Chapter 1

Hydrology, mechanics, and geomorphic implications of erosion by subsurface flow

Thomas Dunne
Department of Geological Sciences, University of Washington, Seattle, Washington 98195

GEOMORPHIC SIGNIFICANCE OF SUBSURFACE FLOW

Hydrologic processes drive erosion by water; therefore, the study of groundwater geomorphology requires a review of the hydrology of subsurface flow and the physical constraints on the erosion that it accomplishes. The principles of groundwater hydrology are described in many textbooks (e.g., Freeze and Cherry, 1979), and only those aspects most specifically related to erosion by flowing groundwater will be reviewed here. The role of subsurface water in mass wasting will be reviewed in Chapter 2. However, in the past 20 years there has been an explosion of research on shallow subsurface flow, with implications for the understanding of erosion processes and landforms. Kirkby and Chorley (1967) and Kirkby (1978) pointed out that the evolution of concepts describing shallow subsurface flow and its relation to surface runoff has implications for understanding erosion processes. This recent research will also be reviewed in the present chapter. The review of erosion mechanisms will be restricted to processes operating in rocks which, although subject to chemical weathering, are less soluble than those producing karst landscapes. The mechanisms and geomorphic role of solution are reviewed in Chapters 6 and 7.

GROUND-WATER STORAGE AND MOVEMENT

Recharge

Although some groundwater originates as connate or juvenile water, the vast majority is meteoric, and the flux of water across the water table is defined in hydrology as the groundwater recharge. Thus, most geomorphological effects of groundwater require recharge by infiltration and percolation to the water table, even if the recharge was remote from the current time and place of interest (Lloyd and Farag, 1978). The magnitude of recharge is limited first of all by the amount of rain that falls or snow that melts at rates lower than the infiltration capacity of the soil. Water delivered to the soil at higher rates runs over the surface as Horton overland flow and is responsible for surface erosion, as discussed in other geomorphological texts (Carson and Kirkby, 1972; Selby, 1982), which review the climatic, edaphic, and biotic controls on infiltration, runoff, and the resulting sheetwash and gully erosion. Some of this runoff may accumulate later in valley floors and recharge the groundwater, particularly in arid landscapes. For many years the geomorphological literature concentrated almost entirely on the action of surface runoff, except in soluble rocks. However, it has long been reported in early literature from Indonesia reviewed by Roessl (1950) and the writings of Hursh (1944) and his colleagues from the forested southern Appalachians, for example, that, over large areas of Earth's surface, infiltration capacities almost always exceed rainfall intensity, and virtually all runoff travels underground.

This subsurface flow percolates through a shallow (vadose) zone, which may be saturated periodically but is unsaturated (i.e., the voids are only partly filled with water) most of the time. In this zone the co-existence of air and water in void spaces is responsible for menisci and the resulting negative pressures (capillary effects) in the pore water. At greater depths lies the saturated zone where all voids are filled with water. In coarse-grained porous media and jointed rocks, the pore-water pressures in this saturated zone are positive with respect to atmospheric pressure, and the boundary between the saturated and unsaturated zones is defined as the water table, where the water pressure is equal to atmospheric pressure. In fine-grained porous media, however, capillary effects fill pores with water to some height above the water table, forming a capillary fringe or tension-saturated zone. In this case, the water table is still defined as the level at which atmospheric pressure exists in the pore water; alternatively, it is defined as the height to which water stands in a well perforated throughout its length. Because water can only enter a well below the water table where pressures exceed atmospheric pressure, this zone is sometimes called the phreatic zone. This term, as well as "vadose zone," are falling into disuse, however, because seasonal fluctuations of the water table and the development of perched saturated zones over aquitards at shallow depth (see later) make the definitions confusing and useless.

As water infiltrates the soil, a portion of it is stored in raising the soil-water content, and is later evaporated back to the atmosphere. The difference between the precipitation input and evaporative loss sets the magnitude of the groundwater recharge, which therefore reflects climatic, edaphic, and biotic factors that are themselves linked. However, many aspects of erosion resulting from subsurface flow are controlled not only by the total annual recharge, but by short-term rates of recharge during rainstorms, melts, or wet seasons. Freeze (1969) used a mathematical model to examine the conditions that favor recharge of the groundwater and, therefore, water-table rise. Recharge is enhanced by: (1) rainfalls or melt of low intensity (relative to the saturated hydraulic conductivity of the soil) and long duration rather than short, intense storms; (2) shallowness of the unsaturated zone above the water table; (3) wet antecedent conditions; and (4) soils with high conductivity, low rate of increase of moisture content with pressure, or high moisture content over a considerable range of pressure-head values.

Percolation

The vertical motion of water in a homogeneous and isotropic soil is described by Darcy’s Law (Rubin, 1966):

\[ I(z) = -K(w) \frac{\partial \Phi}{\partial z} = -K(w) \left[ \frac{\partial \Psi}{\partial z} - 1 \right] \]

(1)

where \( I(z) \) is the vertical flux rate at any depth, \( z \); \( K(w) \) is the hydraulic conductivity, a function of the moisture content, \( w \); \( \Phi \) is the total hydraulic head and \( \Psi \) is the pressure head, both functions of \( w \). A complete list of the symbols used in this chapter and their definitions is found in Appendix 1. As the soil is wetted to progressively greater depths, the vertical moisture gradient, and therefore the pressure gradient, \( \partial \Psi / \partial z \), declines, and \( I(z) \) converges on \( K(w) \). Thus, the maximum value that \( I(z) \) can attain after a long period of rainfall (typically 10 to 60 minutes) is equal to the maximum value that \( K \) can attain, which is the saturated hydraulic conductivity. If the vertical flux rate is imposed at the surface by a fixed rainfall intensity less than this saturated conductivity, the moisture content, \( w \), will increase until \( K(w) \) is sufficient to pass the water at the applied rate. The exact flow path depends on the degree to which the hydraulic conductivity is homogeneous and isotropic (Zaslavsky and Rogowski, 1972), but in many simple stratigraphic conditions is vertical. The rainfall intensity controls the flux rate at the surface from the beginning of the storm and at increasing depths later in the storm as the wave of moisture of constant \( w \) penetrates deeper into the soil. At the end of the rainstorm, the surface influx stops, but water continues to drain out of the upper layers of the soil, so that the wave of vertical moisture flux is damped as water moves to lower levels (Youngs, 1958; Dunne, 1978; Sakura, 1983).

A succession of rainstorms generates waves of percolation, each becoming damped as it travels deeper into the unsaturated zone. Inter-storm differences of intensity, duration, and the intervening evaporation usually mask this simple picture, but it is useful for illustration. At sufficient depth, such as in the saturated zone in a deep, permeable aquifer, one might observe only slow, seasonal changes in recharge, representing the highly damped influence of many rainstorms. In other cases, a saturated zone at a depth in the range 1 to 10 m may be recharged by a single large rainstorm, but many hours after the end of the storm and at a rate less than the maximum infiltration rate (for examples see Dunne, 1970, p. 27; 1978, p. 287). Open fractures in bedrock or long, large openings in soil profiles may reduce both the delay and damping of these single waves of recharge.

Regional groundwater flow

When the percolating water reaches the saturated zone, its movement continues to reflect the spatial variation of the hydraulic potential, as expressed in the general form of Darcy’s Law:

\[ Q = -K \nabla \Phi \]

(2)

where \( Q \) is the specific flux (discharge per unit cross-sectional area), which in an isotropic homogeneous medium is parallel and opposite to the gradient of the hydraulic head \( \Phi \). The saturated hydraulic conductivity, \( K \), is constant at saturation (in most rocks and soils), but may vary spatially and with direction as a result of lithologic and pedologic characteristics. The hydraulic head is the sum of the elevation head, \( z \), measured with respect to some arbitrary datum, and the pressure head, \( \Psi \), measured as the height to which water rises inside a piezometer. Surfaces or contours of equal \( \Phi \) can thus be defined, and are called equipotentials. Equation 2 implies that flow lines cross these equipotentials at angles that depend on the degree to which \( K \) is homogeneous and isotropic but are 90° for the simplest case.

Boundary conditions such as the presence of underlying aquitards, surface topography, and the locations of drainage outlets in stream depressions, as well as heterogeneous and anisotropic hydraulic properties of the aquifer resulting from lithologic, structural, and stratigraphic characteristics, cause the flow lines to diverge from the vertical and to follow arcuate paths to a drainage outlet. Figure 1 indicates the arcuate flow paths perpendicular...
Implications of erosion by subsurface flow

Figure 2. Regional groundwater flow in an isotropic homogeneous aquifer beneath a sinusoidal topography with drainage outlets at different elevations (Toth, 1963).

to equipotential lines in a homogeneous isotropic aquifer beneath a simple periodic topography, underlain by an impermeable boundary. The underlying aquifer and the vertical flow divides beneath the streams and ridge crests form the boundary conditions, which together with the uniform distribution of recharge along hillslopes, force the flow lines to curve. For simple boundary conditions and uniform hydraulic characteristics, there are analytical methods and computer methods (Wang and Anderson, 1982).

If the topography has a regional slope, with drainage lines at different elevations, the regional groundwater body develops a nested set of flow subsystems (Fig. 2), as illustrated by Toth’s (1962, 1963, 1966) mathematical analysis and interpretation of field measurements. As the amplitude of the topography increases, more of the groundwater flow is included in local systems between the ridge and the nearest depression, and less travels long distances to regional drainage lines. Freeze and Witherspoon (1967) used finite-difference computations to illustrate how geology can complicate these simple patterns.

In most discussions and applications of Darcy’s Law it is assumed that the water migrates through intergranular pores at velocities that keep the flow behavior well within the viscous range. Fractured rocks, on the other hand, allow flow through cracks, joints, fissures, and large solutional cavities. If the fractures are dense, narrow, and of roughly uniform size, Darcy’s Law is usually adequate for describing the flow if it is slow enough and if the hydraulic properties are specified for a sufficiently large, representative volume, although strong heterogeneities and anisotropies may be imposed by fracture densities and orientations. Where groundwater flows through only a few large, widely spaced fissures, the continuum assumption implicit in Darcy’s Law is not valid, and it may be necessary to consider the role of individual conduits. If the fractures are sufficiently wide, flow behavior deviates from the linear behavior specified by Equation 2, and new empirical flow laws are presently being investigated for such complex, rough conduits. Unfortunately, for many aspects of groundwater geomorphology in massively jointed rocks or conduits in soils, such non-Darcian aspects of the flow are paramount, and modeling based on Darcy’s Law is not applicable.

Shallow subsurface flow

In many landscapes, the hydraulic conductivity of the near-surface zone decreases from the topsoil to the unweathered parent material. This decrease may be gradual or abrupt, and is not necessarily monotonic, although for the sake of simplicity it will
be treated as such here. As the downward-percolating rainwater encounters zones of diminishing vertical hydraulic conductivity, the moisture content of successive layers must be raised to ever-increasing levels to pass the water at the applied percolation rate. At some depth, moisture content will be raised to saturation. At this depth a “perched” zone of saturation develops, either above an obvious impeding horizon, or simply at some depth (varying between rainstorms) at which the saturated conductivity is less than the rainfall intensity (Fig. 3). At these depths, water is diverted laterally and percolates downslope through the colluvium or bedrock to stream channels. The thickness of the resulting saturated zone and the pressure of water in it depend on the vertical percolation rate, initial moisture content, hydraulic properties of the medium, the gradient of the impeding layer, and hillslope length.

Kirkby and Chorley (1967) called this subsurface flow “throughflow,” which Chorley (1978, p. 374) defined as “downslope flow of water occurring physically within the soil profile, usually under unsaturated conditions except close to flowing streams, occurring where permeability decreases with depth.” Whipkey and Kirkby (1978) give a thorough discussion of throughflow under saturated conditions, which is of greater significance for geomorphology. Because the definition of throughflow is no more specific than the older term “interflow,” which had been rejected by hydrologists as being poorly defined (e.g., Hewlett and Hibbert, 1967, p. 275), I will use only the term “subsurface flow” and will designate whether the flow is saturated or unsaturated, and whether the flow path lies through soil or rock at shallow depth or deeper, whenever issues are important.

