Keywords
reactive transport, porous media, dissolution kinetics, linear stability, wormhole formation, pore-scale processes, numerical methods

Abstract
We review theoretical and computational research, primarily from the past 10 years, addressing the flow of reactive fluids in porous media. The focus is on systems where chemical reactions at the solid–fluid interface cause dissolution of the surrounding porous matrix, creating nonlinear feedback mechanisms that can often lead to greatly enhanced permeability. We discuss insights into the evolution of geological forms that can be inferred from these feedback mechanisms, as well as some geotechnical applications such as enhanced oil recovery, hydraulic fracturing, and carbon sequestration. Until recently, most practical applications of reactive transport have been based on Darcy-scale modeling, where averaged equations for the flow and reactant transport are solved. We summarize the successes and limitations of volume averaging, which leads to Darcy-scale equations, as an introduction to pore-scale modeling. Pore-scale modeling is computationally intensive but offers new insights as well as tests of averaging theories and pore-network models. We include recent research devoted to validation of pore-scale simulations, particularly the use of visual observations from microfluidic experiments.
1. INTRODUCTION

The term reactive transport refers to the coupling of chemical reactions with mass transport, usually in aqueous solutions. In systems where mixing is incomplete, transport of reactants and products can play an important role in determining the overall reaction rates. Reactive transport modeling is frequently applied to subsurface hydrology: to predict the dispersal of contaminants from fossil fuel extraction, chemical processing, and mining; to optimize fossil fuel extraction by using acid injection to open pore spaces in the rock; and to design strategies to remove carbon dioxide from the atmosphere by mineralization to carbonates. In this review, we focus on one particular component of reactive transport modeling, namely how chemical reactions with the rock matrix generate feedback loops that can lead to highly localized dissolution.

It has long been known that aqueous ions, transported by flowing water, can dramatically reshape the environment over long periods of time. In his seminal work, *Principles of Geology*, first published in 1830, Charles Lyell remarked that “to the solvent power of water, surcharged with carbonic acid, and percolating various winding rents and fissures, we may ascribe those innumerable subterranean cavities and winding passages which traverse the limestone in our own and many other countries” (1, p. 211). Lyell popularized the uniformitarian principle, emphasizing gradual changes subject to invariant physical and chemical laws, to explain geological forms observed today. Reactive transport modeling is an attempt to make these ideas quantitative and predictive.

Although applications of reactive transport are most frequently found in the geological sciences, its core components—fluid flow, mass transport, and chemical reactions—fall squarely within the domain of chemical engineering. We find engineering applications of the same principles within catalysis (2), bioengineering (3), and soil engineering (4).

2. REACTIVE TRANSPORT AT THE DARCY SCALE

In this section, we lay out governing equations for a minimal model of reactive transport that is nevertheless capable of describing feedback between transport and chemical reactions:

\[
\begin{align*}
\partial_t \phi + \nabla \cdot \mathbf{u} &= 0, \quad \mathbf{u} = -\frac{K}{\mu} \nabla p, \\
\partial_t (\phi \mathbf{c}) + \nabla \cdot (\mathbf{u} \mathbf{c}) &= \nabla \cdot \mathbf{D} \cdot \nabla \mathbf{c} - R, \\
c_{\text{sol}} \partial_t \phi &= R.
\end{align*}
\]

The essence of the Darcy formulation is that the specific geometry of the pore space is replaced by a porosity field \( \phi(x, t) \), which is the fluid volume fraction in the neighborhood of \( x \) at time \( t \). The velocity field \( \mathbf{u}(x, t) \) is usually referred to as the Darcy (or superficial) velocity and represents the volumetric flux in a small volume around the point \( x \). The permeability \( K \) is assumed to be a local property, dependent on volume fraction and the structure of the pore space (see Section 3.1). If the porosity is varying sufficiently slowly, the Darcy velocity satisfies the equation for an incompressible fluid: \( \nabla \cdot \mathbf{u} = 0 \). Similar considerations can be used to derive balance equations for the reacting species. Schematically, advection of concentration \( \frac{d(\phi \mathbf{c})}{dt} \) is balanced by dispersion \( \nabla \cdot \mathbf{D} \cdot \nabla \mathbf{c} \) and reaction \( R \). The dispersion tensor depends on porosity and fluid velocity (5), even for isotropic materials, but at sufficiently low flow velocities, \( \mathbf{D} \approx D_0 \phi \mathbf{I} \), where \( D_0 \) is the molecular diffusion coefficient.
The reaction rate $R$ (in mol m$^{-3}$ s$^{-1}$) describes the consumption or production of ions (per unit volume) by the porous mineral; it depends on the chemical kinetics at the mineral–fluid interfaces and on the exposed surface area of the mineral. For example, dissolution of limestone under acidic conditions can be described by a single reaction:

$$\text{CaCO}_3 + \text{H}^+ \rightarrow \text{Ca}^{2+} + \text{HCO}_3^{-}.  \quad 4.$$  

The rate $R = k_D \alpha_{\text{H}^+} s$ is the product of a (temperature-dependent) rate constant (in mol m$^{-2}$ s$^{-1}$), the activity of the protons (dimensionless), and the specific surface area $s$ (in m$^{-1}$). However, unless the reaction rate is small, ionic concentrations at mineral surfaces can differ substantially from their bulk solution values. Reactive transport modeling provides a means of connecting surface activities (or concentrations) to bulk solution values. Loss of mineral volume by dissolution reactions, such as Equation 4, often leads to greatly enhanced permeability, by creating preferential flow paths referred to as wormholes or solution pipes. They play a key role in shaping the landscape, resulting in striking forms that occur repeatedly in different regions of the earth, and in a variety of geotechnical applications: acidization, self-propping, carbon sequestration, well-bore capping, and dam stability (see Section 2.2).

### 2.1. Length and Time Scales

Reactive infiltration of a porous material is characterized by two important length scales (Figure 1): the downstream penetration length $l_d$, over which the infiltrating solution equilibrates with the rock as a result of chemical reactions, and the upstream length $l_u$, which is the width of the diffusive boundary layer in the fully dissolved region:

$$l_d = \frac{D_0 \varphi \bar{u}}{u_0} \frac{2}{\sqrt{1 + 4H}} - 1, \quad l_u = \frac{D_0 \varphi u}{u_0}.  \quad 5.$$

**Figure 1**

Concentration and porosity profiles in a dissolving porous matrix. The reaction front, located at $\zeta(t)$, marks the rightmost place where the matrix is fully dissolved ($\varphi = \varphi_u$). The concentration profile decays with different length scales, $l_u$ and $l_d$, in the upstream $[x < \zeta(t)]$ and downstream $[x > \zeta(t)]$ regions.

The ratio of these two length scales is determined by the ratio of porosities and the reactive transport parameter, $H = D_0 \varphi_0 k \phi_0 / u_0^2$; here, $u_0$ and $s_0$ are the mean velocity and reactive surface area in the initial sample, and the reaction rate $k$ is in meters per second. When reaction rates are large or the flow rate is small ($H \rightarrow \infty$), the solution is in equilibrium with the mineral and the reactant penetration length is negligible in comparison to the upstream length: $l_d \ll l_u$ (6, 7). Conversely, $H \rightarrow 0$ corresponds to a convection-dominated process (7, 8), where the concentration is constant all the way to the front: $c_u = c_m$.

In laboratory systems the size of the sample creates an additional length scale $L$, which is typically not present in natural systems (or is much larger than $l_d$ and $l_u$). Specifically, when $L$ is comparable to or smaller than $l_d$, dissolution is nearly uniform.

In natural systems, diffusive length scales ($l_u$) are in the range from millimeters to meters, since typical groundwater velocities are small: $u_0 \approx 10^{-10} – 10^{-7}$ m s$^{-1}$. The downstream length also depends on the product of reaction rate and specific surface area, and both are highly variable. Specific surface areas in rocks span at least five orders of magnitude: $10^2$ m$^{-1} – 10^7$ m$^{-1}$ (9). Reaction rates range from $k = 10^{-11}$ m s$^{-1}$ for silicates, $k \gg 1$) and systems where diffusion is negligible ($H \ll 1$).

