes, such as isolation of toxic wastes. The inverse problem is highly significant in low-permeability environments because it may be the only way of estimating parameter values appropriate at large scale. Few analyses have been posed as detection problems, in which one attempts to understand past stresses on the system from its current condition. However, it may be a useful approach in some instances; an example is detection of cumulating mechanical strains as evidenced by fluid pressures.

Despite such a classification, the problems, in reality, are closely interrelated. One may attempt to sort out the flow processes which are responsible for current conditions in a low-permeability system. This will lead to a better understanding of the flow history, the chemical evolution of the groundwaters, and the relation of the groundwaters to the geologic evolution of the system. As a corollary, if the history can be understood, it may also be possible to improve predictions of behavior for significant periods of time in the future.

Having considered experimental investigations of low-permeability flow in the first part of this paper, we are in a position to examine their results in relation with efforts to analyze large flow systems. We earlier recognized experimentation and testing which addressed two types of questions. One line of inquiry concerns the nature of flow laws and constitutive relations controlling groundwater movement in tight media, while the other accepts the premises of standard mathematical models of groundwater flow and seeks to evaluate the parameters in these models. The extension of small-scale observations to large-scale analysis may be considered in a parallel fashion.

Let us first examine our understanding of the nature of low-permeability flow, as derived from experimentation, in relation with large-scale theoretical analyses to date. The following broad generalization can be drawn:

1. Hydraulic gradients appear to cause flow in accordance with Darcy’s law at all observed gradients, but realistically small gradients have not been tested. In the absence of direct observations of flow behavior at reasonable in situ gradients, most investigators, whether conscious of the problem or not, have assumed that Darcy’s law forms an adequate basis for analysis. In view of the evidence at present, this is the more conservative approach. However, the readiness with which this premise is accepted undoubtedly stems from the demonstrable success of Darcy’s law in more permeable media over the range of in situ gradients. A few analyses have been made, which will be mentioned, using relations other than Darcy’s law.

2. Transient flow under experimental conditions is well described by the same mathematical descriptions widely used for more permeable media, one form of which is expressed by (1). The underlying model assumes linear elastic deformation of the solid skeleton in response to changes in effective stress. This is a poor assumption over the long periods of interest in large-scale systems. It has, nonetheless, formed the basis for most analyses. This can be laid to its successful application in conventional groundwater studies, familiarity of descriptions such as (1), and the suite of solutions available. There has also been a lack of consensus on the need for a more complex model and the form it should take.

3. Nonhydraulic coupled flow is often experimentally significant in argillaceous media. With few exceptions, however, quantitative analyses of flow in argillaceous environments have ignored osmosis and other coupled flow phenomena. There are at least three possible reasons for this. First, there is a great deal of uncertainty related to extrapolating laboratory experience to field scales and subsurface environments, leaving unanswered the question of its significance in situ. Second, ignoring coupled flow has often been a matter of expedience; it significantly complicates the analysis, is imperfectly understood, and has data requirements which are difficult to meet. Third, abundant experience with more permeable media, in which coupled flow is demonstrably unimportant, may have increased acceptance of the assumption of insignificant coupling.

Next, consider the relation between experimentally measured parameters and their large-scale counterparts. There is a need for estimates of parameters which describe large volumes of rock over long periods of time if large-scale flow is posed as a prediction or detection problem. Aside from estimates of regional K which can be made in certain steady flow systems and indirect techniques such as geochemical studies, data from short-term tests involving relatively small volumes of rock provide the only information about these properties. For lithologic types or for specific formations these data suggest ranges of values for the large-scale parameters.

Experimental data are helpful in suggesting minimum permeability values for large volumes of rock because fractures and other discontinua tend to increase the aggregate value. Extrapolation to field scale thus involves an uncertain size scale dependency for K. Brace [1980, 1984] has compiled laboratory and in situ permeability measurements for argillaceous media and crystalline rocks. His compilation suggests that in argillaceous media, K can range as small as $10^{-15}$ m/s and in crystalline rocks as small as $10^{-16}$ m/s. He notes that while K may not increase drastically with scale in argillaceous media, it often does in crystalline rocks because they can maintain transmissive fractures. Walder and Nur [1984] have argued, however, that large volumes of low-permeability crystalline rock can exist at depth because fractures fill or heal.

For transient flow, the quantity $K/s$, known as hydraulic diffusivity and here denoted $K(L^2/T)$, is more significant. Few comparisons of measured $K$ in tight media are available. Therefore I have compiled from various sources measured values of $K$ for a variety of tight lithologies; these are shown in Figure 3 plotted against depth or equivalent effective stress. The bulk of these data are from transient laboratory tests discussed earlier. A few were determined using in situ techniques. The difficulty of reliably determining $s$, and therefore $K$ by experimental techniques should be borne in mind when
Fig. 3. Measured hydraulic diffusivity ($k$) for a variety of low-permeability media plotted against effective stress during the measurement and the equivalent depth. Most data were obtained from laboratory tests. Data are presented for London clay [Vaughan and Walbancke, 1973], Tertiary clay [Wolff, 1970], clay aquifers in California [Neuman and Witherspoon, 1972], Bearpaw Shale and Morden Shale [Balasubramonian, 1972], Fort Union Shale [Smith and Redlinger, 1953], Cucaracha Formation and Culebra Formation [Lutton and Banks, 1970], Pierre Shale [Bredehoeft et al., 1983; C. E. Neuzil, unpublished data, 1986], Eleana argillite [Lin, 1978], and Westerly Granite and Creighton Gabbro [Trimmer et al., 1980]. D. Trimmer kindly supplied additional data necessary for computing the diffusivity of the granite and gabbro.

Fig. 4. Plot comparing experimental values of vertical compressibility computed from representative porosity-depth relations in sedimentary sequences. The shaded region encloses all experimental data from sources cited by Neuzil [1985]. Curve A was computed from porosity data for the Gulf Coast presented by Dickinson [1953]; curve B is from a representative porosity-depth relation suggested by Cathles and Smith [1983]. The convention that effective stress increases at approximately $1.3 \times 10^4$ Pa/m was used to compute curves A and B and to plot the experimental data on an equivalent depth scale.

Interpreting these data. Figure 3 suggests that experimental $k$ for argillaceous media often lies in the range of $10^{-9}$ to $10^{-7}$ m$^2$/s, while that of intact crystalline rocks tends to be somewhat larger because of their small specific storage and lies in the range of $10^{-6}$ to $10^{-3}$ m$^2$/s.

The relation between these data and the appropriate large-scale values is more equivocal than in the case of hydraulic conductivity; unlike $K$, these data may not indicate minimum values for large volumes. While the size scale dependence of $K$ is important, the uncertainty associated with laboratory values of $S_e$ is mainly its time scale dependence. This is an artifact of the choice of transient flow model, as elaborated below.

The parameter $S_e$ is defined using elastic constants. However, it appears that (1) geologic media deform viscoelastically over long periods, and (2) the equivalent, or effective, elastic compressibility needed to account for the long-term viscoelastic deformation can be significantly larger than the experimental compressibility. Viscoelastic deformation is known experimentally [e.g., Šukljje, 1969] and from settlement structures, [Bjerrum, 1967] but extrapolation of this observed behavior much beyond a human time scale is uncertain. A better indication of the potential significance of viscoelastic effects is provided by geological evidence. Porosity-depth profiles permit computation of effective long-term compressibility if one assumes that porosity decrease with depth is due solely to one-dimensional mechanical compression. Figure 4 is a plot comparing vertical compressibility computed from representative porosity profiles in sedimentary sequences with experimental values for sedimentary media obtained from published sources. Figure 4 suggests that effective compressibility over geologic time may exceed experimental values by 1 to 3 orders of magnitude. The effective long-term compressibilities computed in this fashion may be too large because the porosity changes in situ can also reflect extraneous effects such as diagenesis and horizontal strain. Nonetheless, the data strongly suggest that there is significant viscoelastic deformation over geologic time in sedimentary media; if we choose
to apply models based on elastic behavior, the corresponding effective values of $S_e$ can therefore be substantially larger than experimental values. Thus the data in Figure 3 should be viewed with the understanding that the effective values of $\kappa$ may be either larger or smaller. Loading efficiencies, which are also defined with elastic constants for the solid skeleton, are probably also significantly greater over long periods of time than laboratory data suggest.

Analyses of large-scale low-permeability systems are limited to varying degrees by the problems discussed above. The following sections consider analysis of large-scale systems with steady state, transient, and nonhydraulic flow, respectively. These are followed by a discussion of the problems of site specific studies.

### Steady Flow in Low-Permeability Environments

Attainment of steady flow in a low-permeability flow system is favored by a stable geologic setting, small $S_e$ and small dimensions. An important instance where steady flow is often assumed to be occurring is in strata confining unconfined aquifers. Indeed, the earliest recognition of low-permeability formations seems to have been in their role in steady state flow systems as confining layers. Chamberlin [1885] and Darton [1909] were among the first to recognize the significance of steady state leakage through confining layers.

Walton [1960, 1965] was probably the first to attempt to quantify steady state leakage through confining layers. To do this, he analyzed transient pumping test data to characterize the vertical permeability of the confining layers. His results can be questioned because he used an analysis [Hantush and Jacob, 1955] which did not allow for confining layer storage. However, his work provided an early indication of the large fraction of the flow which can leak through confining layers.

Steady state flow presents the possibility of determining the regional $K$ of the tight units. The flow in the system is determined by the geometry of the aquifer and confining layer, the boundary conditions, aquifer conductivity, and leakage through the confining layer. Thus, if one knows the head in the aquifer, leakage through the confining layer (and thus $K$) may be estimated if the other properties are known.