In the simplest case of steady-state percolation down a long, straight hillslope, the slope of the water table in Figure 3 is approximately equal to the slopes of the ground surface and the base of the saturated zone. The flow lines, thus, are approximately parallel to the water table, but the saturated hydraulic conductivity (K) may vary with height (h) normal to the base of the saturated zone. Darcy's Law reduces to

\[ q_i = \int_{h_0}^{h} J \, dx = \sin \alpha \int_{h_0}^{h} K(h) \, dh \]

where \( q_i \) is the downslope discharge of subsurface flow per unit width of hillslope, \( J \) is the recharge rate, \( x \) is the horizontal distance, \( \alpha \) (\( \approx \theta \), the hillslope angle) is the angle of the water table, and \( h_w \) is the thickness of the saturated zone. Thus, the saturated layer thickens downslope, at a linear rate if \( K \) is constant with depth:

\[ h_w = \frac{Jx}{K \sin \theta} \]

and more slowly if \( K \) varies, for example, with \( h^2 \):

\[ h_w = \left( \frac{3Jx}{K \sin \theta} \right)^{1/2} \]

Since flow lines parallel to the slope imply equipotentials normal to the slope (for an isotropic and homogeneous soil), the distributions of hydraulic head and pore pressure within the saturated zone can be obtained as shown in Figure 4. Beven (1982) and Iida (1984) demonstrated the prediction of water-table elevations on planar hillslopes at any time before steady state is reached. Iida also illustrated the effects of nonplanar hillslopes on the water-table and pore pressure. Dietrich and others (1986) used the Iida method to calculate pore pressures in colluvium for their study of the effect of subsurface flow on mass wasting.

These methods predict that the saturated thickness increases monotonically along the slope, and take no account of the drawdown of the water table by drainage out of a seepage face at the
Implications of erosion by subsurface flow

Water table

Flow line

Ground surface

Total head = elevation head - pressure head

\[ \Phi_A = h_w \cos \theta - 0 \]

\[ \Phi_B = 0 - h_w \cos \theta \]

Figure 4. Relation of hydraulic head and pressure head to flow lines and water-table elevations in a homogeneous isotropic soil on an infinite slope. \( \Phi \) is the hydraulic head.

downstream end of the flow system. A very approximate solution for the shape of the steady-state water table above a sloping impermeable boundary under uniform recharge, \( I \), can be obtained by using the Dupuit-Forchheimer assumptions (Freeze and Cherry, 1979) that the aquifer is homogeneous and isotropic, flow lines are horizontal, and the hydraulic head gradient is equal to the slope of the water table. Then, in Figure 5, the discharge per unit width at some horizontal distance, \( x \), is

\[ q_s = Ix = -K\tau \frac{dt}{dx} \]

where \( \tau \) is the saturated thickness of the aquifer and \( x \) is the elevation of the water table. It is not necessary in this analysis that \( x = 0 \) be set directly beneath the surface drainage divide if there is a stratigraphic control on the recharge area.

Since \( z - \zeta = (L-x) \tan \beta \)

\[ q_s = Ix = -KLx \frac{dt}{dx} \]

and the slope of the water table is given by

\[ \frac{dt}{dx} = \frac{I}{KL \tan \beta - Kz \tan \beta} \]  

(4)

Except for the case of \( \beta = 0 \), which implies that the water table is parabolic, the solution of the differential equation is complicated and does not yield much physical insight to this writer. However, for any \( \beta \), Equation 4 indicates that the water-table slope, and therefore the magnitude of the hydraulic head gradient (\( i \)) at the base of the hillslope (\( x = L \)) where \( z = z_L \) (the height of the seepage face), is

\[ i = \frac{dt}{dx} = -\frac{IL}{Kz} \]  

(5)

The Dupuit-Forchheimer assumption of horizontal flow lines is a useful approximation only where the water-table gradient is small, and thus Equation 4 is particularly inaccurate for predicting the water table near the seepage face where it predicts unrealistically large water-table gradients. However, it incorporates the essential physics of the steady-state water-table response to recharge, and thus is used here for illustration of the effects of the major controls on the hydraulic-head gradient near the seepage face. It indicates that the hydraulic-head gradient at the seepage face is directly proportional to the recharge rate and drainage area, and inversely proportional to hydraulic conductivity and height of the seepage face, which in Figure 5 is equated to the depth of streamflow. The gradient of the underlying flow base, \( \beta \), is not included in the steady-state form of Equation 4 at \( x = L \), but this is partly a limitation imposed by the assumption of horizontal flow lines. In the field, the magnitude of \( \beta \) sets a lower limit on the angle of the water table at the seepage face. Also, the time required for attaining steady state, and therefore the peak discharge and water-table gradient, in shorter events would increase with \( \beta \). Iida (1984) showed that in any recharge event shorter than the equilibration time for subsurface flow in a shallow colluvial aquifer of uniform depth

\[ q_{pk} = \frac{IK}{p \sin \beta \cos \beta \tau} \]  

(6)

where \( q_{pk} \) is the peak discharge per unit width, \( p \) is the porosity, and \( \tau \) is duration of constant recharge entering the saturated layer. Beven (1982) added the duration of unsaturated percolation to the analysis.

If the soil and rock are sufficiently deep and permeable to conduct all the infiltrated water to the stream channel, the flow path lies entirely below the ground surface. Field experiments describing such subsurface flow were summarized by Dunne

Figure 5. Steady-state water table under uniform recharge, \( I \), using the Dupuit-Forchheimer assumption of horizontal flow lines over a planar aquiclude.
(1978) and by Whipkey and Kirkby (1978), who summarized the hydrological principles governing the results. More recently there has been some debate about the relative contributions of rapid flow through large voids, called macropores (see later), and more diffuse, slower flow through intergranular and interpedal voids (Pearce and others, 1986; Sklash and others, 1986). The answer, suggested by soil physics (Beven and Germann, 1981, 1982) and by experiment (Horton and Hawkins, 1965), is likely to be that the relative importance of macropore flow increases with rainfall intensity, antecedent soil moisture, and the linearity of the moisture content—tension—conductivity relation, and that it decreases with increasing depth of soil. In turn, these parameters vary with physical geography, and the field studies of Pearce and colleagues. Jones (1981), Gilman and Newson (1980), and others are currently mapping the geography of these relative contributions. The issue can become very important for understanding erosion by subsurface flow in some regions, as described in later sections of this chapter.

On footslopes and in valley bottoms where the water table is close to the ground surface at the beginning of the storm (Fig. 6), the tension-saturated zone may intersect the ground surface, and the addition of only a small amount of water to the surface converts the soil-water pressures in this zone to positive values, so that the water table rises rapidly. As it rises, the water table slopes steeply toward the channel, driving an increase in subsurface flow to the channel, often through a highly permeable soil or sediment with abundant macropores. Ragan (1968) and Dunne (1978) demonstrated by piezometric, water-table, and soil-moisture measurements that such rapid rises of the regional and local water table outcropping along a valley floor or footslope could cause the water table to rise in a ridge-like pattern, sloping both toward and away from the stream. Sklash and Farvolden (1979) and Sklash and others (1986) have provided isotopic evidence of the relative contributions of runoff from this source to storm runoff, while Sakura and Taniguchi (1983) and Gillham (1984) have studied the process experimentally.

There is a limit to the amount of subsurface flow that can pass downhill within a soil profile. The limit is governed by the hillslope gradient, the hydraulic conductivity, and the soil depth, as expressed in Equation 3. When $h_w$ coincides with the soil surface, this limit is reached, and any downslope decrease in this capacity or further additions of moisture by rainfall must be accommodated by overland flow (Fig. 6). Downslope decrease in the capacity of the soil to transmit subsurface flow causes water to emerge from the ground surface as return flow (Musgrave and Holtan, 1964; Dunne and Black, 1970b), also referred to as exfiltration (Freeze, 1974). The water emerging from this fully saturated soil is augmented by direct rainfall or melt onto the saturated zone, and the sum is referred to as saturation overland flow (Kirkby and Chorley, 1967). This form of runoff is most likely to occur where topographic or stratigraphic conditions force flow to converge, or where gradient, soil thickness, or hydraulic conductivity decrease (Kirkby and Chorley, 1967; Dunne and Black, 1970a, b; Tanaka, 1982; Wilson and Dietrich, 1987). Hence, the emergence of shallow subsurface flow is most common on footslopes, in swales, and on thin soils with low conductivity. The emergence is associated with a lateral or downward component of the hydraulic head gradient in the water, exerting an oppositely directed stress on soil particles, which may thus be eroded and carried downslope by the saturation overland flow. The erosion mechanisms will be considered later in this chapter.

Subsurface flow in soils and near-surface sediments above the regional water table also includes some non-Darcian components as a result of large passages, which modify the flow and pressure fields and in turn are altered by them. Such passages include root holes, the burrows of soil fauna, shrinkage cracks, cracks around landslide blocks, and tectonic joints. The origin and hydraulic characteristics of "macropores" in soils have been described by Jones (1981) and Beven and Germann (1982). Figure 7, developed from an original example in the latter paper, summarizes the influence of macropores on overland and subsurface flow and their interaction. Some larger passages develop as a result of the subterranean erosion processes described later, and may attain diameters of up to tens of meters (Parker and others, 1964; Carey, 1976). In these passages, the flow process is turbulent streamflow, and is similar to conduit flow in large solutional cavities described in Chapter 7. The complex geometry of these tortuous conduits has so far defeated any study of their hydraulics, but an indication of their roughness is given by a measurement of flow speed in an elliptical soil pipe (Tanaka and others 1982, p. 34) from which a value of approximately $4 \text{ m}^2\text{s}^{-1}$ can be calculated for the Chezy C-value for the smoothness of pipe.
implications of erosion by subsurface flow

The pressure at which water can flow from a porous medium and enter a macropore or a conduit such as an open joint depends on the size of the passage, but for the large cracks and root holes of interest here, water in the porous medium must be under positive pressure relative to atmospheric pressure. Thus, overland flow, which occurs over a saturated soil surface, may enter such a passage, falling freely or trickling slowly down the sides, depending on its surface velocity and the morphology of the passage. If it falls freely, it may travel far down the macropore before entering the porous medium. If the flow trickles slowly down the margins of the passage through an unsaturated soil or rock, a portion of it will be drawn into the unsaturated matrix by the resulting pressure gradient (Horton and Hawkins, 1965; Bouma and others, 1978). For flow to continue along the macropore, the influx must exceed the lateral loss, which decreases through time as the matrix becomes wet and the pressure gradient decreases. Thus, Heede (1971) noted that intense rainstorms in Colorado did not generate flow in tunnels; only longer snowmelt periods could wet the tunnel margins sufficiently to allow discharge.

Water that infiltrates the porous medium under unsaturated conditions migrates within the matrix, avoiding macropores; but if the soil becomes saturated, even locally, either from the surface, or above some horizon that impedes vertical drainage and forms a perched zone of saturation, or at the roof of the macropore if vertical percolation supplies water to that point faster than unsaturated percolation can convey it around the cavity, the pressure gradient is reversed, and water flows from the matrix to the macropore (Fig. 7), along which it can travel much faster than the matrix flow. By this means, large passages provide paths through which water can by-pass slow flow through the soil matrix and, especially important, through horizons of low conductivity. Even thin, discontinuous zones of low conductivity in the soil or rock may force water into passages of various size and origin. Pierson (1983) has demonstrated experimentally that the local increase in conductivity associated with a macropore may cause an abrupt rise in pore pressure at the end of a blocked passage. Such a pressure increase could lead to slope failure or seepage erosion under favorable conditions.

Groundwater discharge

Subsurface flow discharges from the ground surface where the topography is intersected by the water table in an unconfined aquifer, or lies below the piezometric surface of a confined aquifer that has become connected to the surface by some breach (such as a fault) in the overlying aquiclude. Thus, topography, geologic structure and stratigraphy, and the recharge rate and size of the flow system determine the spatial distribution of seepage. The terms "seepage zone" or "seepage face" are applied to the area of ground surface over which the discharge occurs. These zones may be active perennially or episodically, depending on the timing of recharge and the geometry of the flow system as described above.

If one walks along a seepage zone, it is usually possible to discern variations in local discharge rates along a contour. Where the discharge is sufficiently concentrated, the feature is informally called a spring. Fetter (1980), p. 165–167) illustrates some of the topographic and geologic conditions that localize springs. Large springs, usually occurring at a stratigraphic and/or structural inhomogeneity, are easy to recognize, and the fact that they distort the near-surface flow field and drain large volumes of aquifer cause them to dry out the neighboring hillside, but there is a spectrum of spring magnitudes that grades imperceptibly into seepage zones.