**Converting reaction rates**: Dissolution rate constants ($k_D$) are usually reported in terms of the mass flux at unit activity (mol m$^{-2}$ s$^{-1}$); for acid reactions, $k_D$ should be divided by 1,000 mol m$^{-2}$ to convert to a rate constant $k$ (in m s$^{-1}$).
to \( k = 10^{-6} \text{m s}^{-1} \) for dissolution of gypsum (both in water). Even faster rates (up to \( 10^{-3} \text{m s}^{-1} \)) are attainable in acidization of carbonate reservoirs by HCl solutions \((10,11)\). This leads to a wide range of penetration lengths in nature, from submillimeter scales to meters \((7)\). However, only the downstream length scale is observable in the geological record of mineral compositions.

Dissolution occurs with a characteristic time scale \( t_d = (k s_0 \gamma)^{-1} \), where \( \gamma = c_{\text{in}}/c_{\text{sol}} \) is the ratio of aqueous ion to mineral concentrations. Typically \( \gamma < 10^{-3} \) and the time scales for dissolution are large, much larger than either concentration or velocity relaxation times. For dissolution of carbonates by meteoric waters, \( t_d \) is of the order of days to years, while silicate systems dissolve over hundreds of thousands of years. This means that the time derivatives in Equations 1 and 2 can be neglected without noticeable error, which is a significant theoretical and computational simplification.

### 2.2. Interface Stability and the Evolution of Forms

Whenever a more mobile phase invades a less mobile one, the interface between the phases becomes unstable. A paradigm problem of this kind is viscous fingering \((12)\), where the less viscous fluid displaces the more viscous one. However, from Darcy’s law (Equation 1) we recognize that an alternative way of increasing the mobility of the invading phase is by modifying its permeability. This is precisely what happens during dissolution of porous or fractured rock. The dispersion relation, linking the instability growth rate \((\omega)\) with the wave number of the perturbation \((k)\), is linear in \( k \) at long wavelengths:

\[
\omega(k) \sim \frac{M_u - M_d}{M_d + M_u} k, \tag{6}
\]

where \( M = K(\varphi)/\mu \) is the mobility of the regions upstream and downstream of the front (Figure 1). Equation 6 applies to both viscous fingering and dissolution, but the mechanisms for short-wavelength stabilization of the fronts are different. In viscous fingering, large-\( k \) perturbations are stabilized by surface tension \((13)\). However, in reactive infiltration, the stabilization mechanism depends on the dissolution regime, characterized by the parameter \( H \) (Equation 5).

In the thin-front limit \((H \gg 1)\), short wavelengths are stabilized by diffusion and the instability wavelength is comparable to \( l_u \). A more subtle mechanism is in place in the convective regime \((H \ll 1)\), where there is a region of partially dissolved material ahead of the dissolution front. The diffuse interface makes the mobility contrast less sharp and weakens the instability \((14)\). It leads to a peak in the growth rate over length scales comparable to the front thickness, \( l_d \), while keeping \( \omega(k) \) positive as \( k \to \infty \) \((15)\).

To summarize these insights, let us imagine a natural system with fixed \( k, s_0, \) and \( D_0 \) but a variable flow rate \( u_0 \). Then, at very low fluid velocities (large \( H \)), the instability is controlled by \( l_u \), which decreases as \( u_0 \) increases. For larger flows (small \( H \)), the system becomes controlled by \( l_d \), which increases with \( u_0 \). Consequently, there is a minimum instability wavelength (as a function of \( u_0 \)), which occurs when \( H \approx 1 \) \((7)\).

In the later stages of dissolution, undulations in the front are transformed into well-defined finger-like channels or wormholes, which rapidly advance into the medium. As dissolution proceeds, these fingers interact, competing for the available flow, and eventually the growth of the shorter ones ceases. Figure 2a–c shows an evolving fracture aperture field, and Figure 2d shows the flow field at the latest time. Clearly, there are many more fingers visible in the aperture maps (Figure 2a–c) than in the flow map (Figure 2d) because the longer wormholes suck flow from the shorter ones \((16)\). The pressure gradient in a
long wormhole is steeper than in the surrounding shorter ones, because the flow rate is higher. In the upstream part of the fracture, a short wormhole is therefore at a higher pressure than a long one, so the flow in the surrounding matrix is directed toward the long wormhole. Downstream, the situation is reversed, with the region around the tip of the long wormhole at a higher pressure than the surrounding medium, so flow is directed away from the wormhole; this can be observed near the tips of the wormholes in Figure 2d. Shorter wormholes, deprived of any flow (and therefore reactant), quickly become saturated with dissolved ions and stop growing. On the other hand, the longer wormholes again compete for the flow between themselves, leading to the emergence of a hierarchical, scale-invariant structure, with many short wormholes and only a few long ones (Figure 2c). Numerical simulations (17,18) suggest that the distribution of wormhole lengths follows a power law, \( N(L) \sim L^{-\alpha} \), where \( N(L) \) denotes the number of wormholes longer than \( L \). For an initially homogeneous system, the exponent is between 1 and 1.2 (16–18).

Field examples of dissolutional fingering include solution pipes (Figure 3a), which are formed in the epikarst zone close to the surface (19,22). They exhibit scale invariance with an exponent close to 1, as observed in numerical simulations (18). However, it is not always straightforward to measure the full length of the pipe, since their tips are usually buried. In Figure 3a,b the length of the solution pipe has been exposed by quarrying. Despite their similar appearance, the two field examples in Figure 3 were most likely formed under quite different conditions. The solution pipes in Canunda National Park, South Australia (Figure 3a), were formed in the vadose zone, where the soil is partially water-saturated, whereas those in Smerdyna, Poland, are thought to have been formed during glacier melting, when the water table was much higher, and were most likely fully saturated.

The spontaneous appearance of solution pipes (or wormholes) is a beautiful example of self-organization, demonstrating that not all geological structures need historical precursors. Ongoing discussions about the role of initial conditions in determining geological pattern formation are particularly vivid in the analysis of solution pipes (Figure 3). Studies have suggested that preexisting heterogeneities in rock or soil—animal burrows (23), surface cracks (22,24), or vegetation (22)—cause the flow to focus in particular areas. On the other hand, numerical models (18,25) suggest that dissolution patterns are largely insensitive to initial conditions. Inhomogeneities in the primitive rock do affect the length and time scales over which wormholes emerge (18,26,27), but in a large enough system, the naturally growing instability eventually overwhelms the initial perturbation (18).

Several experiments on flow focusing (Figure 3c,d) have been performed in relation to acidization (10,20), which is used in the petroleum industry to enhance oil and gas

---

Figure 2

(a–c) Time evolution of a fracture aperture field. The competition between growing wormholes gives rise to a hierarchical pattern, with many short wormholes and a few long ones. (d) Flow field around the advancing tips. Red shading indicates the regions of highest aperture or flow, followed by green, blue, and black.

Figure 3

Examples of unstable growth. (a) Soil-filled solution pipes at Canunda National Park, South Australia (19). (b) Solution pipes in Smerdyna, Poland. The longest wormhole in the photograph is approximately 5 m. (c) Wormholes (30 cm long) produced during a carbonate acidization experiment (20). (d) Wormholes in a microfluidic experiment dissolving a synthetic gypsum fracture (21). Panel a courtesy of Ken Grimes, La Trobe University. Panel b courtesy of Dr. Lukasz Uzarowicz, Warsaw University of Life Sciences–SGGW.
production by increasing the permeability of the rock. The shapes of the wormholes formed during acidization depend strongly on the flow, with more conical, smoother wormholes forming at lower flow rates and highly ramified, tortuous wormholes appearing at high flow rates. Long, thin (dominant) wormholes, formed at intermediate flow rates, are the most effective for petroleum engineering, since they minimize the acid required for a given increase in permeability.

Estimates of wormhole growth rates and, in particular, the so-called breakthrough time (the moment when the longest wormhole reaches the outlet of the system) are crucial for a number of geotechnical problems. These include risk assessment of potential leakage of sequestered carbon dioxide, safety of dam sites in soluble rocks, risk of catastrophic ground subsidence due to solutional widening of fractures, and danger of water seepage into toxic waste repositories.