This approach has recently been used in a number of studies of tight confining layers. Bredehoft et al. [1983] studied the Cretaceous shales confining the Dakota aquifer in South Dakota. Quasi-three-dimensional numerical simulations of the flow in the aquifer-confining layer system over an area of $3 \times 10^5$ km$^2$ indicated that most of the flow through the aquifer system occurred as leakage through the confining layers. Further, the hydraulic head in the aquifer was strongly controlled by this leakage. Bredehoft et al. [1983] found that the regional $K$ of the composite confining layer was $6 \times 10^{-11}$ m/s and that for the individual shales it ranged from a low of $5 \times 10^{-12}$ m/s at depth to $2 \times 10^{-9}$ m/s near the surface. Interestingly, these regional values were as much as a factor of $10^3$ larger than laboratory and in situ determinations of $K$ of the shale. Bredehoft et al. [1983] and Neuizil et al. [1984] interpreted this result as indicating fracture enhancement of the regional $K$. Wirojanagud et al. [1984] and Sengor and Fogg [1984] analyzed the flow system in the Palo Duro basin, Texas, using quasi-three-dimensional and vertical two-dimensional simulations, respectively. Like Bredehoft et al. [1983], they concluded that leakage through an extensive confining layer, in this case composed of evaporites, constituted a large fraction of the flow in the system. The simulations provided an estimate of $10^{-12}$ m/s for the regional $K$ of the evaporites. This is close to the value suggested by small-scale testing, Belitz [1985] analyzed flow in the Denver Basin which contains a thick, low-permeability sequence of Cretaceous shales. Because of the nature of the flow system there it was possible only to determine a maximum regional $K$ of approximately $3 \times 10^{-13}$ m/s for the sequence. Butler [1984] simulated flow in the Williston Basin and obtained estimates of confining layer $K$. However, his estimates may be inaccurate because of an assumption of steady state conditions soon after development of the aquifers.

As valuable as these regional determinations are, they represent such large volumes of rock that behavior at intermediate scales is problematical. While Bredehoft et al. [1983] show the probable presence of fractures which enhance the shale permeability, Neuizil et al. [1984] suggest that large blocks of lower-permeability shale may exist between the fractures. They hypothesize that fractures may dominate the regional leakage across the shales at near-steady state conditions while much more slowly responding transient conditions exist in the blocks.

The determination that steady state flow indeed exists in a large volume of tight rock is not easy to make. Bredehoft et al. [1983] argued the likelihood of steady state flow conditions in the Dakota aquifer confining layers on the basis of their geologically stable setting. The Denver Basin, simulated by Belitz [1985] as being at steady state, has in fact often been interpreted as being transient [e.g., Otiman, 1984]. In these instances, as in most, knowledge of hydraulic head within the low-permeability units is insufficient to make firm interpretations. One can only demonstrate, as Belitz [1985] did, that observed conditions are consistent with a steady state system.

### Persistence of Transient Flow

One of the more interesting and challenging aspects of low-permeability environments is transient flow on a large scale. Many of the “anomalous” pressures which occur worldwide in a variety of geologic settings appear to be manifestations of very old transient conditions in low-permeability environments.

As noted earlier, analyses of large-scale transient flow in tight environments have usually been based on well-established descriptions applied in permeable media. In order to understand the various analyses using this approach, their relations to one another, and their limitations, it is necessary to consider a fairly general description of transient flow in an elastic porous medium. Biot [1941] first presented such equations accounting for the complete coupling between stress and strain in the porous skeleton and fluid pressure in the pores. Rice and Cleary [1976], using the results of Nur and Byerlee [1971], generalized Biot’s equations for the case of compressible grains. In more familiar notation [see van der Kamp and Gale, 1983], their equations may be written as

$$\nabla \cdot K \nabla h = S_e \frac{\partial h}{\partial t} - (\alpha - \alpha_e) \frac{\partial s_i}{\partial t} - \frac{1}{v} \frac{\partial v}{\partial t} \tag{6}$$

and

$$\nabla^2 s_i = \gamma \lambda \nabla^2 h \tag{7}$$

The quantity $S_e$ is a three-dimensional specific storage [$L^{-1}$] [van der Kamp and Gale, 1983]. It and other notations are defined at the end of the paper.

The term $(\partial v/\partial t)/v$ accounts for a variety of processes which result in an actual or apparent change in fluid volume. These
include temperature changes, fluid release or uptake during
diagenetic reactions, and porosity changes due to strain of the
porous skeleton, pressure solution, and other causes. The in-
clusion of this term is intended merely as a reminder of the
multiplicity of processes which are not usually considered in
permeable environments but which may play a role in causing
transient flow when the permeability is small. The stress
change term, $\partial \sigma / \partial t$, which affects flow through porosity
change, could have been lumped under this term. Equations
(6) and (7) assume that the solid grains and matrix are per-
fectly elastic and the strains are small. For nonhomogeneous
fluids they can be formulated in terms of pressure rather than
hydraulic head.

Examination of the relation expressed by (6) affords a qual-
titative understanding of one of the most significant aspects
of low-permeability environments, namely the great longevity of
transient flow and the consequent ability of slow geologic pro-
cesses to profoundly affect flow. This is best seen when (6) is
nondimensionalized; in this way, systems with differing values
of $K$, $S$, and different dimensions can be reasonably compared.
It is helpful at this point to note that the quantity $K/S$ is a
three-dimensional hydraulic diffusivity, here denoted $k[L^2/T]$.

As an example, consider the effect of stress changes, $\partial \sigma / \partial t$,
caused by geologic processes. The nondimensional version of
the stress term in (6) may be approximated as $\partial (k \partial \sigma / \partial t)$. In
this form it is clear that even when rates of stress change are
small, the term can be numerically significant if $k$ is sufficiently
small and $l$ is sufficiently large. Studies of one-dimensional
systems [Gibson, 1958; Neuzil and Pollock, 1983; Walder and
Nur, 1984] indicate that the effect of stress changes on flow
becomes significant when this term is of the order of 0.1. Sugges-
ted rates of stress change caused by erosion [Neuzil and
Pollock, 1983], deposition [Bredehoeft and Hanshaw, 1968], or
tectonic activity (M. D. Zoback, personal communication,
1984), together with values of $k$ suggested by the one-
dimensional values in Figure 3, indicate that the term can
significantly exceed 0.1 in reasonably large regions ($l \approx 100$ m).

If a geologic process maintaining transient flow ceases, the
flow evolves to a steady state. The response time or time to
attain steady state may be referenced to dimensionless time,
$t^*$, another quantity appearing in the nondimensional form of
(6); it is defined by $t^* = t/k^2$. Although exact behavior in any
instance depends on the domain geometry and boundary con-
ditions, some generalization is possible. Analytical solutions to
the transient flow equation (see, for example, the transient
flow solutions of Bredehoeft and Hanshaw [1968], Hanshaw
and Bredehoeft [1968], and Hsieh et al. [1981] and consoli-
dation theory in the work by Šukije [1969, p. 140]) indicate that
transient behavior persists until $t^*$ is of the order of 1. We
can thus compute the approximate time involved as $t = t^* k^2 /\kappa = t^* k^2 /k^2$. Again using $k$ values suggested by Figure 3 and
assuming reasonable dimensions ($l \approx 100$ m), we arrive at
values of $t$ ranging up to $10^6$ yr and greater. This illustrates
that transient flow response times may be comparable with
geologic time scales.

As the square of $l$ appears in $t^*$, the dimensions of the
low-permeability region are particularly important. If $l = 1000$
$\text{m}$ rather than 100 $\text{m}$, the computed times are 100 times
greater. Conversely, if transmissive fractures are present, and $l$
is smaller, the response time is drastically reduced. The signifi-
cance of size is particularly well illustrated on a large scale by
shale units in the South Caspian basin described by J. D.
Bredehoeft and R. D. Djervanshir (unpublished manuscript,
1986). They presented data from a region which has been
subjected to relatively homogeneous depositional loading
which shows that the transient excess pressures in the shale
become more pronounced as the average thickness of the
shales increases.

These simple computations illustrate the necessity of con-
sidering the effects of slow geologic processes on groundwater
flow in low-permeability environments. Of course, flow in rel-
tively permeable regions must also change as the geologic
framework evolves in time; however, in these the flow adjusts
so readily that it remains quasi-steady (or fully steady from
a human viewpoint) through all but the most rapid distur-
bances. Only in low-diffusivity regions does fully transient
flow develop. The significance of low permeability in this con-
text was stated by Dickenson [1953] and was discussed by
Hubbert and Rubey [1959] in their classic paper on thrust
faulting. However, the quantitative relation between rates of
diagenetic change, hydraulic diffusivity, and transient flow
seems to have been first appreciated and explored in a pair of
papers by Bredehoeft and Hanshaw [Bredehoeft and Hanshaw,
1968; Hanshaw and Bredehoeft, 1968]. In these papers a vari-
ty of geologic processes potentially capable of inducing long-
lived transient flow were analyzed.

The interaction of geologic processes and transient flow can
be broadly generalized. Some geologic processes affect the
flow by altering conditions at the boundaries of the low-
permeability region. For example, uplift or erosional exposure
of outcrops can alter hydraulic head at a boundary. Other
processes, such as stress and temperature changes and related
diagenetic effects, act to produce, in effect, distributed fluid
sinks or rises as expressed in the last two terms of (6). Excess or
deficient heads are generated which either slowly dissipate or
are maintained if the generating process continues. Flow into or
out of the low-permeability region over long periods of time can
occur as a result, a point of interest for the confinement of waste.
This is illustrated diagrammatically in Figure 5 for the case of
vertical flow.

It appears that in nature, processes causing changes which
act as effective fluid sources are most common. As an exam-
ple, burial metamorphism may cause porosity reduction and fluid
generation but has no counterpart during erosional exhaus-
tion. Thus one would expect excess transient pressures to be
more common than deficient transient pressures, which indeed
seems to be the case [Fertl, 1976].

The interaction between geologic processes and flows is cen-
trally important to the consideration of transient flow phe-
nomena. Large-scale quantitative analyses which incorporate
it have usually addressed problems involving depositional
burial or erosional exhaustion and associated physical and
chemical phenomena. Therefore the following discussion will
be centered on these processes and the approaches devised to
analyze them.

Effects of geologic processes: Simple burial and den-
dation. Burial and denudation are perhaps the most ubiqui-
tous and significant geologic processes affecting flow in
low-permeability environments; few areas escape their effects
over geologically significant periods. They are relatively ame-
table to analysis because the stress and temperature history can
be directly related to overburden history. Also, erosion and sedi-
mentation are often laterally extensive, enabling simplification
of the analysis to one dimension vertically. A great deal of
insight has been gained through analysis of these processes
because nearly all the significant types of geologic effects capa-
bale of affecting flow, such as stress and temperature changes
diagenesis, can be involved. In particular, sediment
burial has been studied because it occurs in areas such as the Gulf Coast, where significant excess transient pressures have been encountered during petroleum exploration.