Seepage is rarely uniform along a contour because inhomogeneities and anisotropy of aquifer properties, particularly porosity and hydraulic conductivity, cause concentrations of flow within the aquifer and near the seepage face. The most obvious of these heterogeneities result from fractures that can have a range of widths from micro-cracks to master joints and solution conduits. The effect of these properties on groundwater flow fields is beginning to receive study, but their consequences for flow near seepage faces has not yet been examined. It is now known, however, that a formation previously characterized as spatially uniform is better thought of as having random spatial variations of hydrogeologic parameters (Freeze, 1975). Thus, a "homogeneous" formation has highly correlated frequency distributions of conductivity and porosity, the former usually being lognormally distributed and the latter being normally distributed. The formation is considered to be homogeneous if these frequency distributions do not change through space. If they do vary spatially, and especially if they are multimodal, the formation is considered to be heterogeneous. If the properties are strongly autocorrelated spatially (e.g., if master joints are spaced uniformly), the concentrations of flow will be strong. Ineson (1963), for example, showed that large differences of transmissivity (the product of saturated thickness and hydraulic conductivity) occurred along the strike of...
the Chalk aquifer in southeastern England. The formation is a soft limestone with systems of fine fractures, but small faults and zones of concentrated fracturing along the axes of gentle anticlines transverse to the regional dip locally increase the transmissivity more than tenfold. These maxima have a roughly uniform spacing of 10 to 15 km, and valleys coincide with them.

Howard (1986b) incorporated random variability of the hydraulic conductivity into his two-dimensional mathematical model of groundwater flow and resulting erosion, but he did not utilize any spatial autocorrelation parameter to reflect repetitive aquifer properties such as joint systems. Nevertheless, the alteration of the shape of the seepage boundary by erosion in his mathematical and laboratory models caused flow concentrations with a characteristic spacing as the headward-eroding channels competed in draining water out of the aquifer (see later).

Local increases in groundwater flux and exfiltration can also result from the presence of buried features such as an anticline or other protuberance of low-conductivity rock buried by more permeable sediments. Such disturbances of the flow field often cause sufficient groundwater concentration to affect features such as springs, seepage zones, or channel networks, and the presence of buried structures is often recognized from such drainage features during map or photo interpretation. Concentrated sources of groundwater inflow can also produce concentrations of outflow nearby, although the diffusion of groundwater tends to reduce the concentration over greater flow distances. For example, near Bananal in southeastern Brazil, I have seen large gullies (vogorocos), which are eroding headward through a valley fill (Coelho Netto and others, 1987) branch and head toward concentrations of subsurface flow emanating from bedrock hollows partially filled with deep colluvium on the valley sides.

In a later section of this chapter it will be emphasized that the flow direction and speed, and therefore the hydraulic potential gradient in the outflow zone, are important characteristics governing the mechanics of erosion by water emerging from the ground. Construction of flow nets for regional or local flow fields (using piezometric measurements or the graphical or mathematical modeling techniques illustrated by Cedergren, 1967; Freeze and Cherry, 1979; and Wang and Anderson, 1982) indicates that these flow speeds and directions, as well as the location and extent of the discharge area, depend on: recharge rate; topography; and the stratigraphy, thickness, lateral extent, and hydraulic properties of rock formations. Of particular importance is the transient response to high recharge rates, which is exceedingly complex to compute (Freeze, 1971) for realistic landscape geometries, and thus is usually defined by piezometric measurements. The timing and magnitude of the increased hydraulic gradients in the discharge zone depend on the timing and magnitude of widespread water-table rise in the recharge zone. Local flow systems with shallow water tables and high hydraulic conductivity respond quickly, and the highest hydraulic gradients in the discharge zone occur during relatively intense, individual rainstorms. By contrast, discharges from large flow systems and those in less permeable rocks respond only slowly to recharge, so that a long rain or snowmelt season may be required to cause widespread water-table rise and the maximum hydraulic gradients at the outlet of such a system.

The geometry of the outflow boundary, as well as stratigraphic or structural inhomogeneities in the aquifer near the boundary, strongly affect the flow and pressure fields in the critical region where emerging groundwater has a potential for erosion. Figures 1 and 11 indicate that wherever the water table outcrops at a concave boundary, flow will converge on the concavity. It will be established later in this chapter that such a flow concentration produces an increase in hydraulic head gradient and a concentration of seepage erosion. At a three-dimensional concave boundary such as a stream head, the concentration is particularly strong. Rulon and others (1985) have used numerical modeling to examine the development of several seepage faces on slopes in which layers with low hydraulic conductivity force a horizontal component on the flow field. The hydraulic-head distributions and water-table configurations depend strongly on the position of the impeding layers and their hydraulic properties relative to those of the intervening aquifers. The horizontal component of flow increases with the recharge rate and with the ratio of the conductivities of aquifer to aquitard, and it increases as the highest aquitard lies at greater depth in the section.

Stratigraphic inhomogeneities may cause the development of "perched" saturated zones above the regional groundwater body. These perched zones may be saturated perennially or ephemeral, and in the latter case they may be difficult to recognize, despite their hydrological and geomorphological significance. However, the distinction between perched, unconfined, or confined groundwater bodies is sometimes only an artifact of the choice of measurement point (Davis and DeWiest, 1966, p. 44-45), and the mechanics of flow and erosion in perched aquifers is exactly the same as in deeper, regional groundwater bodies. Therefore, the following discussion of erosion mechanics makes no distinction between the vadose and phreatic zones, which themselves are poorly defined and confusing terms, or between perched, confined, or unconfined aquifers. For example, in a zone that is usually defined to be vadose, erosion may occur only during wet years when a perched water table develops there, or when the regional water rises to that elevation, establishing phreatic conditions.

**EROSION BY SUBSURFACE FLOW**

**Introduction**

Subsurface water emerging either at a free face (such as a stream bank, gully head, or cliff) or along a valley floor or hillside has a potential for erosion if the outflow is sufficient to mobilize particles and then overcome local flow resistance and maintain the velocity necessary to transport mobilized particles away from the site. Such erosion must begin from the downstream end of a subterranean flow path and extend headward, either underground or at the surface. This is required because sediment eroded along
a flow path must be evacuated downstream along a previously formed conduit or channel.

Many studies of erosion by subsurface flow have involved detailed examination of channel or valley networks, but little consideration of process. The following text reviews the status of knowledge about the processes responsible for erosion by groundwater and the hydrological constraints on these processes. Attention is paid to the important questions of how the erosion is accomplished and what governs the relation between process and form in a spectrum of landscapes with varying degree of effectiveness of erosion by groundwater relative to erosion by surface agents.

**Terminology**

Subsurface flow can erode in two ways: (1) through the development of a critical body force or drag force that entrains particles in water seeping through and out of a porous medium, causing either liquefaction or Coulomb failure; and (2) through the application of a shear stress to the margins of a macropore, which may have originated independently of the water flow. Parker (1963, p. 104–106) made a distinction between these two processes contributing to the formation of conduits. He also concluded that the first is characteristic of unconsolidated materials, whereas the latter is found in consolidated materials, especially where cracks provide the main avenues for water flow and where the critical shear stress for erosion is lowered by the physical chemistry of the interstitial cement or pore water. The critical stress required for entrainment in either case may be reduced by weathering caused by the exfiltrating water, but landform development through the process referred to as seepage weathering (Higgins, this volume) requires that the weathered residuum be removed by one of the erosion processes referred to above, even if in the limit the mobilizing process is simply mass failure as the seepage intensity declines to a low value.

There is confusion in the geomorphological terminology referring to these processes. For example, Dunne (1970, 1980) followed the engineering literature (e.g., Terzaghi, 1943; Zaslavsky and Kassif, 1965) in referring to the first of the above-mentioned processes as “piping,” which caused “spring sapping” (i.e., undermining of the spring head when a critical seepage force is generated). Higgins (1982, 1984) made a distinction between “piping” and “sapping,” but his use of “piping” followed that of soil scientists, agronomists, and hydrologists (e.g., Fletcher and others, 1954; Jones, 1981), who used the term either without reference to a particular erosion process or with reference to processes unrelated to seepage forces. Chorley (1978, p. 370) simply defined piping as “the formation of natural pipes in soil or other unconsolidated deposits by eluviation or other processes of differential subsurface erosion.” In the papers edited by Bryan and Yair (1982), the term “piping” is also commonly used to imply only pipe formation, and often by the second of the erosion processes mentioned above, which had formerly been called “tunnel erosion” (Bennett, 1939; Buckham and Cockfield, 1950).

The uncertainty is well illustrated in the literature reviewed by Jones (1981).

Terminology that causes so much confusion of the important issue of process should be abandoned. I suggest that the two erosion processes referred to above be called, respectively, seepage erosion (following Hutchinson, 1968, p. 691) and tunnel scour. If “pipe formation” and “tunnel erosion” must be used, they should be employed only in the morphogenetic sense used by Chorley (1978), rather than specifying a process. I suggest that the term “piping” no longer be used. Sapping should be restricted to its dictionary definition as “the undermining of a foundation by digging or eroding” (Concise Oxford Dictionary, 1951), or “the extension of a trench from within the trench itself” (Webster’s New Collegiate Dictionary, 1981). Thus, sapping can be induced by seepage erosion or tunnel scour, and by mass failure or undercutting by a stream. If seepage is concentrated at a spring, trench-like extension or spring sapping is likely; if the seepage is more diffuse or if springs are close together, an entire hillside may gradually retreat by gentle sapping of its footslope. The terms “sapping” and “pipe formation” give no sense of the process responsible for the undermining or the conduit erosion.

The processes of seepage erosion and tunnel scour are not mutually exclusive. For example, seepage erosion is likely where percolation converges on the head of a biogenic macropore (Jones, 1987, p. 232). Parker (1963) described how seepage erosion of gully walls could produce a conduit that could be rapidly enlarged by corrosion and caving of the walls and roof. In the limit, however, the two processes of seepage erosion and tunnel scour are governed by different mechanics and therefore different controls on the resulting morphology and density of channels. Both seepage erosion and tunnel scour can cause retreat of hillslopes by basal erosion or the formation of tunnels or pipes. These conduits are ephemeral, however, on a geological time scale, and they eventually collapse, causing hillside retreat or channel formation with new bounding hillslopes.

**Mechanics of seepage erosion**

Seepage erosion is the entrainment of soil or rock resulting from water flowing through and emerging from a porous medium. Individual grains or large masses of soil or fractured rock may be involved. The mobilization may occur by flowage if the fluid stresses cause the particles to lose their frictional strength, or by Coulomb failure allowing sliding or avalanching. The latter situation is more likely on steep slopes where seepage erosion may trigger or grade into landsliding or debris flow (Iverson and Major, 1986).

The entrainment process may be preceded by a long period of weathering concentrated at the seepage face, reducing the cohesion of the geologic material. Although there have been few detailed instrumental studies of the phenomenon, even casual field observation of seepage faces often reveals more thorough mineral decomposition there than in the surrounding rock (Howard, 1986a). These weathering processes are usually necessary for
reducing the cohesion of intact rock before sapping can occur by seepage erosion or small mass failures. The exfiltrating water is responsible both for concentrating the weathering and for provoking mass failure and seepage erosion, and therefore factors controlling the magnitude and direction of flow and pore-pressure gradients are important to all of these processes. Laitly (1983) made a microscopic study of the deposition by exfiltrating water of calcite within the pores of the weakly cemented Navajo Sandstone in southern Utah. The calcite eventually wedges apart the sand grains, so that the strength of the weathered zone depends on cohesion within the vesicular calcite. This layer of calcrete-separated grains thickens until it can no longer adhere to the stronger sandstone, and it falls off as spalls several centimeters thick. The spalling can be triggered by frost-action, or by the weight of icicles formed by freezing groundwater discharge, or by the seepage pressure itself. This type of detailed study, which is relevant to dry climates where evaporation at the seepage face can bring percolating solutions to saturation, should be complemented by similarly detailed studies of weathering in other climatic and lithologic environments. Sharp (1976) made qualitative observations of frost action on rocks in a zone of groundwater discharge. The growth of ice crystals lifted some rock particles away from the seepage face, while others forming in the interstices of the rock reduced its hydraulic conductivity, and would therefore be expected to increase the seepage pressure and favor spalling. Instrumental studies of concentrated chemical decomposition in seepage zones remain to be done.

Whether acting on decomposed or unaltered geologic materials, seepage erosion results from exfiltrating water imposing a force associated with a hydraulic head gradient across a grain or a larger volume of rock or soil. If seepage is responsible for concentrated weathering, it is also imposing a force on the weathering particles. Therefore, the slope angle required for the particles to collapse or the degree of strength reduction required to allow the particles to fall away are reduced in the presence of a seepage force. As the seepage intensity diminishes (for example, where the exfiltrating water evaporates at the free face), seepage erosion grades into pure seepage weathering with removal by gravitational mass wasting. However, most mass wasting is affected to some degree by seepage.