2.3. Fracture Dissolution

Dissolution of rough fractures can be cast into the form of a Darcy problem by introducing the fracture aperture field, \( h(x,y,t) \). After integrating over aperture, we obtain analogs of Equations 1–3 for the evolution of the fluid flux \( q(x,y) = \int_0^h udz \), velocity-averaged concentration \( \bar{c}(x,y) = |q|^{-1} \int_0^h ucdz \), and aperture (26, 28):

\[
\nabla \cdot q = 0, \quad q = -\frac{h^3}{12\mu} \nabla p, \quad \text{7.}
\]

\[
\nabla \cdot (q\bar{c}) = \nabla \cdot (D_0h \cdot \nabla \bar{c}) - 2k_{\text{eff}}\bar{c}, \quad \text{8.}
\]

\[
c_{\text{sol}} \partial_t h = 2k_{\text{eff}}\bar{c}. \quad \text{9.}
\]

We note that \( \bar{c} \), appearing in the the diffusional flux, is in fact the average concentration rather than the velocity-weighted one (28, 29), but it is common to ignore this distinction, as we do here. Fracture dissolution tends to be convection dominated \((H \ll 1)\) so that the diffusional term is small (28). Again, we have assumed an acidization reaction (Equation 4) where the concentration of aqueous ions is small in comparison to \( c_{\text{sol}} \), so that the time dependence in Equations 7 and 8 can be neglected. Here, the reaction rate constant includes a correction for diffusion across the fracture aperture, \( k_{\text{eff}}^{-1} = k^{-1} + 2h/DSh \), where \( Sh \approx 8 \) is the Sherwood number for parallel plates. Karst formation kinetics \((k \sim 10^{-7} \text{ m s}^{-1})\) are nevertheless reaction limited \((k_{\text{eff}} \approx k)\) until the fracture aperture is of the order of centimeters. However, acidization of calcite \((k \sim 10^{-3} \text{ m s}^{-1})\) is almost always controlled by proton diffusion to the fracture surface \((k_{\text{eff}} \approx DSh/2h)\). Note the important distinction between this example and the discussion in Section 2.1. Here, the added hindrance to dissolution comes from ion transport (by diffusion) across the aperture; it can be an important effect in fractures or wormholes (30), because the aperture is larger than a typical pore size. In Section 2.1, the issue is whether reactant can be convected to the front fast enough to keep up with its consumption. Unfortunately, both situations are frequently referred to as reaction-limited versus transport-limited kinetics, without further clarification.

A comparison of Equations 7–9 with the corresponding equations for two-dimensional (2D) porous media, Equations 1–3, shows striking similarities, but there is a significant difference in the way that the dissolution profile evolves. From an initially homogeneous distribution, the porosity at the inlet increases monotonically until all the soluble material is dissolved. After that, a front of fully dissolved rock propagates downstream at a constant
velocity (31). It is generally assumed that instabilities in porous rocks are superposed on a planar propagating front (6,7), although in fact an instability can also be generated off the initial uniform porosity field (31). By contrast, in fracture dissolution there is no limit on the local permeability and no invariant front; the aperture $h$ can increase without bound.

A somewhat different situation arises if the amount of soluble material in the fracture is limited: for example during dissolution of carbonate or silica cements lining the walls of an otherwise insoluble fracture or vein. The removal of cement can increase the flow between fracture and matrix during the extraction of shale gas. Interestingly, such fractures show intense wormhole formation (21), which allows for increased gas recovery. During hydrofracturing, wormholes can not entirely maintain permeability after the fluid pressure is released, reducing the need for proppant (32,33).

Instabilities in fracture dissolution were first discovered by numerical simulations (26) and later confirmed by experiments (35, 36). Linear stability analysis shows that fractures are even more unstable to dissolution than porous media (15), where a minimum wavelength is needed for unstable growth (37). However, fracture dissolution is almost always unstable (28,38), and intense wormholing is to be expected (see Section 2.2). The onset of a reactive infiltration instability has been observed in a microfluidic fracture system, with good agreement between theoretical predictions of the wavelength of the initial instability and experimental observations (39).

Wormholing is the driving force behind cave formation in karst systems (15,26,40), and cave conduits are classical manifestations of nonuniform dissolution. Caves are initiated along fractures and bedding planes, which have quasi-2D structure (Figure 4a), while the mature cave is almost always a system of pipelike conduits (Figure 4d). Figure 4b shows closely spaced wormholes at the inlet, while the quarried face in Figure 4a is more distant and only the longest wormholes survive. Figure 4c shows a circular conduit emerging from a fracture, but on a much smaller scale (micrometers as opposed to meters). An interpretation of these figures and their connection to models of hierarchical growth can be found in Reference 18.

Figure 4
Examples of nonuniform fracture dissolution accompanied by wormhole formation. (a) An exposed fracture (black arrows) at a limestone quarry in Smerdyna, Poland. Two wormholes (approximately 10 cm in diameter) have developed along the fracture. (b) A dissolutionally enlarged styolite with several wormholes in a limestone wall of Mitzpe Ramon quarry, Israel. (c) Wormhole at the outlet of a fractured limestone core (34). (d) A conduit in a phreatic cave in Dan yr Ogof, Swansea Valley, South Wales. The white arrows mark the initial fissure. Panel b courtesy of Pawel Kondratiuk, University of Warsaw. Panel c courtesy of María García Ríos and Linda Luquot, Technical University of Catalonia. Panel d courtesy of Brendan Marris, Dudley Caving Club.

These observations may provide insight into how conduits can penetrate so deeply into limestone formations. A simple estimate of the penetration length of aggressive (undersaturated) solutions infiltrating a fracture system indicates $l_p < 1$ m, meaning that limestone caves should not exist at all (41). A possible resolution of this paradox is the sharp drop in the dissolution rate of CaCO$_3$ near saturation (42). However, reactive infiltration instabilities offer a simpler explanation (15); flow in spontaneously formed conduits (Figure 2) is several orders of magnitude larger than the initially uniform flow distributed across the fracture plane, and the penetration length increases accordingly. Numerical simulations (38) show that highly unsaturated solutions can penetrate deeply into fractures as wormholes.
develop (18).

At the laboratory scale, flow focusing can be observed outside of the typical wormholing regime of flow rates (43, 44). An interesting discovery was that even in the uniform dissolution regime \( (l_p \gg L) \) preferential flow paths emerged. These flow paths were not nucleated from the homogeneous matrix but were preexisting, aided by the short length of the sample \( (L = 0.012 \text{ m}) \). The pathways increased in diameter all along their lengths, instead of nucleating at the inlet and propagating downstream, as is the case with larger samples. Nevertheless, these flow paths also competed with one another, with the faster-growing ones eventually capturing most of the flow. Similar dynamics, with flow paths emerging along the length of the sample, are encountered when dissolution is driven by a spatially dependent solubility; for example in hypogene karst systems or in melt extraction (45, 46).

Real-time optical imaging of a dissolving fracture (35) has been used to provide detailed aperture data from a known and reproducible initial condition. This is an ideal experimental benchmark for simulations attempting to capture the physics of fracture dissolution, a precursor to more recent microfluidics experiments (39). Numerical simulations (47, 48) showed wormholes developing at the same locations as in the experiment, despite the strongly nonlinear nature of the instability. A similar numerical model was later used for a broader investigation of fracture dissolution (49).

Experimental advances in studying fracture dissolution (34, 50, 51) are being supported by numerical simulations with more complex chemistry (52, 53). However, these models still utilize a depth-averaged approximation (Equations 7–9), which does not entirely suffice as wormholes develop. Recently, a 2D model for fracture dissolution was tested against three-dimensional (3D) simulations (38). The initial aperture evolution was well reproduced by the depth-averaged model, but once wormholes develop, the flow field in them is too large by approximately a factor of two, which artificially enhances wormhole competition. 3D effects are also important at intersections between fracture planes, where a fully 3D model of the flow is needed to capture the correct mixing (54).

