CHAPTER 2

# Processes and mechanics of erosion

Soil erosion is a two-phase process consisting of the detachment of individual soil particles from the soil mass and their transport by erosive agents such as running water and wind. When sufficient energy is no longer available to transport the particles, a third phase, deposition, occurs.

Rainsplash is the most important detaching agent. As a result of raindrops striking a bare soil surface, soil particles may be thrown through the air over distances of several centimetres. Continuous exposure to intense rainstorms considerably weakens the soil. The soil is also broken up by weathering processes, both mechanical, by alternate wetting and drying, freezing and thawing and frost action, and biochemical. Soil is disturbed by tillage operations and by the trampling of people and livestock. Running water and wind are further contributors to the detachment of soil particles. All these processes loosen the soil so that it is easily removed by the agents of transport.

The transporting agents comprise those that act areally and contribute to the removal of a relatively uniform thickness of soil, and those that concentrate their action in channels. The first group consists of rainsplash, surface runoff in the form of shallow flows of infinite width, sometimes termed sheet flow but more correctly called overland flow, and wind. The second group covers water in small channels, known as rills, which can be obliterated by weathering and ploughing, or in the larger more permanent features of gullies and rivers. A distinction is commonly made for water erosion between rill erosion and erosion on the land between the rills by the combined action of raindrop impact and overland flow. This is termed interrill erosion. To these agents that act externally, picking up material from and carrying it over the ground surface, should be added transport by mass movements such as soil flows, slides and creep, in which water affects the soil internally, altering its strength.

The severity of erosion depends upon the quantity of material supplied by detachment over time and the capacity of the eroding agents to transport it. Where the agents have the capacity to transport more material than is supplied by detachment, the erosion is described as detachment-limited. Where more material is supplied than can be transported, the erosion is transport limited.

The energy available for erosion takes two forms: potential and kinetic. Potential energy ($PE$) results from the difference in height of one body with respect to another. It is the product of mass ($m$), height difference ($h$) and acceleration due to gravity ($g$), so that

$$PE = mhg \tag{2.1}$$

**Table 2.1** Efficiency of forms of water erosion

| Form | Mass* | Typical velocity ($m\,s^{-1}$) | Kinetic energy† | Energy for erosion‡ | Observed sediment transport§ ($g\,cm^{-1}$) |
|---|---|---|---|---|---|
| Raindrops | $R$ | 6.0 | $18R$ | $0.036R$ | 20 |
| Overland flow | $0.5R$ | 0.01 | $2.5 \times 10^{-5}R$ | $7.5 \times 10^{-7}R$ | 400 |
| Rill flow | $0.5R$ | 4¶ | $4R$ | $0.12R$ | 19,000 |

* Assumes rainfall mass of $R$ of which 50 per cent contributes to runoff.
† Based on $\frac{1}{2}mv^2$.
‡ Assumes that 0.2 per cent of the kinetic energy of raindrops and 3 per cent of the kinetic energy of runoff is utilized in erosion.
§ Totals observed in mid-Bedfordshire, England, on an 11° slope, on sandy soil, over 900 days. Most of the energy of raindrops contributes to soil particle detachment rather than transport.
¶ Estimated using the Manning equation of flow velocity for a rill, 0.3 m wide and 0.2 m deep, on a slope of 11°, at bankfull, assuming a roughness coefficient of 0.02.

which, in units of kg, m and $m\,s^{-2}$ respectively, yields a value in Joules. The potential energy for erosion is converted into kinetic energy ($KE$), the energy of motion. This is related to the mass and velocity ($v$) of the eroding agent in the expression

$$KE = \frac{1}{2}mv^2 \qquad (2.2)$$

which, in units of kg and $m\,s^{-1}$, also gives a value in Joules. Most of this energy is dissipated in friction with the surface over which the agent moves so that only 3–4 per cent of the energy of running water and 0.2 per cent of that of falling raindrops is expended in erosion (Pearce 1976). An indication of the relative efficiencies of the processes of water erosion can be obtained by applying these figures to calculations of kinetic energy, using eqn 2.2, based on typical velocities (Table 2.1). The concentration of running water in rills affords the most powerful erosive agent but raindrops are potentially more erosive than overland flow. Most of the raindrop energy is used in detachment, however, so that the amount available for transport is less than that from overland flow. This is illustrated by measurements of soil loss in a field in mid-Bedfordshire, England. Over a 900-day period on an 11° slope on a sandy soil, transport across a centimetre width of slope amounted to 19,000 g of sediment by rills, 400 g by overland flow and only 20 g by rainsplash (Morgan et al. 1986).

## 2.1 Hydrological basis of erosion

The processes of water erosion are closely related to the pathways taken by water in its movement through the vegetation cover and over the ground surface. During a rainstorm, part of the water falls directly on the land, either because there is no vegetation or because it passes through gaps in the plant canopy. This component of the rainfall is known as direct throughfall. Part of the



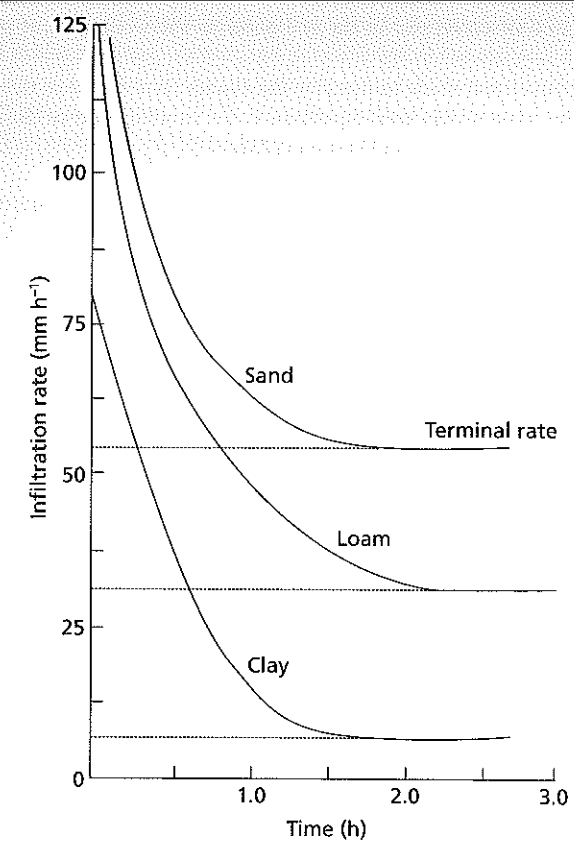



**Fig. 2.1** Typical infiltration rates for various soils (after Withers & Vipond 1974).

rain is intercepted by the canopy, from where it either returns to the atmosphere by evaporation or finds its way to the ground by dripping from the leaves, a component termed leaf drainage, or by running down the plant stems as stemflow. The action of direct throughfall and leaf drainage produces rainsplash erosion. The rain that reaches the ground may be stored in small depressions or hollows on the surface or it may infiltrate the soil, contributing to soil moisture storage, to lateral movement downslope within the soil as subsurface or interflow or, by percolating deeper, to groundwater. When the soil is unable to take in more water, the excess contributes to runoff on the surface, resulting in erosion by overland flow or by rills and gullies.

The rate at which water passes into the soil is known as the infiltration rate and this exerts a major control over the generation of surface runoff. Water is drawn into the soil by gravity and by capillary forces, whereby it is attracted to and held as a thin molecular film around the soil particles. During a rainstorm, the spaces between the soil particles become filled with water and the capillary forces decrease so that the infiltration rate starts high at the beginning of a storm and declines to a level that represents the maximum sustained rate at which water can pass through the soil to lower levels (Fig. 2.1). This level, the infiltration capacity or terminal infiltration rate, corresponds theoretically to the saturated hydraulic conductivity of the soil. In practice, however, the infiltration capacity is often lower than the saturated hydraulic conductivity because of air entrapped in the soil pores as the wetting front passes downwards through the soil.

Various attempts have been made to describe the change in infiltration rate over time mathematically. One of the most widely used equations is the modification of the Green and Ampt (1911) equation proposed by Mein and Larson (1973):

$$i = A + \frac{B}{t} \tag{2.3}$$

where $i$ is the instantaneous rate of infiltration, $A$ is the transmission constant or saturated hydraulic conductivity of the soil, $B$ is the sorptivity, defined by Talsma (1969) as the slope of the line when $i$ is plotted against $t$, and $t$ is the time elapsed since the onset of the rain. This equation has been found to describe well the infiltration behaviour of soils in southern Spain (Scoging & Thornes 1979) and Arizona (Scoging et al. 1992) but Bork and Rohdenburg (1981), also working in southern Spain, obtained better results with the equation proposed by Philip (1957):

$$i = A + \frac{B}{\sqrt{t}} \tag{2.4}$$

while Gifford (1976) found neither equation satisfactory for semi-arid rangelands in northern Australia and Utah. Kutílek et al. (1988) tested both equations against field measurements obtained with double-ring infiltrometers and found that neither fitted the data well, giving errors of between 10 and 59 per cent when used to estimate saturated hydraulic conductivity. One reason for the error is the failure to predict infiltration correctly under conditions of surface ponding when the soil develops a viscous resistance to air flow. Morel-Seytoux and Khanji (1974) developed the following equation to allow for this:

$$i = \frac{k_s}{\beta}\left(1 + \frac{(\theta_t - \theta_i)(H_0 + \Delta\psi)}{I}\right) \tag{2.5}$$

where $k_s$ is the saturated hydraulic conductivity; $\beta$ is a viscous correction factor, which varies in value between 1.1 and 1.7, depending on soil type and ponding depth but averages 1.4; $\theta_i$ is the initial soil moisture content by volume; $\theta_t$ is the actual volumetric moisture content of soil in the zone between the ground surface and the wetting front; $H_0$ is the depth of ponded water; $\Delta\psi$ is the change in $\psi$ between the soil surface and the wetting front; $\psi$ is the difference in pressure between the pore-water and the atmosphere; and $I$ is the total amount of water already infiltrated. As a result of including the viscous correction factor, eqn 2.5 predicts lower infiltration rates than either eqn 2.3 or eqn 2.4.

Infiltration rates depend upon the characteristics of the soil. Generally, coarse-textured soils such as sands and sandy loams have higher infiltration rates than clay soils because of the larger spaces between the pores. Infiltration capacities may range from more than 200 mm h$^{-1}$ for sands to less than 5 mm h$^{-1}$ for tight clays (Fig. 2.1). In addition to the role played by the inter-particle spacing or micropores, the larger cracks or macropores exert an important influence over infiltration. They can transmit considerable quantities of water so that clays with well defined structures can have infiltration rates that are much higher than would be expected from their texture alone. Infiltration behaviour on many soils is also rather complex because the soil profiles are characterized by two or more layers of differing hydraulic conductivities; most agricultural soils, for example, consist of a disturbed plough layer and an undisturbed subsoil. Many soils on construction sites comprise a heavily compacted subsoil covered by a thinner and less compacted



topsoil. Local variability in infiltration rates can be quite high because of differences in the structure, compaction, initial moisture content and profile form of the soil and in vegetation density. Field determinations of average infiltration capacity using infiltrometers may have coefficients of variation of 70–75 per cent. Eyles (1967) measured infiltration capacity on soils of the Melaka Series near Temerloh, Malaysia, and obtained values ranging from 15 to 420 mm h$^{-1}$, with a mean of 147 mm h$^{-1}$.

According to Horton (1945), if rainfall intensity is less than the infiltration capacity of the soil, no surface runoff occurs and the infiltration rate equals the rainfall intensity. If the rainfall intensity exceeds the infiltration capacity, the infiltration rate equals the infiltration capacity and the excess rain forms surface runoff. As a mechanism for generating runoff, however, this comparison of rainfall intensity and infiltration capacity does not always hold. Studies in Bedfordshire, England (Morgan et al. 1986) on a sandy soil show that measured infiltration capacity is greater than 400 mm h$^{-1}$ and that rainfall intensities rarely exceed 40 mm h$^{-1}$. Thus no surface runoff would be expected, whereas, in fact, the mean annual runoff is about 55 mm from a mean annual rainfall of 550 mm. The reason runoff occurs is that these soils are prone to the development of a surface crust. Two types of crust can be distinguished. Where a crust forms *in situ* on the soil, it is termed a structural crust; where it results from the deposition of fine particles in puddles, it is called a depositional crust (Boiffin 1985). As shown by studies on loamy soils in north-east France, crusting can reduce the infiltration capacity from 45–60 to about 6 mm h$^{-1}$ with a structural crust and 1 mm h$^{-1}$ with a depositional crust (Boiffin & Monnier 1985; Martin et al. 1997). Reductions in infiltration of 50 (Hoogmoed & Stroosnijder 1984) to 100 per cent (Torri et al. 1999) can occur in a single storm. The importance of crusting and sealing was also emphasized by Poesen (1984), who found that infiltration rates were higher on steeper slopes where the higher erosion rate prevented the seal from forming.