Writers concentrating on the role of seepage pressure in eroding and undermining hillslopes have expressed this effect as a body force acting on some representative volume of the porous medium to reduce its internal resistance to rupture (e.g., Terzaghi, 1943, p. 258; Zastawsky and Kassell, 1965; Iverson and Major, 1986). Kochel and others (1985), emphasizing the continuity between seepage erosion and the fluvial transport that must evacuate the mobilized sediment to maintain a channel, considered seepage erosion as the result of the fluid drag of emergent groundwater applying a torque to particles at the ground surface. Both the body force and drag force are proportional to the near-surface gradient of the hydraulic potential so that the numerical results are equivalent. However, some of the characteristics of the resulting failure and detailed form of the channel head and seepage face may differ, depending on whether the erosion comprises fluid particulate transport or mass failure, including debris-flow initiation. Both forms of seepage erosion can be observed at seepage faces, particularly at channel heads, but little fieldwork has been done to define the conditions under which each occurs.

Terzaghi (1943) defined the critical conditions required for the simplest case of instantaneous seepage erosion (he called it piping) of a relatively large mass of cohesionless material by the vertical component of seepage beneath a horizontal surface. In Figure 8, the specific water flux, \( q \), is represented by a vector that lies in the \( x-z \) plane and makes an angle \( \lambda \) with the vertical. This seepage is driven by a gradient in the potential energy per unit weight of water, which is called the hydraulic head \( \Phi \). The potential gradient creates a force per unit area of the porous medium normal to the flow line indicated by \( q \). However, in this simplest case the surface is horizontal, and we are concerned with vertical forces that act to move particles outward from that surface. The force per unit area of the horizontal cross section \( \Delta x \Delta y \) is provided by the vertical component of the water flux, and thus the vertical component of the hydraulic head gradient. With reference to Figure 8, the force associated with the vertical flux is

\[
\frac{\text{Force}}{\text{Area}} = \frac{\rho_l g (\Phi_2 - \Delta \Phi) - \rho_l g \Phi_1 \Delta x \Delta y}{\Delta x \Delta y}
\]

where \( \Delta \Phi \) is given by \( \cos \lambda (\Phi_1 - \Phi_2) \). This seepage force is distributed through the volume of the body, to yield

\[
\frac{\text{Force}}{\text{Volume}} = \frac{\rho_l g \Delta \Phi \Delta x \Delta y}{\Delta x \Delta y \Delta z} = -\rho_l g \frac{\partial \Phi}{\partial z} \quad \text{as } \Delta z \to 0
\]

where \( \partial \Phi / \partial z \) is the vertical component of the hydraulic head gradient.

The vertical seepage force is resisted by the immersed weight of the grains, aggregates, or larger quantity of rock or soil in the control volume:

\[
\frac{\text{Force}}{\text{Volume}} = (\rho_l - \rho_i) g (1 - \rho)
\]

where \( \rho \) is the porosity. If the force due to the upward vertical component of the seepage exceeds this immersed weight, the particles will be buoyed until the effective stress, and therefore
effective strength, decline to zero: the material will undergo static liquefaction. Thus, the critical condition required for liquefaction is:

\[ -\rho g \frac{d\phi}{dz} \geq (\phi_0 - \phi) g (1 - p) \]

or, if the magnitude of the hydraulic head gradient is represented by \( i_z \) which has a vertical component, \( i_z \), then

\[ i_z \geq \frac{(\phi_0 - \phi)}{\rho g} (1 - p) \]

(7)

In turn, the magnitude of \( i \) is related to the seepage discharge by

\[ i = \frac{Q}{K} \]

(8)

If the flow field were three-dimensional, then \( i \) would be equal to \( |\nabla \phi| \) in Equation 2.

When the condition described in Equation 7 occurs, the material may dilate and increase its permeability rapidly over a relatively large area, or a few smaller grains may be lifted first, and outflow may converge on the resulting cavities, increasing the water flux and hydraulic gradient and extending the cavities to form ephemeral pipes (Källén, 1977). The material loses all of its frictional strength and deforms by flowing. Figures 1 and 6 show that vertical components of seepage and potential gradients may exist in valley floors and on footslopes. Sakura and others (1987) have measured equipotential fields and hydraulic gradients in laboratory model and showed that liquefaction and valley-head failure occurred when the normal component of the hydraulic gradient exceeds a critical value. During the snowmelt season in Vermont, I have observed highly turbid water seeping from valley floors, and have measured vertical components of pore-pressure gradients of the same magnitude as those predicted by Terzaghi's theory to be required for the observed occurrence of erosion by emergent groundwater. In these cases the dislodged grains were carried away by saturation overland flow and by channel flow, but if large masses of material suddenly lose their effective strength in this way a debris flow may result. In either case, liquefaction proceeds downward from the surface because it is there that the effective strength of the material declines to zero (Bear, 1972, p. 188).

If the geologic material in Figure 8 were cohesive, an extra force would be acting across its base to resist the vertical separation. Zaslavsky and Kassif (1965) pointed out that seepage erosion would then occur as a tensile failure as cracks parallel to the surface become pressurized with the fluid. The balance of forces in Equation 7 would then be altered to

\[ i_z \rho g \geq (\phi_0 - \phi) g (1 - p) + \frac{c}{\Delta x} \frac{\Delta y}{\Delta z} \]

(9)

where \( c \) is the cohesion per unit area. The critical hydraulic potential gradient thus depends on the thickness of the volume that eventually separates. Near the surface, \( \Delta z \) may be so small relative to \( c \) that the difference in hydraulic head (and therefore the seepage force) is insufficient for erosion. With increasing depth (\( \Delta z \) in Fig. 8), the volume being mobilized by the seepage force (per unit volume), \( i_z \rho g \), becomes so large that its movement cannot be resisted by the tensile strength of the material, which remains constant as \( \Delta z \) increases. Failure occurs at a critical depth

\[ \Delta z_c = \frac{c}{i_z \rho g - (\phi_0 - \phi) g (1 - p)} \]

(9a)

However, if \( c \) is large, realistic values of \( i \) lead to thick failed layers, which would be unlikely to flow. On the other hand, as \( \Delta z_c \) declines to the size of fragments that commonly flow, the seepage gradient would have to be unrealistically high to cause liquefaction of material with a significant cohesion. Thus, although Equation 9a describes failure, it seems likely that the liquefaction of cohesive materials requires that the cohesive bonds first be weakened considerably by weathering, gravity, or other forces near the seepage face, and the erosion process depends mainly on the condition summarized in Equation 7. The resistance of cohesive materials to seepage erosion is probably the reason for Parker's (1963, p. 107) observation that tunnel scour is the most common mode of conduit formation in consolidated materials.

Static liquefaction may also be impeded by plant roots, which can be viewed as resisting the vertical separation in the manner of a cohesion, or as providing an extra vertical load when roots tied laterally or vertically to stable ground are extended by dilution of the ground surface. This effect may be the reason why artesian pore pressures and swelling of the ground surface have been noticed before colluvium liquefied to produce debris flows in some forested bedrock hollows.

Iverson and Major (1986) generalized the analysis of seepage forces to account for static liquefaction and Coulomb failure due to water flowing in any direction within colluvium on hillslopes. Thus, they took into account the conditions that produce mass failure, debris-flow initiation, and seepage erosion, which grade into one another and extend channels. Their analysis was strictly limited to homogeneous, isotropic, and cohesionless soils subject to steady, uniform seepage on a planar hillside, but it illustrates important conditions that are relevant to interpreting most field situations. Iverson and Major analyzed the influence of seepage on liquefaction of a hillside defined in Figure 9a, where the seepage force is proportional to the hydraulic gradient, defined in terms of its magnitude \( i \) and the angle \( \lambda \), which it makes with the surface normal. Liquefaction occurs under the condition summarized in Equation 7, which requires that

\[ i = \frac{i_z}{\cos (\lambda + \theta)} > \frac{(\phi_0 - \phi)}{\rho g} \frac{(1 - p)}{\cos (\lambda + \theta)} \]

(10)

The magnitude \( i \) required for seepage erosion by liquefaction is therefore a minimum where \( \lambda = -\theta \), which implies vertically upward seepage. However, if \( i \) is sufficiently large, liquefaction can occur at other combinations of \( \lambda \) and \( \theta \), as indicated by Equation 10 and summarized graphically in the original paper.
Iverson and Major showed that the direction and magnitude of the seepage gradient also affect the susceptibility of a hillslope to Coulomb failure, and that, even when water is emerging from the surface, this form of mass failure often preempts liquefaction or occurs when the hydraulic head gradient is insufficient for liquefaction. They demonstrate that the flow of water through colluvium and the attendant head gradient alter both the effective stress, which controls the frictional resistance, and the alongslope driving force. If the flow vector is pointed outward and downward, for example, it reduces the frictional resistance and increases the downslope stress, whereas downslope flow parallel to the surface increases the driving stress but not the normal stress, yielding the usual “infinite-slope” stability equation (Carson and Kirkby, 1972, p. 156). Other combinations are obvious from Figure 9, or from Equation 3 of the original paper. Modification of the force balance of saturated colluvium on a hillslope to take into account the effects of seepage magnitude and direction led Iverson and Major to summarize the condition for Coulomb failure of a cohesionless material as:

$$i > \left( \frac{\sin(\phi - \theta)}{\sin(\lambda - \phi)} \right) \frac{(\rho_i - \rho)}{\rho_i}$$

(11)

where $\phi$ is the angle of internal friction of the porous medium. One particularly interesting result of this equation is that a specific (not vertical) seepage direction ($\lambda = 90 - \phi$) requires the minimum seepage magnitude to provoke failure, or yields the minimum stable hillslope angle for a given magnitude of $i$ and $Q$.

The original paper investigates all conceivable combinations of hydrologic and geotechnical parameters of the model.

Kochel and others (1985) and Howard and McLane (1988), on the other hand, emphasized the continuity between seepage erosion and fluvial transport. They constructed a two-dimensional sapping chamber (Fig. 10) with walls 5 cm apart. The intensity of groundwater flow through the sand hillslopes constructed in the tank could be altered by adjusting the water levels in the reservoir at the rear of the tank and at the outflow, as well as the position of the screen separating the sand from the reservoir, which sets the length of the flow system. The authors reported that sapping over the seepage zone involved an intimate mixture of seepage erosion and overland transport. Seepage erosion dominated at the upper end of the seepage zone where the surface slope was steep, hydraulic gradients in the emergent groundwater were strong, and the surface flow was weak. Removal of sediment from this zone triggered mass failure from the zone upslope. Further downstream, the hydraulic gradient decreased in magnitude, but the overland flow and resultant fluvial transport increased. Sapping was most intense at the upper end of the seepage face, causing headward erosion by slumping of sand, which then had to be transported downstream before the backwearing could continue. The fluvial transport required the development of a surface gradient proportional to the amount of sediment eroding from the model landscape, and inversely proportional to the water discharge and thus to the average hydraulic gradient through the hillslope.

Kochel and others (1985) and Howard and McLane (1988) analyzed the interaction of surface runoff and emergent groundwater by considering the balance of torques (Fig. 9b) acting on a particle that must roll out of an intergranular “pocket” over a rotation point. They considered four forces to be acting on the particle: gravity ($F_g$), cohesion ($F_c$), the drag imposed by surface runoff parallel to the slope ($F_w$), and the drag imposed by emerg-
Implications of erosion by subsurface flow

The forces are defined with respect to particle characteristics as:

\[ F_t = C_1 \left( \rho_s - \rho_l \right) g D^3 \]  \hfill (12)

\[ F_s = C_2 \rho_s g \frac{i}{D} \]  \hfill (13)

\[ F_n = C_3 \tau D^2 \]  \hfill (14)

\[ F_z = C_4 D^2 \]  \hfill (15)

where \( C_1, C_2, \) and \( C_3 \) represent the effects of particle shape and packing, and \( C_4 \) requires a specific measurement for cohesive soils. The boundary shear stress \( \tau \) exerted on the grains by surface flow and the hydraulic gradient \( i \) must also be measured or estimated independently from hydrologic considerations.