Stress effects: If the processes considered are laterally extensive, and lateral strains can be ignored, (6) and (7) may be simplified to a single equation. Using Biot's [1941] constitutive relation between stress, strain, and fluid head, this can be shown to be

\[ \frac{\partial^2 h}{\partial z^2} = \frac{\partial h}{\partial t} - \frac{\zeta}{\gamma_f} \frac{\partial S}{\partial t} \]  

A similar equation was first presented by van der Kamp and Gale [1983]. One-dimensional hydraulic diffusivity, \( \kappa \), is defined as \( K/S \), where \( S \) is one-dimensional specific storage and \( \zeta \) is the one dimensional loading efficiency. Equation (8) completely describes the coupling between stress and fluid flow in one dimension. The expression used in many transient laboratory test techniques, (1), is a form of (8); as used here, the storage parameters \( S \) and \( \zeta \) correspond exactly, in definition, to those obtained in one dimensional tests.

For problems involving change of overburden it has also been found helpful to cast stress in terms of overburden thickness and to use excess pressure in place of pressure [Gibson, 1958] or, in this instance, excess head in place of head. Following Neuzil [1985], (8) may thus be written as

\[ \kappa \frac{\partial^2 h'}{\partial z^2} = \frac{\partial h'}{\partial t} - C \frac{\partial L}{\partial t} \]  

where \( C \) is a dimensionless coefficient incorporating the loading efficiency and defined by

\[ C = \left[ \gamma' \zeta - \gamma_f \right] / \gamma_f \]  

and \( \partial L/\partial t \) is the erosion or deposition rate, \( L \) being the land surface elevation. This form of the equation simplifies the problem by ignoring variations in sediment specific weight, \( \gamma \), with depth. For most geologic problems this is held to be a reasonable simplification because most of the variation in \( \gamma \) occurs close to the surface.

The resulting formulation, (9), describes the transient groundwater flow in response to depositional loading \( (\partial L/\partial t > 0) \) or erosional unloading \( (\partial L/\partial t < 0) \) in a convenient fashion. If \( C \) is positive, (9) predicts that deposition \( (\partial L/\partial t > 0) \) will produce excess heads and erosion \( (\partial L/\partial t < 0) \) will produce subnormal heads. The opposite is true when \( C \) is negative. The creation of "fossil" pressures or "paleopressures," which has sometimes been discussed [e.g., Bradley, 1975; Fertl, 1976; Tóth and Millar, 1983], is conceived as resulting from a process which is described by (9) when \( C \) is negative.

Whether "fossil" pressures, in the sense discussed here, actually exist is questionable; data presented by Neuzil [1985] indicate that \( C \) is rarely negative. In general, the larger the absolute value of \( C \), the more pronounced the effects of stress change associated with deposition or erosion on the fluid pressures.

The preceding discussion implicitly assumes that the process being considered acts on a sequence which is "normally" pressured initially. This need not be the case; for example, it is entirely possible for excess fluid pressures to exist, as a result of earlier depositional or tectonic processes, when an episode of erosion begins. The tendency for erosion to reduce fluid pressures would then be superimposed on the preexisting fluid pressures. If the erosion were not sufficiently vigorous, it could fail to reduce the pressures to hydrostatic or lower values; excess pressures would persist. Likewise, deposition could occur above a "subnormally" pressured sequence and fail to raise pressures above hydrostatic. In any event, the effects of processes are superimposed on the preexisting fluid pressures.

The process described by (9) has been qualitatively appreciated for some time, particularly among some petroleum geologists. Faced with many examples of "anomalously" high pressures encountered during drilling, they cited the effect of increasing overburden weight as a possible cause. Dickinson [1953] was perhaps the first to suggest sedimentary loading of a low-permeability sequence as a cause of overpressures found in the Gulf Coast. Hubbert and Rubey [1959] also suggested that it may sometimes cause the high pore pressures they implicated in thrust faulting.
If the grains are assumed to be incompressible ($\kappa = 0$) and one further assumes $x_e \gg \kappa \rho f$, (9) can be written as

$$\frac{\partial^2 h}{\partial x^2} = a \frac{\partial T}{\partial t} - b$$

(11)

Gibson [1958] first derived (11) and solved it analytically for the case of a layer which thickens at a constant rate. Bredehoft and Hanshaw [1968] recognized that the solution could be applied to the geologic problem of continuous sedimentation in a basin and used it to evaluate loading as a cause of abnormally high pressures in the Gulf of Mexico. Their results can be viewed as a solution to the inverse problem; they concluded that if the sediments had a $\kappa$ of $10^{-8}$ m$^2$/s or smaller, sedimentary loading by itself could cause the excess pressures often encountered in the Gulf Coast. Somewhat later, and apparently independently, Smith [1971, 1973] examined the problem of sedimentary loading using an equation equivalent to (11) but derived in terms of porosity. Bishop [1979] used (11) to analyze pressure generation in tight sediments being buried beneath permeable strata. Keith and Rimstidt [1985] followed an approach similar to Smith’s and also derived their equation in terms of porosity. Smith, Bishop, and Keith and Rimstidt all concluded that sedimentary loading alone is capable of generating and maintaining the excess pressures in the Gulf Coast environment.

Both Smith’s [1971, 1973] development and that of Keith and Rimstidt [1985] accounted for dimension changes due to vertical strain of the porous matrix as well as changes in matrix compressibility with effective stress. Finite strain in vertically loaded media has also been analyzed in the context of soil mechanics, most notably by Gibson et al. [1967]. A comprehensive discussion of finite strain analyses is given by Pane [1981]. For the geologic problems considered here, neglect of dimension changes due to porous matrix strain does not introduce significant error.

Erosional unloading may also explain subnormal pressures in some instances as speculated by Russell [1972] and Dickey and Cox [1977]. Neuil and Pollock [1983] solved (11) numerically for the case of erosion at a constant rate and concluded that subnormal transient pressures could be generated. They also suggested that unloading presents the possibility that pressures will drop into the negative range or that desaturation may occur. Considering the problem in the context of soil mechanics, Koppula and Morgenstern [1984] obtained analytical solutions for (11) which described erosion.

**Thermal effects:** Temperature changes may be expected as deposition or denudation displaces rocks through the geothermal gradient and heat is convected with flowing groundwater. Rocks undergoing burial are heated and then, if denudation of the overburden occurs, are cooled. Barker [1972] used a very simplified analysis, which allowed for no solid strain or fluid flow, as a basis for suggesting that thermal expansion of water may cause anomalous pressures. Controversy over the relative importance of thermal and loading effects ensued, and several qualitative arguments for various views were presented [Bradley, 1975, 1976; Dickey, 1976; Chapman, 1980; Plumley, 1980].

Thermal changes cause thermoelastic responses in the solid matrix which act to change the pore volume and thus the fluid pressures, as well as the state of stress. Concurrently, thermal expansion of the fluid alters the fluid pressure, which affects the state of stress in the solid. As Palciauskas and Domenico [1982] point out, the resulting problem is quite complex; however, they suggested that thermomechanical effects in the solid can be ignored. Considering the magnitude of thermal expansivity of water, minerals, and rocks, Delaney [1982] also argued that the thermal effect on the solid could be ignored for porosities exceeding 0.01. That is, the expansion of water will dominate the thermal effects. For nonisothermal conditions under these assumptions the flow and deformation is the rock are described by (6) and (7). In this case, (6) is written as

$$\nabla \cdot KV h = S_f \frac{\partial h}{\partial t} - \frac{a - a_f}{a_f} \frac{\partial \sigma_t}{\partial t} - \frac{a_f}{V} \frac{\partial T}{\partial t}$$

(12)

where $\alpha_f$ is the thermal expansivity of the pore water [$^\circ$C$^{-1}$]. The last term in (12) is a specific version, for nonisothermal conditions, of the term $(\partial h/\partial t)/v$ found in (6).

It will be noted that (12) and equations which follow are written in terms of head; buoyancy effects are thus ignored. Although investigators such as Domenico and Palciauskas [1979] wrote their equations in terms of pressure, they too ignored buoyancy effects. Indeed, in one-dimensional analyses in which free convection cannot occur, buoyancy has little effect on the flow. The one-dimensional simplification is not particularly restrictive in this regard because free convection is likely only in relatively permeable settings. However, Blanchard and Sharp [1985] have argued that free convection can occur in thick sequences with permeability as small as that corresponding to a $K$ of $10^{-9}$ m/s. This suggests that free convection may be possible in the most permeable settings which we are concerned in this review.

The movement of heat and groundwater in porous media are coupled phenomena. Heat is transported by moving groundwater as well as by conduction, and temperature changes affect the flow through the last term in (12) and through buoyancy changes. Coupled flow of heat and groundwater have been analyzed by Bredehoft and Papadopoulos [1965], Domenico and Palciauskas [1973], Sharp and Domenico [1976], Sharp [1976], Garven and Freeze, [1984a, b], and Bethke [1985], among others. Studies by these investigators have served to demonstrate that while groundwater flow may noticeably affect temperature distributions in low-permeability systems, conductive heat transport is dominant over convective heat transport. That is, in low-permeability systems, groundwater flow does not significantly alter the temperature distribution from that produced by conduction alone. This can be seen by the approximately linear thermal profiles computed by Sharp and Domenico [1976] and the approximately linear thermal profiles present in the Gulf Coast [Boedner et al., 1985].

While Bodner et al. [1985] correctly argue that convective perturbations of temperature may be valuable indicators of flow, the preceding observations have led to the use of justifiable simplifications of the coupled system. Thus Domenico and Palciauskas [1979] and Palciauskas and Domenico [1980] assumed that, during burial or exhumation, slow displacement along the geothermal gradient occurs without significantly disturbing it. The change in temperature at a point in the sediment column can therefore be directly related to the erosion or deposition rate by the geothermal gradient. Under this assumption and in one dimension, (12) and (7) become, in terms of excess head $h'$,

$$\frac{\kappa}{c} \frac{\partial^2 h'}{\partial x^2} = \frac{\partial h'}{\partial t} - \frac{\partial L}{\partial t}$$

(13)

$$C = \frac{\gamma - \gamma_f}{\gamma_f} + \alpha_{Tf} \frac{mG}{S_f}$$

where $C$ is now given by

$$C = \frac{\gamma - \gamma_f}{\gamma_f} + \alpha_{Tf} \frac{mG}{S_f}$$
A similar equation, for the more restrictive assumptions that \( x_f = 0 \) and \( \zeta = 1 \), was first obtained by Domenico and Palciauskas [1979]. A comparable equation was also presented by Walder [1984]. Equation (13) becomes identical with (9) in the absence of thermal effects (when \( x_f \) or \( G \) is zero). Domenico and Palciauskas [1979] and Palciauskas and Domenico [1980] discussed the solution of (13) in a simpler form of \( C \) which they assumed to have a constant, representative value. Their results indicated that the thermal contribution to pressure was significant and, in fact, might permit pressures tending to exceed lithostatic. They surmised that this would lead to inelastic dilation of the rock.