The presence of stones or rock fragments on the surface of a soil also influences infiltration rates but in a rather complex way depending on whether the stones are resting on top of the surface or are embedded within the soil. Generally, rock fragments protect the soil against physical destruction and the formation of a crust, so that infiltration rates are higher than on a comparable stone-free bare soil. However, on soils that are subject to crusting, a high percentage stone cover can produce a worse situation; a 75 per cent cover of rock fragments embedded in a crusted surface on a silt-loam soil reduced infiltration rates to 50 per cent of those on a stone-free soil (Poesen & Ingelmo-Sanchez 1992).

The important control for runoff production on many soils is not infiltration capacity but a limiting moisture content. When the actual moisture content is below this value, pore water pressure in the soil is less than atmospheric pressure and water is held in capillary form under tensile stress or suction. When the limiting moisture content is reached and all the pores are full of water, pore water pressure equates to atmospheric pressure, suction reduces to zero and surface ponding occurs. This explains why sands that have low levels of capillary storage can produce runoff very quickly even though their infiltration capacity is not exceeded by the rainfall intensity. Since hydraulic conductivity is a flux partly controlled by rainfall intensity, increases in intensity can cause conductivity to rise so that, although runoff may have formed rapidly at a relatively low intensity, higher rainfall intensities do not always produce greater runoff. This mechanism explains why infiltration rates sometimes increase with rainfall intensity (Nassif & Wilson 1975). Bowyer-Bower (1993) found that, for a given soil, infiltration capacity was higher with higher rainfall intensities because of their ability to disrupt surface seals and crusts that would otherwise keep the infiltration rate low.

Once water starts to pond on the surface, it is held in depressions or hollows and runoff does not begin until the storage capacity of these is satisfied. On agricultural land, depression

**Table 2.2** Surface roughness (*RFR*) for different tillage implements compared to other expressions of random roughness

| Implement | Roughness (*RFR*) ($cm\,m^{-1}$) | Random roughness (*RR*) (mm) |
|---|---|---|
| Moldboard plough | 30–33 | 33–48 |
| Chisel plough | 24–28 | 17–26 |
| Cultivator | 15–23 | 6–15 |
| Tandem disc | 25–28 | 18–26 |
| Offset disc | 32–35 | 38–51 |
| Paraplough | 32–35 | 10 |
| Spike tooth harrow | 17–23 | 8–15 |
| Spring tooth harrow | 25 | 18 |
| Rotary hoe | 21–22 | 12–13 |
| Rototiller | 23 | 15 |
| Drill | 20–21 | 10–12 |
| Row planter | 13–22 | 5–13 |

*Note:* The term, *RFR*, is essentially an index of the tortuosity of the soil surface. An alternative and widely used descriptor of the roughness of the soil surface is random roughness (*RR*, mm), defined as the standard deviation of a series of surface height measurements (Currence & Lovely 1970). There is a good correlation between *RR* and *RFR* which can be expressed by (Auerswald personal communication):

$\ln RR = 0.29 + 0.099RFR,\ r = 0.995,\ n = 27$

$RFR = -1.77 + 9.25 \ln RR,\ r = 0.912,\ n = 27$

Surface roughness, expressed by *RR*, declines over time as a function of cumulative rainfall ($R_c$):

$RR(t) = RR(0)e^{-\alpha R_c}$

where *RR(t)* is the random roughness at time (*t*), *RR*(0) is the original random roughness after tillage and $\alpha = 2.8 \times 0.3S_i$, where $S_i$ is the silt content of the soil (0–1) (if $\alpha \geq 0$, $\alpha$ is set to −1) (Alberts et al. 1989).
*Source:* after Auerswald, personal communication.

storage varies seasonally depending on the type of cultivation that has been carried out and the time since cultivation for the roughness to be reduced by weathering and raindrop impact. Table 2.2 gives typical values of depression storage (*DS*; mm) for surfaces produced by different tillage implements, based on their roughness index (*RFR*; $cm\,m^{-1}$) (Auerswald, personal communication):

$$DS = 0.14e^{0.04RFR} \quad (2.6)$$

$$RFR = \frac{L_A - L_0}{L_A} \times 100 \quad (2.7)$$

where $L_0$ is the straight-line distance between two points along a transect of the soil surface and $L_A$ is the actual distance measured over all the microtopographic irregularities.



Surface roughness and therefore depression storage decline over time through weathering and raindrop impact. Auerswald (personal communication) developed the following relationship to express the decline in roughness as a function of the cumulative kinetic energy of rainfall:

$$RFR(t) = RFR_0 e^{-0.7\sqrt{KE(t)}} \quad (2.8)$$

where *RFR*(*t*) is the roughness at a certain time, $RFR_0$ is the initial roughness and *KE*(*t*) is the accumulated kinetic energy of the rain at time (*t*). Depression storage also varies with the soil with clay soils having 1.6–2.3 times the storage volume of sandy soils. The roughness values given in Table 2.2 relate to soils with about 20 per cent clay. These base values ($RFR_{base}$) can be adjusted to give roughness for different clay contents ($RFR_{CC}$) using the relationship:

$$RFR_{cc} = RFR_{base}(0.4 + 0.025CC) \quad (2.9)$$

where *CC* is the percentage clay content of the soil. This relationship is valid for clay contents up to 25 per cent; for higher clay contents it is recommended to use the 25 per cent value.

## 2.2 Rainsplash erosion

The action of raindrops on soil particles is most easily understood by considering the momentum of a single raindrop falling on a sloping surface. The downslope component of this momentum is transferred in full to the soil surface but only a small proportion of the component normal to the surface is transferred, the remainder being reflected. The transfer of momentum to the soil particles has two effects. First, it provides a consolidating force, compacting the soil; second, it produces a disruptive force as the water rapidly disperses from and returns to the point of impact in laterally flowing jets. Whereas the impact velocity of falling raindrops striking the soil surface varies from about $4\,m\,s^{-1}$ for a 1 mm diameter drop to $9\,m\,s^{-1}$ for a 5 mm diameter drop, the local velocities of these jets are about twice these (Huang et al. 1982). These fast-moving water jets impart a velocity to some of the soil particles and launch them into the air, entrained within water droplets that are themselves formed by the break-up of the raindrop on contact with the ground (Mutchler & Young 1975). Thus, raindrops are agents of both consolidation and dispersion.

The consolidation effect is best seen in the formation of a surface crust, usually only a few millimetres thick, which results from clogging of the pores by soil compaction and by the infilling of surface pore spaces by fine particles detached from soil aggregates by the raindrop impact. Studies of crust development under simulated rainfall show that crusts have a dense surface skin or seal, about 0.1 mm thick, with well oriented clay particles. Beneath this is a layer, 1–3 mm thick, where the larger pore spaces are filled by finer washed-in material (Tackett & Pearson 1965). That raindrop impact is the critical process was shown by Farres (1978), who found that, after a rainstorm, most aggregates on the soil surface were destroyed, while those in the lower layer of the crust remained intact, even though completely saturated. A tap of these aggregates, however, caused their instant breakdown. This evidence indicates that although saturation reduces the internal strength of soil aggregates, they do not disintegrate until struck by raindrops.

The actual response of a soil to a given rainfall depends upon its moisture content and, therefore, its structural state and the intensity of the rain. Le Bissonnais (1990) describes three possible responses:

- If the soil is dry and the rainfall intensity is high, the soil aggregates break down quickly by slaking. This is the breakdown by compression of air ahead of the wetting front. Infiltration capacity reduces rapidly and on very smooth surfaces runoff can be generated after only a few millimetres of rain. With rougher surfaces, depression storage is greater and runoff takes longer to form.
- If the aggregates are initially partially wetted or the rainfall intensity is low, microcracking occurs and the aggregates break down into smaller aggregates. Surface roughness thus decreases but infiltration remains high because of the large pore spaces between the microaggregates.
- If the aggregates are initially saturated, infiltration capacity depends on the saturated hydraulic conductivity of the soil and large quantities of rain are required to seal the surface. Nevertheless, soils with less than 15 per cent clay content are vulnerable to sealing if the intensity of the rain is high.

Over time, the percentage area of the soil surface affected by crust development increases exponentially with cumulative rainfall energy (Govers & Poesen 1985), which, in turn, brings about an exponential decrease in infiltration capacity (Boiffin & Monnier 1985). Crustability decreases with increasing contents of clay and organic matter since these provide greater strength to the soil. Thus loams and sandy loams are the most vulnerable to crust formation.

Studies of the kinetic energy required to detach one kilogram of sediment by raindrop impact show that minimal energy is needed for soils with a geometric mean particle size of 0.125 mm and that soils with geometric mean particle size between 0.063 and 0.250 mm are the most vulnerable to detachment (Fig. 2.2; Poesen 1985). Coarser soils are resistant to detachment because of the weight of the larger particles. Finer soils are resistant because the raindrop energy has to overcome the adhesive or chemical bonding forces that link the minerals comprising the clay particles. The wide range in energy required to detach clay particles is a function of different levels of resistance in relation to the type of clay minerals and the relative amounts of calcium, magnesium and sodium ions in the water passing through the pores (Arulanandan & Heinzen 1977). Overall, silt loams, loams, fine sands and sandy loams are the most detachable. Selective removal of particles by rainsplash can cause variations in soil texture downslope. Splash erosion on stony loamy soils in the Luxembourg Ardennes has resulted in soils on the valley sides becoming deficient in clay and silt particles and high in gravel and stone content, whereas the colluvial soils at the base of the slopes are enriched by the splashed-out material (Kwaad 1977). Selective erosion can affect soil aggregates as well as primary particles. Rainfall simulation experiments on clay soils in Italy show that splashed-out material is enriched in soil aggregates of 0.063–0.50 mm in size (Torri & Sfalanga 1986).

The detachability of soil depends not only on its texture but also on top soil shear strength (Cruse & Larsen 1977), a finding that has prompted attempts to understand splash erosion in terms of shear. The detachment of soil particles represents a failure of the soil by the combined mechanism of compression and shear under raindrop impact, an event that is most likely to occur under saturated conditions when the shear strength of the soil is lowest (Al-Durrah & Bradford 1982). Generally, detachment decreases exponentially with increasing shear strength. Broadly linear relationships have been obtained, however, between the quantity of soil particles detached by raindrop impact and the ratio of the kinetic energy of the rainfall to soil shear strength (Al-Durrah & Bradford 1981; Torri et al. 1987b; Bradford et al. 1992).



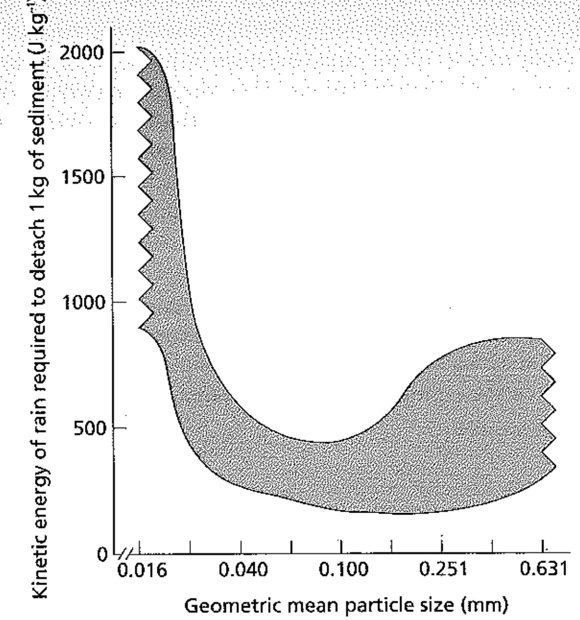



**Fig. 2.2** Relation between geometric mean particle size of the soil and the rainfall energy required to detach 1 kg of sediment. Shaded area shows range of experimental values (after Poesen 1992).

Rain does not always fall on to a dry surface. During a storm it may fall on surface water in the form of puddles or overland flow. Studies by Palmer (1964) show that as the thickness of the surface water layer increases, so does splash erosion. This is believed to be due to the turbulence that impacting raindrops impart to the water. No increase in splash erosion with water depth has been observed, however, on sandy soils (Ghadiri & Payne 1979; Poesen 1981). There is, however, a critical water depth beyond which erosion decreases exponentially with increasing water depth because more of the rainfall energy is dissipated in the water and does not affect the soil surface. Laboratory experiments have shown that the critical depth is approximately equal to the diameter of the raindrops (Palmer 1964) or to one-fifth (Torri & Sfalanga 1986) or one-third (Mutchler & Young 1975) of the diameter. These differences in the value of critical depth are due to the different experimental conditions used in the experiments, particularly the soils, which ranged from clays to silt loams, loams and sandy loams.