At incipient motion, the rotational point makes an angle, \( \phi \), with the normal to the hillslope surface about the center of the particle to be moved (Fig. 9b). The average value of this angle for the loosely packed, granular material is approximately equal to the angle of internal friction used in the formulations based on mass wasting (see above). The balance of torques as erosion begins is defined by:

\[ F_z D = F_w D \cos \phi + F_s D \cos (\theta - \phi) - F_t D \sin (\phi - \theta) \]  \hfill (16)

The authors then showed that for cohesionless sediment with an angle \( \phi \) and negligible surface flow, the maximum stable slope for the flow depends once again on the direction \( \nu \) and magnitude of the hydraulic gradient. Thus, for cohesionless sediment with an angle \( \phi \) and negligible surface flow, failure occurs when:

\[ i > \frac{C_1 \left( \rho_s - \rho_l \right) \sin (\phi - \theta)}{C_2 \rho_s \cos (\phi + \theta - \phi)} \]  \hfill (17)

The magnitude of \( i \) required for such failure is minimized when \( \nu = (\phi - \theta) \), which is equivalent to the condition in Equation (11) because, between Figures 9a and b, such a condition implies that \( \lambda = 90^\circ \), which also minimizes \( i \). The results of Kochel and others (1985) agree with those of Iverson and Major (1986) to the effect that a specific, nonvertical seepage vector will provide the critical condition governing erosion and retreat of the hillslope or spring head. This seepage vector is governed, in turn, by the geometry of the groundwater flow net, surface topography, hydraulic conductivity, and recharge rate.

Morphological expression of seepage erosion

Introduction. Other chapters in this volume describe the characteristics of landforms generated by seepage erosion. However, there are some generalizations about morphology that follow from the principles of groundwater flow and seepage erosion mechanics reviewed in this chapter. These generalizations will be outlined here. Many recent studies have been motivated by a search for terrestrial analogs of Martian landscapes where, it is hypothesized, groundwater has been an important landforming agent. Thus, many investigators have focused their attention on terrestrial landscapes where the geomorphic role of groundwater is relatively free from complications by other influences. The field studies have concentrated on gently dipping aquifers in highly permeable or arid zones where surface erosion and shallow mass wasting are relatively slow, and where there is relatively strong structural control of landform evolution by major vertical joints, bedding planes, and strong contrasts in hydraulic conductivity. The landscapes studied include beaches (Higens, 1982), the Colorado Plateaus (Laity and Malin, 1985; Howard, 1986a), Hawaiian volcanoes (Baker, 1986; Kochel, 1986), and the Libyan desert (Maxwell, 1982).

It is likely that the geomorphic effectiveness of groundwater is also widespread outside of these "model" landscapes, but that in other regions groundwater interacts with the effects of hillslope erosion, or is not quite so strongly controlled by structural features. Thus, some landscape characteristics interpreted by Laity and Malin (1985), Howard (1986a), and others to be diagnostic of the landforming role of groundwater may simply be the effects of groundwater in particular geologic settings, such as gently dipping, massive, permeable rocks with little weathered residuum. These characteristics include: elongate channel networks; theater-headed valleys; short, stubby tributaries with large junction angles; relatively constant valley width from source to mouth; and strong structural control of channel orientation, wall morphology, and longitudinal profiles. They may be found only in landscapes formed by emerging groundwater in certain geologic settings, but seepage erosion may also develop channel networks without these characteristics in other geologic regions. For example, Dunne (1970, 1980) proposed that less distinctive Appalachian channel networks, with only slight structural control (Newell, 1970), were also formed as a result of groundwater emergence that promoted seepage erosion of weathered rock and colluvium. DeVries (1976) proposed that the stream network of the Netherlands was formed as a result of the outcropping of regional water tables as groundwater invaded Pleistocene deposits after deglaciation. Carlson (1963, p. C5) proposed a similar reason for the occurrence of streams. In these cases the geologic structure, climate, and hillslope erosion processes favored the development of a landscape without some of the diagnostic characteristics listed above.

Channel initiation and network growth. If a groundwater flow field, of any extent or depth, is subject to a perturbation as the water emerges from the ground surface, there will be a tendency to focus weathering and erosion, and to notch the surface, creating the potential for elongation of the notch into a channel and eventually a valley. The perturbation may result from erosion of the boundary by some other process, such as river erosion (Fig. 11a), or it may result from the establishment of a new or larger groundwater flow field within a geologic deposit with irregular boundaries formed by some other agent, such as a glacier, a river, or wind. DeVries (1976) described, for example, the expansion of groundwater seepage zones in Pleistocene fluvial
and eolian sediments as precipitation increased during the Holocene. The flow lines converge on the re-entrants in the boundary as shown in Figures 1 and 11a. Convergence of flow on the initial or eroded, concave boundary increases the water flux, \( Q \) in Equation 8, and therefore the hydraulic gradient and seepage forces in Equations 7, 10, 11, and 17. Thus, the probability of attaining the critical conditions for seepage erosion during a particular storm or season is increased by disturbance of the boundary.

Dunne (1970, 1980) also proposed that spatial variation in the hydrogeological properties of the aquifer could cause either local increases of specific discharge, \( Q \), or local reductions in resistance to weathering and seepage erosion. For example, there are autocorrelated spatial variations of both hydraulic conductivity and porosity in even homogeneous isotropic aquifers, as previously described. These variations result from random and periodic differences in the factors controlling petrogenesis, diagenesis, and deformation, which cause variations in grain size, pore fillings, and in fracture widths and spacing. Other spatial variations in mineralogic composition result in variable resistance to chemical decomposition and therefore in changes of porosity, hydraulic conductivity, and tensile strength.

Each of the spatial variations in hydrogeologic properties affects the flow field, causing local concentrations of discharge at a seepage face. The most obvious examples of this effect occur where master joints provide narrow zones of high and anisotropic conductivity, which localize outflow and generate major springs (Schick, 1965; Campbell, 1973; and Fig. 12), but in many other field situations it is difficult to identify visually the hydrological parameter responsible for the concentration, and as far as I know, detailed instrumental measurements that might identify the effective spatial patterns have not been made. On a larger scale, however, Ineson (1963) identified small faults and fracture zones along the axes of gentle anticlines as the primary controls on the distribution of transmissivity in the English Chalk. The local concentration of outflow either accelerates weathering, lowering the resistance to seepage erosion, or increases the hydraulic gradient, increasing the seepage force (Equations 7, 10, 11, and 17). If seepage erosion occurs, it perturbs the boundary, and again increases the flow convergence (Fig. 11b), increasing the probability of further seepage erosion. As the head of the embayment retreats into the land surface or hillslope, this flow convergence will increase (Fig. 11c), setting up a positive feedback loop that affects the growth of a new valley.
The initial seepage erosion may produce only a subterranean conduit rather than a valley. Howard (1986b, p. 134) observed the formation of temporary conduits in his experiments on seepage erosion, presumably as a result of cohesion imposed by intergranular water in the tension-saturated zone. It is difficult to point unequivocally to published reports of conduits formed by seepage erosion alone, because most reports do not consider the pipe-forming process in detail, and perhaps because cohesion resists seepage erosion, as indicated by Equation 9. However, it is likely that in some cases the heads of conduits are extended by seepage erosion, even if their roofs are maintained by cohesion and they are eroded by tunnel scour along most of their courses. Zaslavsky and Kassif (1965, p. 315-316) emphasized that survival of an arch over a pipe or tunnel requires cohesive strength, which presumably could be provided by both interparticle forces and plant roots. However, over long periods of geologic time, as the land surface is lowered above them, the roofs of conduits are ephemeral and they collapse to form gullies and valleys. In many situations, pipes do not form: the channel or valley forms as a direct result of seepage erosion and mass wasting into the sapped zone.

It is a common observation that spring heads or valleys formed by seepage erosion at a range of scales down to the size of laboratory models (e.g., Kochel and others, 1985; Howard, 1986b) exhibit a roughly uniform spacing. In some cases (e.g., the valleys studied by Campbell [1973], Howard [1986a], and Layte and Malin [1985] on the Colorado Plateaus or some of the large springs draining basalts on the Columbia and Snake River Plateaus), it is obvious that the uniformity is imposed by the roughly equal spacing of master joints. However, there have been no formal quantitative studies of the relation between the hydrogeologic properties, the length scales of patterns in these properties, and the average spacing of channels formed by seepage.

The most general influence on the relatively uniform spacing of channels, occurring even in the absence of structural controls, is the competition of growing channel networks for subsurface flow. Figure 11c indicates that convergence of flow lines toward indentations or alcoves in a boundary or at the heads of channels is compensated by divergence between these features. Divergence of the flow lines causes a decrease in the local water flux, and therefore a lower hydraulic gradient and probability of erosion, stabilizing the surface between the indentations. The growing indentations along the boundary compete with one another in draining groundwater from the aquifer, and this competition sets the drainage density of the consequent streams in a manner that has not yet been derived mathematically. However, on the basis of the foregoing discussion of mechanics it seems reasonable to speculate that high recharge rates, low hydraulic conductivity, high porosity, and susceptibility to weathering would all tend to increase drainage density, as would the presence of a steeply sloping aquiclude beneath the aquifer (see below).

The extending valley eventually intersects another zone where the combination of hydrogeologic characteristics causes a sufficiently strong concentration of discharge to cause seepage erosion on the valley wall (Fig. 11c). Alternatively, the consequent stream draining the valley may undercut the valley wall, creating an embayment (Figure 11a). Either perturbation causes a headward-sapping tributary to form, along which the process can be repeated to generate a branched hierarchical drainage network (Fig. 13).

If neither the flow field nor distribution of rock properties is affected by strong heterogeneities or anisotropy, one might expect the junction angle to depend on the ratio of the general gradient of the water table (along x-x' in Fig. 11c) to the lateral perturbation of the water-table slope imposed by the formation of the consequent stream, assuming that the resultant of these two components is sufficiently steep to trigger seepage erosion. Both of these components of the local hydraulic gradient depend on the recharge rate and hydraulic conductivity, and the lateral component depends also on the spacing and time-varying depth of inci-

![Diagram](image_url)
sion of the consequent valleys. If all of these variables were held constant except for the general gradient of the water table along $x=x'$, and if the resultant local gradient were sufficiently steep, the junction angles would become more acute as the gradient increased. Such a high gradient could be imposed, for example, by a sloping aquiclude beneath the formation being sapped. On the other hand, various degrees of preferred orientation can be imposed on junction angles, channel reaches, and valley bends if either the flow field or the distribution of weaker zones is controlled by rock structure. Schick (1965), Campbell (1973), and Laity and Malin (1985, p. 211–216) have documented examples of the control of stream and valley orientations by rock fractures. Figure 14 shows some examples of interpretations by Laity and Malin (1985). Figure 12 shows another example.

The possibilities for structurally controlled valley networks range from rectangular, trellised patterns to dendritic patterns, in which the influence of structure can only be recognized through systematic map and field measurements. For example, Newell (1970) conducted a series of joint surveys in calcareous schist in the region that includes the drainage basin in Figure 12. He plotted poles to joint surfaces on an equal-area graticule and contoured the results to define sets of high-angle joints. He then measured hillslope gradients and aspects and found a rough parallelism between the preferred orientation of the stream valleys orthogonal to the slopes and that of the fracture pattern. Newell (1970, p. 111) interpreted the parallelism as being due to the concentration of groundwater flow and chemical weathering.

Ideally, such flow in bedrock would proceed with greatest ease along the set of joint planes most nearly normal to the force equipotential surface of the water table. Under some preexisting land surface, leaching along such preferred zones would gradually lower the bottoms of valleys and reorient them with respect to the trend of the joints being followed. The resulting cover of weathered residuum provided the course of least resistance to headward eroding streams. Proceeding from the initial topography, low areas were lowered more rapidly than the contiguous slopes and divides. Under these differential rates a fabric of valleys and relict interfluves developed which reflect the structural permeability of the underlying bedrock.

Although it is not clear that leaching alone would lower the surface, rather than producing a cohesionless residuum with low porosity, chemical weathering would reduce the resistance of the rock, and thus Newell's regional interpretation agrees well with the interpreted role of groundwater based on the instrumental observations of pore pressures, seepage erosion, and local topography in the small basin shown in Figure 12.

Whether or not their orientations are controlled by rock structure, the headward growth of channels progressively disrupts the two-dimensional groundwater field shown in Figure 11, producing a network of valleys (Fig. 13). As the number of spring heads increases, each spring drains groundwater from a smaller area. The spring or valley heads compete for groundwater, and if one is at a lower elevation than its neighbors, it can grow faster and leave them dry as a result of subsurface piracy. Reduction of the drainage area diminishes the groundwater flux and therefore the hydraulic gradient (Equation 8). The frequency of groundwater recharge events large enough to cause seepage erosion of weathered rock will decrease as the drainage area diminishes, and eventually only weak colluvium traveling to the valley head will be removed.