Treatments of the problem of burial and exhumation are distinguished largely by differences in the assumed relative importance of water \( (x_f) \) and grain compressibility \( (x_g) \) and water thermal expansivity \( (x_{fT}) \) as compared with rock compressibility \( (x_r) \). The assumptions employed by various investigators can be expressed as different forms of the dimensionless coefficient \( C \) multiplying \( dL/dt \); this provides a convenient method of summarizing and comparing approaches. Several of the analyses cited are compared in this manner in Figure 6, which plots \( C \) versus depth.

Also plotted in Figure 6 and shown as dashed lines are computed values of \( C \) using the more general form given by (14). One of the dashed lines (curve 5) represents values computed using "in situ" \( x_r \) derived from a representative porosity-depth relation presented by Dickinson [1953] for the Gulf Coast (see Figure 4, curve A). The other (curve 6) represents \( C \) values computed using laboratory-measured \( x_r \) for a variety of sedimentary lithologies. The dashed lines and line 4 are largely distinguished from the other treatments by the fact that the variation of \( x_r \) with effective stress is accounted for.

Examination of Figure 6 suggests that the various simplifications cause the treatments to vary significantly from each other and from the more complete formulation represented in curves 5 and 6. The differences are particularly great at depth. It is of interest to note the difference between curves 4 and 6. These are identical formulations except that thermal effects were ignored in arriving at curve 4. The differences between them therefore represent the possible contribution of thermal effects. Curve 6 suggests that \( C \) increases significantly below approximately 2500 m (8000 ft). This result from marked decreases in measured \( x_r \) (and thus \( x_f \)) with burial and a consequent increase in the final term of (14). These data suggest that fluid pressures in low-permeability sediments undergoing burial can increase significantly below a certain depth. There is, in fact, an often-observed tendency to approach lithostatic values in a "transition zone" after several thousand feet of depth, as, for example, in the data presented by Dickinson [1953] and Sharp [1976] for the Gulf Coast. This is generally thought to reflect lithology controlled permeability changes [e.g., Bishop, 1979; Bodner et al., 1985; Keith and Rimstidt, 1985]. However, curve 6 provides an explanation for the tendency of the transition to occur at a certain depth without recourse to lithology-related permeability changes. It suggests that matrix strain is the dominant pressure-producing mechanism at shallow depths, while fluid thermal expansivity is the dominant mechanism at greater depths.

A different interpretation is provided by curve 5, computed using estimates of effective long-term compressibility, here denoted \( x'_r \). Curve 5 suggests that one may ignore water and grain compressibility and, except at great depth, the thermal expansion of water. Under these circumstances, fluid thermal expansion probably would be unable to generate the pressures in excess of lithostatic hypothesized by Domenico and Palciauskas [1979]. This is reflected in results of analyses presented by Walder [1984] and Keith and Rimstidt [1985]. Walder used \( x'_r \) values similar to those used to compute curve 5, and Keith and Rimstidt based their analysis on observed porosity-depth relations; both analyses suggested that thermal pressuring would be small.

However, as pointed out earlier, porosity change with depth may not be a good indicator of \( x'_r \). Processes other than mechanical deformation may change porosity. Walder and Nur [1984] proposed that pressure solution of minerals at grain contacts and precipitation in the pore space is an important porosity-decreasing process in crystalline rocks at depth. Thus the actual value of \( x'_r \) may be smaller than estimates obtained from porosity-depth relations, such as those shown earlier in Figure 4. Any actual difference between values of \( x_r \) and \( x'_r \) is due to time dependent or viscoelastic deformation of the solid skeleton. Herein lies a difficult problem facing analysts of long-term transient flow. The use of flow models based on assumed elastic behavior of the medium introduces the time scale dependence of specific storage and loading efficiency through the compressibility of the matrix. An additional and perhaps more fundamental problem is whether the behavior of certain systems cannot be analyzed using an elastic model regardless of the choice of parameters. Indeed, there is a range of possible behavior which is depicted schematically in Figure 7.

As indicated in Figure 7, short-term experimental behavior is mostly in the range where deformation is essentially elastic. Viscoelastic deformation may be clearly in evidence, as secondary consolidation, for example, but often at longer time
scales than the flow experiment. At the other extreme, in large flow systems with a very slow transient flow response, viscoelastic deformation should be quite significant during the flow. However, it would probably occur over shorter time scales than the flow and thus be additive with, and indistinguishable from, the elastic deformation. Walden [1984] suggested this possibility on the basis of theoretical analysis of viscoelastic deformation of clays observed in laboratory tests.

Between extremes of short- and long-term transient flow responses, and represented by the center column of Figure 7, lies a range wherein viscoelastic deformation and the flow response occur on comparable time scales. Bjerrum [1967] provides a description of this type of behavior in response to loading. In this range the mechanical interaction between fluid and solid is more complex than represented by (6) and (7), and it may be necessary to incorporate a rheological model of the porous solid into the analysis. Equations (6) and (7) are derived by assuming that a linear elastic (Hookean) constitutive law governs deformation in response to changes in effective stress, with viscoelastic deformation insignificant during transient flow; a standard transient flow description applies with experimental values of $S_i$ and loading efficiency. (b) Viscoelastic deformation significant but on time scale comparable with that of transient flow; a complex transient flow description may be necessary. (c) Viscoelastic deformation significant but on time scale smaller than that of transient flow; a standard transient flow description applies with experimental values of $S_i$ and loading efficiency may be significantly different from those obtained experimentally.

To briefly summarize the preceding discussion, both short- and long-term transient flow can probably be satisfactorily described with a linear elastic relation for solid deformation in response to effective stress changes. However, the elastic constants applicable at long time scales will be different from those applicable to or measured in the short term. At intermediate time scales a linear elastic description may be inadequate. It is unclear whether flow at intermediate time scales is sufficiently affected by this complexity to warrant a more sophisticated flow model such as that of Gibson and Lo [1961]. It is also unclear what time scales are “intermediate” in this context. These questions await further research.

A further complication is presented by the difference between compressibility during compression, $a_{re}$, and expansion, $a_{re'}$, which is manifested as the permanent set suffered by rocks and soils when compressed and allowed to rebound [Jaeger and Cook, 1969; Scott, 1963]. Both $a_{re}$ and $a_{re'}$ can be straightforwardly measured in the laboratory. Often, such tests suggest that $a_{re'}$ exceeds $a_{re}$ by approximately an order of magnitude. Domenico and Palciauskas [1979] suggested that the same approximation applies over geologically significant periods. If so, thermal effects may figure much more prominently when effective stress decreases, as when erosion removes overburden.

In experimental soil mechanics, $a_{re}$ is often found to be approximately equal to $a_{re'}$ in “overconsolidated” media which have borne higher effective stresses in the past [e.g., Scott, 1963, p. 175]. Reddy [1970] found this also to be true for low-permeability lenses which he studied in situ. However, there is no reason to expect that a similar relationship prevails over geologically significant time periods. Estimates of long-term expansion compressibility, $a_{re'}$, are therefore less readily made than estimates (as in Figure 4) of $a_{re}$. As a result, few attempts have been made to estimate $a_{re'}$. Cacasgrande [1949] and Peterson [1958] used geotechnical data and geological evidence to study the response of the Cretaceous Bearpaw Shale to past erosional unloading. Their results indicate that $a_{re'}$ is significantly larger than $a_{re}$. Extrapolation of more recent laboratory data by Lo et al. [1978] suggests that the ultimate rebound in shales exceed the short-term rebound by 1 to 2 orders of magnitude.

The mechanical behavior of rock is thus problematic but critical for analyzing the fluid flow. To some degree, this is true even in relatively permeable media. However, the shortcomings of existing models are most pronounced for the very prolonged flow transients in environments where permeability is small.

It is necessary that these problems be recognized and addressed. Current understanding of deformation in geologic materials should be incorporated in analyses. Further experimental work can refine what is known about time dependent deformation over relatively short periods; these data may provide useful bases for extrapolation to long periods. This is particularly true for behavior under stress decrements, for which few experimental data exist. Study of indirect geologic evidence should provide further guidance. For example, the lithospheric structure under certain seamounts led Watts and Cochran [1974] to suggest that elastic stresses in these rocks were retained over tens of millions of years. Analogous approaches may provide information in other settings and lithologies.

**Physicochemical effects:** The long periods of time which are of interest here also dictate consideration of physicochemical phenomena in addition to mechanical and thermomechanical phenomena discussed above. During burial metamorphism, various chemical processes may cause solution, precipitation, and reactions involving the solid and fluid phases. Gas or liquid sources and rock structure and porosity changes can result.

Because argillaceous media are an important low-permeability environment, burial metamorphism of clays is of particular interest. It has long been recognized that clay mineralogy often changes with depth in thick argillaceous sequences and that it differs in old and young shales. Powers [1967] and Burst [1969] suggested that these data show the smectite to illite transformations occurred with burial, driven by heating. The process releases bound water, and the sup-
Buried metamorphism is probably the single most important process causing transient flow for which an adequate quantitative treatment is not available. Uncertainty about effects of stress and temperature changes lies mainly in the mechanical response of the porous matrix; the uncertainty surrounding the effects of burial metamorphism and diagenesis lies in the processes themselves. Part of the difficulty results from the fact that changes lumped under this term represent several diverse lithology-specific processes. While it may be possible to estimate rates of smectite transformation [Walder, 1984], for example, the associated restructuring of the porosity and resulting changes in effective stress and fluid pressure are difficult to account for. Rates of porosity change from pressure solution are equally difficult to estimate; the problem is complicated by the coupling between pressure solution and pore pressure through the effective stress. Moreover, the total suite of processes occurring may be considerably more complex than those yet considered. Incomplete understanding of these phenomena thus limits the ability to analyze prolonged flow phenomena. We need to be able to relate rates of pore fluid production and porosity change to rates of change of temperature, effective stress, and total stress. High-temperature and high-pressure experiments with natural clays and shales, in which strain and fluid pressure are monitored, could provide a basis for obtaining the needed relationships.