Experimental studies show that the rate of detachment of soil particles with rainsplash varies with the 1.0 power of the instantaneous kinetic energy of the rain (Free 1960; Quansah 1981) and with the square of the instantaneous rainfall intensity (Meyer 1981). The detachment rate ($D_r$) on bare soil can be expressed by equations of the form:

$$D_r \propto I^a s^c \tag{2.10}$$

$$D_r \propto KE^b s^c e^{-dh} \tag{2.11}$$

where $I$ is the rainfall intensity (mm h$^{-1}$), $s$ is the slope expressed in m m$^{-1}$ or as a sine of the slope angle, $KE$ is the kinetic energy of the rain (J m$^{-2}$) and $h$ is the depth of surface water (m). Although

2.0 is a convenient value for $a$, the value may be adjusted to allow for variations in soil texture using the term $a = 2.0 - (0.01 \times \%\text{ clay})$ (Meyer 1981). Similarly, the value of 1.0 for $b$ may be varied from 0.8 for sandy soils to 1.8 for clays (Bubenzer & Jones 1971). Values for $c$ are in the range of 0.2–0.3 (Quansah 1981; Torri & Sfalanga 1986), also varying with the texture of the soil (Torri & Poesen 1992). It should be remembered that the slope term in this equation refers to the local slope for a distance equivalent to only a few drop diameters from the point of raindrop impact – for example, that on the side of a soil clod – and not the average ground slope. Thus, for practical purposes, the slope term is often omitted from calculations of soil particle detachment. A value of 2.0 is convenient for $d$ as representative of a range of values between 0.9 and 3.1 for different soil textures (Torri et al. 1987b).

In contrast, average ground slope is important when considering the overall transport of splashed particles. On a sloping surface more particles are thrown downslope than upslope during the detachment process, resulting in a net movement of material downslope. Splash transport per unit width of slope ($T_r$) can be expressed by the relationship:

$$T_r \propto I^j S^f \tag{2.12}$$

where $j = 1.0$ (Meyer & Wischmeier 1969) and $f = 1.0$ (Quansah 1981; Savat 1981). There is some evidence to suggest that the value of $f$ decreases on steeper slopes; Mosley (1973) gives a value of 0.8 and Moeyersons and De Ploey (1976) a value of 0.75 where slope angles rise to 20 and 25° respectively. Foster and Martin (1969) and Bryan (1979) found that splash transport increases with slope angle to reach a maximum at about 18° and that on steeper slopes $f$ becomes negative.

These relationships for detachment and transport of soil particles by rainsplash ignore the role of wind. Windspeed imparts a horizontal force to a falling raindrop until its horizontal velocity component equals the velocity of the wind. As a result, the kinetic energy of the raindrop is increased. Not surprisingly, detachment of soil particles by impacting wind-driven raindrops can be some 1.5–3 times greater than that resulting from rains of the same intensity without wind (Disrud & Krauss 1971; Lyles et al. 1974a). Wind also causes raindrops to strike the surface at an angle from vertical. This affects the relative proportions of upslope versus downslope splash. Moeyersons (1983) shows that where the angle between the falling raindrop and the vertical is 20°, net splash transport is reduced to zero for slopes of 17–19° and has a net upslope component for gentler slopes. Where the angle between the falling raindrop and the vertical is 5°, zero splash occurs on a slope of 3°.

Since splash erosion acts uniformly over the land surface its effects are seen only where stones or tree roots selectively protect the underlying soil and splash pedestals or soil pillars are formed. Such features frequently indicate the severity of erosion. Splash erosion is most important for detaching the soil particles that are subsequently eroded by running water. However, on the upper parts of hillslopes, particularly those of convex form, splash transport may be the dominant erosion process. In Calabria, southern Italy, under forest and under scattered herb and shrub vegetation, splash erosion accounts for 30–95 per cent of the total transport of material by water erosion (van Asch 1983). In Bedfordshire, England, splash accounts for 15–52 per cent of total soil transport on land under cereals and grass but only 3–10 per cent on bare ground (Morgan et al. 1986). As runoff and soil loss increase, the importance of splash transport declines, although very low contributions of splash to total transport were also measured in Bedfordshire under woodland because of the protective effect of a dense litter layer. Govers and Poesen (1988) found that although raindrop impact detached 152 t ha$^{-1}$ of soil over one year on a bare loam soil on a 14° slope in Belgium, splash transport accounted for only 0.2 t ha$^{-1}$ of the soil loss. The most important contribution of splash erosion was to deliver detached particles to overland flow, which was the main agent of sediment transport in the interrill areas.



## 2.3 Overland flow

Overland flow occurs on hillsides during a rainstorm when surface depression storage and either, in the case of prolonged rain, soil moisture storage or, with intense rain, the infiltration capacity of the soil are exceeded. The flow is rarely in the form of a sheet of water of uniform depth and more commonly is a mass of anastomosing or braided water courses with no pronounced channels. The flow is broken up by stones and cobbles and by the vegetation cover, often swirling around tufts of grass and small shrubs.

### 2.3.1 Hydraulic characteristics

The hydraulic characteristics of the flow are described by its Reynolds number ($Re$) and its Froude number ($F$), defined as follows:

$$Re = \frac{vr}{\nu} \tag{2.13}$$

$$F = \frac{v}{\sqrt{gr}} \tag{2.14}$$

where $r$ is the hydraulic radius, which, for overland flow, is taken as equal to the flow depth and $\nu$ is the kinematic viscosity of the water. The Reynolds number is an index of the turbulence of the flow. The greater the turbulence, the greater is the erosive power generated by the flow. At numbers less than 500 laminar flow prevails and at values above 2000 flow is fully turbulent. In laminar flow, each fluid layer moves in a straight line with uniform velocity and there is no mixing between the layers, whereas turbulent flow has a complicated pattern of eddies, producing considerable localized fluctuations in velocity, and a continuous interchange of water between the layers. Intermediate values are indicative of transitional or disturbed flow, often a result of turbulence being imparted to laminar flow by raindrop impact (Emmett 1970). The Froude number is an index of whether or not gravity waves will form in the flow. When the Froude number is less than 1.0, gravity waves do not form and the flow, being relatively smooth, is described as tranquil or subcritical. Froude numbers greater than 1.0 denote rapid or supercritical flow, characterized by gravity waves, which is more erosive.

Field studies of overland flow in Bedfordshire reveal Reynolds numbers less than 75 and Froude numbers less than 0.5 (Morgan 1980a). Flows with Reynolds numbers less than 40 and Froude numbers less than 0.13 were observed by Pearce (1976) in the field near Sudbury, Ontario. In various field experiments on semi-arid hillsides in the Walnut Gulch Experimental Watershed, Arizona, Froude numbers in overland flow were consistently less than 0.5 even though Reynolds numbers ranged from 100 to 1200, depending upon local variations in flow depths due to stones and microtopography (Parsons et al. 1990).

### 2.3.2 Detachment of soil particles by flow

The important factor in the above hydraulic relationships is the flow velocity. Because of an inherent resistance of the soil, velocity must attain a threshold value before erosion commences.

Basically, the detachment of an individual soil particle from the soil mass occurs when the forces exerted by the flow exceed the forces keeping the particle at rest. Shields (1936) made a fundamental analysis of the processes involved and the forces at work to determine the critical conditions for initiating particle movement over relatively gentle slopes in rivers in terms of the dimensionless shear stress ($\Theta$) of the flow and the particle roughness Reynolds number ($Re^*$), defined respectively by:

$$\Theta = \frac{\rho_w u_*^2}{g(\rho_s - \rho_w)D} \tag{2.15}$$

$$Re^* = \frac{u_* D}{\nu} \tag{2.16}$$

where $\Theta$ is known as the Shields number, $\rho_w$ is the density of water, $g$ is the acceleration of gravity, $\rho_s$ is the density of the sediment, $D$ is the diameter of the particle and $u_*$ is the shear velocity of the flow, expressed as:

$$u_* = \sqrt{grs} \tag{2.17}$$

When the value of $Re^*$ is greater than 40 (turbulent flow), the critical value of $\Theta$ for particle movement assumes a constant value of 0.05. Unfortunately, this value does not hold when, as is the case with overland flow, the particles are not fully submerged or the flow has Reynolds numbers in the laminar range. Studies with rock fragments in shallow flows suggest that $\Theta_c$ is about 0.01 in value (Poesen 1987; Torri & Poesen 1988). Other research (Govers 1987; Guy & Dickinson 1990; Torri & Borselli 1991) indicates that the Shields number consistently overpredicts the hydraulic requirements for particle movement. This implies that the initiation of movement is not solely a phenomenon of fluid shear stress but is enhanced by other factors. Among those not accounted for by the Shields number are the effects of raindrop impact on the flow, the angle of repose of the particle in relation to ground slope, the strong influence of gravity as the slope steepness increases, the cohesion of the soil, changes in the density of the fluid as sediment concentration in the flow increases and abrasion between particles moving in the flow and the soil beneath.

Since the above approach has not proved entirely satisfactory, a more empirical procedure has been adopted based on a critical value of the flow's shear velocity for initiating particle movement. As can be seen in Fig. 2.3 (Savat 1982), for particles larger than 0.2 mm in diameter, the critical shear velocity increases with particle size. A larger force is required to move larger particles. For particles smaller than 0.2 mm, the critical shear velocity increases with decreasing particle size. The finer particles are harder to erode because of the cohesiveness of the clay minerals of which they are comprised, unless they have been previously detached and, as a result, lost their cohesion, in which case they can then be moved at very low shear velocities. In practice, the critical shear velocities required to erode soil may differ from those shown in Fig. 2.3 because the latter are derived for surfaces of uniform particle size. With mixed particle sizes, the finer particles are protected by the coarser ones so that they are not removed until the shear velocity is great enough to pick up the larger particles. Counteracting this effect, however, is the action of rainsplash, which may detach soil particles and throw them into the flow.

Once the critical conditions for particle movement are exceeded, soil particles may be detached from the soil mass at a rate that is dependent on the shear velocity of the flow and the unit discharge (Govers & Rauws 1986). The direct application of this relationship is only valid,



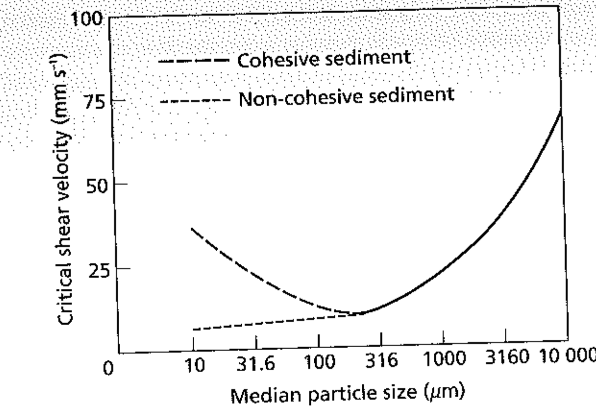



**Fig. 2.3** Critical shear velocity in turbulent water flow for soil particle detachment as a function of particle size (after Savat 1982).

however, if the shear velocity is exerted solely on the soil particles, which implies that the resistance to the flow is entirely due to grain resistance. This situation is only true for completely smooth bare soil surfaces. In practice, resistance due to the microtopographic form of the soil surface and the plant cover is usually more important and grain resistance may be as little as 5 per cent of the total resistance offered to the flow (Abrahams et al. 1992). Since it is difficult to determine the level of grain resistance, only very generalized relationships can be developed for describing detachment rate ($D_f$). These depend on a simple relationship between detachment and flow velocity.

The velocity of flow is dependent upon the depth or hydraulic radius ($r$), the roughness of the surface and the slope ($s$). This relationship is commonly expressed by the Manning equation:

$$v = \frac{r^{2/3} s^{1/2}}{n} \tag{2.18}$$

where $n$ is the Manning coefficient of roughness. Equation 2.18 assumes fully turbulent flow moving over a rough surface. Using the continuity and Manning's velocity equations, Meyer (1965) showed that:

$$v \propto s^{1/2} Q^{1/3} \tag{2.19}$$

for constant roughness conditions, where Q is discharge or flow rate. Assuming that the detachment rate varies with the square of the velocity (Meyer & Wischmeier 1969):

$$D_f \propto Q^{2/3} s^{2/3} \tag{2.20}$$

Quansah (1985) obtained experimentally higher exponent values, however, for a range of soil types from clay to sand:

$$D_f \propto Q^{1.5} s^{1.44} \tag{2.21}$$

Equations 2.20 and 2.21 both relate only to the action of water flow over the soil surface. Quansah (1985) found that the exponents decreased in value when the flow was accompanied by rainfall to give:

$$D_f \propto Q^{1.12} s^{0.64} \tag{2.22}$$

indicating that raindrop impact inhibits the ability of flow to detach soil particles.