2.4. Darcy-Scale Simulations

Reactive transport simulations at the Darcy scale are now well established, with codes available to simulate chemical reactions between multiple components distributed among several phases. A comprehensive and readable introduction can be found in (55). General-purpose reactive transport codes originated more than 30 years ago (56–58) and have been under constant development ever since. Progress has been documented in several review articles (59–63).

Most reactive transport codes use a sequential algorithm to decouple Equations 1–3. The critical component is the solution of a generalization of Equation 2, which for a saturated aqueous phase can be written as

\[
\partial_t (\phi c_k) + \nabla \cdot (\mathbf{u} c_k) = \nabla \cdot \mathbf{D} \phi \cdot \mathbf{u} c_k + \sum_{j=1}^{N_R} \nu_{jk} R_j (c),
\]

where \( R_j \) is the rate of reaction \( j \) \( (N_R \) reactions in all \) and \( \nu_{jk} \) is the stoichiometric number of component \( k \) in reaction \( j \). Typically the dispersion coefficient for each species is assumed to be the same; otherwise charge imbalances will be generated by differential diffusion (61, 64).

Implicit time integration is used in reactive transport simulations to maintain stability.
over the wide range of time scales in the problem. Typically, this is a first-order backward difference (implicit Euler) formula; at time \( t_{n+1} \),
\[
\frac{\partial c_k^{n+1}}{\partial t} \approx \frac{(c_k^{n+1} - c_k^n)}{\Delta t}.
\]
An unconditionally stable second-order difference is also possible, without additional computational cost, but requiring an additional vector of concentrations \( c^{n-1} \). After spatial (finite-volume) and temporal discretization, these equations can be approximated by nonlinear algebraic equations:
\[
\sum_{P'} A_{PP'}^c c_{P',n+1}^{P'} = \sum_{j=1}^{N_R} \nu_{kj} R_j(c^{P,n+1}) = b^P,
\]
where \( A \) is a matrix representation in the finite-volume basis of the convection-diffusion operator (the first three terms in Equation 10). Each control volume (or cell), identified by the index \( P \), has a set of neighboring cells, identified by \( P' (P \in P') \); for second-order accuracy the concentration must be evaluated at the center of mass of each control volume. The matrix operates on the vector of unknown concentrations at time \( t^{n+1} \), while the vector \( b^P \) involves concentrations at the current step \( t^n \) and possibly the previous step as well.

Time scale separation allows the porosity in \( A \) to be taken as constant over the time step \( \phi_{P,n} \), while the velocity field \( \mathbf{u}_{P,n+1} \) is already determined at \( t_{n+1} \). If \( \mathbf{u} \) depends on \( c \), as in electro-osmosis, for example, an additional outer iteration can be used to correct the velocity field. The same procedure can also be used to solve the steady-state problem.

There is frequently a wide separation of time scales between different chemical reactions. Reactions in the aqueous phase are usually rapid, while reactions with minerals tend to be slow. The kinetics of fast reactions can be replaced by equilibrium conditions on the ion–activity products, reducing the number of concentrations that need to be solved (57, 65). Only primary species need to be transported (Equation 10), while the secondary concentrations are determined from the primary species concentrations within each cell. Many alternative schemes for making the partitioning have been developed (66,67), aiming for greater generality and more-automated implementations.

Two primary algorithms are used to solve discretized reactive transport problems, such as Equation 11, with many variants of each. The global implicit approach (GIA) solves the fully coupled equations by using Newton–Raphson iteration (57,65,68). It requires a large Jacobian matrix of partial derivatives, \( J_{kk'}^{PP'} \); identical components in neighboring cells are coupled by the transport operator \( (A) \), while different components in the same cell are coupled by reactions. Starting with values from the previous time step, concentrations within each cell are iterated to convergence using the Jacobian matrix to linearize the residual (69). Improved convergence is obtained by iterating the logarithm of the concentrations (57,62). The Jacobian can be determined numerically, eliminating the need for explicit differentiation of complex formulas for the reaction rates. However, it requires \( N_P \) evaluations of the \( N_R \times N_R \) matrix of reaction rates at every iteration, which is computationally expensive. However, quasi-Newton methods replace repetitive evaluations of the Jacobian matrix by a much faster iterative update. Modern solvers typically prefer the BFGS (Broyden–Fletcher–Goldfarb–Shanno) algorithm to the older DFP (Davidon–Fletcher–Powell) method.

The main alternative to GIA is the sequential iteration approach (SIA) (56,70), which is an operator-splitting method that separates the transport and reaction terms in Equation 11. Equation 11 is solved iteratively using concentrations from the previous iteration to calculate the reaction term. SIA is simple and does not require a Jacobian matrix, but in some instances the convergence is poor (68). A possible compromise between GIA and SIA is to solve for each concentration sequentially but allow a change in concentration of
the target species to modify the reaction rate. In this case, only the diagonal elements of the Jacobian need to be recalculated. We note that a segregated solution strategy, with nonlinear iterations within a subspace, is typical in general-purpose finite-volume codes such as OpenFOAM.

Many of the most heavily used reactive transport codes have been summarized in a review (62), including tables of features, functionality, and numerical methods. Several benchmarking studies comparing different reactive transport codes have been made (71–74), including a special issue of *Computational Geosciences* (75). While these studies found minor disagreements between different codes, the overall predictions are remarkably similar.

A particularly interesting recent development is the Reaktoro geochemical package (76,77; see [http://www.reaktoro.org/](http://www.reaktoro.org/)), which solves chemical equilibrium and kinetics problems within a batch-reactor framework. Reaktoro does not include a transport module, but it has elegant Python and C++ interfaces that allow it to be coupled to a variety of transport codes such as FEniCS (77) or OpenFOAM (78). Currently, it can utilize the thermodynamic and kinetic data from the SUPCRT92 (79), PHREEQC (80), and GEMS (81) databases.

Spatial discretization of the advection term $\nabla \cdot (uc)$ in Equation 2 can lead to instabilities in the solution if the concentration on a face is linearly interpolated from the neighboring cell centers. A bounded solution can be guaranteed (at least in one dimension) if all coefficients in the transport matrix ($A_{PP}'$) are positive, which can be assured by taking concentrations from the cell upwind of the face (the direction opposite to the fluid flux). However, the resulting scheme is only first-order accurate, with an additional numerical diffusion $\sim h^2/\Delta t$ on top of molecular diffusion. Nevertheless, most reactive transport codes use an upwind scheme because it is robust and stable, but even stationary solutions can be significantly in error. The most successful numerical schemes are based on nonlinear flux limiters, with the face value formed by a combination of upwind and central differencing (82). The flux limiter switches from central differencing (for smooth fields) to upwind on the basis of the changes in slope of the field. However, the flux limiter has to be nonlinear to assure boundedness of the solution, so that the transport matrix ($A$) is also nonlinear, although still sparse. There are many formulas for the flux limiter, but no consensus on which is the best in all cases (82).

3. REACTIVE TRANSPORT AT THE PORE SCALE

Advances in X-ray imaging, particularly tabletop micro-CT (computed tomography), have spurred advances in computational techniques to take advantage of the detailed descriptions of a pore space that are now available. In contrast to Darcy-scale models, where the pore space is characterized by a porosity field, in a pore-scale model each space point $r$ belongs either to a mineral phase or to the fluid phase. The mineral grains are assumed to be impermeable to flow, although diffusion of ions (but without convection) is possible if the grains themselves contain small connected pores (83).

A minimal model for pore-scale dissolution can be represented by the following equations:
\[ \nabla \cdot (\mathbf{u} \mathbf{u}) + \nabla (p/\rho) = \mu \nabla^2 \mathbf{u}, \quad \nabla \cdot \mathbf{u} = 0, \]  
\[ \nabla \cdot (\mathbf{uc}) = D_0 \nabla^2 c, \]  
\[ D_0 (\mathbf{n} \cdot \nabla)c_s = k \mathbf{c}, \]  
\[ c_{os} \partial_t \mathbf{r}_s = -D_0 (\mathbf{n} \cdot \nabla)c_s \mathbf{n}. \]

The velocity field \( \mathbf{u}(\mathbf{r}, t) \) satisfies the Navier–Stokes equation, with no-slip boundary conditions on all the grain surfaces. Once again, the large time scale separation (\( \gamma \ll 1 \)) between transport and mineral removal (Equation 15) allows for a stationary solution for both the velocity (Equation 12) and concentration fields (Equation 13). Only molecular properties—mass density (\( \rho \)), viscosity (\( \mu \)), diffusion coefficient (\( D_0 \)), and reaction rate constant (\( k \))—enter into the governing equations at the pore scale. Constitutive equations for permeability, hydrodynamic dispersion, and specific surface area are not necessary; instead, these macroscopic properties follow naturally from the microscopic physics in the specific pore geometry.