*Processes in two and three dimensions.* The theoretical shortcomings discussed in the previous section apply equally or perhaps more so to two- and three-dimensional analyses which we now consider. While one-dimensional analyses are extremely useful, they are limited in application. It is likely that in most settings the state of stress in rocks evolves in a more complicated fashion than implied by the assumption of negligible horizontal strain in an elastic medium. Indeed, even under the ideal conditions assumed for the one-dimensional analysis some horizontal strain may result from vertical movement normal to the earth’s curved surface [Haxby and Turcotte, 1976]. In many real settings, moreover, rocks are significantly affected by tectonic and topography related stress changes.

The associated rates of stress change, $\partial \sigma_i / \partial t$, are problematic but can be relatively large. Rates of tectonic stress change as high as $10^3$ to $10^6$ Pa/yr are not unusual (M. D. Zoback, personal communication, 1984). While this is the high end of the range, it is significantly larger than the $10^{-5}$ Pa/yr rate of overburden stress change associated with representative denudation and erosion rates [Bredehoefl and Hanshaw, 1968; Neuzil and Pollock, 1983].

Thus tectonic stress may often be expected to affect fluid pressures and flow in tight environments. Tectonic squeezing has been speculated as the cause of “abnormal” pressures in California [Watts, 1948; Berry, 1973], and North Dakota [Finch, 1969]. Because of the uncertainties involved, however, little quantitative analysis of the problem has been possible. Indeed, the problem of tectonic stress may, in many instances, be most profitably posed as a detection problem. In such a context, fluid pressures would be viewed as cumulative stress phenomena. We need to be able to relate rates of pore fluid change to rates of change of temperature, effective stress, and total stress. High-temperature and high-pressure experiments with natural clays and shales, in which strain and fluid pressure are monitored, could provide a basis for obtaining the needed relationships.

The theoretical shortcomings discussed in the previous section apply equally or perhaps more so to two- and three-dimensional analyses which we now consider. While one-dimensional analyses are extremely useful, they are limited in application. It is likely that in most settings the state of stress in rocks evolves in a more complicated fashion than implied by the assumption of negligible horizontal strain in an elastic medium. Indeed, even under the ideal conditions assumed for the one-dimensional analysis some horizontal strain may result from vertical movement normal to the earth’s curved surface [Haxby and Turcotte, 1976]. In many real settings, moreover, rocks are significantly affected by tectonic and topography related stress changes.

The associated rates of stress change, $\partial \sigma_i / \partial t$, are problematic but can be relatively large. Rates of tectonic stress change as high as $10^3$ to $10^6$ Pa/yr are not unusual (M. D. Zoback, personal communication, 1984). While this is the high end of the range, it is significantly larger than the $10^{-5}$ Pa/yr rate of overburden stress change associated with representative denudation and erosion rates [Bredehoefl and Hanshaw, 1968; Neuzil and Pollock, 1983].

Thus tectonic stress may often be expected to affect fluid pressures and flow in tight environments. Tectonic squeezing has been speculated as the cause of “abnormal” pressures in California [Watts, 1948; Berry, 1973], and North Dakota [Finch, 1969]. Because of the uncertainties involved, however, little quantitative analysis of the problem has been possible. Indeed, the problem of tectonic stress may, in many instances, be most profitably posed as a detection problem. In such a context, fluid pressures would be viewed as cumulative stress meters providing information about past tectonic activity.

Some insight into the problem has been gained by extension of the simpler one-dimensional analysis. Domenico and Palciauskas [1979] and Palciauskas and Domenico [1980] considered the case where the ratio of horizontal to vertical stress is arbitrary, but constant over time. Their results demonstrated that tectonic squeezing enhances burial overpressuring, while tectonic extension subdues it. Walder [1984] considered...
two cases in which horizontal strain was incorporated in a source term in a one-dimensional flow equation. The first case he considered was erosion followed by isostatic uplift as envisioned by Haxby and Turcotte [1976]. For the case of erosion, Walder’s results suggest that underpressuring, resulting from vertical and horizontal expansion, and cooling, should occur. The second case Walder considered was that of an accretionary sediment wedge at a convergent plate boundary. Under the simplifying assumption that lateral strain dominates over other effects, he derived an equation similar to (6). However, rather than a term accounting for the rate of stress change, \( \partial \sigma_r / \partial t \), Walder’s equation included a term with the strain rate imposed by the relative plate motion. With the relations implied by the equation he estimated maximum permeabilities at which representative strain rates would significantly perturb pressures. The values he obtained correspond to \( K \) of \( 10^{-11} \) to \( 10^{-12} \) m/s. As these values fall within a reasonable range for fine-grained sediments, Walder’s results support the notion of strain-induced excess pressures in these environments.

Besides simplifying mechanical behavior, one-dimensional analyses force the flow of fluids to be vertical, an approximation which may sometimes be unrealistic [Bethke, 1985]. In particular, when laterally continuous permeable units are present, this simplification may be too restrictive [Magara, 1976; Walder, 1984]. J. D. Bredehoef and R. D. Djevanshir (manuscript in preparation, 1986) discuss a basin in the Caspian region in which lateral flow in sands appears to be quite significant. Sharp [1978] attempted to overcome this problem by solving a one-dimensional equation of the form of (12) coupled with a heat transport equation in a series of one-dimensional profiles to simulate depositional loading and thermal effects in the Ouachita basin. These solutions provided a quasi-two-dimensional simulation of the basin in a vertical plane.

A somewhat different approach has recently been applied to the two-dimensional problem of sedimentary loading in a basin. It uses a statement of fluid mass conservation in a deforming porous medium (such as used in the derivation of (6)). The relation can be expressed as

\[
\frac{\partial p}{\partial t} = -\frac{1}{\rho} \nabla \cdot (\rho \mathbf{q}) - \frac{1}{1-n} \frac{\partial n}{\partial t} + n \frac{\partial T}{\partial t} \tag{15}
\]

(see, for example, Domenico and Palciauskas [1979], equation 4). The right-hand terms describe fluid flow, porosity change, and thermal expansion of the fluid, respectively. Analyses considered thus far have developed the expression further by relating porosity change to stress and fluid pressure. Thus, for example, Domenico and Palciauskas (1979) used Biot’s (1941) constitutive relations to determine that under their assumptions,

\[
\frac{1}{1-n} \frac{\partial n}{\partial t} = -x \left( \frac{\partial \sigma_r}{\partial t} - \frac{\partial p}{\partial t} \right) \tag{16}
\]

Substitution of (16) in (15) leads to an equation of the form of (12). However, (16) or any similar relation includes all assumptions about long-term mechanical behavior of the solid and requires, as discussed above, knowledge of \( \partial \sigma_r / \partial t \), the total stress changes through time. The changes in horizontal as well as vertical stress must be known, or special conditions, such as no horizontal strain, must be assumed to constrain them. In addition, by omitting a term to describe diagenetic porosity changes, (16) implicitly assumes them to be insignificant unless they are embedded in the medium compressibility, \( x \). Thus, in attempting to express porosity changes in terms of other variables much of the uncertainty related to large-scale transient behavior is introduced.

If the porosity changes with time in the region of interest can be ascertained independently, these problems are bypassed, and (15) itself or a comparable equation can be solved. This approach has been used by Bethke [1985] to simulate flow in a hypothetical basin fill under continuing sediment input. Bethke considered a slowly subsiding continental basin in which significant transient pressure was unlikely to have been generated; he was thus able to argue that observed porosity-depth relations could be used to determine the temporal changes of porosity, \( \partial n / \partial t \), in a sediment packet undergoing burial.

In the cases of interest here, where transient excess pressures are generated, \( \partial n / \partial t \) is problematical; the normal porosity decrease is disrupted because the excess pressures do not permit the usual increase in effective stress with depth [e.g., Magara, 1969]. In these instances, Bethke suggested what is the only feasible approach, namely, using relations which incorporate the same types of assumptions as (16). By doing this the advantages are lost, and the method becomes equivalent to those described earlier. Indeed, a relation similar to (16) and predetermined \( \partial \sigma_r / \partial t \) must be incorporated, in some fashion, in any analysis of stress-induced transient flow.

Cathles and Smith [1983] followed a similar line of reasoning to compute flow from a compacting basin. They used the simplifying assumption that the loss of fluid and accompanying porosity decrease occurred late in basin development and over a short period of time through fractures.

These analyses, together with one-dimensional studies, have begun to address the problem of the concurrent evolution of geologic and groundwater flow systems. An unanswered question at this juncture concerns the significance of the fluid flow-stress coupling in two and three dimensions. Both (6) and (7) are necessary to fully describe transient flow. The one-dimensional analyses described earlier satisfy these relations exactly for the conditions they assume. However, the two-dimensional analyses described have ignored the conditions expressed by (7). While studies have addressed the implications of ignoring (7) in the context of aquifer development [Gambolati, 1974] and certain foundation engineering problems [Schiffman et al., 1969], none have addressed its significance in situations where all-around total stress varies in time. As we have seen, such stress changes can be quite pronounced on the time scales of interest. Also, the fully coupled flow-stress description may have importance for an understanding of how the state of stress evolves and its relation to deformation.

A related question concerns viscoelastic deformation of the rock, which will be different in response to horizontal and vertical stresses. Inelastic deformation may relieve superimposed tectonic stresses [Domenico and Palciauskas, 1979], but the stress due to overburden load remains regardless of the extent of deformation. Indeed, time dependent deformation apparently causes horizontal stresses to tend toward equality with vertical stresses [Domenico and Palciauskas, 1979]. The resulting interactions between stress, elastic and viscoelastic strain, and flow are complex but possibly important in a variety of settings. An important area for research now would seem to be investigation of appropriate two- and three-dimensional models for long-term solid matrix deformation and methods for determining parameters for the models.