Based on the findings from laboratory experiments by Meyer and Monke (1965), who observed that the rate of detachment depended on the amount of sediment already in the flow, Foster and Meyer (1972) proposed that the term $D_f$ in the above equation applies to the detachment capacity rate that occurs only when the flow is clear. Under other conditions, $D_f$ depends on the difference between the actual sediment concentration in the flow ($C$) and the maximum concentration that the flow can hold ($C_{max}$):

$$D_f \propto (C_{max} - C) \tag{2.23}$$

This implies that the detachment rate declines as sediment concentration in the flow increases and that when the maximum sediment concentration is reached, the detachment rate is zero. Merten et al. (2001), however, found that detachment continued to occur for a short time after maximum sediment concentration was attained even though deposition was taking place, indicating that the flow takes time adjust to changing sediment loads. Laboratory studies on sandy loam soils by Kamalu (1993) showed that the detachment rate for flow without rainfall remained at the capacity rate up to the time that maximum sediment concentration was reached, at which point detachment ceased. Thus, there is still considerable uncertainty about the nature of the mechanisms involved in soil particle detachment by flow.

### 2.3.3 Transport of soil particles by flow

Once sediment has been entrained within the flow, it will be transported until such time as deposition occurs. Meyer and Wischmeier (1969) proposed that the transporting capacity of the flow ($T_f$) varies with the fifth power of the velocity, so that:

$$T_f \propto Q^{5/3} s^{5/3} \tag{2.24}$$

This compares well with the following relationships derived respectively by Carson and Kirkby (1972) and Morgan (1980a) from a consideration of the hydraulics of sediment transport:

$$T_f = 0.0085 Q^{1.75} s^{1.625} D_{84}^{-1.11} \tag{2.25}$$

$$T_f = 0.0061 Q^{1.8} s^{1.13} n^{-0.15} D_{35}^{-1} \tag{2.26}$$

where $D_{84}$ and $D_{35}$ define the particle size of the surface material at which respectively 84 and 35 per cent of the grains are finer. All these equations relate to the action of overland flow on its own, whereas, in practice, the flow is usually accompanied by rainfall. The interaction with raindrop impact causes a slight rise in the value of the exponents for discharge and slope. Laboratory experiments by Quansah (1982) with overland flow and rainfall combined gave the relationship:

$$T_f \propto Q^{2.13} s^{2.27} \tag{2.27}$$



Thus, while, as seen by eqns 2.21 and 2.22, the detachment capacity of flow is reduced by raindrop impact, its transporting capacity is enhanced (Savat 1979; Guy & Dickinson 1990; Proffitt & Rose 1992). The degree of enhancement depends on the resistance of the soil, the diameter of the raindrops and the depth and velocity of the flow. Govers (1989) found that high sediment concentrations could increase velocity by up to 40 per cent, especially at low discharges and flow depths. His experiments, however, were carried out for flow without rain, whereas Guy et al. (1990) found that the impact of rain decreased flow velocity by about 12 per cent.

Govers (1990) investigated three different types of equations, based on grain shear velocity, effective stream power and unit stream power for describing the transport capacity of overland flow, defined as the maximum sediment concentration that can be carried. For ease of use, the relationships based on unit stream power were preferred since this is simply the product of slope and flow velocity. He found that:

$$C_{max} = a(sv - 0.4)^b \tag{2.28}$$

where $a$ and $b$ are empirical coefficients dependent on grain size. Everaert (1991) confirmed the above equation for flows without simultaneous rainfall, obtaining values of $b$ from 1.5 to 3.5 for particles with median grain diameters ($D_{50}$) of 33 and 390 $\mu$m respectively. The impact of rainfall had a negligible influence on the relationship for fine particles but reduced the exponent for coarser particles to 1.5, indicating that rainfall diminishes the ability of overland flow to transport coarse material.

Instead of trying to define transport capacity only in terms of flow properties, some researchers have attempted to relate transport capacity to the maximum sediment concentration that a flow can carry when a balanced condition exists between detachment and deposition (Rose et al. 1983; Styczen & Nielsen 1989). The rate of deposition ($D_p$) is:

$$D_p = v_s \cdot C \tag{2.29}$$

where $v_s$ is the settling velocity of the particles (Proffitt et al. 1991). Torri and Borselli (1991) took data from the experiments of Govers (1990) and obtained a good agreement between the transport capacity of the flow estimated from a balance-based approach and that estimated from eqn 2.28, indicating that the latter is a reasonable expression of the transport capacity of flow.

Given the rather shallow depths of overland flow, the considerable role played by surface roughness and the generally low Reynolds and Froude numbers, it can be proposed that most of the sediment transported is derived by raindrop impact and that, except on steep slopes or on smooth bare soil surfaces, grain shear velocity rarely attains the level necessary to detach soil particles. Since, as seen earlier, particles between 0.063 and 0.250 mm in size are the most detachable by raindrop impact and, from Figs 2.2 and 2.3, it can be seen that the most detachable particles by flow are within the 0.1–0.3 mm range, the sediment carried in overland flow is deficient in particles larger than 1 mm and enriched in finer material. Thus, over time, areas of erosion on a hillside will become progressively sandier and areas of deposition, particularly in valley floors, will be enriched with clay particles.

The trend towards increasing sandiness in eroded areas is also brought about by another mechanism. Most of the sediment splashed into the flow is moved only relatively short distances before being deposited. Since deposition is a particle-size selective process, with the coarser particles being deposited first, the deposited layer becomes progressively coarser (Proffitt et al. 1991) and, as seen in section 2.2, may develop into a depositional crust. Less of the finer material is then exposed to erosion. This mechanism can take place even within an individual storm so that detachment is highest at the beginning of the storm and transport capacity is reached very quickly.

Plots of the relationship between sediment transport by overland flow and discharge, as measured in the field, do not always conform with those expected from the research described above. Work in Bedfordshire, England (Jackson 1984; Morgan et al. 1986), and in southern Italy (van Asch 1983) shows that sediment transport varies with discharge raised by a power of 0.6–0.8. The similarity of this value to that in equations for bed load transport in rivers implies that the transport process is dominantly one of rolling of the particles over the soil surface as bed load. Kinnell (1990) considers that the sediment component contributed to overland flow by raindrop impact is moved as bed load and the component contributed through detachment by the flow itself is moved as suspended load. This implies that sediment transport may be better expressed by eqn 2.20 on low slopes where soil particle detachment is solely by rainsplash but by eqn 2.27 on steeper slopes or with higher flow velocities when particle detachment by flow also takes place. It is also likely that the process is extremely dynamic, so that the most relevant equation for describing sediment transport continually changes through time.

### 2.3.4 Spatial distribution

The dynamic nature of the process is even more apparent when the spatial extent and distribution of overland flow over a hillside is considered. Horton (1945) described overland flow as covering two-thirds or more of the hillslopes in a drainage basin during the peak period of the storm. He viewed overland flow as being the result of the rainfall intensity exceeding the infiltration capacity of the soil, with the following pattern of distribution over land surface. At the top of the slope is a zone without flow, which forms a belt of no erosion. At a critical distance from the crest sufficient water accumulates on the surface for flow to begin. Moving further downslope, the depth of flow increases with distance from the crest until, at a further critical distance, the flow becomes concentrated into fewer and deeper flow paths, which occupy a progressively smaller proportion of the hillslope (Parsons et al. 1990). Hydraulic efficiency improves, allowing the increased discharge to be accommodated by a higher flow velocity. Nevertheless, the hydraulic characteristics of the flow vary greatly over very short distances because of the influence of bed roughness associated with vegetation and stones. As a result, erosion is often localized and after a rainstorm the surface of a hillside displays a pattern of alternating scours and sediment fans (Moss & Walker 1978). Eventually, the flow breaks up into rills. That overland flow occurs in such a widespread fashion has been questioned, particularly in well vegetated areas where such flow occurs infrequently and covers only that 10–30 per cent of the area of a drainage basin closest to the stream sources (Kirkby 1969a). Under these conditions its occurrence is more closely related to the saturation of the soil and the fact that moisture storage capacity is exceeded, rather than infiltration capacity. Although, as illustrated by the detailed studies of Dunne and Black (1970) in a small forested catchment in Vermont, the saturated area expands and contracts, being sensitive to heavy rain and snow melt, rarely can erosion by overland flow affect more than a small part of the hillslopes.

Since most of the observations testifying to the power of overland flow relate to semi-arid areas or to cultivated land with sparse plant cover, it would appear that vegetation is the critical factor. Some form of continuum exists, ranging from well vegetated areas where overland flow occurs rarely and is mainly of the saturation type, to bare soil where it frequently occurs and is of the Hortonian type. Removal of the plant cover can therefore enhance erosion by overland flow. The change from one type of overland flow to another results from more rain reaching the



ground surface, less being intercepted by the vegetation and decreased infiltration as rainbeat and deposition of material from the flow cause a surface crust to develop. Exceptions to this trend occur in areas where high rainfall intensities are recorded. Hortonian overland flow is widespread in the tropical rain forests near Babinda, northern Queensland, where six-minute rainfall intensities of 60–100 mm h$^{-1}$ are common, especially in the summer, and the saturated hydraulic conductivity of the soil at 200 mm depth is only 13 mm h$^{-1}$. As a result, a temporary perched water table develops in the soil soon after the onset of rain, subsurface flow commences and this quickly emerges on the soil surface (Bonell & Gilmour 1978).

Where runoff rates are relatively high over most of the hillside, overland flow, or, more strictly, the combined action of overland flow and raindrop impact as interrill erosion, can be the dominant erosion process on the upper and middle slopes, with deposition of material as colluvium on the footslopes. This appears to be true for many agricultural areas on non-cohesive soils. On loose, freshly ploughed soils on colluvial deposits on 18–22° slopes in Calabria, Italy, van Asch (1983) found that overland flow accounted for 80–95 per cent of the sediment transport. On unvegetated sandy soils in Bedfordshire with an 11° slope, it accounts for 50–80 per cent (Morgan et al. 1986). On a loam soil on a 14° slope in northern Belgium it accounts for 22–46 per cent of the total soil loss, with rates ranging from 24 to 100 t ha$^{-1}$ (Govers & Poesen 1988). Interrill processes can also be the main agent of erosion on well vegetated slopes if the rainfall is very high.

## 2.4 Subsurface flow

The lateral movement of water downslope through the soil is known as interflow. Where it takes place as concentrated flow in tunnels or subsurface pipes its erosive effects through tunnel collapse and gully formation are well known. Less is known about the eroding ability of water moving through the pore spaces in the soil, although it has been suggested that fine particles may be washed out by this process. Pilgrim and Huff (1983) measured sediment concentrations as high as 1 g l$^{-1}$ in subsurface flow through a silt-loam soil on a 17° slope under grass in California in storms of 10 mm h$^{-1}$ intensity or less. The material, uniformly fine with particles ranging from 4 to 8 $\mu$m in diameter, was being detached by raindrop impact at the surface and then moved by the flow through the macropores in the soil. Under tropical rain forest on 8–14° slopes at Pasoh, Negeri Sembilan, Malaysia, where erosion rates are very low, the material removed as dissolved solids in the subsurface flow can amount to 15–23 per cent of the total sediment transport (Leigh 1982).

Subsurface flow is enhanced where subsurface drainage systems have been installed, which can then serve as important pathways for sediment movement. On a 30.6 ha catchment with silty clay loam soils at Rosemaund, Herefordshire, flows through tile drains account for up to 50 per cent of the annual sediment loss of 0.8 t ha$^{-1}$ (Russell et al. 2001). More important than the sediment concentrations, however, are the concentrations of base minerals, which can be twice those found in overland flow. Essential plant nutrients, particularly those added by fertilizers, can be removed, thereby impoverishing the soil and reducing its resistance to erosion. In the Syvbroek catchment, Denmark, 58 per cent of the total phosphorus delivered annually to the water course comes from subsurface drains (Hansen 1990).

## 2.5 Rill erosion

As indicated earlier, it is widely accepted that rills are initiated at a critical distance downslope, where overland flow becomes channelled. The break-up of overland flow into small channels or microrills was examined by Moss et al. (1982). They found that, in addition to the main flow path downslope, secondary flow paths developed with a lateral component. Where these converged, the increase in discharge intensified particle movement and small channels or trenches were cut by scouring. Studies of the hydraulic characteristics of the flow show that the change from overland flow to rill flow passes through four stages: unconcentrated overland flow; overland flow with concentrated flow paths; microchannels without headcuts; and microchannels with headcuts (Merritt 1984). The greatest differences exist between the first and second stages, suggesting that the flow concentrations within the overland flow should strictly be treated as part of an incipient rill system. In the second stage, small vortices appear in the flow and, in the third stage, develop into localized spots of turbulent flow characterized by roll waves (Rauws 1987) and eddies (Savat & De Ploey 1982). At the point of rill initiation, flow conditions change from subcritical to supercritical (Savat 1979). The overall change in flow conditions through the four stages seems to take place smoothly as the Froude number increases from about 0.8 to 1.2, rather than occurring when a threshold value is reached (Torri et al. 1987b; Slattery & Bryan 1992). For this reason, attempts to explain the onset of rilling through the exceedance of a critical Froude number have not been successful and additional factors have had to be included when defining its value. Examples are the particle size of the material (Savat 1979) and the sediment concentration in the flow (Boon & Savat 1981).

Greater success has been achieved relating rill initiation to the exceedance of a critical shear velocity of the runoff (eqn 2.17). Govers (1985) found that on smooth or plane surfaces, where all the shear velocity is exerted on the soil particles, the sediment concentration in the flow increased with shear velocity more rapidly once a critical value of about 3.0–3.5 cm s$^{-1}$ was reached. At this point, the erosion becomes non-selective regarding particle size, so that coarser grains can be as easily entrained in the flow and removed as finer grains. A value of about 3.5 cm s$^{-1}$ for the critical shear velocity only applies to non-cohesive soils or soils, which, because they are highly sensitive to dispersal or to liquefaction, resemble loose sediments. Rauws and Govers (1988) proposed that, except for soils with high clay contents, the critical shear velocity for rill initiation ($u_{*crit}$) is linearly related to the shear strength of the soil ($\tau_s$) as measured at saturation with a torvane:

$$u_{*crit} = 0.89 + 0.56\tau_s \tag{2.30}$$

Using a shear vane, the equivalent equation is (Brunori et al. 1989):

$$u_{*crit} = 0.9 + 0.3\tau_s \tag{2.31}$$

An alternative but similarly conceived approach relates rill initiation to a critical value of the ratio between the shear stress exerted by the flow ($\tau$) and the shear strength of the soil ($\tau_s$) measured with a shear vane. When $\tau/\tau_s > 0.0001$–$0.0005$, rills will form (Torri et al. 1987a). In all these relationships, it should be stressed that the shear velocity or shear stress is applied wholly to the soil particles and should be strictly known as the grain shear velocity or the grain shear stress.



Once rills have been formed, their migration upslope occurs by the retreat of the headcut at the top of the channel. The rate of retreat is controlled by the cohesiveness of the soil, the height and angle of the headwall, the discharge and the velocity of the flow (De Ploey 1989a). Downslope extension of the rill is controlled by the shear stress exerted by the flow and the strength of the soil (Savat 1979). Shear stress also determines the rate of detachment of soil particles by flow within the rill, which can be broadly described by an equation of the type (Foster 1982):

$$Df = K_r(\tau - \tau_c) \tag{2.32}$$

where $K_r$ is a measure of the detachability of the soil and $\tau_c$ is the critical shear stress for the soil.

The transport capacity of rill flow can be approximately represented by eqns 2.27 or 2.28. Govers (1992) found experimentally that flow velocity in rills could be related to the discharge by the relationship:

$$v = 3.52Q^{0.294} \tag{2.33}$$

and that this gave better predictions than the Manning equation (eqn 2.18). This was because over a range from 2 to 8° slope had no effect on flow velocity; neither did the grain roughness or the surface form of the soil. Govers (1992) therefore modified eqn 2.28 by replacing the velocity term with eqn 2.33 to read:

$$C_{max} = a(3.52Q^{0.294} - 0.0074)Q \tag{2.34}$$

where $a$ is dependent upon the grain size of the sediment and the value of 0.0074 is interpreted as the critical value of unit stream power. Although this equation expressed the maximum sediment concentration that can be carried by runoff in a rill, the actual sediment concentration and, therefore, the erosion may vary considerably from this. This is because the supply of sediment to the rill is not solely dependent upon the detachment of soil particles by the flow. Instead, the rill has to adjust continually for pulse influxes of sediment due to wash-in from interrill flow on the surrounding land, erosion and collapse of the head wall and collapse of the side walls. Mass failure of the side walls can contribute more than half of the sediment removed in rills, particularly when heavy rains follow a long dry period during which cracks have developed in the soil (Govers & Poesen 1988).

Since raindrop impact increases the transport capacity of the flow and, through the detachment of soil particles, causes higher sediment concentrations, Savat (1979) argued that the interaction of rainfall with the flow would enhance the probability of rilling. Quansah (1982) and Dunne and Aubry (1986), however, found that the particles detached by rain filled in the microchannels as fast as they could form, so that rilling was inhibited. It appears that the two sets of processes compete so that either the microchannels are short-lived because they drain away the overland flow, become laterally isolated and fill in, or the concentration of flow increases its erosive power and the channels deepen, widen and migrate both upslope and downslope.

Since rill flow is non-selective in the particle size it can carry, large grains, even rock fragments up to 9 cm in diameter (Poesen 1987), can be moved. Meyer et al. (1975) found that 15 per cent of the particles carried in rills on a 3.5° slope of tilled silt loam were larger than 1 mm in size and that 3 per cent were larger than 5 mm. On a 4.5° slope of bare untilled silt loam, 80 per cent of the sediment transported in rills was between 0.21 and 2.0 mm in size and most of the clay particles were removed as aggregates within this size range (Alberts et al. 1980).

As expected from its considerable erosive power, rill erosion may account for the bulk of the sediment removed from a hillside, depending on the spacing of the rills and the extent of the area affected. Govers and Poesen (1988) found that the material transported in rills accounted for 54–78 per cent of the total erosion. In a cloud burst on 23 May 1958 near Banská Bystrica, in the Czech Republic, rills accounted for 70 per cent of the total sediment eroded (Zachar 1982). These figures contrast with the situation in mid-Bedfordshire, England, where rills contributed only 20–50 per cent of the total erosion (Morgan et al. 1986) and in the Hammerveld-1 catchment in the Belgian loess, where rills accounted for one-third of the erosion over a three-year period (Vandaele & Poesen 1995).

## 2.6 Gully erosion

Gullies are relatively permanent steep-sided water courses that experience ephemeral flows during rainstorms. Compared with stable river channels, which have a relatively smooth, concave-upwards long profile, gullies are characterized by a headcut and various steps or knick-points along their course. These rapid changes in slope alternate with sections of very gentle gradient, either straight or slightly convex in long profile. Gullies also have relatively greater depth and smaller width than stable channels, carry larger sediment loads and display very erratic behaviour, so that relationships between sediment discharge and runoff are frequently poor (Heede 1975a). A widely recognized definition used to separate gullies from rills is that gullies have a cross-sectional area greater than 1 $m^2$ (actually 929 $cm^2$) (Poesen 1993). Gullies are almost always associated with accelerated erosion and therefore with instability in the landscape.

### 2.6.1 Gully formation

At one time it was thought that gullies developed as enlarged rills but studies of the gullies or *arroyos* of the southwest USA revealed that their initiation is a more complex process. In the first stage small depressions or knicks form on a hillside as a result of localized weakening of the vegetation cover by grazing or fire. Water concentrates in these depressions and enlarges them until several depressions coalesce and an incipient channel is formed. Erosion is concentrated at the heads of the depressions, where near-vertical scarps develop over which supercritical flow occurs. Some soil particles are detached from the scarp itself but most erosion is associated with scouring at the base of the scarp, which results in deepening of the channel and the undermining of the headwall, leading to collapse and retreat of the scarp upslope. Sediment is also produced further down the gully by bank erosion. This occurs partly by the scouring action of running water and the sediment it contains and partly by slumping of the banks. Between flows sediment is made available for erosion by weathering and bank collapse. This sequence of gully formation, described by Leopold et al. (1964) in New Mexico, is shown in Fig. 2.4.

Not all gullies develop purely by surface erosion, however. Berry and Ruxton (1960), investigating gullies in Hong Kong that formed following clearance of natural forest cover, found that most water was removed from the hillsides by subsurface flow in natural pipes or tunnels, and when heavy rain provided sufficient flow to flush out the soil in these, the ground surface subsided, exposing the pipe network as gullies. Numerous studies record the formation of gullies by pipe collapse in many different materials and climatic environments. The essential requirements are steep hydraulic gradients in a soil of high infiltration capacity through macropores but



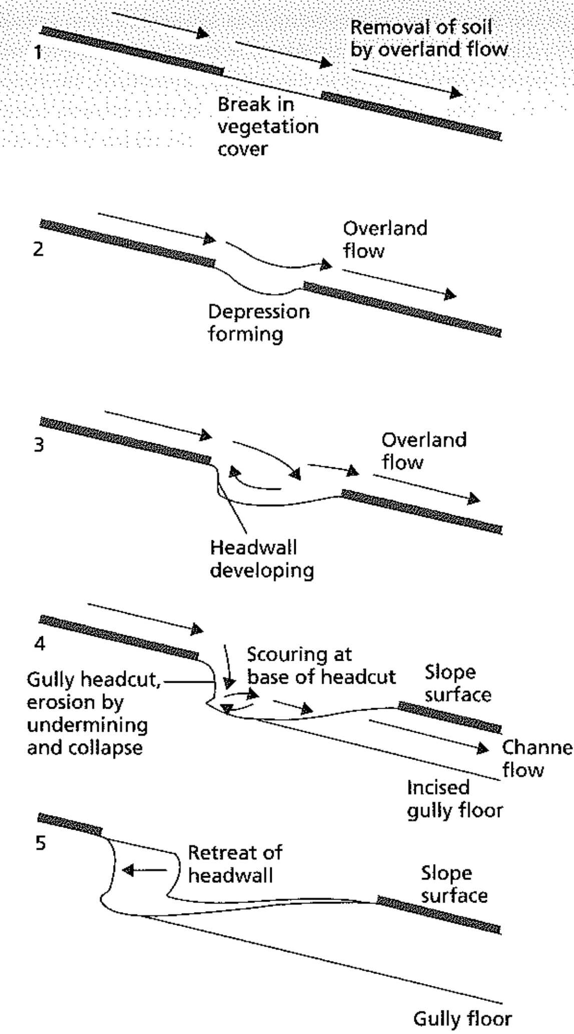



**Fig. 2.4** Stages in the surface development of gullies on a hillside (after Leopold et al. 1964).

low intrinsic permeability, so that water does not move readily into the matrix (Crouch 1976; Bryan & Yair 1982). Suitable soils include those prone to cracking as a result of high sodium absorption, shrinkage on drying or release of pressure following unloading of overlying material.

Tunnel erosion has been widely reported in many hilly and rolling areas of Australia, where it is associated with duplex soils. These are soils characterized by a sharp increase in clay content between the A and B horizons so that the upper layer, 0.03–0.6 m in depth, varies from a loamy sand to a clay loam and the lower layer ranges from a light to heavy clay. According to Downes (1946), overgrazing and removal of vegetation cover cause crusting of the surface soil, resulting in greater runoff. This passes into the soil through small depressions, cracks and macropores but, on reaching the top of the B horizon, moves along it as subsurface flow. Localized dispersion of the clays in areas of subsurface moisture accumulation is followed by piping. Heavy summer rains cause the water in the pipes to break out on to the surface. Eventually, the roofs of the pipes collapse and gullying occurs. In the Loess Plateau of China, tunnel erosion contributes 25–30 per cent of the catchment sediment yield (Zhu 2003). Most of this relates to the development of new tunnels rather than enlargement of existing ones and takes place in single storms with return periods of 50 years or more. In a storm of 107 mm over 7.5 hours, 67 new tunnels were created; another storm of 37 mm in 115 minutes resulted in the formation of 123 new inlets. Sediment concentration in tunnel flow ranges from 8.2 to 893.2 $g\,l^{-1}$, which is very similar to that in channel flow (Zhu et al. 2002). Once formed, the tunnels continue to erode both during and between storms through earth falls, slumps, water erosion and roof collapse; in some cases, slumps and earth falls lead to their temporary blockage.

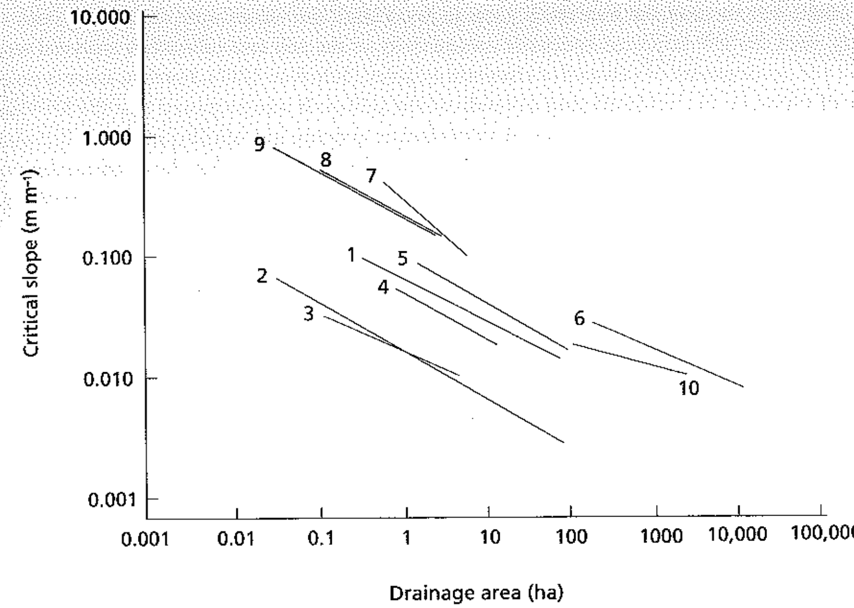
A third way in which gullies are initiated is where linear landslides leave deep, steep-sided scars, which may be occupied by running water in subsequent storms. This type of gully development has been described in Italy by Vittorini (1972) and in central Värmland, Sweden, by Fredén and Furuholm (1978). In the latter case, a 3–20 m deep, 20–40 m wide and 100 m long gully formed in glaciofluvial deposits following the 1977 spring snowmelt, which caused the removal, as a mass flow, of 20,000 $m^3$ of saturated silt and sand in less than three days.

### 2.6.2 Threshold conditions

The main cause of gully formation is too much water, a condition that may be brought about by either climatic change or alterations in land use. In the first case, increased runoff may occur if rainfall increases or if less rainfall produces a reduction in the vegetation cover. In the second case, deforestation, burning of the vegetation and overgrazing can all result in greater runoff. If the velocity or tractive force of the runoff exceeds a critical or threshold value, gullying will occur. The threshold values, however, can show a very wide range. For example, $\tau_c$ varies from 3.3 to 32.2 Pa on cultivated silt-loam soils in Belgium and from 16.8 to 74.4 Pa on cultivated stony sandy loams in the Alentejo area of Portugal (Nachtergaele 2001).

Where the exceedance of the threshold relates to changes in climate or land use, the threshold is described as extrinsic (Schumm 1979) because the changes are external to the processes operating within the gully. Attempts to relate gullying solely to changes in external factors have not proved entirely successful, however, because not all gullies in an area appear to respond in the same way. In order to explain the onset of instability in one gully while its neighbours remain stable, Schumm (1979) examined the role of intrinsic thresholds, which are related to the internal working of the gully. From a review of studies in Wyoming, Colorado, New Mexico and Arizona, a discriminant function was established between stable and unstable conditions in terms of the size of the catchment area ($a$), which controls discharge, and the channel slope ($s$), which controls the velocity of runoff. When, for a given catchment area, channel slope exceeds a critical value, incision occurs, creating a channel characterized by one or more head scarps. Subsequent scouring causes the gully to become very active: the channel widens, deepens and extends headwards. Over time, the channel slope is reduced, promoting a consolidation phase as the gully stabilizes, the channel fills in, the sides and head wall become flatter and vegetation regrows. Deposition steepens the slope again and triggers a new phase of gullying. Thus gullies pass through successive cycles of erosion and deposition. It is not uncommon for the head of a gully to be extremely active while the lower section of the gully is stabilizing or for gullies to contain a sequence of alternating stable and unstable sections.






1, Central Belgium; 2, Central Belgium; 3, Portugal; 4, France; 5, United Kingdom (South Downs); 6, Colorado, USA; 7, Sierra Nevada, USA; 8, California, USA; 9, Oregon, USA; 10, New South Wales, Australia

**Fig. 2.5** Relationship between critical slope and drainage area for development of gullies (after Poesen et al. 2003).

Begin and Schumm (1979) and Moore et al. (1988) have established critical $s$–$a$ relationships to the effect that gullies form when:

$$sa^b > t \tag{2.35}$$

where $t$ is the threshold value. Threshold values are higher for non-cultivated land than for cropland and also vary with the type of vegetation cover, differences in soil structure and soil moisture and type of tillage (Poesen et al. 2003). The threshold values also depend on the methodology. The lines plotted in Fig. 2.5 are best-fit regression lines passing through data obtained for gullied catchments. As a result, some gullied catchments will plot below the lines. Begin and Schumm (1979), however, proposed defining the line below which gullied catchments did not occur and which could therefore be interpreted as defining the condition at which all valley floors were stable; there will, however, be some ungullied catchments, which will plot above the line. More recently, Morgan and Mngomezulu (2003) showed that discriminant functions relating $s$ to $a$ could be used to separate gullied from ungullied catchments in four areas of Swaziland. Out of a sample of 201 catchments, only 8 per cent were incorrectly classified by this method.

The value of $b$ in eqn 2.35 is generally interpreted in relation to the processes operating in the catchment. Values >0.2 are associated with erosion by surface runoff and those <0.2 as indicating subsurface processes and mass movement (Montgomery & Dietrich 1994; Vandekerckhove et al. 2000). For three of the four study areas in Swaziland, the values were <0.2, which was surprising, since the low saturated hydraulic conductivity of the soil and subsurface material would inhibit the development of subsurface channels or pipes (Scholten 1997). Further, there is no evidence of piping in the headwalls and sidewalls of the gullies. It was proposed that the low values may reflect the increasing importance of groundwater seepage in contributing to undermining and collapse of the headwalls in the later stages of gully evolution after the channels have cut down to bedrock. This change in dominant process over time was observed in similar gully systems in Madagascar (Wells & Andriamihaja 1991).

### 2.6.3 Valley floor gullies

Valley floor gullies generally take the form of ephemeral gullies developed in topographic swales in the landscape where runoff concentrates during heavy rains. They occur particularly where the surrounding hillslopes are convexo-concave, most of the land is under arable farming, the soils are either freshly tilled and loose or crusted and peak discharges reach several cumecs ($m^3 s^{-1}$) (Poesen 1989). They can also form when runoff from either exceptional rainfall or snowmelt occurs over frozen subsoils (Øygarden 2003). The channels can be several metres deep but generally they are limited in depth to no more than 25–30 cm by an underlying plough pan. They are essentially a surface phenomenon formed when the tractive force exerted by the runoff exceeds the resistance of the soil. Once formed, however, tractive force plays only a minor role in their further development, for which the main mechanism is the retreat of the headwall. In loess soils, this occurs mainly by slab failure as the wall is undermined by plunge-pool erosion and basal sapping, and vertical tension cracks develop on the slope above. Although head wall collapse may be the most important source of sediment, as in the gullies in western Iowa (Bradford & Piest 1980), measurements southwest of Sydney, Australia, show that erosion of the sidewalls subsequent to retreat of the headcut is much more important in contributing sediment to the gullies than the process of incision by the headcut itself (Blong et al. 1982).

Even though flow along the floor of the gully may contribute very little sediment compared to head wall and side wall processes, it is still important in removing sediment from the gully, without which the gully would fill in and stabilize. Stabilization can often be temporary, however, with the channels reforming in the same place in subsequent storms. Provided the material is flushed out, the gully can continue to retreat headwards until the area upslope contributing runoff to the gully head decreases sufficiently for the *s–a* threshold to fall below the critical value (Nachtergaele et al. 2002).

### 2.6.4 Valley side gullies

Valley side gullies develop more or less at right angles to the main valley line where local concentrations of surface runoff cut into the hillside, subsurface pipes collapse or local mass movements create a linear depression in the landscape. Valley side gullies may be continuous – that is, they discharge into the river at the bottom of the slope – or discontinuous, fading out into a depositional zone and not reaching the valley floor. Once formed, they can grow upslope by headward retreat and downslope by incision of the channel floor.

Among the most spectacular valley side gullies are the *lavakas* found in the deeply weathered hills of the central part of Madagascar. They have developed on convex hillsides with basal slopes of 30–45° adjacent to wide flat-floored valleys. Under the average rainfall of 1000–2000 mm per year and maximum mean monthly temperatures of 18–26°C, the underlying metamorphic rocks have been chemically weathered to form a 5–25 m deep cover of saprolite. According to Wells and



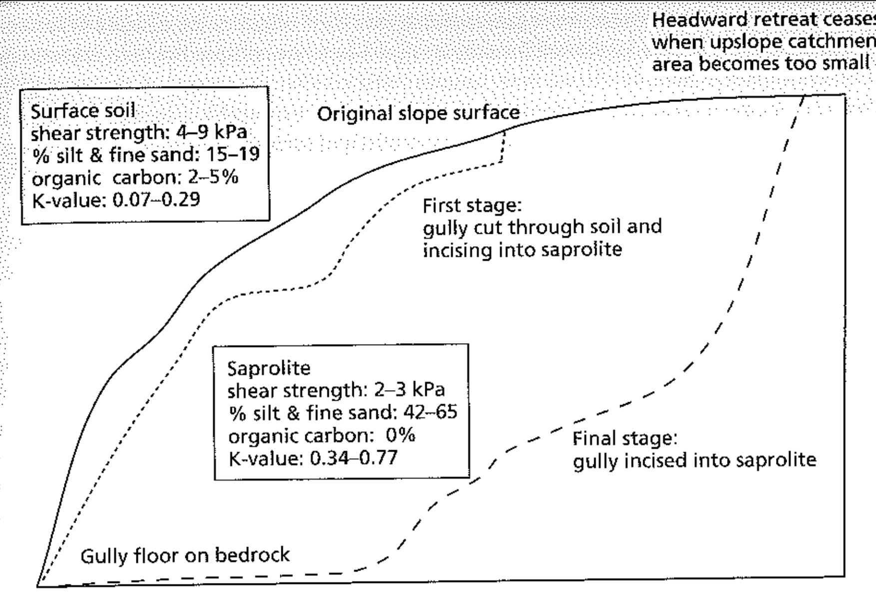



**Fig. 2.6** Gully formation on deeply weathered saprolite in Swaziland. Note the high erodibility of the saprolite compared to the surface soil.

Andriamihaja (1991), gullies form where the vegetation is destroyed, probably by grazing, or along paths and tracks. Infiltration capacities in these areas are reduced to <50 mm $h^{-1}$ compared with >300 mm $h^{-1}$ on the surrounding land. With ten-minute rainfall intensities of 80–100 mm $h^{-1}$, sometimes 200–500 mm $h^{-1}$, runoff is quickly produced with sufficient force to develop rills within about 20–30 m of the crest, just where the slope steepens rapidly into the convexity. The rills enlarge over time, develop head scarps and cut down into the underlying saprolite. Once this material is reached, the rate of downcutting, particularly in the head area, is very rapid until the channel reaches the groundwater zone above the bedrock. Thereupon, downcutting ceases but the gully erosion continues to expand by sapping, undermining and collapse of the headwall and side walls (Wells et al. 1991). Since this occurs in the area of greatest relief, the gully develops its typical pear-shaped plan form, broadest towards the top of the slope and with a narrow outlet downslope. In this stage, upslope runoff contributes little, if anything, to the gully growth, whereas localized concentrations of groundwater enhance headward bifurcation and cause the perimeter of the gully to take on a feather-edged appearance when viewed in plan. Figure 2.6 shows the development of similar gully systems in Swaziland. Compared to the surface soil, the underlying saprolite has a high content of silt and very fine sand, no organic matter and a low shear strength (Scholten 1997), which means it has a much higher erodibility. These characteristics, combined with the depth of material into which downcutting can occur, explain why gully erosion is so rapid once the more resistant surface soil has been breached. Similar patterns of valley side gully development have been observed in the Appalachian Piedmont, southeast USA (Ireland et al. 1939), southeast Brazil (de Meis & de Moura 1984) and Zimbabwe (Whitlow & Bullock 1986).

In the examples just described, there is no evidence of piping. Yet piping is often associated with valley side gullies, even though the relationship can be complex. At Springdale in the

Riverina area of New South Wales, some gullies clearly owe their origin to the collapse of tunnels, as evidenced by the remnants of the soil forming bridges where the collapse has been incomplete, while others appear to have developed from tunnels initiated after the cutting of the main gully. Developed along permeable zones between layers of relatively impermeable material, they depend on the incision of the gully to expose the permeable seams and form the hydraulic gradient along which water can move (Crouch et al. 1986).

Valley side gullies can often evolve into badlands, particularly where a fragile environment has been disturbed by unwise land use (López-Bermúdez & Romero-Diaz 1989). Marl and clay soils are particularly susceptible to badland development, as are soils with high contents of gypsum, since these are prone to piping. The processes responsible for badlands depend on how the soil responds to increases in moisture. Where the soil seals quickly, the badlands develop by surface erosion through rills and gullies. Where the soil has a high infiltration rate, water collects within the soil, resulting in instability and rapid shallow mass movements (Bouma & Imeson 2000; Moretti & Rodolfi 2000). Once formed, badlands take a long time to heal and many of those found in the Mediterranean, although very spectacular, record rather low rates of erosion today and are relics of former periods of active gullying. Those of the Guadix Basin in southeast Spain were probably developed around 2000 BC (Wise et al. 1982). In contrast, in areas of active badland development annual erosion rates can exceed 260 $\mathrm{t\,ha^{-1}}$ (Cervera et al. 1990).

### 2.6.5 Erosion rates

Although gullies can remove vast quantities of soil, gully densities are not usually greater than 10 $\mathrm{km\,km^{-2}}$ and the surface area covered by gullies is rarely more than 15 per cent of the total area (Zachar 1982). This results in a considerable contrast between the erosion rate for an individual gully and its contribution to the overall soil loss of an area. Rates of headwall extension can be very rapid for relatively short periods of time. Measurements on the Mbothoma Gully system, Swaziland, showed very rapid retreats of between 2.5 and 6.3 $\mathrm{m\,yr^{-1}}$ from 1947 to 1961, followed by a slowing down to 0.13–0.51 $\mathrm{m\,yr^{-1}}$ for the period 1961–1980 (WMS Associates 1988). A new gully opposite Mbothoma developed in the 1960s and up to 1990 eroded headwards at a rate of 14 $\mathrm{m\,yr^{-1}}$; the rate then decreased to 5 $\mathrm{m\,yr^{-1}}$ between 1990 and 1998 (Sidorchuk et al. 2003). Erosion from a gully developed on arable land near Cromer, Norfolk, England in 1975 was estimated at 195 $\mathrm{t\,ha^{-1}}$ (Evans & Nortcliff 1978) and erosion from a winter runoff event in southern Norway in January 1990 exceeded 100 $\mathrm{t\,ha^{-1}}$ in many gullies (Øygarden 2003). In gullies near Bathurst, New South Wales, soil loss from the side walls alone amounted to 1100 $\mathrm{t\,ha^{-1}}$ over a three-year period beginning April 1984 (Crouch 1990a). These figures contrast with annual rates for whole catchments for which typical figures include 3–5 $\mathrm{t\,ha^{-1}}$ for the gullied watersheds at Treynor, Iowa (Bradford & Piest 1980), 3–16 $\mathrm{t\,ha^{-1}}$ for the *lavakas* in Madagascar (Wells & Andriamihaja 1991) and 6.4 $\mathrm{t\,ha^{-1}}$ for a watershed near Gilgranda, New South Wales, of which 60 per cent came from gully heads (Crouch 1990b).

In a review of worldwide data, Poesen et al. (2003) showed that gully erosion can contribute between 10 and 94 per cent of overall soil loss from an area, with values between 30 and 75 per cent being typical. The contribution of gullies to total erosion is therefore not easily predictable. It depends on the characteristics of the storm, the topography of the catchment and the land cover at the time the storm occurs. In the 94-hectare Blosseville catchment, Normandy, France, a rain storm of 60 mm on 26 December 1999 with a maximum six-minute intensity of 55 $\mathrm{mm\,h^{-1}}$ caused erosion of 10 $\mathrm{t\,ha^{-1}}$. Some 93 per cent of the catchment had less than 20 per cent vegetation cover. Ephemeral gullies contributed 24 per cent of the erosion. A 60 mm storm with a maximum six-minute intensity of 105 $\mathrm{mm\,h^{-1}}$ in the same catchment on 9 May 2000 produced a much lower



erosion of 1 $\mathrm{t\,ha^{-1}}$ because some 73 per cent of the catchment had a vegetation cover greater than 60 per cent. However, 83 per cent of the erosion occurred in ephemeral gullies (Cerdan et al. 2002a).

## 2.7 Mass movements

Although mass movement has been widely studied by geologists, geomorphologists and engineers, it is generally neglected in the context of soil erosion. Yet Temple and Rapp (1972) found that in the western Uluguru Mountains, Tanzania, landslides and mudflows are the dominant erosion processes. They occur in small numbers about once every ten years. The quantity of sediment moved from the hillsides into rivers by mass movement is far in excess of that contributed by gullies, rills and overland flow. Further, less than 1 per cent of the slide scars are in areas of woodland, 47 per cent being on the cultivated plots and another 47 per cent on land lying fallow. The association of erosion with woodland clearance for agriculture is thus very clear. Further evidence for this is provided by Rogers and Selby (1980) in respect of shallow debris slides on clay and silty clay soils derived from greywackes in the Hapuakohe Range, south Auckland, New Zealand. Clearance of the forest for pasture causes a decline in the shear strength of the soil over the five- to ten-year period needed for the tree roots to decay. As a result, landslides under pasture are triggered by a storm with a return period of 30 years, whereas a storm with a 100-year return period is required to produce slides beneath the forest. The close relationship between mass movement and water erosion is illustrated by the studies of the 'bottle slides' in the Uluguru Mountains (Temple & Rapp 1972; Lundgren & Rapp 1974), which develop in areas of large subsurface pipes as a result of the flushing out of a muddy viscous mass of debris and subsequent ground collapse.

The stability of the soil mass on a hillslope in respect of mass movement can be assessed by a safety factor ($F$), defined as the ratio between the total shear strength ($\sigma$) of the soil material along a given shear surface and the amount of stress ($\tau$) developed along that surface. Thus

$$F = \frac{\sigma}{\tau} \tag{2.36}$$

The slope is stable if $F > 1$ and failure occurs if $F \leq 1$. For the simple case of shallow or translational slides, $F$ can be defined as:

$$F = \frac{c' + (\gamma z - \gamma_w h)\cos^2\theta \tan\phi'}{\gamma z \sin\theta \cos\theta} \tag{2.37}$$

where $c'$ is the effective cohesion of the soil, $\gamma$ is the unit weight of soil above the slide plane, $z$ is the vertical depth of soil above the slide plane, $\gamma_w$ is the unit weight of water, $h$ is the height of the groundwater (piezometric) surface above the slide plane, $\theta$ is the slope angle and $\phi'$ is the effective angle of internal friction. Applications of the equation to slides in the Hapuakohe Range show that the value of $F$ is particularly sensitive to changes in $c'$ and $z$. Thus, control measures should be directed at influencing these (Rogers & Selby 1980).

Mass movements, in the varied forms of creep, slides, rock falls and mudflows, are given detailed treatment in numerous books (Sharpe 1938; Zaruba & Mencl 1969; Brunsden & Prior 1984; Anderson & Richards 1987), together with details of how eqn 2.37 can be modified for application to different types of slope failure.

The main factor in wind erosion is the velocity of moving air. Because of the roughness imparted by soil, stones, vegetation and other obstacles, wind speeds are lowest near the ground surface. A plane of zero wind velocity can be defined at some height ($z_0$) above the mean aerodynamic surface. Above $z_0$, windspeed increases exponentially with height so that velocity values plot as a straight line on a graph against the logarithmic values of the height (Fig. 2.7). The change in velocity with height is expressed by the relationship (Bagnold 1941):

$$\bar{v}_z = \frac{2.3}{k} u_* \log\left(\frac{z}{z_0}\right) \tag{2.38}$$

where $\bar{v}$ is the mean velocity at height $z$, $k$ is the von Kármán universal constant for turbulent flow, assumed to equal 0.4 for clear fluids, and $u_*$ is the drag or shear velocity.

Although the movement of soil particles can be related to a critical wind velocity, many workers have attempted to define the conditions more precisely in terms of a critical value of the dimensionless shear stress. Using eqn 2.15 and substituting $\rho_a$ (the density of air) for $\rho_w$ (the density of water), the critical value of $\Theta$, the Shields parameter, for initiating soil particle movement approximates 0.01 (Bagnold 1941), which is much lower than that obtained for water. The difference may be due to the very great difference in the density of the fluids relative to the particle density. A sand grain in air is about 2000 times more massive than the surrounding fluid, whereas it is only about 2.6 times more massive than water (Bagnold 1979). As a result, much higher shear velocities are required to move particles by air and the initial particle motion is more violent. This violence is rapidly transmitted to neighbouring particles, causing a general chain reaction of motion. At this point, grains are dislodged and entrained in the flow relatively easily and therefore at relatively low Shields numbers (Iversen 1985).

Although the Shields coefficient has been successfully applied to particle movement in air (Bagnold 1951; Iversen 1985), for most soil conservation work it is sufficient to relate the detachment of soil particles by wind to a critical value of the shear velocity, using it as a surrogate measure of the drag force exerted by the flow. Shear velocity is directly proportional to the rate of increase in wind velocity with the logarithm of height and is therefore the slope of the line in Fig 2.7(d). Its value can be determined by measuring the wind speed at two heights but, in practice, by assuming $v = 0$ at height $z_0$, it can be obtained by measuring the speed at one height and applying the formula derived from eqn 2.37:

$$u_* = \frac{k}{2.3} \cdot \frac{\bar{v}_z}{\log(z/z_0)} \tag{2.39}$$

For use in estimates of sediment transport by wind, the velocities should be measured within 0.2 m of the ground surface (Rasmussen et al. 1985; Mulligan 1988).

Bagnold (1937) identifies two threshold velocities required to initiate grain movement. The static or fluid threshold applies to the direct action of the wind. The dynamic or impact threshold allows for the bombardment of the soil by grains already in motion. Impact thresholds are about 80 per cent of the fluid threshold velocities in value. In addition to detaching soil particles at a lower threshold velocity, the detachment potential of sediment-laden air is further enhanced by increases of 8–18 per cent in shear velocity close to the ground surface (Sørensen

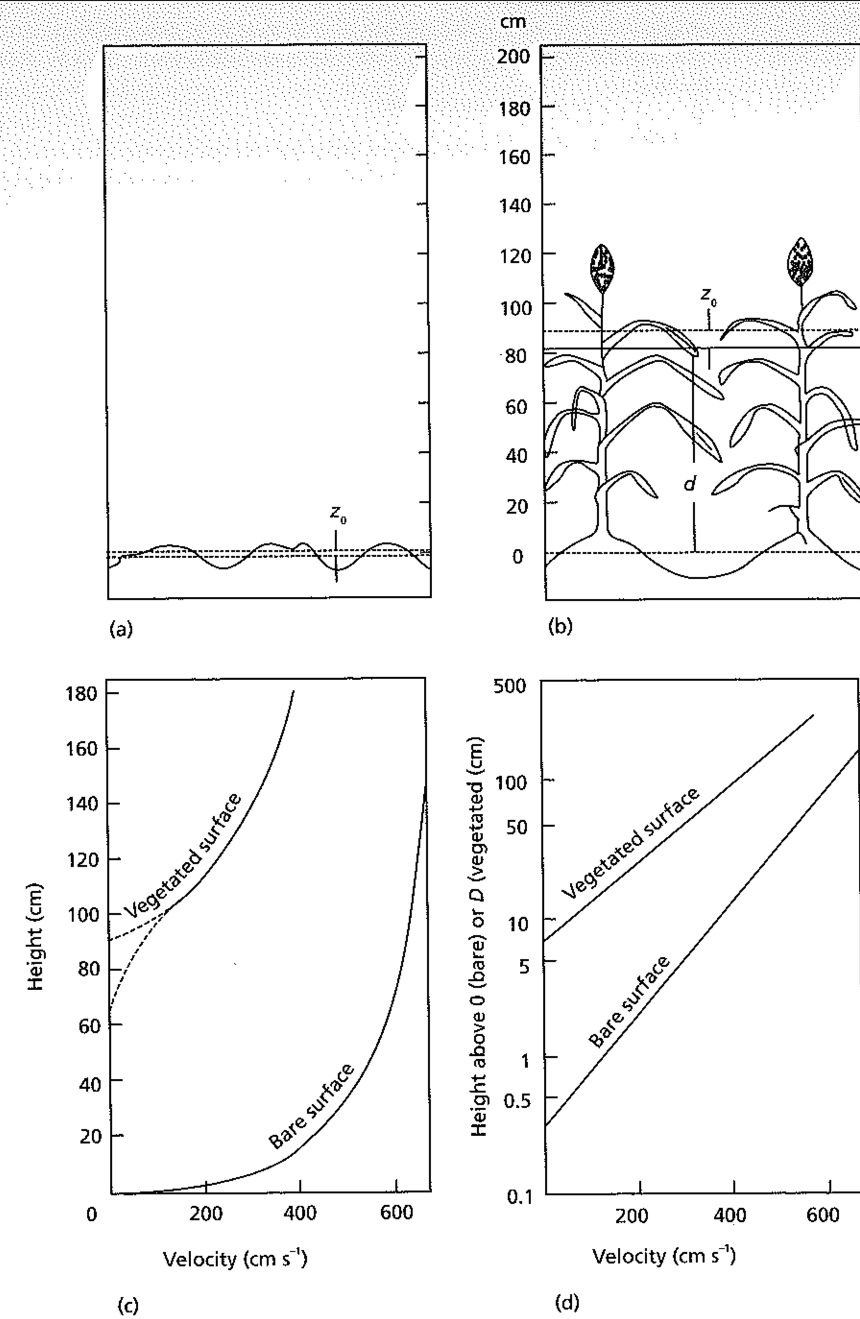



**Fig. 2.7** Wind velocity near a soil surface. (a) Zero wind velocity occurs at a height $z_0$, which lies above the height of the mean aerodynamic surface and below the high points. (b) A crop cover raises the height of the mean aerodynamic surface by a distance $d$ and also increases the value of $z_0$ so that the plane of zero wind velocity ($d + z_0$) occurs at a height that is equal to about 70 per cent of the height of the plants. (c) Wind velocity profiles above a bare surface and a vegetated surface plotted with linear scales. (d) Wind velocity profiles plotted with a logarithmic scale for height (after Troeh et al. 1980).

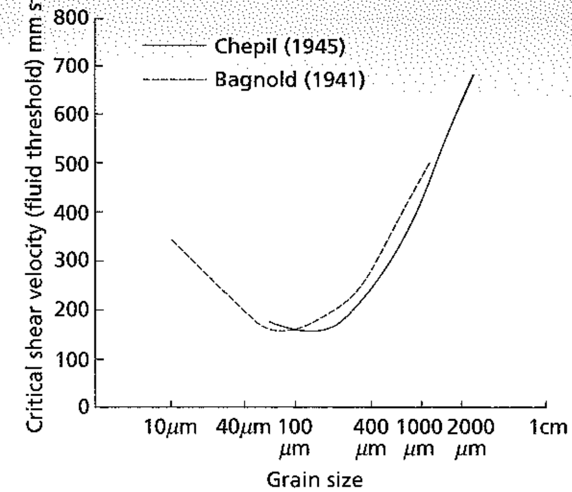



**Fig. 2.8** Critical shear velocities for wind for erosion as a function of particle size (after Savat 1982).

1985; Dong et al. 2002). This arises through the addition of grain-borne shear stress to the air flow. The critical shear velocities vary with the grain size of the material, being least for particles of 0.10–0.15 mm in diameter, and increase with both increasing and decreasing grain size (Chepil 1945). The resistance of the larger particles results from their size and weight. That of the finer particles is due to their cohesiveness and the protection afforded by surrounding coarser grains (Fig. 2.8).

Once in motion, the transport of soil and sand particles by wind takes place in suspension, surface creep and saltation. Suspension describes the movement of fine particles, usually less than 0.2 mm in diameter, high in the air and over long distances. Surface creep is the rolling of coarse grains along the ground surface. Saltation is the process of grain movement in a series of jumps. Initially, drag and lift forces cause a particle to rise with a vertical ejection velocity that is about twice the shear velocity of the air (Willetts & Rice 1985). Drag with the surrounding air quickly reduces the vertical velocity, which is also opposed by the settling velocity of the particle. Once in the air the particle takes on a horizontal velocity, imparted by the wind, so that, while settling back to the ground, it is blown forwards. The result is an overall particle movement or jump comprising a vertical rise to a maximum height at which the settling velocity exceeds the vertical velocity, followed by a falling path at an angle of 6–12° to the horizontal. Individual jump lengths for coarse grains vary from about 60 to 400 mm, increasing with the shear velocity (Sørensen 1985; Willetts & Rice 1988). On striking the ground, the impact energy of the saltating grain is distributed into a disruptive part that causes disintegration of the soil and a dispersive part that imparts a velocity to other soil particles and launches them into the air (Smalley 1970). In a soil blow, between 55 and 72 per cent of the moving particles are carried in saltation.

The rate at which soil particles are dislodged from a bare soil surface was found in wind tunnel experiments (Sørensen 1985; Jensen & Sørensen 1986) to follow the relationship:

$$D_{fa} \propto u_*^{2.8} \tag{2.40}$$

where $D_{fa}$ is the rate of detachment of soil particles in moving air. Although Willetts and Rice (1988) produce comparable data from their experiments, they caution that the overall data base is too small to determine the value of the exponent in eqn 2.39. It probably lies, however, between 2.0 and 3.0.

In contrast, general agreement exists among researchers that the transport capacity of moving air varies with the cube of the shear velocity. Based on this relationship and considering the transport of grains as representing a transfer of momentum from the air to the moving particles, Bagnold (1941) developed the following equation for determining the maximum sediment discharge per unit width ($T_{fa}$):

$$T_{fa} = C(d/D)^{1/2}(\rho_a/g)u_*^3 \tag{2.41}$$

where $C$ is an experimentally derived parameter relating to grain size, $d$ is the average diameter of the material, $D$ is a standard grain diameter of 0.25 mm, $\rho_a$ is the density of the air and $g$ is the gravitational constant. Among the more commonly used expressions of eqn 2.40 is that proposed by Lettau and Lettau (1978):

$$T_{fa} = C(\rho_a/g)u_*^3(1 - u_t/u_z) \tag{2.42}$$

where $T_{fa}$ is in g m$^{-1}$ s$^{-1}$, $C$ is a constant related to grain size and typically is about 6.5 in value, $u_t$ is the impact threshold velocity for particle movement (m s$^{-1}$) and $u_z$ is the wind velocity at height $z$.

It should be stressed that, at best, eqns 2.41 and 2.42 provide estimates of the maximum sediment transport rate (transport capacity) that can occur. They will not necessarily predict the actual rate of sediment transport. Where the soil surface is very resistant, the surface has become armoured by rock fragments or a vegetation cover is present, the equations will overpredict. Under these conditions, sediment transport will vary with the shear velocity of the wind raised to a power of less than 2.0. Further, sediment transport is a highly dynamic process reflecting the considerable short-term fluctuations in wind velocity (Sterk et al. 1998). Periods of sediment transport alternate with periods of no transport, even over a duration as short as ten minutes. During individual wind storms, activity may occur for only 16–21 per cent of the time. The fluctuations in sediment transport most probably relate to short-term changes in the turbulent structure of the storm.

Wind erosion impoverishes the soil and also buries the soil and crops on surrounding land. Although, as already seen, the most erodible particles are 0.10–0.15 mm in size, particles between 0.05 and 0.5 mm are generally selectively removed by the wind. Chepil (1946) found that areas of wind-blown deposits are enriched by particles in the 0.30 and 0.42 mm range. Resistance to wind erosion increases rapidly when primary particles and aggregates larger than 1 mm predominate. If erosion results in armouring of the surface so that more than 60 per cent of the surface material is of this size, the soil is almost totally resistant to wind erosion. Although saltation is the most important process from the viewpoint of soil erosion, wind erosion through the movement of dust particles in suspension can give rise to additional effects of contamination of food and water, aggravation of respiratory diseases, health risks associated with the transport of pathogens and interference with switches in machinery. Wind erosion also creates problems of visibility for road, rail and air transport.

Box 2

## Initiation of soil particle movement

The initial movement of a soil particle by running water or wind takes place when the forces tending to move the particle exceed the forces resisting movement. In simple terms, there are three forces acting on the particle: the vertical or weight force moving the particle vertically downwards towards the ground surface, the lift force tending to make the particle rise vertically and the drag force exerted by the flow moving the particle horizontally along the surface (Fig. B2.1). The relative proportions of the downward, lift and drag forces depend upon the slope angle. On a horizontal surface the weight force predominates and on a steeply sloping surface the lift and drag forces are dominant.

Considering initially only the weight and drag forces, it is possible to define the critical condition at which a particle will move forwards, rotating about its point of contact with its neighbouring particle in a downslope or downwind direction (point A in Fig. B2.1). The condition is defined as the point at which the two forces are equal. Thus:

$$\text{weight force} = \text{drag force} \qquad \text{(B2.1)}$$

$$g(\rho_s - \rho)\frac{\pi}{6}d^3 \tan\phi \cdot \frac{d}{2}\sin\phi = \tau_c \cdot \frac{d}{2}\sin\phi \qquad \text{(B2.2)}$$

where $g$ is the gravity force, $\rho_s$ is sediment density, $\rho$ is the fluid density, $d$ is the particle diameter, $\phi$ is the angle of repose of the soil particles and $\tau_c$ is the critical value of the drag force ($\tau$) for particle movement. For running water, $\tau = \gamma rs$ in which $\gamma$ is the specific weight of the fluid, $r$ is the hydraulic radius and $s$ is the slope. For air flow, $\tau = \rho_a u_*^2$, in which $\rho_a$ is the density of the air and $u_*$ is the shear velocity of the flow, defined as the rate of change in velocity with height. Rearranging the equation to determine $\tau_c$ and considering the degree of packing of the grains as an additional factor gives:

$$\tau_c = \eta g(\rho_s - \rho)\frac{\pi}{6}d \tan\phi \qquad \text{(B2.3)}$$

where $\eta$ is a packing index.

The lift force arises from two components:

- differences in flow velocity between the top and bottom surfaces of the particle leading to a pressure gradient which encourages the particle to rise, which is known as the Bernoulli effect;
- turbulent eddies within the flow producing localized flow velocities close to the ground surface acting in an upwards direction.

The Bernoulli effect occurs in the following way. After a grain has been rolling along the ground for a short distance, the velocity of the air at any point near the grain is made up of two components, one due to the wind and the other to the spinning of the grain. On the upper side of the grain both components have the same direction but on the lower side they are in opposite directions. As a result of greater velocity on the top surface of the grain, the pressure there is reduced, while pressure at the lower surface increases. This difference in static pressure produces a lifting force and when this is sufficient to overcome the weight of the grain, the grain rises vertically.

Once a soil particle has been detached from the soil mass and entrained in the flow, it is carried forwards until the flow velocity falls below the settling velocity of the particle. The settling velocity ($v_s$) can be calculated from Stokes' Law:

$$v_s = \frac{2gr^2(\rho_s - \rho)}{9\mu} \qquad \text{(B2.4)}$$

where $r$ is the radius of the particle assuming it to be spherical and $\mu$ is the viscosity of the fluid.

Figure B2.2 shows the relationship between the critical velocity for entrainment and the fall or settling velocity for particles of different sizes based on studies made in rivers (Hjulström 1935). This shows that a soil particle of 0.01 mm requires a flow of 60 cm s$^{-1}$ to detach it but it is not deposited until the flow velocity falls below 0.1 cm s$^{-1}$. In contrast, a coarse particle of 1.0 mm diameter needs a velocity of 25 cm s$^{-1}$ to detach it and it will be deposited when the velocity falls below 7 cm s$^{-1}$. The difference between the erosion and fall velocities is much less for the coarser particles, which means that, once entrained, they are carried for only relatively short distances.

Since settling velocity is related to particle size, during deposition the coarser particles are deposited first, with progressively finer

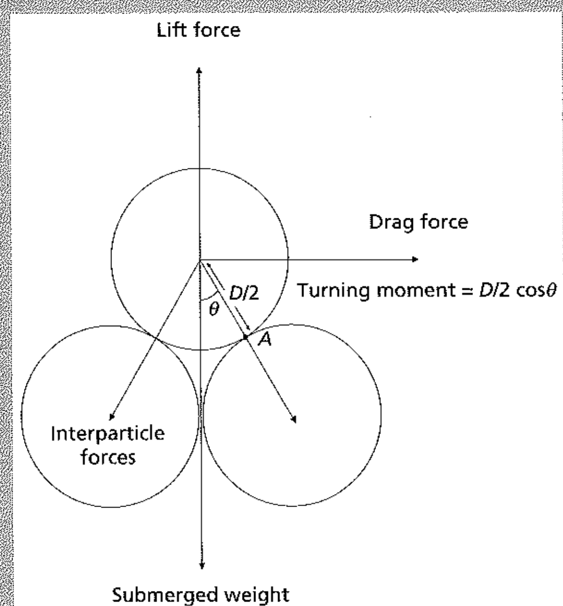



**Fig. B2.1** Forces involved in the initiation of particle movement. When movement begins the uppermost particle will move against the underlying particle, turning at contact point A.

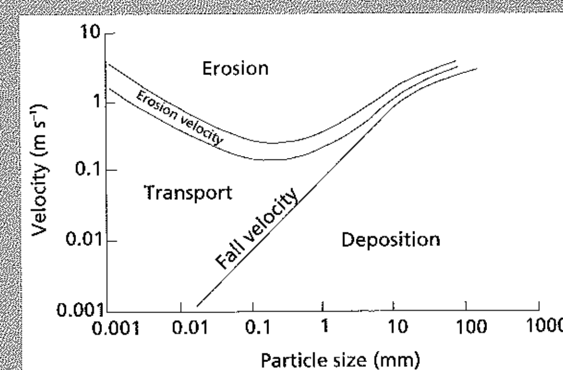



**Fig. B2.2** Critical water velocities for erosion, transport and deposition of sediment as a function of particle size (after Hjulström 1935).

*Continued*

grains dropping out of the flow as the flow velocity continues to decline. As a result, deposits from slope wash on concave slopes at the foot of hillsides tend to comprise mainly sands, grading into silts and clays with increasing slope length and decreasing slope angle. On many slopes, silts and clays can be transported across the footslope and into the adjacent river. However, these relationships can sometimes be offset by the redetachment of the deposits and the trapping of fine particles in the wake of coarser ones, with the result that the transport of coarse material over fines is enhanced (Beuselinck et al. 2002). Deposits of wind-blown material comprise mainly silts and sands, moved in saltation, whereas the clay particles are carried long distances in suspension in the atmosphere.