Chemical reactions enter into pore-scale dynamics as boundary conditions (Equation 14), balancing the diffusive flux of ions to the mineral surface \( D_0 (\mathbf{n} \cdot \nabla)c_s \) with the consumption of the same ions by the chemical reaction. In these equations, we use the subscript \( S \) to indicate the mineral surface. A key advantage of pore-scale modeling is that the concentration entering into the rate equation is well defined (Equation 14), unlike the bulk concentration, which is difficult to connect to the surface value. The motion of the boundary follows from the dissolution reaction; it is always normal to the surface (Equation 15). We have again supposed a typical acidization reaction (Equation 4) with a surface reaction rate constant \( k \), but other chemistry can be represented in a similar fashion (78).

### 3.1. Averaging Theories: From Pore Scale to Darcy Scale

Equations for Darcy-scale reactive transport (Equations 1–3) can be derived from the pore-scale dynamics (Equations 12–15) by upscaling methods, including volume averaging (84–87), moment expansions (88), and multiscale analysis (89,90). The key assumption is that there is a clear separation between macroscopic variations in porosity and microscopic ones. In the problem illustrated in Figure 2, the macroscopic length \( L \) would be related to the wormhole diameter or the instability wavelength; it is similar to the grid spacing in a fully resolved Darcy-scale simulation. On much smaller scales \( l \), there are rapid variations in porosity crossing from grain to grain, but on scales larger than \( l \), the material is assumed to be homogeneous. Then a macroscopic grid block is comprised of a large number \( (L^3/l^3) \) of statistically equivalent unit cells. By upscaling, one can derive average equations on length scales \( L \) from solutions of pore-scale equations for a single unit cell (volume \( l^3 \)).

Multiscale analysis can be used to establish conditions for a valid upscaling (90,91), using the ratio of microscopic \( (l) \) and macroscopic \( (L) \) lengths, \( \epsilon = l/L \), to determine bounds on the Péclet \( (Pe = u_0 L/D_0) \) and Damköhler \( (Da_{II} = kL/D_0) \) numbers. The advection-diffusion equation (Equation 13) can only be upscaled when \( Pe < \epsilon^{-2} \); otherwise, transport is advection dominated at the pore scale and cannot be homogenized (90). Reactions impose additional restrictions on the pore-scale parameters \( k \) and \( u_0 \) (91): \( Da_{II} < 1 \) and \( Da_{II} < \epsilon Pe \). The first constraint \( (Da_{II} < 1) \) means that we cannot apply upscaling to rapidly dissolving minerals such as CaCO\(_3\) under acid attack \( (k \approx 10^{-3} \text{ m s}^{-1}) \) (92,93) or CaCO\(_3\) and CaSO\(_4\) under neutral conditions \( (k \approx 10^{-6} \text{ to } 10^{-5} \text{ m s}^{-1}) \) (93,94). In all these
cases, the homogenization condition $D_{HH} < 1$ cannot be satisfied for any length scale $L > 1$ mm. Nevertheless, acidization rates for many minerals (95) are sufficiently small ($k_D < 10^{-6}$ mol m$^{-2}$ s$^{-1}$) to permit homogenization on at least the meter scale. Minerals with upscalable dissolution rates, say, $k < 10^{-9}$ m s$^{-1}$, require flow rates in excess of $10^{-6}$ m s$^{-1}$ to be upcaled from pores (100 μm) to cores (0.1 m) ($\epsilon = 10^{-3}$). Flow velocities in laboratory experiments usually exceed this value, but in nature, such flow rates (30 m/year) are uncommon in carbonate rocks. Because reactive transport occurs over a very broad range of conditions, the validity of an upscaling needs to be assessed on an individual basis, taking into account both flow and reaction rates.

3.2. Pore-Scale Experiments

Experimental investigations are uncovering small-scale physical and chemical processes that can influence the macroscale evolution of rocks. For example, leakage of water along fracture networks (96) or electro-osmotic transport of ions through nanopores (97) can influence the emergence of geological forms on the kilometer scale. Multiscale analysis of field measurements pose significant challenges for reactive transport modeling. For example, the speed of propagating alteration fronts have been difficult to reconcile with standard reactive transport models (98). The observed front velocity was two orders of magnitude faster than expected, even assuming the fluids were always in equilibrium with the mineral.

Laboratory studies in model systems are a promising route to improving pore-scale models, and the KBr → KCl replacement reaction has been widely used as a paradigm for investigating pore-scale processes (99). Scanning electron microscope images show that an intricate connected pore structure develops in the region of the replacement front (100). Even though there is no externally imposed flow, the porous front develops fingers that are reminiscent of wormhole formation (101). A possible convective mechanism is electro-osmosis, driven by the differential diffusion of Br$^-$ and Cl$^-$ ions (102); the replacement reaction generates gradients in the anion concentrations between the solid (KBr) and the bulk solution (KCl). Modeling not only the ion transport (102) but also the evolving porosity presents a significant computational challenge.

Recently, several new microfluidic technologies have been developed to study reactive mixing (103, 104) and evolution of porosity (105, 106). Microfluidic devices made from natural materials have been created specifically to study pore-scale processes in minerals. Innovative fabrication methods were used to take cross sections of a fractured core that had been imaged by X-ray tomographic microscopy and create slices of different thickness etched into either a shale or a glass matrix. The displacement of water by supercritical CO$_2$ was sensitive to the different roughness of the two materials, emphasizing the dependence of fracture–matrix interactions on surface properties. Replacement of SrSO$_4$ by BaSO$_4$ has also been investigated in a microfluidic experiment (107). A BaCl solution was injected into a porous matrix of celestine crystals (SrSO$_4$), and the replacement was monitored by Raman imaging. The change in mineral phase was modeled by the microcontinuum method (108, 109; see Section 3.4.3.2). Regions of high flow in the simulation were spatially correlated with the experimentally observed void space.

Multiphase reactive transport is particularly challenging to simulate (110), and microfluidic experiments will play a key role in further development. A novel experiment used circular channels containing a porous matrix of calcite, together with two dead-end pores at the center, to investigate dissolution by saturated brine (111); both single-phase and
multiphase (containing gaseous CO$_2$) flows were studied. Such experiments can provide critical tests of more complex pore-scale models.

### 3.3. Evolving Grain Shapes from Microfluidic Experiments and Simulations

An important goal of pore-scale modeling is to provide a firm foundation for upscaled models. As a result, it is essential to know that the pore-scale simulation is accounting for all the important physics. In the past decade, experimental measurements of effluent concentrations have formed a basis for validating pore-scale methods and codes (112–115). Nevertheless, a more strict validation should focus on the local dissolution at each point on the mineral surface, rather than the overall dissolution rate. In fact, the size and shape of an evolving grain, as observed in microfluidic experiments, have been successfully predicted by pore-scale simulations (109, 116, 117). Figure 5a shows a comparison of simulated grain shapes with observations from a microfluidic experiment (117); in this case, there are no fitting parameters in the comparison. Going forward, it will be interesting to investigate more complex dissolution kinetics, such as calcite dissolution at higher pH (118), and more complex geometries, such as arrays of soluble disks. 3D printing technologies (119) might be used to create experimental samples. However, since small amounts of additive were found to reduce the rate constant for CaSO$_4$ dissolution by a factor of five (117), the polymer binders used in 3D printing may also cause large variations in the dissolution rate constant.