The problem may be even more complex than these arguments

\[
\frac{\partial \sigma_r}{\partial t} = \frac{\partial p}{\partial t} - x \left( \frac{\partial \sigma_r}{\partial t} - \frac{\partial p}{\partial t} \right) \tag{16}
\]
mments suggest; other forms of inelastic behavior have also been postulated in low-permeability environments. Several authors who have discussed sedimentary deposits [Sharp, 1978; Domenico and Palciauskas, 1979; Palciauskas and Domenico, 1980; Cathles and Smith, 1983] and crystalline rocks [Walder, 1984; Walder and Nur, 1984] have hypothesized that fracturing may cause sporadic periods of increased flow and relief of excess pressures by enhancing the permeability. In some instances, notably as described by Walder [1984] and Walder and Nur [1984], the process is envisioned as being reversible. The fractures close or heal, reducing permeability to its former low value and initiating another cycle of fracturing.

Palciauskas and Domenico [1980] distinguished between macroscopic fracturing and dilatancy which they considered to result from microfracturing. They described macroscopic fracturing as analogous to hydraulic fracturing in wells, occasioned when the fluid pressure exceeds the least principal stress and tensile strength of the medium. This presumably was the mechanism envisioned by Sharp [1978] and Cathles and Smith [1983]. Domenico and Palciauskas [1979] proposed it specifically as a response to fluid pressures when they begin to exceed the lithostatic load.

Dilatancy is a porosity and bulk volume increase known from experimental work in sedimentary rocks [Handin et al., 1963] and crystalline rocks [Brace et al., 1966]. Domenico and Palciauskas [1979] and Palciauskas and Domenico [1980] proposed that dilatation is proportional to deviatoric stress and included a term for it in their version of (12). They suggested that significant permeability increases accompany dilatation.

The mechanical behavior in these processes is not well understood, particularly in the long term. As Palciauskas and Domenico [1980] themselves acknowledge, we do not yet have an experimentally based constitutive relation for dilatation. Further, there are no data concerning the effect of dilatation on permeability. Resolution of these problems is necessary to evaluate the role of such inelastic phenomena in low-permeability settings.

An interesting geologic setting, exemplifying the complex behavior discussed above, is found in accretionary wedges at convergent plate boundaries. Containing large volumes of sediment, these features are only now being explored. It has been suggested that excess pressures created in the sediments by the tectonic strains and overthrust loading facilitate further thrust faulting [von Huene and Lee, 1983; Westbrook and Smith, 1983].

Changes in hydraulic boundary conditions. Besides the types of changes already described, geologic processes affect flow in tight environments by altering the hydraulic boundary conditions. The geometry of rock bodies and hydraulic heads at their boundaries is changed by processes such as tectonic folding and tilting, faulting, and erosional exposure of aquifers. This occurs in any tectonically active region, whether undergoing tectonic extension, such as developing rift zones, or compression. The possibilities are varied and difficult to generalize; it is perhaps useful to consider a specific example.

Töth [1978] describes an aquifer in Alberta which was probably exposed by valley erosion approximately $3 \times 10^3$ years ago. Exposure presumably caused a relatively rapid head decline in the aquifer as it became unconfin ed. Underlying the aquifer is a 500-m-thick confining layer with a hydraulic diffusivity estimated by Töth as $5 \times 10^{-9}$ m$^2$/s. Had flow in the system originally been at steady state, the head decline in the aquifer would have acted like a nearly instantaneous change at the confining layer boundary. The resulting one-dimensional flow in the confining layer can be analyzed with a solution to (1) for these boundary conditions. The solution is graphed by Hanshaw and Bredehoeft [1968, Figure 2], from which it can be seen that transient flow should persist, in this case, for approximately $5 \times 10^5$ years; the groundwater flow in the underlying confining bed may therefore still be in a transient state.

An analogous situation occurs in formations confining aquifers when the aquifer head declines because of development. Recently, Bredehoeft et al. [1983] showed that during development of the Dakota artesian aquifer, most of the water released from storage came from the adjoining tight shales. Moreover, the properties of the confining Cretaceous shales [Bredehoeft et al., 1983] suggest that transient leakage to the aquifer should continue for approximately $10^6$ years.

Nonhydraulic Flow in Large-Scale Systems

Two distinct flow phenomena which may be classed as nonhydraulic were considered in the earlier discussion of small-scale experimental work, namely, non-Darcian flow in response to hydraulic gradients and coupled flow in response to nonhydraulic driving forces. In this section we consider the implications of these phenomena for flow in geologic systems on a large scale.

Speculations on the effects of non-Darcian flow. The applicability of Darcy's law, as we have seen, is usually a fundamental assumption of large-scale flow studies. This is entirely appropriate in view of the lack of convincing evidence to the contrary. However, it is also well to recall that the applicability of Darcy's law, in the media of interest and under in situ hydraulic gradients, is strictly an assumption (see Figure 2).

A small number of investigators have considered the effects of specific types of non-Darcian behavior. Because of the lack of consistent experimental evidence for non-Darcian behavior of any kind, analyses based on an assumed specific non-Darcian flow law can only be regarded as highly speculative. Florin [1951] and Elnagger et al. [1971] analyzed consolidation, and Pascal [1981] analyzed the response to hydraulic stresses with a threshold gradient. They showed that under the flow laws they assumed, apparently transient head distributions would become "frozen" and represent a new steady state. Schmidt and Westmann [1973] analyzed consolidation using a nonlinear flow law. Remson and Gorelick [1982] proposed considering the problem more generally and discussed the potential importance of non-Darcian flow in waste repository design. They argued that the existence of a threshold gradient for flow, for example, would render waste contaminant analyses more conservative than realized, leading to overdesign of repositories.

Geologic membranes in large-scale systems. In contrast with non-Darcian flow, coupled flow in certain geologic materials is clearly demonstrable experimentally. However, speculation about coupled flow in the subsurface predates experimental work on geologic materials. According to Hanor [1983], the problem of accounting for the origin of subsurface brines led to speculations by Russell [1933] and De Sitter [1947] that ultrafiltration by shales was operative. Attempts to explain certain ore deposits led Mackay [1946] to publish similar speculations. Subsequently, as experimental evidence for significant coupled flow in certain geologic materials began to accumulate, numerous studies addressed the possibility of large-scale osmosis and ultrafiltration in the subsurface.

Most discussions of membrane phenomena in the subsurface paralleled experimental experience and considered the in-
teration of reservoirs (aquifers) and membranes (confining layers). By experimental criteria, the presence of salt-bearing groundwater separated from fresher groundwater by a shale or similar medium provides the conditions for osmotic flow of the fresh water toward the salty. If the salty water occurs in a reservoir surrounded by tight shales, it was reasoned that the osmosis may produce observable excess pressures there. Berry [1959], Hanshaw [1962], and Hanshaw and Hill [1969] suggested this mechanism to explain highs in the piezometric surface in the San Juan Basin region of Utah, Colorado, and New Mexico, and Berry and Hanshaw [1960] interpreted high heads in the San Joaquin Valley, California, as being due to osmosis. Hanshaw and Zen [1965] proposed osmosis as a mechanism for generating excess pressures and facilitating thrust faulting.

The role of ultrafiltration in the subsurface has also been widely discussed. Berry [1966, 1967] suggested that ultrafiltration helped create brines in the Imperial Valley, California, and Hitchon et al. [1971] argued that the composition of sedimentary formation fluids in western Canada should also be explained this way. White [1965] discussed the relative abundances of various ions in brines in connection with their different mobilities in membranes. Based on their experimental studies of filtering efficiencies for different ions [Kharaka and Berry, 1973], Kharaka and Berry [1974] interpreted the geochemistry of formation fluids in the Kettleman Dome area of California as indicating that extensive ultrafiltration had occurred. Graf [1982] has reviewed these and other studies. More recently, Gautier et al. [1985] summarized the geochemical criteria that they consider indicative of membrane filtration. They indicated that on this basis, filtration is indicated in Central Valley, California, the Alaskan North Slope, and the Gulf of Mexico.

As noted above, these discussions are distinguished by the fact that they conceptualize the system as being discontinuous and usually consider effects in aquifers adjacent to membranes without considering behavior in the membrane itself. This may partly reflect the influence of experimental work which has been limited to reservoir-membrane interaction. Perhaps a more important reason, however, is that membrane effects in the subsurface are likely to be detected only in permeable reservoirs; these are the source of nearly all data on subsurface fluid chemistry and pressure.

The question of behavior within geologic membranes, in the laboratory as well as in situ, has received much less attention. The data for investigations of this sort are difficult to obtain. One such investigation is that by Marine [1974] and Marine and Fritz [1981], who measured high hydraulic heads within the low-permeability fill of the Triassic Dunbarton Basin, South Carolina. They interpreted their data as indicating that the high heads occur within the fill as a result of osmotic flow from surrounding permeable rocks with fresher waters.

All of the studies cited are mainly qualitative. Few quantitative analyses of in situ coupled flow have been made. Exceptions include analyses by Bredelhoef et al. [1963, 1983] which combined quantitative flow analyses with simple models of ultrafiltration. However, even these studies adopted the simplification of ignoring osmosis and considered only steady flow conditions.

A more general quantitative description of coupled flow in the subsurface can be obtained with the constitutive relations between flows and forces expressed by the phenomenologic equations (2) combined with statements of mass and energy conservation. The important flows in geologic environments are those of pore fluid, solute, electric charge, and, in certain environments, heat. Electrical effects usually result from chemical forces and flows and thus are effectively embedded in the chemical effects. Thus a useful simplification is to consider only flows of pore fluid and solute.


Mitchell et al. [1973] [see also Mitchell, 1976, pp. 370-373] solved Greenberg's equations numerically for a representative problem involving a layer of clay or shale. The simulation results provide important insight into possible transient flow behavior in large membranes. The shale was considered to lie adjacent to aquifers in which the salt concentration abruptly increased to a value that remained constant. Values for the coefficients were derived from the experimental data for clay cakes discussed earlier. Osmotic flow out of the shale produced a pattern of low hydraulic heads similar to that indicated for other processes and qualitatively depicted in Figure 38. Thus in practice, osmotic effects would be difficult to distinguish on the basis of hydraulic head measurements alone. Because the shale in the simulation did not act as a perfectly efficient filter, it was slowly invaded by the salt. Substantial concentration differences within the shale were indicated. Thus measurement of pore fluid chemistry variations together with hydraulic head measurements within a shale would be more diagnostic of coupled flow than head measurements alone. Membrane behavior is a distinct possibility if solute concentration and hydraulic heads both decrease or both increase inward from the boundaries.