This limiting condition for channel extension by seepage erosion of a weathered residuum can be illustrated approximately by generalizing the Dupuit-Forschheimer analysis in Equation 4 to the case of steady-state seepage from a conical drainage area in colluvium mantled bedrock converging on the head of a channel (Fig. 15). The limitations of this analysis were discussed previously, but again it includes the essential physics of the way in which the drainage area, recharge rate, and hydraulic conductivity control the hydraulic head gradient at the seepage face.

In some increment of radial distance, $dr$, the drainage area converging on the 1-m strip of the perimeter of a channel head generates an increase in discharge

$$dq = I \cdot r \, dr$$

where $r$ is the radial distance and $I$ is the convergence angle of the drainage area (in radians). Integrating this differential equation, and taking into account the boundary condition that $q = 0$ at $r = r_D$, yields the specific discharge at any distance, $r$:

$$q_r = -\frac{I \cdot r}{2} \left( r_D^3 - r^3 \right)$$

where the first negative sign on the right-hand side implies that
Implications of erosion by subsurface flow

Figure 14. Relations between geologic structure, valley morphology, and network pattern in the Colorado Plateau. Pattern and form depend primarily on the direction of groundwater flow in the Navajo Sandstone relative to the orientation of the valley (assuming a more or less orthogonal joint set). A and B are developed on a monocline: A has a strongly asymmetric network and B has a stubby, branched, symmetric network. C has a more elongate, symmetrical network because the gently plunging syncline focuses more groundwater on the heads of the tributaries. In D, groundwater flows away from valley walls and streamflow is generated by overland flow, which during its convergence on the channel erodes tapered valley heads. (Source: Laity and Malin, 1985.)

discharge increases in the direction of decreasing \( r \). The Dupuit-Forchheimer approximation yields

\[
q_x = -K \epsilon \left( \frac{r}{r^2} \right) \frac{dx}{dt}
\]

where, in Figure 15

\[
\zeta = z - r \tan \beta.
\]

Combination of the three previous equations yields the water-table slope:

\[
\frac{dx}{dt} = \frac{1}{2K} \left[ \frac{1}{r(t - r \tan \beta)} \right] - \frac{1}{2K} \left[ \frac{r}{t - r \tan \beta} \right], 
\]

which does not have a simple solution, except for the case where \( \beta = 0 \):

\[
r^2 - r_t^2 = \frac{1}{K} \left[ r_t^2 \ln \left( \frac{r_t}{r_1} \right) - \frac{(r_t - r_1)}{2} \right].
\]

Since

\[
r_t = \frac{1}{r_1},
\]

the steady-state specific discharge at the channel head is

\[
q_x = -\frac{I_0}{2} \left( r_t^2 - r_1^2 \right).
\]

This discharge entering the stream channel would determine the depth of streamflow, and therefore the height of the seepage face, \( z_s \), in Equation 19. For example, if the 1-m perimeter of the channel head has a drainage area with a convergence angle of 30°, the channel width (\( b \)) would be 0.27 m (Fig. 16). If the gradient of this channel (\( s \)) is 0.02 and its Manning roughness (\( n \)) is 0.05, the depth of flow, and therefore \( z_s \) at steady-state outflow, would be

\[
\frac{z_s}{b} = \left[ \frac{q_s}{b \cdot s_0} \right] = 1.18 \left[ \frac{q_s}{b \cdot s_0} \right].
\]

where \( q_s \) is given by Equation 20 with \( \epsilon = \pi / 6 \). Substitution of Equation 21 into Equation 18 for \( r = r_s \) yields an approximate expression for the hydraulic head gradient at the seepage face:

\[
i = \frac{dx}{dr} \bigg|_{r=r_s} = \frac{1}{2K} \left[ r_s \left( \frac{1}{r_s} \right) \left( \frac{1}{t - r_s \tan \beta} \right) - \frac{1}{2K} \left[ \frac{r_s}{t - r_s \tan \beta} \right] \right].
\]

This gradient is equivalent to \( i \) in Equations 11, and 17, and
Figure 15. Steady-state water table under uniform recharge on a conical drainage area with a slope \( \beta \) and convergence angle \( \epsilon \) using the Dupuit-Forchheimer approximation of horizontal flow lines.

Therefore determines whether seepage erosion by Coulomb failure will occur. Equation 22 cannot predict liquefaction directly because the Dupuit-Forchheimer assumption precludes upward flow components, but it still indicates which factors govern the magnitude one would expect for those upward components if they were to be defined by graphical or numerical methods. Conversely, Equation 22 and its subsidiary equations summarize how the drainage area (represented by \( \epsilon \) and \( r_p \)) must diminish to prevent seepage erosion for a given set of the other parameters in the hydrologic and erosion equations. The equation therefore summarizes the controls on the drainage area required for channel initiation, which in turn controls the drainage density, the governing factors of which will be discussed in the next section. Equation 22 is used only as a summary of the interactions that govern drainage density. It cannot be used quantitatively because of the limitations of the Dupuit-Forchheimer model, and because the equation does not include such potential influences as changes in hydraulic conductivity due to weathering as the seepage face is approached.

If recharge events competent to cause seepage erosion of even colluvium occur only infrequently, the head of the valley may be partially filled and evolve into an unchanneled swale, floored by a zone subject to saturation overland flow or, if the colluvium is deep enough, by a well-drained slope generating only subsurface stormflow. This accumulation can then be scoured from the bedrock and a channel reestablished during rare seepage erosion (Dietrich and others, 1987, Fig. 4). Thus, the drainage density may vary over time, but the long-term valley head is localized by seepage erosion in the weathered bedrock. If the upper end of the valley is sufficiently steep, the seepage erosion will occur in the form of landsliding, and the headward sapping of the channel grades into the repeated failure and refilling of funnel-shaped bedrock hollows, as described by Dietrich and Dunne (1978). Dietrich and others (1986) have shown how the stable drainage density of channels forming at the downslope end of long, trough-shaped hollows in California can be rationalized in terms of the role of groundwater seepage in landslide mechanics.

Little work has yet been devoted to mathematical modeling of the evolution of valley networks by seepage erosion. Howard (1986b) has made the most progress in this direction. He developed a finite-difference model of steady two-dimensional flow through a porous medium with a spatially random hydraulic conductivity. As flow emerged from a scarp face, the local rate of erosion was assumed to be proportional to the specific discharge rate, \( q \), minus a critical discharge rate. The erosion altered the boundary and therefore the flow field, and the computation was repeated for the next time step. The model simulates the evolution of valley networks and groundwater flow fields, although the mechanics of seepage erosion are not represented explicitly. It illustrates that even where aquifer properties are randomly distributed, a roughly uniform spacing is imposed on the valley networks by competition for groundwater flow between the lengthening valleys.

**Drainage density.** Until a rigorous mathematical model is developed for seepage erosion resulting from two- (or three-) dimensional, transient groundwater flow, it will not be possible to specify in an exact, quantitative manner the controls on the drainage density of channels formed by this process. However, by reviewing the components of Equation 22 in conjunction with those of Equations 7, 10, 11, and 17 describing seepage erosion, it

**Circular arc**

![Diagram of a circular arc with a drainage area and channel](image)

**Drainage area**

**Channel**

Figure 16. Drainage of a conical swale to a channel head with an upstream perimeter of 1 m and a width of 2 \( \epsilon \) sin (\( \epsilon/2 \)) m.
is possible to suggest the most probable controlling factors, even if the exact functional relations cannot yet be defined. The drainage density, which is related in a rather complex way to the drainage area required for channel initiation, is represented in Equation 22 by \( r_D \), \( r_S \), and by \( c \), implicit in \( z_b \).

The most obvious control involves the amount of water (\( J \)) available for seepage erosion. If the rock is significantly weakened by weathering before seepage erosion occurs, the total annual runoff may be an important control through its influence on rock decomposition (seepage weathering). However, the removal of the weathered residuum by seepage erosion is likely to occur during some shorter period of high discharge and hydraulic gradient. Thus, total wet-season rainfall or spring snowmelt might be the critical variable in large groundwater flow systems. The deeper and more extensive the flow system controlling the stream network density, the more sustained would the water input have to be to generate seepage erosion at the outlet. Shallower and smaller flow systems develop critical flow rates and hydraulic gradients in shorter, more intense rainstorms. In Figure 12, the bedrock flow system feeding the springs at A, C, and D and upwelling beneath the channel at F generated maximum pore pressures during the spring snowmelt and sustained them for several weeks; in the shallower flow systems in soils at sites B and E, peak hydraulic gradients directed into the slope occurred in response to daily snowmelt cycles and large rainstorms (Dunne, 1970).

It is difficult to isolate which hydrologic input is the dominant control of channel formation and drainage density by statistical analysis of morphometric data because of the high degree of correlation among various measures of rainfall intensity over different time scales at one location. The relevant measure probably varies with the geometry and hydrogeological parameters (mainly hydraulic conductivity and porosity) of the flow system, since these are the factors affecting the transient response of groundwater flow (Freeze and Cherry, 1979, p. 65). The fact that drainage density increases with some measure of precipitation in regions where runoff is dominated by subsurface flow has been established by Chorley (1957), Chorley and Morgan (1962), Cotton (1964), and DeVries (1976).

The intensity and form of weathering and its effect on resistance to seepage erosion (\( c, \phi, p, \) and \( \rho_s \) in Equations 7, 9, and 10) must also affect channel formation and drainage density. These processes are controlled in turn by climate, hydrology, and lithology, as discussed in Chapter 1. The hydrogeologic characteristics of the aquifer are important in setting the drainage density. For a fixed rate of groundwater recharge (\( J \)) per unit area of catchment, the hydraulic gradient (\( i \)) at a seepage face intensifies as the hydraulic conductivity (\( K \)) declines (Equation 22), decreasing the critical discharge and therefore drainage area necessary for seepage erosion. The effect of spatial variations in hydraulic conductivity was also emphasized earlier as a factor controlling gradients of hydraulic head (Rulon and others, 1985).

Finally, relief and gradient should be important controls on drainage density. In a homogeneous isotropic aquifer the water table is a subdued replica of the ground surface (Fig. 1), and strong local relief causes high hydraulic gradients between channels and the intervening drainage divides (Freeze and Cherry, 1979, p. 193–199). Also, a sloping aquiclue beneath an aquifer steepens the hydraulic gradient for a fixed recharge rate and conductivity (Equations 18 and 22). Equations 10, 11, and 17 indicate that the critical hydraulic gradient required for seepage erosion also decreases as the hillslope angle increases. But this angle is not, in general, independent of the other controls discussed here. However, it is possible to conclude that both relief and the dip of the aquiclue should be positively correlated with drainage density. DeVries (1976) pointed out that on Pleistocene sandy deposits in the Netherlands, where he interprets the stream networks to result from the outcropping of groundwater, drainage density is inversely proportional to relief. This result does not conflict with the assertions made above because DeVries does not call on seepage erosion to form the channels and valleys. Although there is no explicit description of process mechanics in his paper, he seems to interpret the stream net as resulting from surface sediment transport (by \( F_W \) in Equation 16) wherever the recharge rate, hydraulic conductivity, and initial surface geometry have caused the water table to outcrop. For a fixed recharge rate and hydraulic conductivity, low relief and gentle gradients would increase the area of the seepage zone and the number of depressions intersected by the water table.

Channel networks or scarp retreat. Another problem relates to whether concentrated seepage erosion at a spring head will produce a valley or simply sap the hillside along a broad, irregular front to produce an escarpment. As channels or valleys are eroded headward, as shown in Figure 11c, the intervening, unchanneled surface may also retreat as a result of mass failure or undercutting by waves or streams. Up to this point it has been assumed that the former retreat rate far exceeds the latter, so that a channel network forms. This result is more likely where there are strong lateral gradients in hydraulic conductivity, porosity, or susceptibility to weathering, and where there is some dominant wavelength to the spatial pattern (such as regularly spaced master joints) so that flow concentrations are strong. Concentration of erosion at the valley heads is also favored if the structural and lithologic properties of the rock make it susceptible to a radical increase in the frequency and magnitude of sapping in response to the increase in discharge resulting from flow concentration.