![Figure 5](https://www.annualreviews.org/annual/reviews/2023/figure5.png)

Shapes of a dissolving gypsum chip. (a) Images from a microfluidic experiment. The red outlines indicate the shapes from a finite-volume simulation at the same time as the photograph. (b) Sequence of shapes of a dissolving disk in a two-dimensional potential flow from finite-volume (red curve) and conformal-mapping (black dashed curve) calculations. Shapes are shown at equal intervals [0.05, 0.15, ..., 0.95] of the time to complete dissolution; the axes in panel b are in units of the initial radius of the disk.

In addition to the underlying physics, we must be concerned with the accuracy and robustness of the numerical implementations, particularly with regard to the difficult problem of interface tracking (Section 3.4). A recent comparison of several different codes (116) showed a good level of coherence in predicting reactant concentration and grain shapes, although the agreement was by no means perfect. One thing lacking in such comparisons is a reference calculation, which we suggest could be provided by complex analysis. Although it imposes some limitations on the physics—2D potential flow and a constant concentration boundary condition—conformal mapping offer an essentially exact calculation of a nontrivial dissolution problem (4). The evolving shape of the grain is represented by a time-dependent mapping from a fixed domain in the complex plane (a circle, typically). The problem then reduces to finding a differential equation for the time evolution of the mapping. We have recently shown that precise agreement between finite-volume simulations and conformal mapping can be maintained to the point where the solid vanishes (120). Figure 5b shows a comparison of the shapes. At the final time, the leading edge of the dissolving disk in the finite-volume simulation is only 1% in advance of the conformal-mapping result. This is an encouraging indication of the degree of accuracy with which a moving interface can be tracked (see Section 3.4.1.1).
3.4. Pore-Scale Simulation Methods

Pore-scale modeling seeks to overcome limitations of Darcy-scale models by avoiding the uncertainties associated with local constitutive equations and the more fundamental limitations imposed by upscaling (Section 3.1). Here, the key computational problem is how to best represent the mineral–fluid interface and how to update it as the mineral dissolves. Two methods are commonly used to characterize the pore space: either as a field $\alpha(r)$, where $\alpha$ indicates the volume fraction of mineral (or fluid), or as a function $S(r) = 0$, defining the interface between mineral and fluid. Representing the mineral distribution by a phase field ($\alpha$) has several advantages: It is robust and can be implemented on a uniform structured mesh, which merges naturally with tomographic images of the pore space. However, the interface must either be spread out over several cells, or approximated as a sequence of steps. Alternatively, a sharp interface can be defined by a surface function; for example as an array of connected polygons. We organize our review of methodologies according to how the interface is defined and tracked (Sections 3.4.1–3.4.3), and how the field equations are solved (Section 3.4.4). We also briefly review pore-network models (Section 3.4.5), which represent the geometry in a statistical sense rather than as an explicitly connected pore space.

3.4.1. Interface tracking. Interface tracking requires that the mineral–fluid interface is defined by a bounding curve (2D) or surface (3D), with an equation of motion to describe its time evolution. Examples of interface tracking include conforming (body-fitted) meshes, level sets, and immersed-boundary methods. Although these are all considered sharp-interface models, the interface in level-set and immersed-boundary methods is not commensurate with the underlying mesh, and as a result the fields must be artificially smoothed over distances of the order of the mesh size. Locally adaptive mesh refinement (121) is used to improve the resolution in the region of the interface. Typically, level-set and immersed-boundary methods solve the field equations on a uniform structured mesh, whereas a conforming mesh is almost always unstructured. The additional overhead of an unstructured mesh is sometimes compensated by a more efficient use of grid points.

3.4.1.1. Conforming-mesh methods. In a conforming mesh, a subset of the vertices lie on the mineral–fluid boundaries. This allows for a precise description of the interface between mineral and aqueous phases, with a simple conservative calculation of the flux across each interface; cells near the interface require no special treatment. As the mineral dissolves, boundary points move in accordance with Equation 15, and the flow and transport (Equations 12–14) are solved in the new geometry. This is a simplification of the arbitrary Lagrangian–Eulerian (ALE) method (122) because mesh motion is decoupled from the solution of the field equations by the large time scale separation. In the ALE method, the field equations are solved in a Lagrangian frame, which automatically includes mesh motion. The mesh is then relaxed to prevent entanglements, and the fields are interpolated to the new mesh. Here, the field equations can be solved in an Eulerian frame and no interpolation is required; for each time step in the mesh evolution, the steady-state fields are solved from scratch, although the solution from the previous time step can sometimes be a useful initial condition.

Modern software can construct high-quality meshes derived from piecewise-continuous (triangulated) surfaces. Examples of simulations of reactive transport in porous media using conforming meshes include packed beds of cylinders (104) and spheres (115, 123).
major drawback to a conforming mesh is the evolving boundary. Simply moving the surface points in response to the dissolution flux (109) leads to a rapid degradation of the mesh and large errors in the solution. However, mesh relaxation permits much larger motions of the boundary (38); simulations of fracture dissolution tracked more than 100-fold increases in aperture. The same algorithm was later applied to dissolution of a cylindrical disk (Figure 5). However, mesh relaxation eventually fails for two reasons. First, individual cells become skewed, with cells that are initially rectangular tending to take on diamond shapes. Second, if the dissolving surface develops cusps, this causes large errors in determining the tangent plane around the cusp. A workable solution is to recreate a new volume mesh from scratch, using the current surface points as input (120). This process is almost entirely error free (because the fields are recalculated), but it adds to the computational expense.

3.4.1.2. Level-set methods. Level-set methods represent an interface as the zero surface of the signed distance function from the cell centers: \( s(r) = 0 \). They systematically approximate the geometric properties of the interface (tangents, normal, curvature) and permit topological changes without requiring additional rules. However, they do not intrinsically conserve mass, which is problematic in solutions of charged particles. Moreover, there is a growing (in time) disparity between the signed distance field and the actual distance from the cell center to the interface. The signed distance field must be periodically recalculated to maintain consistency over large interface motions.

Level-set methods have been applied to precipitation and pore clogging (124–126), tracking the permeability and porosity of a small number of evolving pores. A comparison of level sets and voxelated meshes (Section 3.4.2.1) as a means to describe a moving interface is reported in Reference 126. Although the relationship between permeability and porosity was similar, clogging was faster with the voxelated mesh, presumably due to the larger surface area. The evolution of permeability was also much rougher with the voxelated mesh. This finding suggests that the common practice of comparing macroscopic properties, such as permeability–porosity relationships (123), may not be a good indicator of how a method behaves when there is an evolving interface.

3.4.1.3. Immersed-boundary methods. An immersed boundary is a set of connected Lagrangian marker points which move with the local velocity of the interface. As in level-set methods, the interface has to be smoothed over distances of the order of the mesh size, so that the fields can be interpolated between the fluid grid points and the marker points. The method has a similar issue as a conforming mesh, in that the marker points need to be relaxed within their local tangent plane in order to preserve the proper connectivity. Similarly, additional rules are needed for topological changes in the mineral–fluid boundary. There are only a few instances of applications of immersed-boundary methods to reactive transport problems (127).

3.4.2. Interface reconstruction methods. Interface reconstruction methods track an evolving porosity field and construct the interface at each step from its spatial variation. However, reconstruction tends to introduce discontinuities in the interface position and errors in mass conservation.

3.4.2.1. Voxel method. The voxel method (128, 129) is the simplest reconstruction algorithm and is implemented on a regular cubic mesh. Within each cell, the porosity is
updated as a function of time, and a threshold porosity is used to segregate the points into fluid and solid. The interface is constructed from faces separating solid and fluid cells. A moving interface is represented by changing the state of a node from solid to fluid (or vice versa) once the porosity crosses the threshold value. The voxel method is robust and can be applied to highly irregular interfaces. In contrast, surface normals cannot be easily defined, and the mineral surface area does not converge with successive grid refinement.

Despite the apparent simplicity of voxel-based methods, several technical details make them time consuming to develop: when and how to change the state of a node from fluid to solid (or vice versa), what to do with the residual mass and momentum, and how to initialize the fields at a newly created node. There is no a priori answer to any of these questions, and considerable effort has gone into developing and validating different schemes for applying lattice Boltzmann (LB) methods to moving interfaces (128–133). Nevertheless, at present the voxel method remains the most popular methodology for pore-scale reactive transport simulations.

3.4.2.2. Volume-of-fluid method. The volume-of-fluid (VOF) method is the classic interface reconstruction scheme (134), with many variants (135) and wide applicability to multiphase flow problems. Like level sets, the VOF method incorporates adaptive mesh refinement to improve the resolution near the interface. In a typical reactive transport problem, the indicator function is the porosity, which is unity in the fluid phase and zero in the (impermeable) solid. When discretized on a uniform grid, the indicator function has intermediate values (0–1) in the region of the interface. The VOF surface is impossible to advect because it changes discontinuously across the mineral–fluid interface, so it is usually determined by reconstruction from the discretized indicator function. The standard PLIC (piecewise linear interface construction) scheme uses the porosity of surrounding cells to estimate the direction of the surface normal vector \( \mathbf{n} = \nabla \phi / |\nabla \phi| \) and then locates the position of the planar partition along the normal direction that gives the correct fluid volume within the cell. However, this is typically only a first-order scheme, and more considered calculations of the normal vector are needed to obtain a second-order method (136). Reconstruction does not, in general, lead to smooth or unambiguous interfaces. As a result, coupled level-set/VOF methods have been developed for multiphase flow (137). The VOF method (is used to transport mass (and momentum) conservatively, while level-set methods (Section 3.4.1.2) are used to advect the interface. Reactive transport equations have been implemented using these ideas (138).

3.4.3. Diffuse-interface models. In contrast with the sharp-interface methods described above, diffuse-interface models (DIMs) deliberately spread the interface over several grid cells.

3.4.3.1. Phase-field methods. Classical phase-field methods are based on a free-energy functional of the order parameter (or porosity in this case) and require no explicit treatment of the interface; it simply emerges from the free energy. Reactive transport models for precipitation and dissolution (139, 140) have been compared with level-set methods (141), with a high degree of concurrence between the two methods. However, a drawback of phase-field methods (and DIMs in general) is that the surface reaction rate is an output, not an input. It is not clear whether the connection between input parameters and effective reaction rate is entirely independent of geometry, which makes these methods less desirable.
for quantitative calculations, especially those with complex chemistry. This is an issue worthy of further study.

3.4.3.2. Microcontinuum model. An alternative DIM (142) is based on the realization that a Darcy-scale model of a propagating dissolution front will be sharp if the reaction rate is sufficiently high ($H \gg 1$). The underlying assumption is that the solutions are in chemical equilibrium with the mineral, as is always the case for a sharp interface at the Darcy scale, and so the surface rate constant is not relevant. However, at the pore scale it is quite possible to have a sharp mineral boundary but a diffuse concentration, so in this sense the approach is not entirely realistic. A large reaction rate is needed to match the results of an explicit interface model, and it shares the drawback of phase-field methods that the surface reaction rate is not a direct input to the model. But the lack of a clear connection to the pore-scale physics is compensated for by the robustness and simplicity of the method.

An important feature of microcontinuum models is the use of the Brinkman equation to describe the flow (108,109,142). Unlike the Darcy equation, which breaks down in fully dissolved regions, the Brinkman equation transitions smoothly from matrix flow (Darcy) to void flow (Stokes). The connection between surface reaction rate and the volume dissolution of the porous matrix is made by a local estimate of the reactive surface area, $s = \psi|\nabla\phi|$, where the function $\psi$ serves to localize the dissolution in the region of the interface. However, the local surface area is somewhat arbitrary and is not as closely related to the geometry of the surface as in conforming-mesh methods (143), level-set methods (126), or VOF-based methods (138). It has the same strengths (robustness and flexibility) and weaknesses (imprecise and indirect connection between model parameters and physical properties) as other DIMs.

3.4.4. Flow and concentration fields. Discretized equations for flow and reactant transport are typically solved by either finite-volume, LB, or particle-tracking methods. Fine-element methods are infrequently used in reactive transport simulations (87).

3.4.4.1. Finite-volume methods. Finite-volume discretization approximates weak solutions of the field equations. The individual terms in Equations 12–14 can be written as a divergence $\nabla \cdot j_\psi$ of a flux $j_\psi = u\psi - B\nabla\psi$, where $\psi$ can be a scalar field such as $p$ or $c$ or a vector field such as $\mathbf{u}$. Partial differential equations can be converted to algebraic equations term by term, using abstract C++ classes (144): $\int_V \nabla \cdot j_\psi \, dx \approx \sum_f j_{\psi,f} \cdot A_f$, where $A_f$ points in the direction of the normal and has a magnitude equal to the area of the face. The location of the field variables and fluxes ensures that the scheme is conservative and second order. The control volumes can be any convex polyhedron, without the restrictions imposed on finite elements by the choice of basis functions.

Finite-volume methods can also be applied to voxelized meshes (78), which are frequently derived from X-ray tomography measurements of core samples; here, the mesh is a regular array of cubic control volumes. Smoother interfaces can be represented by including cut cells near the mineral interface (138). For a structured mesh of rectangular prisms the finite-volume method reduces to a finite-difference approximation, but with boundary conditions on the faces rather than the nodes. Finite-volume approaches also underpin most Darcy-scale reactive transport codes, such as TOUGHREACT, CRUNCHFLOW, and PFLOTRAN (62).
3.4.4.2. **Lattice Boltzmann methods.** LB methods are based on a discrete molecular mechanics (145), later converted to a discrete kinetic theory by ensemble averaging (146,147). An early application, calculating the permeability of sandstone from X-ray tomography images (148), indicated the potential of the method for flows in porous media. The apparent simplicity of LB models generated significant interest and contributed to the development of a range of applications including multiphase flows (149) and reactive transport models (150).

Despite the successes of LB methods, we believe there are reasons for concern about its application to more complex physics. Even at the simplest level of single-phase incompressible flow, there have been 30 years of publications on technical details such as how to apply a body force (151) or a boundary condition (152), with no clearly discernible best practice. In contrast, for the finite-volume method both issues are uncontroversial. Furthermore, lattice models are inherently plagued by broken translational and rotational symmetries. Approximately Galilean-invariant models can be derived for single-phase flows (153,154), but models for multiphase flows have many more unpleasant artifacts, which have to be eliminated or minimized on a case-by-case basis.

Speed and simplicity are frequently cited as advantages of LB methods, but in recent years much of the machinery of computational fluid dynamics has been introduced into LB models: grid refinement (155), implicit time stepping (156), and interface tracking with VOF (157) or level-set methods (158,159). While these refinements can improve the accuracy and efficiency of the methods, they add considerably to their complexity. Thus, it seems reasonable to ask what advantages LB methods confer to compensate for the overhead of working with 19 fields (D3Q19) versus 4 \((p-u)\) or 7 fields (D3Q7) per component versus 1 \(c\). Moreover, while the field can be unambiguously calculated from the distribution, the reverse is not true; initializing or removing mass as the phase changes is therefore more complicated and uncertain than it is with finite-volume methods.

3.4.4.3. **Particle tracking.** Various forms of particle tracking have been used to solve for the flow and concentration fields. In their simplest form, particles represent point sources of mass that superpose molecular diffusion on top of advection in the local flow field. The key computational advantage of particle tracking is that it does not suffer from numerical dispersion; thus, it is particularly advantageous for simulating reactant transport (160,161), because the Péclet number is typically 1,000 times larger than the Reynolds number \((Re)\). On the other hand statistical errors can be a source of significant uncertainty, although on occasion they can be reduced by a clever choice of algorithm (162). Furthermore, boundary conditions are less clear cut than in grid-based methods, and there are many opportunities for subtle but significant errors (162).

In reactive transport simulations, particle-based methods such as smooth-particle hydrodynamics (163,164) or vortex methods (165) are less popular than grid-based methods. Low-\(Re\) flows, where momentum diffusion is widespread and convection is less important, do not play to the strengths of particle methods, which are better suited to high-\(Re\) flows (165). Velocity fields from smooth-particle hydrodynamics include statistical noise and are not entirely incompressible (166), although a feedback mechanism (pressure splitting) can be used to reduce density fluctuations (167). Particle tracking often involves grid-based computations as well, either as part of the solution (165) or to redistribute the particles more efficiently. Calculations of necessity involve a significant number of neighboring particles and are thus less local than finite-volume methods.
3.4.5. **Pore-network models.** Network models introduce a simplified representation of a porous material, either as a network of interconnected capillaries (10, 17, 30) or as spherical pore bodies connected by cylindrical throats (168–170). To simulate the increase in permeability, the diameter of each network segment is increased in proportion to the local reactant consumption. Pore-network models allow for control of the pore architecture (diameters, lengths, connectivity), tuning it to represent different rocks; this is much harder at the Darcy scale. The fidelity of network models has been significantly improved by linking their initial geometry to pore structures obtained from X-ray tomography (170–173). Pore networks generated from tomographic data have been used to reconstruct the diagenetic history of carbonate samples (168) and to trace the evolution of permeability, dispersion, and effective reaction rate during chemical transformations of rocks (174).

Capillary-network models (10, 17) represent the pore space as connected cylinders (diameter $d$). A sparse system of linear equations can be solved to give the pressure at each node and the volumetric flux $q$ in each capillary. The concentration of reactant decays exponentially along the capillary: $c(x) = c_0 \exp(-\pi k_{\text{eff}} dx/q)$, where the rate constant $k_{\text{eff}}$ is corrected for diffusion across the pore, as described after Equation 9. Assuming complete mixing at the intersections, the concentration at the nodes ($c_0$) can be found by solving another sparse linear system. The local expansion of the capillary follows by analogy with Equation 9: $\partial_t (d/2) = k_{\text{eff}} c/c_{\text{sol}}$. If the pores are short in comparison with the penetration length, the capillaries broaden uniformly along their length. The erosion flux can then be integrated along the capillary, and the resulting dissolved volume distributed uniformly (17). Capillary-network models have been used to study the influence of pore heterogeneity (17, 175) and coupling between dissolution and rock mechanics, including dissolution-induced compaction (176).

Pore-capillary models are created by connecting spherical pores with cylindrical pipes (169, 171, 172). They have been used to study complex geochemical systems related to degradation of well-bore cements (171) and CO$_2$ sequestration (172), including changes in pore and throat diameters in response to dissolution and precipitation reactions. However, a frequently overlooked complication is that pore spaces can merge during dissolution, so that the topology of the network is not constant (17). Omitting pore merging can lead to unrealistic pore structures.

Acidization reactions frequently have a simple chemistry (e.g., Equation 4) but nevertheless are a severe test of the ability of network models to capture the transport of ions across thin boundary layers. Finite-element modeling of individual pores has been used to estimate mass transfer coefficients for the components of a pore-network model, which was then used to simulate acidization (169). However, because the pores remain spherical, the emerging wormholes constitute collections of larger and smaller spheres, with a rather remote resemblance to the real geometries.

4. **CHALLENGES**

There are many areas where reactive transport modeling can have an impact on our understanding of subsurface hydrology and geochemistry. In this section, we outline a selection of topics that seem ripe for further investigation. This is a personal choice within a broad field and is probably unduly biased by problems we have worked on and papers that have influenced our thinking.
4.1. Upscaling and the Evolution of Reactive Surface Area

The concept of reactive surface area (units of m$^{-1}$) is critical to upscaling; it serves to convert a surface dissolution rate into a volume one (177). However, it is most often used as an empirical factor to bring theoretical and simulation results in line with experiments or field measurements. The key problem with the concept of reactive surface area is that it fails to consider the variation in reactant concentration on the mineral surface, which is crucial in predicting dissolution rates when the surface reaction rate is high. Adding to the complexity, transport corrections change as dissolution or precipitation of minerals alters the flow paths in the sample. Detailed pore-scale simulations, coupled with well-controlled experiments, seem to be the most promising path forward.

4.2. Predicting Dissolution Patterns

Reactive transport modeling is increasingly able to predict the evolution of permeability and porosity from tomography data (161). However, predicting the patterns emerging from chemical erosion of fractured or porous rocks (e.g., Figures 3 and 4) remains an interesting theoretical challenge. The first step might be a qualitative prediction of the dissolution regime: face dissolution, wormholing, or uniform dissolution (10, 178). At the field scale this task seems daunting, in view of the large differences between laboratory- and field-scale reaction rates (179), but at the core scale it is within reach of existing numerical models (180). The next step would be statistical predictions of average geometrical properties and effective reaction rates (181, 182). The ultimate goal of a reactive transport model would be to predict the locations and topology of the wormholing pattern on the basis of initial data.

Comparisons between simulations and experiment may be easier in dissolving fractures, where the topography can sometimes be limited to a 2D aperture field (Section 2.3). Wormholing in dissolving cores is harder to predict, because the pore-scale topography is fully 3D. Undoubtedly, the predictive capabilities of numerical models are increasing, but the geometric and chemical complexity of the materials and the strong nonlinear couplings between flow, transport, and reaction make prediction of dissolution patterns a demanding task, particularly at the core and field scales. Interestingly, an empirical model for permeability within a Darcy-scale simulation was able to predict the positions and shapes of emerging wormholes with remarkable fidelity (183). This result might suggest that detailed pore-scale information is not entirely necessary, given a better understanding of the connection between porosity distribution and flow.

Multiphase reactive transport is crucial for understanding surface karst formation. Solution pipes, such as those depicted in Figure 3, can be formed both in the vadose zone (Figure 3a) and in water-saturated conditions (Figure 3b). The latter are usually linked to deglaciation processes (184), where karstification is most intense (185) and the piezometric water level within the ice sheet is above ground. Such pipes can be modeled using Darcy’s law (25). However, pipes formed under subtropical and Mediterranean climates are usually associated with vadose conditions (22), and the effects of partial saturation must be included. Understanding the formation of solution pipes in the vadose zone may require a multiphase model beyond the approximations contained in the Richards equation, which is commonly used to describe flow in partially saturated soils.
4.3. Nanoscale Processes

Processes occurring below the pore scale (<10 µm) pose a significant challenge for modeling. Although it is now possible to quantitatively predict the outcome of simple experiments (109, 117), there are many new processes to consider in order to understand the more complex experiments outlined in Section 3.2. Note that we have selected only a few examples from a much broader experimental literature, those which seemed particularly interesting to us and relevant to the development of reactive transport models. Phenomena of interest that are coming within reach of numerical models include electrostatic interactions, activity corrections, multiphase flows, fracture–matrix interactions, and porosity generation during replacement.

Most reactive transport simulations assume that all aqueous species have the same diffusion coefficient, preventing differential diffusion from creating a local imbalance of charge. If species are to have independent mobilities, then the force due to the electric field from the ions needs to be included (61, 64). One important effect of electrostatic interactions is to reduce the diffusion constant of ionic species at small concentrations. A recent validation study for modeling gypsum dissolution showed a marked improvement in the predicted dissolution rate when activity corrections to the diffusion coefficient were included (117). A practical consequence of differential diffusion is that it can drive convective flows in narrow and dead-end pores, which can be the controlling mechanism in replacement reactions (102). In narrow pores, even neutral species can be convected by diffusio-osmosis (186), which generates a large hydrodynamic slip at the pore walls.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Work at the University of Florida was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division under award number DE-SC0018676. Work at the University of Warsaw was supported by the National Science Center (Poland) under research grant number 2016/21/B/ST3/01373.

LITERATURE CITED

30. Fredd CN, Fogler HS. 1998. Influence of transport and reaction on wormhole formation in
porous media. *AIChE J.* 44:1933–49


93. Plummer LN, Wigley TLM, Parkhurst DL. 1978. The kinetics of calcite dissolution in CO₂–water systems at 5°C to 60°C and 0.0 to 1.0 atm of CO₂. Am. J. Sci. 278:179–216
metamorphism of the continental crust governed by nanoscale fluid flow. Nat. Geosci. 10:685–90
shapes of a dissolving mineral grain: comparisons of simulations with microfluidic experiments. Chem. Geol. 540:119459


28 Ladd & Szmyczak


185. Lauritzen SE, Skoglund RO. 2013. Glacier ice-contact speleogenesis in marble stripe karst. In