The time scales implied by the simulation results are also highly significant. Minimum pressures from osmotic flow were attained at \( t^* \approx 4 \), where \( t^* \) is dimensionless time as previously defined. This response time is comparable to that for purely hydraulic transient flow considered in previous sections. As shown there, response times of this order may represent long periods of time. For a perfectly efficient filter this condition would have represented a new steady state. However, time dependent changes occurred at yet a slower rate as salt slowly leaked into the shale, and salt concentrations and hydraulic heads continued to evolve until \( t^* > 10^4 \). The extreme longevity of these "secondary" transient conditions suggests that in systems containing membranes, the chemically driven flow may rarely reach a fully steady, equilibrium condition. Moreover, secondary transient conditions would be difficult to detect because they change so extremely slowly that even in tight media the flow would be quasi-steady. Bredelhoef et al. [1983] have suggested that just such a situation, the incomplete invasion of shales by dissolved salts, may in part explain the geochemistry of the Dakota aquifer system.

This theoretical work represents a relatively primitive understanding of how geologic membranes may function in actual groundwater systems. The complexity of the flow processes lends uncertainty to mathematical models. Unlike the case of hydrodynamic flow, there is very little experience with application of the equations suggested by Greenberg [1971].
and P. C. Trescott (unpublished manuscript, 1975) with which to judge their appropriateness. Greenberg [1971] notes numerous assumptions of uncertain validity which were necessary to simplify his model to a workable form. The complexity of the geologic environment also poses difficult problems. Rather than being homogeneous, uniform bodies, geologic membranes undoubtedly exhibit inhomogeneity and anisotropy with respect to membrane properties. A single argillaceous formation might consist of several layers of more or less efficient membranes. The entire sediment package might act like several membranes in series, increasing the filtration efficiency of the formation as a whole. Bredehoeft et al. [1983] incorporated this concept in their simulations. Transmissive fractures, on the other hand, can probably be expected to “short-circuit” membranes and prevent significant osmotic flow while permitting movement of solutes.

Many of the conceptual problems related to membranes in groundwater systems are manifestations of the difficulty of extrapolating small-scale experimental experience to a large scale. There is great controversy concerning the significance of coupled flow in groundwater systems, even though the requisite conditions for it, such as significant concentration gradients, are common in the subsurface. Some investigators [e.g., Hanshaw and Hill, 1969; Marine and Fritz, 1981] have argued that the existence of membrane effects in certain locales is indicated by a process of elimination. In other words, certain anomalous hydraulic heads are due to osmosis because they are not explicable by other causes. However, as should have been apparent in earlier discussions, it is extremely difficult to exclude any number of geologic processes as causes of transient flow. While the evidence may suggest large-scale osmosis, it can hardly be considered unequivocal. Likewise, the evidence for in situ ultrafiltration has been questioned by Hanor [1983], who also argued that it apparently has not occurred in Gulf Coast sediments. Studies of sandstone diagenesis in the Gulf Coast by Land [1984] suggest a more open geochemical system than one would expect if ultrafiltration were significant. Graf [1982], on the other hand, has argued for extensive ultrafiltration in subsurface environments such as those in the Gulf Coast. While the existence of significant coupled flow in the subsurface is difficult to demonstrate, its absence is equally difficult to show, and for the same reason; real systems are quite complex with cause and effect not easily determined.

The problem of in situ behavior is further complicated by the difficulty of characterizing the driving forces in the subsurface. The chemistry of natural waters is complex; there are generally both solute composition and concentration gradients in groundwater. To compound the difficulty, chemical coefficients of diverse methods of investigation. The need for laboratory experimental data on coupling coefficients in undisturbed geologic media has already been discussed. Development of in situ testing techniques for low-permeability membrane media could be expected to provide useful data. A general in situ test strategy can be outlined: Waters of different chemistry are pumped into packer-isolated or open boreholes terminated in a suspected membrane unit. Significant osmotic flow would be indicated by a tendency toward different hydraulic heads in the boreholes. A synthesis of laboratory and borehole data might elucidate the scale sensitivity of membrane properties. Theoretical investigations using mathematical models can shed light on the expectable effects of inhomogeneity and anisotropy of coupling coefficients. Again, laboratory work is required to characterize these qualities. Field characterization of formations thought to be acting as membranes, while likely to be difficult, can be expected to provide important answers.

The Field Problem

Ultimately, the problem of interest is that of analyzing the flow in a particular low-permeability system. It will be necessary to characterize current conditions in the system, particularly the hydraulic head, and if membrane phenomena are to be considered, the pore fluid chemistry. These characterizations are difficult in tight media.

Determination of undisturbed hydraulic head. Very few measurements of hydraulic head in low-permeability formations have been made. Generally, anomalous pressures are detected in formations intimately associated with tight formations (which maintain the pressures) but which themselves are permeable enough to permit significant flow when penetrated by drilling. As Smith [1971], Bradley [1975], and Bishop [1979] have noted, pressures in shales are generally inferred from these reservoir pressures rather than measured. As a result, our view of the occurrence of transient conditions is probably biased by the fact that anomalous pressures are usually detected only in permeable reservoirs. For this reason, and to be better able to map hydraulic head distributions, measurements of fluid head within low-permeability formations are often desirable.

Indirect indicators of fluid pressure have been used for some time. Several borehole logging measurements respond in a qualitatively predictable manner in regions of abnormal pressure [Fertil, 1976, p. 226]. Some [e.g., Magara, 1969] have carried the concept further by estimating porosity or other properties from borehole logging data and relating them quantitatively to effective stress and fluid pressure. Difficulties with this approach lie in the empirical nature of the relation between log response and fluid pressure caused by lithologic and other variations. These shortcomings are discussed, for example, by Reynolds et al. [1973] and Pritchett [1980]. However, in certain instances it may be possible to draw reliable inferences about fluid pressure indirectly. The tendency of certain overpressured shales to flow [Musgrave and Hicks, 1968; Pritchett, 1980] presumably indicates fluid pressures near lithostatic load. Along similar lines, Walder and Nur [1984] have suggested that seismic low-velocity zones in the crust may indicate regions of crystalline rock with nearly lithostatic fluid pressures.
Direct measurement of formation fluid pressures using boreholes is difficult largely because of the slow response of tight formations to a hydrodynamic disturbance. The simplest measurement technique, simply letting the fluid level in the borehole reach equilibrium, is usually impractical because of the long time required. Such a procedure, in fact, is similar to an open slug test. Marine [1974, also personal communication, 1984] and Marine and Fritz [1981] have observed the slow return to equilibrium in wells completed in low-permeability Triassic sediments. A period of several years was required to gain an accurate estimate of the predrilling head.

As when conducting slug tests, the process may be hastened by shutting in the borehole, usually with inflatable packers. This procedure, however, brings the attendant problems of eliminating small leaks past the packer or in the piping. Connection via leaks between the shut-in portion and other parts of the borehole may be sufficient to cause the pressure to stabilize at an unrepresentative value. Neuzil and Pollock [1983] avoided this problem by isolating a pressure transducer with a shale slurry which self-consolidated to form a low-permeability plug. Wolff and Olsen [1968] developed a pointed piezometer which could be pushed into soft clays; it was designed to have small storage and thus respond quickly to pore pressure changes. The system was successfully used in the field for small penetrations into a clay confining bed [Wolff, 1970]. The instrument was isolated with an inflatable packer.

Assuming that the section to be tested can be successfully isolated, a sufficient period of time must be allowed for drilling-caused disturbance to dissipate. Generally, part of the disturbance results from the imposition of an arbitrary fluid head at the borehole walls during drilling. The distance this effect penetrates the surrounding rock depends on the period the disturbance is applied; it is described by Hantush's [1964] A function. In the case that the well has been "produced" at a reasonably constant rate for some time, the equilibrium head can sometimes be estimated from early shut-in data using a method developed by Horner [1951]. Anderson and Zoback [1982] used this technique to estimate the equilibrium head in a tight underpressured formation. In principle, industry drill stem tests could be used to measure original formation pressure using this technique [Bredelhoft, 1965]. However, few, if any, measurements of this type have been made in the commercial sector. Another problem is that early shut-in data may be strongly affected by thermal disequilibrium introduced by drilling, as Grisak et al. [1985] found in crystalline rocks.

Stress distortion caused by drilling the hole also disturbs surrounding fluid pressures. An analysis by C. E. Neuzil (unpublished manuscript, 1982) based on the plane-strain solution of Hubbert and Willis [1957] for the mechanical response of elastic media to a borehole indicates that induced pore pressure changes in the vicinity of the borehole will be symmetrical and cancel out. However, experience with boreholes in the Pierre Shale (C. E. Neuzil, unpublished data, 1985) and in a shale in Saskatchewan (G. van der Kamp, personal communication, 1985) indicates that recovery of the hydraulic head in the boreholes was delayed. The delayed response suggests that fluid pressures in the vicinity of the borehole may have been depressed. Inelastic deformation may play an important role in this process. Florence and Schwer [1978] analyzed the mechanical response to a circular hole of an elastic-perfectly plastic medium; plastic dilatation, which would lower fluid pressures, was predicted near the hole. In tight materials this presumably could affect fluid pressure in the borehole for long periods of time.

Thus, once shut-in occurs, flow must not only overcome storage in the shut-in volume, it must also dissipate disturbances induced by drilling. Analysis of the interaction between the fluid flow and the complex mechanical behavior near a borehole is needed in order to assess the validity of current measurements and to devise improved measurement strategies.

In the case of shales and other media which may have membrane properties a further potential complication, in the form of osmotically generated pressures, must be considered. For example, were the borehole filled with water substantially fresher than the pore fluid in a surrounding shale, osmotically driven flow into the shale could result. The apparent equilibrium pressure reached in a shut-in borehole would then be less than the equilibrium pore pressure in the shale. The possibility of such a process makes it advisable to determine the pore water chemistry and to duplicate its gross properties in any water injected into the borehole.

**Determination of pore fluid chemistry.** The porosity of crystalline rocks is so small that, except when they are fractured, recovery of analyzable amounts of pore fluid from them is not practical. Argillaceous media generally have high porosity but yield their pore fluid with difficulty. Relatively small quantities (of the order of tens of milliliters) can be obtained from shale and clay cores by squeezing. H. W. Olsen (personal communication, 1982) has extracted pore fluid samples from Cretaceous shale of approximately 30% original porosity by squeezing under high loads. This approach is not entirely satisfactory because some fraction of the water obtained may not have been mobile under normal conditions; the chemistry of the sample thus may not be representative of the free pore water.

**Overview**

When the phenomena of any class are in general ambiguous, and admit of being explained by different or even opposite theories; if few of those exclusive facts are known, which admit but one or a few solutions, then we have no right to expect much from our endeavors to generalize, except the knowledge of the points where our information is most deficient, and to which our observations ought chiefly to be directed.

*Playfair* [1802, pp. 515–516]

The task of analyzing groundwater flow in general is a difficult one. This is indicated by a retrospective examination by Konikow and Patten [1985] of attempts to simulate flow accurately enough to make usable predictions. They showed that the degree of success was quite varied. Analysis of flow in low-permeability environments can be even more challenging in many respects. In this review I have attempted to analyze the aspects which contribute to the difficulty of quantitative analysis. The major ones are summarized in Table 1. Many of the phenomena alluded to in the tabulation are incompletely understood. Perhaps more fundamentally, the significance of some of the phenomena in subsurface systems is unknown or controversial. The questions raised by these uncertainties are important subjects of future research in order to improve the ability to understand groundwater flow in these environments.

**Historical Bias**

Our present understanding of low-permeability environments reflects how the problem has been approached. Soil mechanics and soil physics have long held flow in low-permeability media, particularly clays, to be of major interest. However, in hydrology and hydrogeology most work in tight media has been preceded by that on more permeable media.
TABLE 1. Factors Adding Uncertainty to Nature of Large Scale Groundwater Flow in Low Permeability Environments

<table>
<thead>
<tr>
<th>Factor</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extrapolation of Darcian relation to in situ hydraulic gradients</td>
<td>Unknown</td>
</tr>
<tr>
<td>Coupled flow (osmosis and ultrafiltration)</td>
<td>Importance in subsurface disputed, potentially capable of causing extremely long lived transient conditions</td>
</tr>
<tr>
<td>Geologic processes occurring on time scales comparable with flow</td>
<td>Geologic processes can apparently be the dominant control of flow in certain settings</td>
</tr>
<tr>
<td>Time scale dependence of storage parameters and size scale dependence of hydraulic and nonhydraulic conductivity</td>
<td>Long term and laboratory $S_h$ may differ by one or more orders of magnitude, large and small scale diffusivity and conductivity by several orders of magnitude</td>
</tr>
<tr>
<td>Three-dimensional coupling between stress and flow (equations (6) and (7))</td>
<td>Unknown, probably most significant where tectonic stresses are important</td>
</tr>
<tr>
<td>Viscoelastic deformation of solid skeleton and groundwater flow on comparable time scales (intermediate range in Figure 7)</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

In addition, collective experience with more permeable media is much greater in these disciplines. Consequently, there appears to be a historical bias attached to the study of tight systems, particularly in a hydrogeological context. This seems to explain, in part, certain current approaches to low-permeability problems. For example, it is ironic that few question the applicability of Darcy’s law in tight media at small hydraulic gradients, for which we lack experimental justification, while most quantitative flow analyses have ignored coupled flow, for which ample experimental evidence exists. The experimental evidence for coupled flow is sufficiently extensive that its failure to influence flow in large systems, if true, requires explanation. Further progress will be helped by recognizing this bias.

Such a bias may also play a role in the readiness with which mathematical models based on elastic theory have been extrapolated to describe transient flow over geologically significant time periods. With successful application of transient flow theory on a regional scale in developed aquifers, applying the theory in large, naturally transient low-permeability systems seems a natural extension.

Fundamental Difficulties and the Scientific Method

The study of groundwater flow has generally been characterized by the advancement of theoretical models whose usefulness or “correctness” is subsequently tested by application. As examples, one may cite Terzaghi’s [1923] theory as a description of consolidation and Jacob’s [1940] equation as a description of transient flow in aquifers. Both contain numerous simplifications and assumptions and yet have been shown, through application, to provide useful descriptions in a wide range of problems. As a more recent counterexample, consider the classic convective-diffusive description [e.g., Ogata, 1970] of solute transport in groundwater; recent work, including experimental studies, have shown it to have important limitations [e.g., Simmons, 1982].

The ability to test conceptual models of flow in low-permeability environments is severely limited by the constraints of time; response on a large scale takes too long to observe. Even if monitoring efforts were to extend over decades, the changes would often be unobservably small. In effect, even under the most ideal conditions for data collection all we can obtain is a “snapshot” of the dynamic behavior. In practical terms this means it is difficult to ascertain the appropriate mathematical descriptions and parameters through experience with application to actual systems. Uncertain size scale dependence of hydraulic and membrane conductivities and uncertain time scale dependence of storage parameters are the manner in which this problem is manifested in the current theoretical framework.

The difficulties here are fundamental; they cannot be readily overcome by technological or analytical advances. One way around the problem is to seek indirect evidence for long-term behavior of porous media by studying the geologic evidence of processes which have been operative for a long time. The nature of the problem is such, however, that significant uncertainty is likely to remain a component of most analyses encompassing large spans of time.

In actual geologic environments one is likely to be faced with a multitude of former and present geologic processes occurring at rates comparable with rates of pressure dissipation in tight rocks. Stress, temperature, chemical, and other effects can all play a part in hydrodynamically stressing the system; the lack of quantitative understanding of these processes presents difficulties. It has so far proved very difficult to identify the significant and insignificant processes with confidence.

Viewing in the context of scientific inquiry [Bredelhoft et al., 1983] one attempts to explain observed conditions by constructing rational physical models of the processes causing them. Rational models for several causal processes can often be proposed. The scientific method therefore requires that the models be testable to premit rejection if they are inappropriate. Unfortunately, low-permeability systems are distinguished by the fact that indicative experimental tests are extremely difficult in practice.

Notation

- $C$ dimensionless coefficient.
- $C_A, C_B$ solute concentration at sample boundaries ($M/L^2$).
- $C_s$ solute concentration at a point ($M/L^2$).
- $E$ electrical potential ($L^2 M/Q T^2$).
- $G$ geothermal gradient (deg/L).
- $h$ hydraulic head ($L$).
- $h'$ hydraulic head above or below hydrostatic ($L$).
- $J_i$ flux of ith component (various dimensions).
- $K$ hydraulic conductivity ($L/T$).
- $K_e$ electroosmotic conductivity ($Q T/M$).
- $L$ ground elevation ($L$).
- $L_{ik}$ coupling conductivity (various dimensions).
- $l$ representative size dimension ($L$).
- $n$ porosity ($L^2/L^2$).
- $q$ groundwater flux ($L^2/T$).
- $S_i$ “one-dimensional” specific storage, equal to $S_i(1 - \beta f)$ ($L^{-1}$).
$S_v$ “three-dimensional” specific storage, equal to
$\gamma[(x-a)/((x-a)+n(\sigma_p-a))]$ \((L^1)\).

$T$ temperature (degrees Celsius).

$t$ time \((T)\).

$t^*$ dimensionless time.

$v$ fluid volume \((L^1)\).

$X_k$ kth thermodynamic driving “force” (various dimensions).

$z$ vertical distance.

$\alpha$ bulk compressibility of porous medium \((LT^2/M)\).

$\alpha_f$ compressibility of groundwater \((LT^2/M)\).

$\alpha_u$ bulk compressibility of solids \((LT^2/M)\).

$\alpha_{ef}$ thermal expansivity of groundwater \((1/\text{deg})\).

$\alpha_v$ vertical compressibility (no horizontal strain) of porous medium, equal to $[(1-v)/[3(1-v)]\alpha (LT^2/M)]$.

$\alpha_{ve}$ vertical compressibility (no horizontal strain) of porous medium for compression only \((LT^2/M)\).

$\alpha_{ve}'$ effective vertical compressibility (no horizontal strain) which accounts for long-term deformation \((LT^2/M)\).

$\alpha_{ve}$ vertical compressibility (no horizontal strain) of porous medium for expansion only \((LT^2/M)\).

$\alpha_{ve}'$ effective vertical compressibility (no horizontal strain) which accounts for long-term expansional deformation \((LT^2/M)\).

$\beta$ “three-dimensional” loading efficiency, equal to $\{(x-a)/(x-a)+n(\sigma_p-a)/\}$ \((\text{dimensionless})\).

$\gamma$ specific weight of saturated medium, equal to $(1-n\eta_2 + n\eta_f) (M/L^1 T^2)$.

$\gamma_s$ specific weight of solids \((M/L^1 T^2)\).

$\gamma_f$ specific weight of groundwater \((M/L^1 T^2)\).

$\gamma^*$ “one-dimensional” loading efficiency, equal to $\{(\beta(1+v))/[3(1-v) -2(1 -\alpha/s)\beta(1-2v)]\}$ \((\text{dimensionless})\).

$K$ hydraulic diffusivity, equal to $K_S/(L^2/T)$.

$\lambda$ dimensionless coefficient, equal to $[2(1-\alpha/s)(1-2v)/[3(1-v)]$.

$\nu$ Poisson’s ratio for the porous medium \((\text{dimensionless})\).

$\sigma_t$ mean total stress \((M/T^2 L)\).

$\sigma_v$ vertical total stress \((M/T^2 L)\).

Acknowledgments. For a paper of this sort, critical review by several of one’s colleagues is especially important. For their very helpful reviews of the manuscript at various stages I wish to thank John Bredehoeft, P. A. Domenico, Youfif Kharaka, Allen Moench, Hal Olsen, Mark Person, and Allen Shapiro. I am also grateful to several anonymous reviewers who provided valuable comments. Finally, I am indebted to the fine staff of the U.S. Geological Survey National Center library for their help on numerous occasions.

References


Dennehy, K. F., and P. A. Davis, Hydrologic testing of tight zones in Delaney, P. T., Rapid intrusion of magma into wet rock: Ground-
of a finite diameter well to an instantaneous charge of water, Resour.
by
Bredehoeft, J. D., The drill-stem test: The petroleum industry's deep
problem of studying abnormally high geofluid pressures in connection with search, exploration, and exploitation of oil and gas
Fritz, S. J., and I. W. Marine, Experimental support for a predictive


Kharka, Y. K., and F. A. F. Berry, The influence of geological mem-


Thes, C. V., The relation between the lowering of the piezometer surface and the rate and duration of discharge of a well using groundwater storage, Eos Trans. AGU, 16, 519-524, 1935.


(Received December 12, 1985; revised March 29, 1986; accepted April 14, 1986.)