However, if the spurs between valleys are aggressively undermined by lateral stream erosion, wave action, or mass wasting, the spurs and valley heads could retreat at approximately the same rate. The land surface then does not become as crenulated as shown in Figure 11c or Figure 13; retreating escarpments develop instead. For example, if the rock is prone to rapid mass failure in the presence of only a small amount of exfiltration over some mechanically weak layer, then cliff retreat may be rapid even where the outflow is diffuse. In such a case one would expect valleys to be wide. Laita and Malin (1985) mapped several examples of this condition (e.g., Fig. 17) and pointed out that the rate of lateral weathering approaches that of headward retreat as
the drainage area of the spring head declines (p. 210). Some sequences of geologic materials are so prone to mass failure where groundwater emerges that a single retreating escarpment cannot be maintained; the seepage zone becomes a chaotic terrain of blocks and connected depressions.

It should also be remembered that there is a set of diffusive erosion processes that tend to fill and erode the pipes, channels, and valleys produced by seepage erosion at a range of scales. For example, soil pipes formed by seepage erosion are filled or disrupted by soil falls from the perimeter of the conduit and by bioturbation due to root growth, tree-throw, and animal burrowing. The conduit survives only if the fallen soil is evacuated by the outflowing water and if diffusion of soil toward the pipe is less than or equal to the volume evacuated. The more probable condition over time scales exceeding decades is that collapse and evacuation prevail and the roof eventually collapses, or that disruptive processes such as tree-throw fill the pipes and new ones develop in the vicinity.

On the scales of channel or valley networks, the growth of the network causes the expansion of eroding hillslopes and therefore an increased flux of sediment (Fig. 18a), which the stream must evacuate if it is to sustain the new valley head. The resulting balance between sediment supply from the eroding hillslopes and the sediment transport capacity of the runoff produced by seepage would stabilize the valley head (Kirkby, 1986). Although I do not know of any documented examples, it is possible in principle that an increased sediment supply from the scattered land margin could cause the accumulation of coarse sediment in the lengthening valley floor (Fig. 18b). If sufficiently thick, this rampart of sediment could eventually load the seepage face, preventing seepage erosion (it would increase $F_2$ in Equation 16), and convey the groundwater emerging from the bedrock downstream without causing erosion. Such a situation is more probable where the developing hillslopes are shedding coarse colluvium.

**Forms of valley heads.** Many investigators (e.g., Laitt and Malin, 1985; Baker, 1986; Howard, 1986a) point to the steep, blunt, amphitheater–head valleys commonly produced by sapping in areas where the effects of seepage erosion have been studied.
Kochel and Piper (1986) propose that detailed morphometric analyses could successfully resolve valleys dominated by sapping from those dominated by "runoff processes" (presumably connected with Horton overland flow). They go on to emphasize both in their laboratory simulations and in terrestrial and Martian valleys the presence of amphitheater heads and alcoves. Sakura and others (1987) also found that the valley head was an amphitheater if only exfiltration occurred and more gradual when erosion by surface flow exceeded the effects of seepage erosion. Thus, sapping by seepage erosion and related mass wasting can produce amphitheater-headed valleys, although Laity and Pieri (1986) have challenged the value of these forms as diagnostic criteria of groundwater sapping on Mars. However, there are many other valley systems in the world that seem to have been generated by seepage erosion in the absence of Horton overland flow, but which have tapered valley heads with gentler gradients than those emphasized in the literature.

The form of the hillslopes bounding and defining a valley head depends on either: (a) stability with respect to mass failure; or (b) the balance between erosion and transport processes tending to undermine and steepen the slope and those tending to fill any incision such as that caused by sapping. For example, if seepage erosion removes the basal support from beneath jointed, gently dipping rock (as Laity, 1983, demonstrated on the Colorado Plateau), the hillslope erodes by slab failure and toppling, and stands at metastable angles close to vertical. Even unconsolidated sediments with low cohesion can stand temporarily at steep angles (Lohnes and Handy, 1968; Kochel and others, 1985). However, if the jointing of the rock, or other factors affecting its cohesion and internal friction cause collapse to lower slope angles, or if soil creep, talus accumulation, and debris-flow deposition spread colluvium toward the center of the valley, a more conical, gently sloping valley head results. The valley head can take on an even gentler slope if the water table and the zone of most intense seepage erosion outcrop at different points along the valley in different years. A range of conditions is summarized in Figure 19.

If seepage erosion is concentrated along the base of the valley head, but has achieved a stable drainage density by transporting colluvium at a rate that can be equaled by weathering and soil creep, raveling, debris flow, or other processes of hillslope retreat, the form of the valley head depends on this balance of sediment transport. The result may be a conical or trough-shaped bedrock hollow from which colluvium is periodically evacuated by debris flows (Dietrich and others, 1986) or a gently sloping swale into which soil is transported by creep between rare catastrophic scouring resulting from seepage erosion. Where the role of seepage erosion is to cause a slight increase in the competence and capacity of saturation overland flow or channel flow (Equation 16, Fig. 10), the gradient at the valley head may be very slight. The shapes of these valley heads grade into the runoff-dominated forms referred to in the literature.

It seems appropriate to conclude that valley heads such as those shown in Figures 12 and 19c are just as typical of the results

Figure 19. Forms of valley heads produced by seepage erosion: (a) amphitheater-head valley produced by intense seepage erosion in rocks with a high angle of internal friction with the fallen blocks removed between failures; (b) gradually converging valley head eroded by less intense seepage in rocks with a low angle of internal friction or in rocks susceptible to rapid weathering and mass wasting; (c) gradually converging valley head in which the locus of seepage erosion moves up and down valley during fluctuations of climate. Dashed lines represent ephemeral channels and different water-table positions during wet and dry periods. Soil creep and debris flows cause the accumulation of long foottongues of colluvium at the valley head where the water table is low, and extension of seepage erosion causes later evacuation of sediment from swales.
of spring sapping by seepage erosion as are the amphitheater heads and alcoves occurring in the Colorado Plateau and Hawaii (Fig. 19a). These latter characteristics, along with some others such as low drainage density and stubby tributaries, are diagnostic only of a more restricted set of conditions, including some combination of gently dipping, vertically jointed, highly permeable rocks overlying single, well-defined aquitards, or and regions, or valleys in which the rate of incision and headward retreat are so high that the valley walls have not yet adjusted to the sudden relief generation, or some other condition that allows little colluvium to accumulate as gently slope footslopes. The geomorphic effectiveness of seepage erosion is more widespread than the presence of the supposedly diagnostic forms popularized by recent literature.

**Mechanics of tunnel scour**

Water may enter a hole or crack from some saturated zone within or at the surface of a soil or sediment (Fig. 7). The initial conduit may have a biogenic origin (for examples see Fletcher and others, 1954; Jones, 1981; Beven and Germann, 1982), or be due to tensional cracking along the margins of gullies (Heede, 1971) and landslide blocks, or to shrinkage in certain clay-rich soils under a seasonal rainfall regime (e.g., Fletcher and others, 1954; Jones, 1971; Crouch, 1976), or to seepage erosion, as described above. Many authors have documented evidence of subsurface flow along cracks and tunnels in sodium-rich soils with a high shrink-swell potential occurring in regions with a strongly seasonal hydrologic regime (Parker, 1963; Heede, 1971; Crouch, 1976; Bryan and others, 1978; Imeson, 1986). These soil conditions provide not only the initial subsurface passageways for runoff and sediment transport, but also the low hydraulic conductivity that promotes saturation and enhances flow in the passages.

Thus, factors that promote surface runoff or the development of saturation at shallow depth in the presence of even initially narrow cracks, rootholes, or burrows will increase flow in these conduits. For example, Brown (1962) emphasized the frequency of reports associating the erosion of tunnels with intense rainfall and with soil packing and devegetation resulting from heavy grazing. On the other hand, if the ped between cracks have a dense vegetation cover and root mat, so that their infiltration and water-holding capacities are higher than the intensities and volumes of most rainstorms in the region, the ped will remain unsaturated, and little or no runoff will enter the inter-pedal cracks or animal burrows.

If surface runoff falls into the conduit, it shears soil at the base in a plunge pool. As the water flows along the conduit, a shear stress is applied to the margins, which may thus be eroded. The intricate geometry of natural conduits precludes a useful computation of the magnitude of this fluid shear. However, it is possible to say in a general way, on the basis of the derivation of the Chézy pipe-flow formula, that the eroding stress is proportional to the square of the average flow velocity, and therefore to the head gradient, the diameter or width of the conduit, and its hydraulic roughness, which in turn depends on the various scales of roughness elements along its path, ranging from grains and aggregates to the tortuosity of the entire conduit. The shear stress can be particularly high in tunnels on a steep hillslope or close to the bank of a stream or gully, where the head gradient is high (Parker, 1963; Heede, 1971).

If the shear stress is sufficiently high, grains are eroded from the margins of the conduit. Erosion of the dry boundary may initially be relatively rapid, but may diminish quickly in cohesive soils (Parthenaides, 1971), and material may soften and fall from the roof and walls as it becomes wet. The subsequent fate of the eroded grains depends on the velocity of flow, and therefore whether the conduit has an outlet from which water can drain freely. Outlets of various origin are shown in Figures 20 and 21. If the conduit is blocked, at the end of a root-hole for example, the eroded sediment will be deposited and the passage will be filled. If the passage is near a free face, such as a riverbank, or on a steep slope, the local increase of pore pressure around the conduit may cause seepage erosion (Fig. 20a).

An open connection downslope to the soil surface and sufficient head gradient can maintain a velocity high enough to transport the eroded sediment. Water can enter the passage from the surrounding matrix (if the latter is locally saturated), or from smaller tributary cracks, or from surface runoff. The outlet may be at a free face or on a gentler hillslope (Fig. 20b). It may intersect such a surface because of headward seepage erosion, or because the original conduit intersected the surface as a biogenic
Implications of erosion by subsurface flow

Figure 21. Conduits of various size intersecting the walls of a gully in thoroughly weathered volcanlastic sediments near Karatu, northern Tanzania.

passage or shrinkage crack. If the passageway began as a shrinkage crack, wetting of the soil surface usually closes the cracks late in the wet season, leaving a tunnel enlarged by erosion (Fig. 22). Seasonal repetition of this process can form networks of tunnels, bridged by an arch of cohesive soil. This cover eventually collapses, producing pits at various points along the tunnel, or continuous gullies with sharp, irregular boundaries (Rubeck, 1928; Löffler, 1977; Swanson, 1983; Fig. 21). However, if the subsurface flow exploits deeper cracks, joints, and bedding planes, the larger subterranean cavens referred to as pseudokarst may develop by the same process of shearing the margins of a conduit through otherwise unsaturated geologic materials (Parker and others, 1964; Carey, 1976).

The critical shear stress required to erode the margins of a tunnel is radically reduced if the geologic material contains a dispersive clay, although this is not a necessity for tunnel scour if other conditions are favorable. When wetted, some earth materials swell and eventually disperse due to a complex set of processes, including hydration and desiccation, fracturing, and osmosis (Holmgren and Flanagan, 1977). Many authors have noted the association between tunnel scour and high concentrations of salts (especially of sodium) in the soil (e.g., Heede, 1971; Crouch, 1976; Bryan and others, 1978; Stocking, 1981; Imeson, 1986). Heede (1971) found that in deeply cracked, montmorillonitic clays along gully banks, tunnel scour did not occur unless the exchangeable sodium percentage exceeded a value of about 1.0.

In the soils studied by Stocking (1981), this critical value was closer to 10, but with some values as low as 4.

Several laboratory studies of the erosive behavior of consolidated earth materials, reviewed by Arulanandan and Heinzen (1977) and Heinzen and Arulanandan (1977), have shown that the shear stress required to initiate erosion is affected by the amount and type of clay, and the concentration of solutes in the pore and eroding fluids. The osmotic potential gradient set up at the surface of the clay due to the concentration gradient between pore fluid and eroding fluid, together with the capillary gradient,

Early wet season

- Shrinkage crack
- Erosion of crack walls
- Fluid shear on bed and walls

Late wet season

- Crack closed by swelling
- Collapsing of wall and roof
- Tunnel

Figure 22. Enlargement of a shrinkage crack by tunnel erosion.
cause flow into the porous medium. Dilute water traveling along conduits is drawn into pores and cracks, causing swelling, a reduction of interparticle bonding forces, fracturing into aggregates, and dispersion of particules. The degree of swelling and dispersion depends on the sodium-absorption ratio and the solute concentration in the pore fluid (Crouch, 1976). As the concentration increases, the critical sodium-absorption ratio required for dispersion increases. The physical chemistry of these effects is described by Shainberg and Letey (1984). Clay mineralogy and soil pH also affect the degree of dispersion, while less soluble cements such as calcium carbonate suppress it. Crouch (1976) summarized studies on the dispersive effects of organic matter and concluded that it usually stabilizes soil structure and reduces the degree of dispersion, but that organic compounds of low molecular weight may exacerbate dispersion in some soils.

Heinzen and Arulanandan (1977) demonstrated that the critical tractive stress for a clay decreases as the solute concentration of the eroding fluid declines relative to that in the pore water. A true dispersive clay is one for which the critical tractive stress equals zero. All that is then required for erosion is a flow velocity sufficient to transport the resulting colloidal suspension. Dispersion can also reduce the critical pore-pressure gradient required for seepage erosion by reducing the cohesion in Equation 9. For example, if dispersible silty sediments, with a sufficiently high conductivity to allow significant saturated outflow, become connected with a source of dilute groundwater, dispersive seepage erosion can be particularly rapid. However, more typical is the experience of Sherard and Decker (1977, p. 5) that most dispersive clays are so impervious that little matrix flow and seepage erosion occur, and in order to initiate erosion it is necessary to provide a path for concentrated leakage. The low conductivity then forces flow from the surface or a shallow permeable horizon into cracks that are rapidly enlarged by dispersion-accelerated shear. Many of the field situations described in the papers edited by Bryan and Vair (1982) appear to be of this type. Imeson and others (1982, p. 68) provide a detailed description of one example.

Stocking (1981) presents a particularly interesting interpretation of results from Zimbabwe, where he suggests that high exchangeable sodium concentrations develop in soils where sodium ions released by feldspar weathering accumulate on clay minerals in a shallow, seasonally saturated soil horizon. The defloccinated clays block soil pores and vertical drainage. However, there is groundwater flow in the now-confined aquifer beneath the impeding layer. If gully incises the confined aquifer, headward erosion begins at the gully wall (presumably by either seepage erosion or tunnel scour) and proceeds beneath the impeding layer to form a tunnel with a low gradient. Once the tunnel is formed, the impeding layer forming the roof is breached, possibly by cracking due to gravitational settling, or because the formation of the tunnel increases the vertical head gradient and therefore the vertical flow through pores or cracks. This vertical flow then erodes vertical tunnels tributary to the horizontal tunnel, weakening the roof and eventually causing its demise.

The presence of tunnels in a porous earth material increases the efficiency with which water and sediment can be evacuated from the landscape. Flow within a crack or a tunnel is faster than percolation through the porous matrix, and the presence of steep or overhanging faces of bare, often dispersive material without the binding action of roots provides conditions that favor rapid erosion and sediment transport. The complexity of interaction between the fluid shear (with or without dispersion) and of gravitational collapse of the tunnel margins often causes rapid fluctuations in sediment concentration in the outflow from tunnels. Temporary blockage of pipes, or their joining, can also cause pulsations in flow. Yair and others (1980) attributed pulsations of flow and sediment concentration in runoff to tunnel collapse and formation of mudflows. Bryan and Harvey (1985) measured rapidly fluctuating sediment concentrations of up to 97,000 mg/l emanating from tunnels in Na-montmorillonite-rich clay badlands. Swanson (1983) compared sediment concentrations in tunnel discharge from a sodium-rich colluvium with synchronous concentrations in overland flow from the grassy surface. Concentrations were roughly correlated with discharge, but those in the overland flow ranged between 100 and 1,100 mg/l, while the tunnel discharge had concentrations of 3000 to 30,000 mg/l and flow at the site of a recent roof collapse a few meters upstream of the headcut had concentrations of 800 to 8,000 mg/l. Swanson and others (1989) found that 95 percent of all the sediment eroded from this small drainage basin during a 25-yr, 6-hr rainstorm originated by tunnel scour.

**Morphological significance of tunnel scour**

The following conditions are necessary for the occurrence of tunnel scour:

1. Saturation at the land surface, above a subsurface horizon, or at the roof of a macropore being heavily recharged by percolation from above.

2. An initial passage into which water from the saturated zone can flow.

3. A sufficiently high shear stress on one boundary of the passage to exceed the critical tractive force of the boundary material. Dispersion of clay minerals may reduce this critical tractive force virtually to zero, and such minerals are more likely to form by weathering in subhumid climates, or in certain lithologies such as marine clays or glass-rich volcanic rocks in a range of climates. Mass failure of walls and roofs may also mobilize sediment.

4. Sufficient transport capacity and therefore sufficient flow and hydraulic gradient throughout the passage to discharge all eroded sediment.

5. Sufficient cohesion to maintain walls and a roof for some time.

The origin of such controlling factors is reviewed above.

The formation of a passage also stimulates the tendency to fill it. Thus, tunnels in forest soils, such as those described by Jones (1971) and Pierson (1983) accelerate the inward diffusion of soil particles by creep, or are obliterated by bioturbation.
Implications of erosion by subsurface flow

Small, seasonally enlarged cracks in cracking-clay soils are filled by swelling during the subsequent wet season. Larger tunnels in dispersive soils tend to fill by collapse of sidewalls and roofs until the ground surface is breached (Heede, 1971; Löffler, 1977; Swanson, 1983).

The survival of tunnels depends on the relative rates of the scouring and filling processes, within the constraint set by the survival of roofs the breaching of which converts the tunnel into a gully. Thus, in a steep, windy, forested environment where trees throw is vigorous, tunnels, especially in well-aggregated soils, would not be expected to survive for long or to grow large, except below the root zone in colluvium-filled bedrock hollows (Pierson, 1983), or in peat layers with low biogenic activity (Gilman and Newson, 1980), or beneath shallow-rooted vegetation in percolines (Jones, 1971). Despite the claim of Jones (1971), I am not aware of any proof that such tunnels usually evolve into stream channels, rather than the ensemble of them being a steady-state feature of the landscape with an average rate of production and destruction. At the other extreme, dispersive sedimentary rocks described by Parker (1963) and Parker and others (1964) may erode along tunnels that are too deep to be disrupted by bioturbation or any other filling process except collapse, which is resisted by the high cohesive strength of the unwetted parts of the formations. Therefore, they can survive for long enough to develop the large intricate features known collectively as pseudokarst, described in Chapter 3.

Whatever their depth, dimensions, and age, tunnels are ephemeral on a geological time scale, and the ground is eventually lowered to them by surface erosion or internal failure. Thus, the dominant geomorphological expression of tunnel scour at the surface results from collapse (Rubey, 1928; Buckham and Cockfield, 1950; Löffler, 1977, p. 158). The resulting channel and valley form tends to be sharp-edged (Fig. 21). Collapsed blocks disperse and are eroded by even shallow channelized runoff. Sinks, vertical shafts, and subhorizontal tunnels occur upstream of the heads of subaerial channels, and the local hydrology and sediment transport may be exceedingly complex (Bryan and Harvey, 1985).

The initial conduits may be localized by very subtle sedimentary structures or stress cracks, so that it is difficult to predict the next phase of tunnel extension. Therefore, there is little prospect of predicting the magnitudes of flows and shear stresses to which the tunnel margins will be subject. This uncertainty, together with the difficulty of specifying a critical tractive force in heterogeneous, dispersive material makes the prediction of a stable drainage density intractable at present. Drainage density should be positively correlated with rainstorm intensity, land surface gradient, local relief, and the dispersibility of the geologic material, and in truly dispersed materials (in which the critical tractive force declines to zero), an intricate network of tunnels should eventually consume the material. However, if the critical tractive force is not zero, and if the processes tending to fill the tunnel are significant, the drainage density is stabilized by the balance between tunnel scour and infilling. Examples of balanced and unstable tunnel systems are described in Chapter 3 and the papers referred to in this section.

SUMMARY

Subterranean water can erode geologic materials: (1) as it flows through and emerges from the matrix; and (2) as it flows along conduits of various origin. The former process is here called seepage erosion; the latter is called tunnel scour. The processes are not mutually exclusive, but the conditions favoring each are different. The morphogenetic significance of seepage erosion in establishing valley systems in weathered bedrock over geological time scales is probably greater than that of tunnel scour, which seems to have a large-scale morphological expression only in a rather restricted, though widespread, range of geologic and climatic conditions. The role of seepage erosion in forming valley systems is probably much more widespread than the restricted set of conditions in which "groundwater sapping" has been studied. Conversely, in some landscapes, such as Hawaii, in which groundwater sapping has been called on as the dominant erosive agent, mechanisms such as undermining by plunge-pool activity have not yet been ruled out. Detailed field studies of the dominant erosion processes have been conducted only at a few localities.

Understanding the mechanics of each erosion process requires study of the governing subsurface hydrology. Important areas of uncertainty remain in:

1. quantitative prediction of flow rates and stresses in subterranean runoff as they are affected by the complex geometries and evolving hydrogeological properties at eroding sites;
2. analysis of the roles of structural features of rocks and soils, which affect both the subsurface flow and resistance to erosion;
3. quantitative prediction of the stabilization of valley or channel networks and their drainage density as a function of the controlling variables for each erosion process;
4. analysis of the response of hillslopes, tunnel margins, and other surfaces formed by expansion of the valley or channel networks in order to understand how the formation of channelized forms stimulates the tendency to fill them.

ACKNOWLEDGMENTS

This chapter was written with the support of a grant from the National Science Foundation, EAR-8313172. It was improved significantly as a result of reviews by R. J. Chorley, W. E. Dietrich, C. G. Higgins, M. J. Kirkby, and J. E. Laity. R. M. Iverson was particularly helpful in clarifying my understanding of seepage erosion in cohesive materials.
APPENDIX 1.

Symbols.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>channel width</td>
</tr>
<tr>
<td>c</td>
<td>cohesion</td>
</tr>
<tr>
<td>g</td>
<td>gravitational acceleration</td>
</tr>
<tr>
<td>N_s</td>
<td>thickness of saturated layer normal to slope</td>
</tr>
<tr>
<td>I</td>
<td>infiltration or recharge rate</td>
</tr>
<tr>
<td>i_z</td>
<td>magnitude of the hydraulic head gradient</td>
</tr>
<tr>
<td>k</td>
<td>saturated hydraulic conductivity</td>
</tr>
<tr>
<td>K(w)</td>
<td>saturated hydraulic conductivity, varying with moisture content</td>
</tr>
<tr>
<td>L</td>
<td>horizontal length of a hillslope</td>
</tr>
<tr>
<td>n</td>
<td>Manning’s roughness coefficient</td>
</tr>
<tr>
<td>p</td>
<td>porosity</td>
</tr>
<tr>
<td>Q</td>
<td>subsurface flux per unit cross-sectional area</td>
</tr>
<tr>
<td>q</td>
<td>subsurface flux per unit width (direction, distance, etc., indicated by subscript)</td>
</tr>
<tr>
<td>r</td>
<td>radial distance</td>
</tr>
<tr>
<td>s</td>
<td>channel gradient</td>
</tr>
<tr>
<td>t</td>
<td>time</td>
</tr>
<tr>
<td>w</td>
<td>moisture content</td>
</tr>
<tr>
<td>z_x, z_y</td>
<td>orthogonal coordinates</td>
</tr>
<tr>
<td>a_z</td>
<td>height of a seepage face</td>
</tr>
<tr>
<td>a_p</td>
<td>angle of water table</td>
</tr>
<tr>
<td>a_β</td>
<td>angle of underlying aquiclude</td>
</tr>
<tr>
<td>c_n</td>
<td>concentration angle of a conical hollow</td>
</tr>
<tr>
<td>l_a</td>
<td>water-table height above aquiclude</td>
</tr>
<tr>
<td>θ</td>
<td>hillslope angle</td>
</tr>
<tr>
<td>λ</td>
<td>angle between seepage vector and surface normal vector</td>
</tr>
<tr>
<td>ψ</td>
<td>angle between seepage vector and horizontal</td>
</tr>
<tr>
<td>ρ_f, ρ_s</td>
<td>densities of fluid, solid</td>
</tr>
<tr>
<td>Φ</td>
<td>total hydraulic head</td>
</tr>
<tr>
<td>φ</td>
<td>angle of internal friction</td>
</tr>
<tr>
<td>Ψ</td>
<td>pressure head</td>
</tr>
</tbody>
</table>

REFERENCES CITED


Implications of erosion by subsurface flow


Tanca, T., 1982. The role of subsurface water infiltration in soil erosion processes: International Association of Hydrological Sciences Publication 137, p. 73-80.


MANUSCRIPT ACCEPTED BY THE SOCIETY NOVEMBER 14, 1989

NOTE ADDED IN PROOF

Since this chapter was written, the field guide edited by Howard, Kochel, and Holt (1986) and referred to several times in the text has been published in modified form as: