CheProf User's Guide

Maarten W. Saaltink

Department of Civil and Environmental Engineering,
Universitat Politècnica de Catalunya

2024

Table of Contents

1 Introduction

CheProf (CHEmical PRocesses Object-oriented with Fortran 2003) is software that performs hydrogeochemical calculations. It is written in Fortran 2003 and uses the technique of Object-Oriented Programming (Rouson, Xia and Xu 2011). It is based on the work of Bea et al. (2009), but it is more generalized and takes better advantage of Fortran 2003 features. CheProf is intended to be coupled to other codes, such as codes for solute transport, thus obtaining codes for reactive transport. To date it has been coupled to some simple 1D codes, explained in this user's guide (chapter 3) and coupling is in progress to CodeBright, a finite element code for thermo-hydro-mechanical models (Olivella, et al. 1996).

CheProf can handle various types of phases, such as aqueous phases (with the extended Debye-Hückel equation for activities), gases, mineral phases and cation exchange surfaces. Reactions between the chemical species of these phases can be assumed at equilibrium or to follow a kinetic approach. Other features, such as the effect of temperature on equilibrium constants and kinetic rate laws, electrostatic adsorption models or the Pitzer equation for calculating activities of aqueous phases, have not (yet) been incorporated, but may be in the (near) future. The main function of CheProf is hydrogeochemical speciation, that is, the calculation of the chemical composition given the chemical system and some constraints. CheProf has its own format to read files with chemical data and information, but it can also read data with formats for Phreeqc (Parkhurst and Appelo 2013).

This user's guide has been divided into different chapters. Chapter 2 explains some theory and concepts necessary for proper use of (codes that use) CheProf. Chapter 3 gives a full description of CheProf input files. It also presents four simple codes that use CheProf: 1) CheProfReadWrite, that may be used for CheProf input preperation, 2) ChemMix, that calculates the chemical chemistry after mixing of two end members, 3) WMA1D, that calculates 1D reactive transport using the Water Mixing Approach, and 4) DSA1D, that also calculates 1D reactive transport but using the global implicit or direct substitution approach. Finally, chapter 4 explains the structure of the code. It is intended to readers who are interested in code development or coupling CheProf to other codes.

2 Theory

2.1 Basic Concepts

Porous media consists of various phases, where we use the term phase in its thermodynamic sense. In thermodynamics, a phase is a portion of the medium with properties such as chemical composition, density, which are different from other phases. These phases are separated through clear boundaries. Each phase consists of one or various chemical species. We define a chemical species as any chemical entity distinguishable from other species due to: 1) its chemical composition, and 2) by the phase at which it is present. A chemical reaction is the process by which some species (reactants) are consumed to yield other species (products). Figure 1 illustrates these concepts through a typical but simple example. There are two solid phases (quartz and calcite) that each consist of only one species ($SiO_2(s)$ and $CaCO_3(s)$, respectively), one liquid (water or aqueous phase) and one gas both of which consist of various chemical species. Note that we consider $H_2O(g)$ and $CO_2(g)$ in the gas phase a different species than $H_2O(aq)$ and $CO_2(aq)$ dissolved in water. The same holds for solid $SiO_2(s)$ and dissolved $SiO_2(aq)$. Further note that we always distinguish the phase from the species (calcite versus $CaCO_3(s)$ and quartz versus $SiO_2(s)$), even if the phase only consists of one species and could be considered the same entity. The reason is that in this way pure phases can be considered a special case of phases with various species and do not require a special treatment.

Phases	Species
Calcite	CaCO3(s)
Quartz	SiO2(s)
Gas	N_2, $CO_2(g)$, $H_2O(g)$
Aqueous	$H_2O(aq)$, H^+, OH^-, Ca^{2+}, HCO_3^-, $CO_2(aq)$, $SiO_2(aq)$

Reactions

$CO_2(g) = CO_2(aq)$
$H_2O(g) = H_2O(aq)$
$H_2O(aq) = H^+ + OH^-$
$CO_2(aq) + H_2O(aq) = H^+ + HCO_3^-$
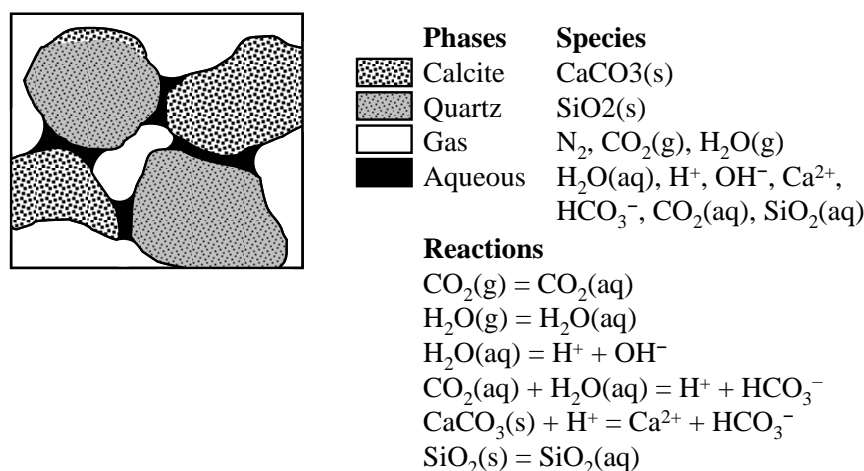$CaCO_3(s) + H^+ = Ca^{2+} + HCO_3^-$
$SiO_2(s) = SiO_2(aq)$

Figure 1. Example of phases, species and reactions in a porous medium at the pore scale.

Concentrations (for establishing chemical composition) can be defined in various ways depending on the type of phase (see Table 1). Each phase has one constraint that has to be met depending on the way its state variables are defined. Although in theory the choice of state variables is arbitrary, in practice it is dictated by thermodynamic theory or laboratory habits on which all thermodynamic data depend. For instance, for aqueous phases one often uses molality (mol per mass of water), because it can be easily fixed by dissolving a weighed amount of salt into a weighed amount of water. We prefer using molarity (mole per volume of phase) as this is the most suitable when dealing with fluxes in porous media. Anyway, one can always calculate the value of one type of state variable from that of another. Table 1 shows the equations to calculate molarity from the state variables. Often these equations can be simplified. For example, for fresh water one can

assume $c = \rho_{aq}x_j$. Note, that the phase constraint for minerals also takes into account that minerals phases are present ($\theta_{min} > 0$) only when the sum of molar fractions equals one. To do so, like Kräutle (2011) we use a complementarity function, φ, defined as:

$$
\begin{aligned}
\varphi(a,b) &= 0 \\
&\Leftrightarrow (a = 0 \text{ and } b \geq 0) \text{ or } (a \geq 0 \text{ and } b = 0) \\
&\Leftrightarrow ab = 0 \text{ and } a \geq 0 \text{ and } b \geq 0
\end{aligned}
\tag{1}
$$

An example of a complementarity function is $\varphi = \min(a,b)$. Typically, minerals only consist of one species, which means that its molar fraction equals one if the mineral is present. If the mineral is not present ($\theta_{min} = 0$), the molar fraction of its only species is not fixed, but may be defined by a mass action law (see next section).

Table 1. Definitions of concentrations normally used for phases. M is molecular mass (kg mol^{-1}), V_{min} is molar volume (m^3 mol^{-1}), z is the charge, ρ is density (kg m^{-3} of phase) and θ is volumetric fraction of phase (m^3 of phase m^{-3} of porous medium). Note the difference between molality (mol kg^{-1}) and molarity (mol m^{-3}).

Phase, α	Chemical state variable, c	Molarity, m	Constraint
Aqueous	Molality (mol kg$_{H2O}{}^{-1}$)	$m_j = \dfrac{\rho_{aq}c_{aq,j}}{1 + \underset{k \neq H_2O}{\sum} c_{aq,k}M_{aq,k}}$	$c_{H2O} = \dfrac{1}{M_{H2O}}$
Gas	Partial pressure (atm)	$m_j = \dfrac{c_{gas,j}}{RT}$	$\displaystyle\sum_k c_{gas,k} = P_{gas}$
Mineral	Molar fraction (mol mol^{-1})	$m_j = \dfrac{c_{min,j}}{V_{min}}$	$\varphi\left(\theta_{min}, 1 - \displaystyle\sum_k c_{min,k}\right) = 0$
Cation exchange surface	Equivalent fraction (eq eq^{-1})	$m_j = \dfrac{c_{exch,j}}{|z_j|}$	$\displaystyle\sum_k c_{exch,k} = 1$

2.2 Chemical equations of state

Given pressure, temperature and chemical composition of a phase we can calculate through equations of state all its other thermodynamic properties such as density, specific enthalpy, specific entropy, specific heat capacity, thermal conductivity, viscosity and molecular diffusion coefficient. Equations require assumptions that can be very different for each case. For instance, the density of the gas in soils can be safely assumed constant ($\rho_{gas} = 1.2$ kg m^{-3}) if the composition, pressure and temperature of the gas normally doesn't change much. If composition or temperature does change and pressure remains small (around 1 atm), the density of gas can be calculated assuming an ideal gas.

$$
\rho_g = \frac{1}{RT}\sum_j c_{gas,j}W_j
\tag{2}
$$

where $c_{gas,j}$ is the partial gas pressure as in Table 1. For cases with high pressure, such as deep CO_2 injection, the density should be calculated by a more complicated equation of state, such as, that of Redlich and Kwong (1949).

An important chemical equation of state is the mass action law for equilibrium reactions. It is formulated as:

3

$$\sum_{j}^{n_s} S_{e,ij} \log a_j = \log k_i \quad i = 1 \ldots n_{re} \tag{3}$$

where S_e are stoichiometric coefficients for equilibrium reactions, k is the equilibrium constant, which may depend on temperature and pressure, and n_s and n_{re} are the number of species and equilibrium reactions, respectively. It can also be written in matrix-vector notation:

$$\boldsymbol{S}_e \log \boldsymbol{a} = \log \boldsymbol{k} \tag{4}$$

The activity, a, can be understood as a concentration that has been corrected for the effects of the interaction of the various chemical species of a mixture on its enthalpy. In case of ideal mixtures or phases, no such effects are assumed and activity equals the chemical state variable of Table 1 ($a_j = c_j$). For non-ideal phases, the correction is expressed by means of an activity coefficient, γ, such that $a_j = \gamma_j c_j$. For aqueous phases, an expression often used is that of Debye and Hückel (1923), extended with $\dot{b}$ term.

$$a_{aq,j} = \gamma_j c_{aq,j} \quad j = 1, \ldots, n_{s,aq}$$

$$\log \gamma_j = \frac{-A z_j^2 \sqrt{I}}{1 + B \mathring{a}_j \sqrt{I}} + \dot{b} I \quad j = 1, \ldots, n_{s,aq} \tag{5}$$

$$I = \frac{1}{2} \sum_{j \neq H_2O} c_{aq,j} z_j^2$$

where $n_{s,aq}$ is the number of aqueous species, I is ionic strength, z is charge, $\mathring{a}$ is ion size, $\dot{b}$ is the Truesdell-Jones parameter and A (= 0.5092 at 25 ^{0}C) and B (= 0.3283 at 25 ^{0}C) are constants that depend on temperature. The activity of the species water (H_2O) is not given by equation (5) and is often assumed to equal unity. However, one can also take into account osmotic and capillary effects (should be checked):

$$a_{H2O} = \left(1 + M_{H2O} \sum_{j \neq H_2O} c_{aq,j}\right)^{-1} \exp\left(\frac{(P_{aq} - P_{ref}) M_{H2O}}{\rho_{aq} R T}\right) \tag{6}$$

where M_{H_2O} is molecular mass of water (0.018 kg mol^{-1}), P_{aq} is the pressure of the aqueous phase (Pa), P_{ref} is a reference pressure (1 atm = 1.01325·10^5 Pa), R is the gas constant (8.314 J mol^{-1} K^{-1}), T is the absolute temperature (K) and ρ_{aq} is the density of the aqueous phase (kg m^{-3}).

For kinetic reactions a reaction rate is expressed through a rate law. There exist many formulations for rate laws. Here we only discuss two. The first is a rate law for the precipitation-dissolution of minerals. A general form is given by Lasaga et al. (1994) with various terms i:

$$\text{if} \left((-\Omega \leq \Omega_{thr} \leq 0) \text{ or } (\Omega_{thr} \geq \Omega)\right) \text{ and } \theta_\alpha \leq 0 \quad r_k = 0$$

$$\text{else} \quad r_k = \sigma_\alpha \sum_i \lambda_i (1 - \Omega^{p_i})^{q_i} \exp\left(-\frac{E_{a,i}}{R}\left(\frac{1}{T} - \frac{1}{298.15}\right)\right) \prod_j a_j^{n_{i,j}} \tag{7}$$

4

$$\log \Omega = \sum_{j}^{n_s} S_{k,j} \log a_j - \log k$$

$$\text{if } \Omega \le 1 \quad \sigma_\alpha = g(\theta_\alpha + \theta_{\alpha 0})^h$$

$$\text{if } \Omega > 1 \quad \sigma_\alpha = g(\theta_\alpha + \theta_{\alpha 0})^h - g(\theta_{\alpha 0})^h$$

where r_k is the reaction rate (mol s^{-1} m^{-3} of porous medium), $\log \Omega$ is the saturation index of the reaction, $\log \Omega_{thr}$ is a threshold saturation index (should be positive when the mineral is a reactant, negative if it is a product), S_k are stoichiometric coefficients for the kinetic reaction, λ are kinetic rate constants at 25 ^{0}C (= 298.15 K), σ_α is the specific area of phase α (m^2 m^{-3} of porous medium), θ_α and $\theta_{\alpha 0}$ is the volumetric fraction and minimum volume fraction of phase α (m^3 m^{-3} of porous medium), E_a is the activation energy (J mol^{-1}), T is the absolute temperature (Kelvin), n_j represents the reaction order with respect to catalyzer species j and p and q are empirical parameters (that are by default 1). A compilation of rate parameters λ, E_a, n, p and q can be found in Palandri and Kharaka (2004). Parameters g and h depend on granular shape. If the mineral is assumed to have a geometry of a slab, the reactive surface is constant. Hence, g is this constant reactive surface and h is 0. If the mineral is assumed to consist of spherical grains, $\sigma_\alpha = \sigma_{\alpha,ref}(\theta_\alpha/\theta_{\alpha,ref})^{2/3}$ where $\sigma_{\alpha,ref}$ and $\theta_{\alpha,ref}$ are reference specific area and volume fraction of mineral phase α, respectively. Hence, $g = \sigma_{\alpha,ref}/(\theta_{\alpha,ref})^{2/3}$ and $h = 2/3$. The minimum volume fraction $\theta_{\alpha 0}$ is used to avoid non-uniqueness and instabilities. Low values can be used.

A second type of rate law is based on Monod kinetics and is applied to oxidation of organic matter by micro-organisms. It catalyzes or inhibits the reactions by a given electron acceptor, if previous electron acceptors in the oxidation sequence are present. It can be formulated as (Van Cappellen and Gaillard 1996):

$$r_k = \theta_\alpha \lambda \prod_i c_{j(i)} \prod_i \frac{c_{j(i)}}{\kappa_i + c_{j(i)}} \prod_i \frac{\kappa_i}{\kappa_i + c_{j(i)}} \tag{8}$$

where r_k is the reaction rate (mol s^{-1} m^{-3} of porous medium), θ_α is the volumetric fraction of phase α (m^3 m^{-3} of porous medium) and λ and κ are kinetic rate constants.

2.3 Degrees of freedom

The number of degrees of freedom is the number of variables that describe a system minus the number of equations of state that constrain it. In order to be able to calculate the properties of the system some equations have to be added. Typically, they are balance equations for mass or energy or equations that fix some variable.

Gibbs' phase rule calculates the degrees of freedom when one wants to determine the thermodynamic properties of all phases of a system without capillarity (that is one pressure for all phases). Table 2 gives the variables and equations of state and their numbers. More variables and equations could be added. However, for every additional thermodynamic property there is an equation of state that writes that property as a function of the variables. Hence, it will not affect the number of degrees of freedom.

Table 2. Variables and equations of state for Gibbs' phase rule. Number of degrees of freedom: $n_f = n_s - n_e - n_p + 2 = n_c - n_p + 2$.

Variable	Number	Equations of State	Number
Pressure, P	1	Phase constraint	n_p
Temperature, T	1	Mass action law	n_e
Concentration, c	n_s		

However, the phase rule says nothing on the amount of a phase in a system. In a porous medium this amount (volume or volume fraction) is important and the number of degrees of freedom changes. Moreover, there can be capillary forces and, therefore, fluid phases (liquids and gas) can have different pressures. Table 3 shows the variables and equations of state for a porous medium within a fixed volume. It distinguishes between fluid phases (liquids and gasses) and solid phases. It does not consider deformations or momentum.
Table 4 does the same for a system with fixed pressures and temperature.

Table 3. Variables and equations of a porous medium with fixed volume. Number of degrees of freedom: $n_f = n_s - n_e + 1 = n_c + 1$.

Variables	Number	Equations of State	Number
Pressures of fluid phases, P_α	n_{fp}	Phase constraints (Table 1)	n_p
Temperature, T	1	Mass action laws	n_e
Concentrations, c	n_s	Retention curves	$n_{fp} - 1$
Volume fractions, θ_α	n_p	$\sum_\alpha \theta_\alpha = 1$	1

Table 4. Variables and equations of a porous medium with fixed pressures and temperature. Number of degrees of freedom: $n_f = n_s - n_e = n_c$.

Variables	Number	Equations of State	Number
Concentrations, c	n_s	Phase constraints	n_p
Volume fractions, θ	n_p	Mass action laws	n_e

2.4 Components

To explain the concept of components it is helpful to use vector and matrix notation. For this, we define a vector of molarities ($\boldsymbol{m}$ in mol m^{-3} of phase) subdivided into molarities for liquid, gas and immobile phases (usually solid phases), indicated by subscript l, g, i:

$$\boldsymbol{m} = \begin{array}{|c|} \hline \boldsymbol{m}_l \\ \hline \boldsymbol{m}_g \\ \hline \boldsymbol{m}_i \\ \hline \end{array}$$

and a vector of reaction rates ($\boldsymbol{r}$ in mol s^{-1} m^{-3} of porous medium) subdivided into rates for equilibrium and kinetic reactions, indicated by subscripts e and k:

$$\boldsymbol{r} = \begin{array}{|c|} \hline \boldsymbol{r}_e \\ \hline \boldsymbol{r}_k \\ \hline \end{array}$$

and a stoichiometric matrix (S), subdivided into matrices for each phase and for equilibrium and kinetic reactions:

$$S = \begin{array}{|c|} \hline S_e \\ \hline S_k \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline S_l & S_g & S_i \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline S_{el} & S_{eg} & S_{ei} \\ \hline S_{kl} & S_{kg} & S_{ki} \\ \hline \end{array}$$

With these definitions we can formulate the mass balances of all species in a closed system without transport:

$$\begin{pmatrix} \dfrac{\partial \theta_l \boldsymbol{m}_l}{\partial t} \\ \dfrac{\partial \theta_g \boldsymbol{m}_g}{\partial t} \\ \dfrac{\partial \boldsymbol{m}_i^{\theta}}{\partial t} \end{pmatrix} = \boldsymbol{S}^T \boldsymbol{r} = \boldsymbol{S}_e^T \boldsymbol{r}_e + \boldsymbol{S}_k^T \boldsymbol{r}_k = \begin{pmatrix} \boldsymbol{S}_l^T \boldsymbol{r} \\ \boldsymbol{S}_g^T \boldsymbol{r} \\ \boldsymbol{S}_i^T \boldsymbol{r} \end{pmatrix} = \begin{pmatrix} \boldsymbol{S}_{el}^T \boldsymbol{r}_e + \boldsymbol{S}_{kl}^T \boldsymbol{r}_k \\ \boldsymbol{S}_{eg}^T \boldsymbol{r}_e + \boldsymbol{S}_{kg}^T \boldsymbol{r}_k \\ \boldsymbol{S}_{ei}^T \boldsymbol{r}_e + \boldsymbol{S}_{ki}^T \boldsymbol{r}_k \end{pmatrix} \tag{9}$$

$$\boldsymbol{m}_i^{\theta} = \begin{pmatrix} \theta_{i1} \boldsymbol{m}_{i1} \\ \theta_{i2} \boldsymbol{m}_{i2} \\ \vdots \end{pmatrix} \tag{10}$$

where θ_α is the volume fraction of phase α (m^3 of phase m^{-3} of porous medium) and $\boldsymbol{m}_i^{\theta}$ is a vector of molarities of species of immobile phases expressed in mol m^{-3} of porous medium. Equation (10) indicates that there can be several immobile phases each with one or more species. Equation (9) can be simplified by multiplying it by the kernel or null space (U) of the equilibrium stoichiometric matrix (S_e) (Steefel and MacQuarrie 1996). By definition $\boldsymbol{US}_e^T = \boldsymbol{0}$. The number of columns of U is equal to the number of species (n_s) and its number of rows is equal to the number of species minus the rank of matrix $\boldsymbol{S}_e$ ($n_s - \text{rank}(\boldsymbol{S}_e)$). Normally the equilibrium reactions can be considered linearly independent of each other, that is, $\text{rank}(\boldsymbol{S}_e) = n_e$ (n_e being the number of equilibrium reactions or rows of $\boldsymbol{S}_e$). The result of the multiplication is a set of mass balances of so-called components. It eliminates the rates of equilibrium reactions ($\boldsymbol{r}_e$) and reduces the number of mass balance equations to $n_s - n_e$:

$$\boldsymbol{U}_l \frac{\partial \theta_l \boldsymbol{m}_l}{\partial t} + \boldsymbol{U}_g \frac{\partial \theta_l \boldsymbol{m}_g}{\partial t} + \boldsymbol{U}_i \frac{\partial \boldsymbol{m}_i^{\theta}}{\partial t} = \boldsymbol{US}_k^T \boldsymbol{r}_k \tag{11}$$

by defining component concentrations $\boldsymbol{w}_\alpha = \boldsymbol{U}_\alpha \boldsymbol{m}_\alpha$, $\alpha = l, g$, $\boldsymbol{w}_i = \boldsymbol{U}_i \boldsymbol{m}_i^{\theta}$ and $\boldsymbol{w} = \theta_l \boldsymbol{w}_l + \theta_g \boldsymbol{w}_g + \boldsymbol{w}_i$, we can rewrite (11) as:

$$\frac{\partial \boldsymbol{w}}{\partial t} = \frac{\partial \theta_l \boldsymbol{w}_l}{\partial t} + \frac{\partial \theta_g \boldsymbol{w}_g}{\partial t} + \frac{\partial \boldsymbol{w}_i}{\partial t} = \boldsymbol{US}_k^T \boldsymbol{r}_k \tag{12}$$
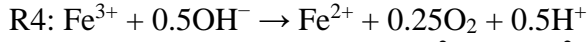
Note that aqueous and gaseous immobile concentrations are expressed in mol m^{-3} of *phase*, where immobile component concentrations are expressed in mol m^{-3} of *porous medium*.

The kernel matrix U is not unique and several methods exist for its calculation. Probably the most practical one is a method that uses the RREF (Reduced Row Echelon Form), a standard method in matrix analysis (Meier 2000) explained in Appendix A. The advantage of this method is that it also calculates the rank of a matrix and generate columns that together constitute an identity matrix and that we can associate to primary species.

To facilitate comprehension, the components and component matrices are illustrated with a small example of Molins et al. (2004). It has the three chemical reactions in equilibrium:

R1: $FeSO_4 \rightarrow Fe^{2+} + SO_4^{2-}$

R2: $H_2O \rightarrow H^+ + OH^-$

R3: $Fe(OH)_3(s) \rightarrow Fe^{3+} + 3OH^-$
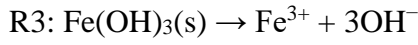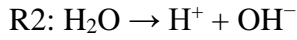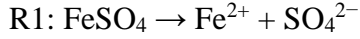
and two kinetic reactions:

R4: $Fe^{3+} + 0.5OH^- \rightarrow Fe^{2+} + 0.25O_2 + 0.5H^+$

R5: $FeS_2(s) + 3.5O_2 + OH^- \rightarrow Fe^{2+} + 2SO_4^{2-} + H^+$

All species are aqueous (or liquid) except the minerals $Fe(OH)_3(s)$ and $FeS_2(s)$ that are solid. There are no gaseous species. The stoichiometric matrix of this chemical system becomes:

$$\boldsymbol{S} = \begin{array}{cccccccccc} H_2O & O_2 & H^+ & OH^- & Fe^{2+} & Fe^{3+} & SO_4^{2-} & FeSO_4 & Fe(OH)_3(s) & FeS_2(s) \\ \end{array}$$

H_2O	O_2	H^+	OH^-	Fe^{2+}	Fe^{3+}	SO_4^{2-}	$FeSO_4$	$Fe(OH)_3(s)$	$FeS_2(s)$	
0	0	0	0	1	0	1	−1	0	0	R1
−1	0	1	1	0	0	0	0	0	0	R2
0	0	0	3	0	1	0	0	−1	0	R3
0	0.25	0.5	0.5	1	−1	0	0	0	0	R4
0	−3.5	1	1	1	0	2	0	0	−1	R5

As there are 10 species and 3 equilibrium reaction, there are $10 - 3 = 7$ components. Application of the method of RREF gives the following kernel matrix:

H_2O	O_2	H^+	OH^-	Fe^{2+}	Fe^{3+}	SO_4^{2-}	$FeSO_4$	$Fe(OH)_3(s)$	$FeS_2(s)$
1	0	0	1	0	0	0	0	3	0
0	1	0	0	0	0	0	0	0	0
0	0	1	−1	0	0	0	0	−3	0
0	0	0	0	1	0	0	1	0	0
0	0	0	0	0	1	0	0	1	0
0	0	0	0	0	0	1	1	0	0
0	0	0	0	0	0	0	0	0	1

$\boldsymbol{U} =$ (above)

The columns of primary species are indicated in grey. Note, that $\boldsymbol{US}_e^T = 0$. However, $\boldsymbol{US}_k^T$ is not:

$$\boldsymbol{US}_k^T = \begin{array}{cc} R4 & R5 \\ \hline -0.5 & -1 \\ 0.25 & -3.5 \\ 1 & 2 \\ 1 & 1 \\ -1 & 0 \\ 0 & 2 \\ 0 & -1 \\ \end{array}$$

The non-uniqueness of the kernel matrix can be further exploited to define components in such a way that some components become decoupled from others. To do so we define a kernel matrix with the following structure:

$$U = \begin{array}{ccc} n_l & n_g & n_i \end{array}$$
$$\begin{array}{|c|c|c|} \hline \multicolumn{3}{|c|}{U_{het}} \\ \hline U_{holc} & 0 & 0 \\ \hline 0 & 0 & U_{hoi} \\ \hline \end{array} \begin{array}{l} n_{het} \text{ heterogeneous components} \\ n_{holc} \text{ homogeneous liquid conservative components} \\ n_{hoi} \text{ homogeneous immobile components} \end{array}$$

where n_l, n_g and n_i are the number of species in liquid, gas and immobile phases. The different components (or rows of blocks of U) have specific characteristics:

− Heterogeneous components consist of species of any phase. They may involve kinetic but not equilibrium reactions. They are represented by the first row of blocks. The block pertaining to these components (U_{het}) must be a kernel of S_e. Taking into account the total number of components or rows of U ($n_s - \text{rank}\, S_e$) and the number of other components, explained below (n_{holc} and n_{hoi}), the number of heterogeneous components is: $n_{het} = n_g + \text{rank}\, S_a + \text{rank}\, S_{ei} - \text{rank}\, S_e$.

− Homogeneous liquid conservative components only consist of liquid species and involve in neither equilibrium nor kinetic reactions. They are represented by the second row of blocks. Blocks pertaining to other phases are zero. The block pertaining to the liquid phase (U_{holc}) must be a kernel of S_a (hence $U_{holc}S_a = 0$). Therefore, the number of liquid conservative components equals the number of liquid species minus the rank of matrix S_a: $n_{holc} = n_l - \text{rank}\, S_a$.

− Homogeneous immobile components only consist of immobile species. They may involve kinetic but not equilibrium reactions. The block pertaining to the immobile species (U_{hoi}) must be a kernel of S_{ei}, because only then equilibrium reactions are eliminated, but not necessarily the kinetic ones. Their number equals the number of immobile species minus the rank of S_{ei}: $n_{hoi} = n_i - \text{rank}\, S_{ei}$.

Matrices U_{holc} and U_{hoi} can be calculated straightforwardly from S_a and S_{ei} by means of RREF. The heterogeneous components (U_{het}) must be linearly independent of U_{holc} and U_{hoi}. To obtain this, U_{het} is the kernel of matrix S_a, of which we have removed the columns referring to primary species defined for U_{holc} and U_{hoi}. For the example this gives the following kernel:

	H_2O	O_2	H^+	OH^-	Fe^{2+}	Fe^{3+}	SO_4^{2-}	$FeSO_4$	$Fe(OH)_3(s)$	$FeS_2(s)$	
	0	0	0	0	1	0	0	1	0	0	
	0	0	0	0	0	1	0	0	1	0	het
$U =$	0	0	0	0	0	0	1	0	−3	0	
	1	0	0	1	−2.5	−3	−1.75	0	0	0	
	0	1	0	0	−0.25	0	1.875	0	0	0	holc
	0	0	1	−1	2	3	−2	0	0	0	
	0	0	0	0	0	0	0	0	0	1	hoi

The columns of primary species are indicated in grey. Note that rows US_k^T referring to homogeneous liquid conservative components become zero:

9

$$\boldsymbol{US}_k^T = \begin{array}{cc} \text{R4} & \text{R5} \\ \begin{array}{|cc|} \hline 1 & 1 \\ -1 & 0 \\ 1 & 2 \\ \hdashline 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \hdashline 0 & -1 \\ \hline \end{array} & \begin{array}{l} \\ \text{het} \\ \\ \\ \text{holc} \\ \\ \text{hoi} \end{array} \end{array}$$

2.5 Speciation

Speciation refers to the calculation of all chemical state variables. Besides the chemical equations of state, it requires a number of other equations (see section 2.3). Often these other equations are derived from mass balances of components (see section 2.4).

2.5.1 Speciation from concentrations of primary species

Given: Temperature, pressure of all fluid phases, volume fractions of all phases (θ_α) and concentrations of $n_c - n_{hoi}$ primary species ($\boldsymbol{c}_{1,given}$), $n_c = n_s - n_e$ being the number of components and n_{hoi} being the number of homogeneous immobile components. The primary species can be of different phases. Units are as stated in Table 1.
Unknowns:
− Concentrations of secondary species, $\boldsymbol{c}_2$ ($n_e + n_{hoi}$)
Equations:
− Mass action laws (n_e), see equation (3) or (4).
− Mass balances of homogeneous solid components (n_{hoi})

$$\sum_{\alpha=i} \theta_{i\alpha} \boldsymbol{U}_{hoi,\alpha} \boldsymbol{m}_{i\alpha}(\boldsymbol{c}_{i\alpha}) - \sum_{\alpha=i} \theta_{i\alpha}^{old} \boldsymbol{U}_{hoi,\alpha} \boldsymbol{m}_{i\alpha}(\boldsymbol{c}_{i\alpha}^{old}) - \sum_{\alpha=i} \boldsymbol{U}_{hoi,\alpha} (\boldsymbol{S}_k)^T \boldsymbol{r}_k(\boldsymbol{c},\theta) \Delta t = 0 \quad (13)$$

Phase constraints are not imposed.

2.5.2 Speciation from aqueous concentrations of components

Given: Temperature, pressure of all phases, and aqueous component concentrations ($\boldsymbol{w}_{given}$ in moles/volume of phase), which will be corrected for reactions with immobile phases.
Unknowns:
− $\boldsymbol{c}$ of all species (n_s)
− θ_α of all mineral and exchange phases
Equations:
− Mass action laws (n_e), see equation (3) or (4).
− Mass balances of all components ($n_c = n_s - n_e$)

$$\begin{pmatrix} \theta_l \boldsymbol{U}_{l,het} \boldsymbol{m}_l(\boldsymbol{c}_l) + \sum_{\alpha=i} \left(\theta_{i\alpha} \boldsymbol{U}_{het,i\alpha} \boldsymbol{m}_{i\alpha}(\boldsymbol{c}_{i\alpha}) - \theta_{i\alpha}^{old} \boldsymbol{U}_{het,i\alpha} \boldsymbol{m}_{i\alpha}(\boldsymbol{c}_{i\alpha}^{old}) \right) \\ \theta_l \boldsymbol{U}_{l,holc} \boldsymbol{m}_l(\boldsymbol{c}_l) \end{pmatrix} - \theta_l \boldsymbol{w}_{given} \quad (14)$$
$$= 0$$

$$\sum_{\alpha=i} \theta_{i\alpha} \boldsymbol{U}_{hoi,\alpha} \boldsymbol{m}_{i\alpha}(\boldsymbol{c}_{i\alpha}) - \sum_{\alpha=i} \theta_{i\alpha}^{old} \boldsymbol{U}_{hoi,\alpha} \boldsymbol{m}_{\alpha}(\boldsymbol{c}_{\alpha}^{old})$$
$$- \Delta t \boldsymbol{U}_{hoi}(\boldsymbol{S}_k)^T \left(\eta \boldsymbol{r}_k(\boldsymbol{c}, \theta) + (1-\eta) \boldsymbol{r}_k(\boldsymbol{c}^{old}, \theta^{old}) \right) = 0 \qquad (15)$$

where l and i refer to liquid and immobile phases and η is a time weighting factor. Note that the homogeneous immobile components of equation (15) by definition have zero aqueous concentrations. Mass balances of these components contain terms for kinetic reactions, which are not considered by mass balances of other components (equation (14)).

– Phase constraints of mineral and exchange phases (see Table 1).

2.5.3 Speciation from aqueous concentrations of components with kinetic terms

It is the same as the previous section 2.5.2, but kinetic terms are added to the mass balances of heterogeneous components. Instead of equation (14), the following equation is required:

$$\begin{pmatrix} \theta_l \boldsymbol{U}_{l,het} \boldsymbol{m}_l(\mathbf{c}_l) + \displaystyle\sum_{\alpha=i} \theta_{i\alpha} \boldsymbol{U}_{het,i\alpha} \boldsymbol{m}_{i\alpha}(\mathbf{c}_{i\alpha}) - \sum_{\alpha=i} \theta_{i\alpha}^{old} \boldsymbol{U}_{het,i\alpha} \boldsymbol{m}_{i\alpha}(\boldsymbol{c}_{i\alpha}^{old}) \\ \theta_l \boldsymbol{U}_{l,holc} \boldsymbol{m}_l(\mathbf{c}_l) \end{pmatrix} -$$
$$\begin{pmatrix} \Delta t \boldsymbol{U}_{het}(\boldsymbol{S}_k)^T \left(\eta \boldsymbol{r}_k(\boldsymbol{c}, \theta) + (1-\eta) \boldsymbol{r}_k(\boldsymbol{c}^{old}, \theta^{old}) \right) \\ 0 \end{pmatrix} - \theta_l \boldsymbol{w}_{given} = 0 \qquad (16)$$

This type of speciation is recommended for the Water Mixing Approach (see section 3.5). Note that by definition mass balances of conservative homogeneous liquid components do not contain kinetic terms.

2.5.4 Speciation from aqueous/immobile concentrations of components

It is the same as section 2.5.2, but the component concentrations in aqueous and immobile phases are used in moles/volume of porous medium. Instead of equation (14), the following equation is required:

$$\begin{pmatrix} \theta_l \boldsymbol{U}_{l,het} \boldsymbol{m}_l(\mathbf{c}_l) + \displaystyle\sum_{\alpha=i} \theta_{i\alpha} \boldsymbol{U}_{het,i\alpha} \boldsymbol{m}_{i\alpha}(\mathbf{c}_{i\alpha}) \\ \theta_l \boldsymbol{U}_{l,holc} \boldsymbol{m}_l(\mathbf{c}_l) \end{pmatrix} - \boldsymbol{w}_{given} = 0 \qquad (17)$$

3 Input file descriptions

3.1 XML

We use XML (Extensible Markup Language) to structure the input data or mark up the data with meta-information. The basic syntax is given below.
Every xml file must start with an xml declaration:
`<?xml version="1.0"?>`
All information is contained within elements. Each element is opened and closed by tags that are defined within triangular brackets < >, e.g.:
`<element1> </element1>`
Closing tags contain a /. One can also open and close the element in one tag:
`<element1/>`
Elements can be nested, e.g.
`<element1>`
   `<element1A/>`
   `<element1B> </element1B>`
`</element1>`
Each element can have one or more attributes, using name/value pairs, e.g.:
`<element1 AtrrName1="AtrrValue1" AtrrName2="AtrrValue2" />`
Attribute names are defined without quotation marks, whereas attribute values must always appear in quotation marks.
Text within <!-- and --> are not read by the code and can be used as comments, e.g.:
`<!-- This is a comment -->`
Although the official xml syntax is case sensitive, this code does not distinguish between lower- and uppercase.

3.2 Basic CheProf input

All codes that use CheProf need input on chemistry. There are two input files: 1) A file for the chemical system, containing all parameters of the chemical equations to be solved (see 3.2.1) and 2) a file for the local chemistries, containing the chemical composition necessary for, e.g., initial and boundary conditions (see 3.2.2).

3.2.1 Chemical System

Element <ChemSys> Chemical System
Optional attribute *Name*: Name of the Chemical system. Default name is 'Chemical system'.
Optional attribute *PhreeqcFile*: Name of an output file of Phreeqc. If it is empty (that is, PhreeqcFile = " "), only the present xml file is used to read the chemical system. If a filename is given, only names of phases and species are read. Other attributes of phases and species are read from the Phreeqc database file, written in *PhreeqcFile* (indicated by "Database file:"). Reactions in the xml file are ignored. Instead, reactions are read from the Phreeqc database file and are all assumed to be in equilibrium. Components are always read from the xml file. The default *PhreeqcFile* is " " (that is, Phreeqc output is not used).
Obligatorily contains one set of phases <PhaseS> (see 3.2.1.1).
Optionally contains one set of reactions <ReactionS> (see 3.2.1.2).
Obligatorily contains one set of components <ComponentS> (see 3.2.1.3).

Obligatorily contains numerical options and parameters <Numerics> (see 3.2.1.4).
Optionally contains output options for chemistry <Output> (see 3.2.1.5).

3.2.1.1 Phase

Element <PhaseS> Set of one or more phases
No attributes.
Obligatorily contains one or more <Phase>.

Element <Phase>
Obligatory attribute *Name*: Name of the phase. Two phases cannot have the same name.
Optional attribute *Mob*: Mobility of the phase. It can be:
   - *Liq*: The phase pertains to the group of phases that move as a liquid.
   - *Gas*: The phase pertains to the group of phases that moves as a gas.
   - *Imm*: The phase pertains to the group of phases that do not move.
   Default Mob is Liq if the phase is named Aqueous, Gas if the phase is named Gas, and Imm otherwise.
Optionally contains one <PhaseModel>. If no PhaseModel is specified, a default model is used (see 3.2.1.1.1).
Optionally contains one or more species <Species>. If no species are specified, the phase consists of one species with its name equal to the name of the phase and with all other attributes equal to default values (see 3.2.1.1.2).
Example:
<phase name="CaCO3(s)"/>
which is equivalent to:
<Phase Name="CaCO3(s)"> <PhaseModel Type="Mineral" MolVol="1.0"/>
   <Species Name="CaCO3(s)" MolMass="0.1000869" Charge="0.0" IonSize="0.0"
b_TJ="0.0"/>
</Phase>

3.2.1.1.1 Phase Model

Element <PhaseModel>
Optional attribute *Type*: Type of phase model. It can be:
   - *Aqueous*: Concentrations are molality (mol/kg water). Ideal activities. Activity of water is 1.0. The constraint of the phase is that the molality of water = (MolMass of water)$^{-1}$.
   - *Aqueous_DH*: Same as *Aqueous*, but activities are calculated with Debye Hückel (equation (5)).
   - *Gas*: Concentrations are partial pressure (atm). Ideal activities. The constraint of the phase is that the sum of partial pressures equals the gas pressure.
   - *Mineral*: Concentrations are molar fraction. Ideal activities. The constraint of the phase is a complementarity function: Min((1 − Sum of mol fractions),VolF) = 0.
   - *Mineral_Ubiq*: Same as *Mineral*, but the mineral is assumed ubiquitous. The constraint of the phase is: 1 − Sum of mol fractions = 0. This type must be used for minerals that precipitate or dissolve kinetically.
   - *Exchange*: Concentrations are equivalent fraction. Ideal activities. The constraint of the phase is that the sum of equivalent fraction equals 1.0.
   Default type is Aqueous if the phase is named Aqueous, Gas if the phase is named Gas, Exchange if the phase is named Exchange and Mineral otherwise.

Optional attribute *Water* (only for *Type = Aqueous* or *Aqueous_DH*): Name of the species that represents water. Default is H2O, if H2O is given in the set of species, and none if H2O is not given in the set of species.

Optional attribute *Dens* (only for *Type = Aqueous* or *Aqueous_DH*): Density of water (kg_{water}/m^3_{phase}). Default is 1000 kg/m^3.

Optional attribute *MolVol*: Only for *Type = Mineral*. Molar volume of the phase (m^3/mol). Default is 1.0 m^3/mol.

3.2.1.1.2 Species

Element <Species>

Obligatory attribute *Name*: Name of the species. Two species cannot have the same name, even if they belong to different phases.

Optional attribute *MolMass*: Molecular mass (kg/mol). Default is deduced from atomic masses of the chemical elements (with proper upper- and lower-case letters) in the name of the species (e.g. "CaSO4:2H2O" gives 0.172171 kg/mol and "Fe(OH)3(s)" gives 0.1068669 kg/mol).

Optional attribute *Charge*: Ion charge. Note that the ion charge of the electron ("e-"), being a fictitious species, should be zero. Default is deduced from the name of the species (e.g., "Na+" gives +1.0, "CO3-2" gives −2.0, "O2" gives 0.0 and "e-" gives zero).

Optional attribute *IonSize*: Ion size for DH (Debye Hückel) activity model (see 3.2.1.1.1 and equation (5)). Default is 0.0.

Optional attribute *b_TJ*: Truesdell-Jones parameter b for DH (Debye Hückel) activity model (see 3.2.1.1.1 and equation (5)). Default is 0.0.

Optional attribute *Fict*: Fictitious species are species that are used in mass action laws, but of which the actual concentrations are zero or negligible and should not be considered in mass balances of components. Typical examples of fictitious species are "e-" in aqueous phases and "X-" in ion exchange phases. *Fict* can be:
- *No*: The species is not a fictitious species.
- *Yes*: The species is a fictitious species.
Default value is "No".

3.2.1.2 Reaction

Element <ReactionS> Set of one or more reactions
No attributes.
Optionally contains one or more <Reaction>.

Element <Reaction>

Optional attribute *Name*: Name of the reaction. Default is a number, which coincides with the order of appearance in the file.

Obligatory attribute *Form*: Text containing the formula of the reaction. It contains "=" "+", "-", stoichiometric coefficients and names of species. The "=" separates reactants (to its left) from products (to its right). If no "=" is used, all are products. If no stoichiometric coefficients are used, a value of 1.0 is assumed. Stoichiometric coefficients of species with the same name are summed. Species must be followed by one or more spaces. All species must have been specified in <Species> (see 3.2.1.1.2).

Optional attribute *LogK*: Log equilibrium constant of the reaction. Default value is 0.0.

Optional attribute *Kin*: It can be:

- *No*: The reaction is assumed to be in equilibrium.
- *Yes*: The reaction is assumed to be kinetic.
  Default value is "No".
Obligatorily (if Kin="Yes") or optionally (if Kin="No") contains one <RateLaw> (see 3.2.1.2.1).
Example:
<reaction name="calcite diss" form="CaCO3(s) + H+ = Ca+2 + HCO3-" logk="-8.3"/>
which is equivalent to:
<Reaction Name="calcite diss" Form="CaCO3(s) + HCO3- = 1Ca+2 + 2 HCO3- - H+" LogK="-8.3" Kin="No"/>

3.2.1.2.1 Ratelaw

Element <RateLaw>.
Obligatory attribute: *Type*: Type of rate law. It can be "Lasaga" or "Monod".

3.2.1.2.1.1 Lasaga rate law

If type="Lasaga", equation (7) of page 4 is applied:

$$\text{if } \big((-\Omega \le \Omega_{thr} \le 0) \text{ or } (\Omega_{thr} \ge \Omega)\big) \text{ and } \theta_\alpha \le 0 \quad r_k = 0$$

$$\text{else} \quad r_k = \sigma_\alpha \sum_i \lambda_i (1 - \Omega^{p_i})^{q_i} \exp\left(-\frac{E_{a,i}}{R}\left(\frac{1}{T} - \frac{1}{298.15}\right)\right) \prod_j a_j^{n_{i,j}} \qquad (7)$$

$$\text{if } \Omega \ge 1 \quad \sigma_\alpha = g(\theta_\alpha + \theta_{\alpha 0})^h$$
$$\text{if } \Omega < 1 \quad \sigma_\alpha = g(\theta_\alpha + \theta_{\alpha 0})^h - g(\theta_{\alpha 0})^h$$

and <RateLaw> has the following additional attributes:
Obligatory attribute *Phase*: Name of the (mineral) phase (α) for volume fraction (θ) and specific surface (σ). It must have been specified in <Phase>.
Optional attribute *Thresh*: Threshold Ω (Ω_{thr}, $\log\Omega$ is the saturation index). Default is 1.0.
Obligatory attribute *GeomConst*: Granular geometry constant (g).
Optional attribute *GeomExp*: Granular geometry exponent (h). Default is 2/3.
Optional attribute *VolF0*: Minimum volume fraction ($\theta_{\alpha 0}$). Default is 0.0.
Optionally contains one or more <Term>.

Element <Term>.
Obligatory attribute *Const*: Kinetic rate constant (λ).
Optional attribute *Eact*: Activation activity (E_a, J Mol^{-1}). Default is 0.0.
Optional attribute *ParInt*: Internal rate law term parameter (p). Default is 1.0.
Optional attribute *ParExt*: Exteral rate law term parameter (q). Default is 1.0.
Optionally contains one or more <Catal>.

Element <Catal> Catalyzer.
Obligatory attribute *Species*: Name of catalyzing species. It must have been specified in <Species>.
Obligatory attribute *Order*: Order of the catalyzer (n).

Example for the rate law of albite dissolution (Palandri and Kharaka 2004):
<Ratelaw Type="Lasaga" Phase="Albite" GeomConst="1.0e4">
  <!-- Acid mechanism -->

```xml
<Term Const="1.35e-10" Eact="65000.0">
   <Catal Species="H+" order="0.457"/> </Term>
<!-- Neutral mechanism -->
<Term Const="9.12e-13" Eact="69800.0"/>
<!-- Base mechanism -->
<Term Const="1.05e-17" Eact="71000.0" ParInt="0.76" ParExt="90.0">
   <Catal Species="H+" order="-0.572"/> </Term>
</RateLaw>
```

3.2.1.2.1.2 Monod rate law

If type="Monod", equation (8) of page 5 is applied:

$$r_k = \theta_\alpha \lambda \prod_i c_{j(i)} \prod_i \frac{c_{j(i)}}{\kappa_i + c_{j(i)}} \prod_i \frac{\kappa_i}{\kappa_i + c_{j(i)}} \tag{8}$$

and <RateLaw> has the following additional attributes:
Obligatory attribute *Phase*: Name of the phase (α) for volume fraction (θ). It must have been specified in <Phase>.
Obligatory attribute *Const*: Kinetic rate constant (λ).
Optionally contains one or more <Fact>.

Element <Fact>.
Obligatory attribute *Type*: It can be:
   - *Lin*: The factor is linear with the concentration of a species.
   - *Cat*: The factor considers a catalyzing species $\left(c/(\kappa + c) \right)$.
   - *Inh*: The factor considers an inhibiting species $\left(\kappa/(\kappa + c) \right)$.
Obligatory attribute *Species*: The name of the species. It must have been specified in <Species>.
Obligatory (if Type = "Cat" or "Inh") attribute *Kappa*: Half saturation or inhibition constant (κ). Note that its unit is defined by the phase of the catalyzing or inhibiting species (e.g., mol/kg if the species pertains to the aqueous phase, see Table 1).

Example for the rate law for denitrification written as:

$$r_k = \theta_{aq} 1.0 \cdot 10^{-6} c_{CH_2O} \frac{c_{NO_3^-}}{2.0 \cdot 10^{-5} + c_{NO_3^-}} \frac{1.0 \cdot 10^{-5}}{1.0 \cdot 10^{-5} + c_{O_2}}$$

```xml
<Ratelaw Type="Monod" Phase="Aqueous" Const="1.0e-6">
   <Fact Type="Lin" Species="CH2O"/>
   <Fact Type="Cat" Species="NO3-" Kappa="2.0e-5"/>
   <Fact Type="Inh" Species="O2" Kappa="1.0e-5"/>
</RateLaw>
```

3.2.1.3 Component

Element <ComponentS> Set of components.
Optional attribute *Method*: It can be:
   - *Manual*: The components are given by the user.
   - *Rref*: The components are calculated from the stoichiometry of the reactions by using Rref (see 2.4).
   Default value is "Manual".

Obligatory (if Method = "Manual") contains one or more <Component>. The number of <Component> must be the number of species minus the number of equilibrium reactions.

Element <Component>. If Method = "Rref", this element is ignored.
Optional attribute *Name*: Name of the component. Default is a number, which coincides with the order of appearance in the file.
Optional attribute *Type*: It can be:
   - *Het*: The component is heterogeneous, that is, it is not homogeneous liquid conservative nor homogeneous immobile.
   - *HolC*: The component is homogeneous liquid conservative, that is, it only contains species of liquid phases and it is not involved in kinetics reactions.
   - *Hoi*: The component is homogeneous immobile, that is, it only contains species of immobile phases.
   Default value is "Het".
Obligatory attribute *Form*: Text containing the formula of the component. It contains "+", "-", coefficients and names of species. If no coefficients are used, a value of 1.0 is assumed. Species must be followed by one or more spaces. All species must have been specified in <Species>.
Example:
<components>
   <component form="Ca+2 + CaCO3(s) " type="het"/>
   <component name="Alkalinity" form="HCO3- - H+ + OH-"/>
   <component name="CEC" form="NaX + 2CaX2" type="hoi"/>
<components>
which is equivalent to:
<ComponentS Method="Manual">
   <Component Type="Het" Name="1" Form="Ca+2 + CaCO3(s)"/>
   <Component Type="Het" Name="Alkalinity" Form="- H+ + 1 HCO3- + 1.0 OH-"/>
   <Component Type="Hoi" Name="CEC" Form="NaX + 2CaX2"/>
<ComponentS>

3.2.1.4 Numerics

Element <Numerics> Numerical options and parameters for chemical speciation calculations.
Optional attribute *Method*: It can be:
   - *NR*: Newton-Raphson is used for the chemical speciation calculations.
   Default value is NR.
Optional attribute *Speciation*: It can be:
   - *FromPrim*: Speciation is done by imposing concentrations of primary species (see 2.5.1).
   - *FromComAqu*: Speciation is done by imposing the aqueous part of component concentrations (see section 2.5.2).
   - *FromComAquKin*: Speciation is done by imposing the aqueous part of component concentrations plus kinetic terms (see section 2.5.3).
   - *FromComAquImm*: Speciation is done by imposing the aqueous plus immobile part of component concentrations (see section 2.5.4).
   Default value is "FromComAqu".
Optional attribute *Tol*: Maximum relative concentration error tolerance. Default is 1.0e-6
Optional attribute *MaxIter*: Maximum number of iterations. Default is 20.

Optional attribute *MaxDiverg*: Maximum number of diverging iterations. Default is *MaxIter*.

Optional attribute *MaxFac*: Maximum correction factor for each iteration. Concentrations cannot increase/decrease more than a factor of MaxFac. Default is 1000.0.

Optional attribute *UseLn*: It can be:
  - *Yes*: The natural logarithm (or ln) of concentrations is used as variable for numerical solution.
  - *No*: Concentrations are used as variable for numerical solution.
  Default value is "Yes".

Obligatory (if speciation="FromPrim") attribute *Primaries*: Text containing the names of the primary species separated by semicolons (;). They must have been specified in <Species>.

Optional attribute *TimeWeight* (only for *Speciation = FromComAqu*): Time weighting factor for kinetic terms (η in equation (15)). Default value is 1.0.

Example:

<numerics primaries="H+; Ca+2"/>

which is equivalent to:

<Numerics Method="NR" Tol="1e-6" Speciation = "FromComAqu"
Primaries="H+; Ca+2"> </Numerics>

3.2.1.5 Output

Element < Output> Output options for writing chemistry to files.

Optional attribute *File*: It is the name of the output file for results of local chemistries. Default is the root of the filename of the chemical system to which *Chem.txt* has been added.

Optional attribute *FileTime*: It is the name of the output file for results of (some) local chemistries versus time. Default is the root of the filename of the chemical system to which *ChemT.txt* has been added.

Optionally contains a <Acti>, <Conc>, <Comp>, <SatI>, <Rate> and <VolF>.

Element <Acti>

Optional attribute *Species*: List of species, of which the −log(activity) or p-value (e.g., pH = −log(activity of H+)) is written. The species in the list must be separated by semicolons (;) and must have been defined in <Species>. If there is no attribute *species*, no activities are written.

Element <Conc>

Optional attribute *Species*: List of species, of which the concentration is written. The species in the list must be separated by semicolons (;) and must have been defined in <Species>. If there is no attribute *species*, no concentrations are written.

Element <Comp>

Optional attribute *Components*: List of components, of which the aqueous, gaseous and immobile concentration is written. The components in the list must be separated by semicolons (;) and must have been defined in <ComponentS >. If there is no attribute *components*, no component concentrations are written.

Element <SatI>

Optional attribute *Reactions*: List of reactions, of which the saturations index is written. The saturation index $= S \log a - \log k$, which equals 0 if the reaction is in equilibrium according to equation (4). The reactions in the list must be separated by semicolons (;) and must have been defined in <ReactionS>. If there is no attribute *reactions*, no saturation indices are written.

Element <Rate>
Optional attribute *Reactions*: List of reactions, of which the rate is written. The reactions in the list must be separated by semicolons (;) and must have been defined in <ReactionS>. If there is no attribute *reactions*, no rates are written.

Element <VolF>
Optional attribute *Phases*: List of phases, of which the volume fraction (m3 of phase/m3 of porous medium) is written. The phases in the list must be separated by semicolons (;) and must have been defined in <PhaseS>. If there is no attribute *Phases*, no volume fractions are written.

Example:
<Output File="Example. Chem.dat">
   <Acti Species="H+; CO2(g); CaCO3(s)"/>
   <Comp Components="Ca; C; Cl"/>
   <VolF Phases="Aqueous; Calcite"/>
</Output>

3.2.2 Local chemistry

Element < LocalChemS> Set of one or more local chemistries
Optional attribute *PhreeqcFile*: Name of an output file of Phreeqc. If it is empty (that is, PhreeqcFile = " "), only the present xml file is used to read the local chemistries. If a filename is given, only names of local chemistries are read. Other attributes (concentrations, phase volume fractions and temperature) are read from the Phreeqc output file. Names of local chemistries in the xml file must coincide with names in the Phreeqc output file. The latter must be contained in lines of the Phreeqc output file starting with "Initial solution" or "Reaction step". It may be necessary to add manually (through a text editor) proper names to the Phreeqc output file. Concentrations of X- (an exchange master species) are not given by the Phreeqc output file. In that case its concentrations are calculated according to (X-) = (NaX)/(Na+), if concentrations of NaX and Na+ are part of the chemical system. The default *PhreeqcFile* is " " (that is, Phreeqc output is not used).
Optionally contains one or more <LocalChem>.

Element < LocalChem>
Optional attribute *Name*: Name of the local chemistry. Default name is a number.
Optional attribute *Temp*: Temperature (^{0}C). Default is 25 ^{0}C.
Optionally contains chemical compositions of one or more <Phase>. Default means default concentrations for phase (see 3.2.2.1) and species of the phase (see 3.2.2.1.1).

3.2.2.1 Phase

Element <Phase>

Obligatory attribute *Name*: Name of the phase. It must have been defined in the chemical system.

Optional attribute *VolF*: If the phase is of the type *Exchange* (see 3.2.1.1.1), it is the cation exchange capacity (eq (m^{-3} of medium)). Otherwise it is the volume (fraction) of the phase (m^3 of phase (m^{-3} of medium)). Default is 0.0.

Optional attribute *Pres*: Pressure of the phase. Default is 1.0 atm.

3.2.2.1.1 Species

Element <Species>

Obligatory attribute *Name*: Name of the species. It must have been defined in the chemical system.

Optional attribute *Conc*: Concentration of the species. The definition of concentration depends on the type of phase (see 3.2.1.1.1). If the phase is of the type *Aqueous* or *Aqueous_DH*, it is the molality (mol of species / kg of water). If the phase is of the type *gas*, it is the partial pressure (atm). If the phase is of the type *mineral*, it is the molar fraction (mol of species / mol of phase). If the phase is of the type *exchange*, it is the equivalent fraction (eq of species / eq of cation exchange capacity). Default is $1.0 \cdot 10^{-10}$.

3.3 CheProfRW

The code CheProfRW just reads xml files for a chemical system and local chemistries (see section 3.2) and writes them again in xml format. This can be useful when one wants to convert a case calculated by, e.g., Phreeqc, into xml files through the *PhreeqcFile* attribute in the chemical system (see 3.2.1) and/or local chemistries (see 3.2.2). The output files contain the full chemical system and/or local chemistries with information obtained from Phreeqc. This information can be checked, adjusted and/or extended for subsequent use by other codes.

CheProfRW requires a small xml file with names of input and output files.

Element <CheProfReadWrite>

Obligatory attribute *FileChemSysIn*: Name of the input file for the chemical system.

Optional attribute *FileChemSysOut*: Name of the output file for the chemical system. If it is empty (that is, FileChemSysOut = " "), no output file is written. The default is " ".

Optional attribute *FileLocalChemIn*: Name of the input file for the local chemistries. If it is empty (that is, FileLocalChemIn = " "), no local chemistries are read. The default is " ".

Optional attribute *FileLocalChemOut*: Name of the output file for the local chemistries. If it is empty (that is, FileLocalChemOut = " "), no output file is written. The default is " ".

Example:
```
<?xml version="1.0"?>
<CheProfReadWrite
FileChemSysIn="Example. ChemSysIn.xml"
FileChemSysOut="Example. ChemSys.xml"
FileLocalChemIn="Example. LocalChemIn.xml"
FileLocalChemOut="Example. LocalChem.xml"/>
```

3.4 ChemMix

The code ChemMix calculates the chemical composition of the mixing of two end members (local chemistries) with various mixing ratios (λ)

$$w_{aq}^m = \lambda w_{aq}^1 + (1 - \lambda) w_{aq}^2 \tag{18}$$

where superscripts, 1, 2 and m refer to end member 1 and 2 and the mixing.
ChemMix requires files for the chemical system and local chemistries (see 3.2) and a small xml file with names of input and output files and a few parameter values. The chemistry of the mixing by doing a speciation as explained in section 2.5.2.

Element < ChemMix>
Obligatory attribute *FileChemSys*: Name of the input file for the chemical system.
Obligatory attribute *FileLocalChem*: Name of the input file for the local chemistries.
Obligatory attribute *IniChem*: Name of the local chemistry that represents the initial condition of the mixing.
Obligatory attribute *EndM1*: Name of the local chemistry that represents the first end member.
Obligatory attribute *EndM2*: Name of the local chemistry that represents the second end member.
Obligatory attribute *nMix*: Number of mixing ratios.

3.5 WMA1D

The code WMA1D uses the Water Mixing Approach of Soler-Sagarra et al. (2022) to solve reactive transport in a simple one-dimensional domain with uniform flow of water. It writes the mass balance equation of (non-homogeneous immobile) components as:

$$\theta_{aq} \frac{\partial w_{aq}}{\partial t} + \theta_{im} \frac{\partial w_{im}}{\partial t} = L(w_{aq}) + U(S_k)^T r_k \tag{19}$$

where $L()$ is an operator representing transport terms. Equation (19) can be discretized in time as follows:

$$\theta_{aq} \frac{w_{aq}^{k+1} - w_{aq}^k}{\Delta t} + \theta_{im} \frac{w_{im}^{k+1} - w_{im}^k}{\Delta t} = L(w_{aq}^k) + U(S_k)^T r_k^{k+\eta} \tag{20}$$

or

$$w_{aq}^{k+1} + \frac{\theta_{im}}{\theta_{aq}} (w_{im}^{k+1} - w_{im}^k) - \frac{\Delta t}{\theta_{aq}} U(S_k)^T r_k^{k+\eta} = w_{aq}^k + \frac{\Delta t}{\theta_{aq}} L(w_{aq}^k) = w_{giv}^{k+1} \tag{21}$$

where superscript k refers to time step. The righthand side of (21) can be discretized as mixing ratios:

$$w_{giv,i}^{k+1} = \sum_j \lambda_{ij} w_{giv,j}^k \tag{22}$$

where subscript i and j refer to cells within the domain. For one-dimensional uniform flow this reduces to:

$$w_{giv,i}^{k+1} = P_e^{-1} w_{giv,i-2}^k + (1 - 2P_e^{-1}) w_{giv,i-1}^k + P_e^{-1} w_{giv,i}^k \tag{23}$$

with Peclet number P_e

$$P_e = \frac{v\Delta x}{D} = \frac{v^2\Delta t}{D} \qquad (24)$$

where v is the porewater velocity ($v = q/\phi$ Darcy flow over porosity), D is the coefficient for diffusion and dispersion, Δx is the cell size and Δt is the time increment. In this scheme we impose that $\Delta x = v\Delta t$. After the mixing the chemistry is calculated from $\boldsymbol{w}_{giv}^{k+1}$ by speciation as explained in section 2.5.3.

WMA1D requires files for the chemical system and local chemistries (see 3.2) and a small xml file with names of input and output files and a few parameter values.
Element <WMA1D>
Obligatory attribute *Flux*: Darcy flux (m/s), q.
Optional attribute *Por*: Porosity, ϕ. Default value is 1.0.
Optional attribute *PeInv*: Inverse Peclet number, P_e^{-1}. Default value is 0.0.
Optional attribute *Length*: Length of the domain (m), L. Default value is 1.0.
Optional attribute *TotTime*: Total time to be simulated (s). Default value is $L\phi/q$.
Optional attribute *NOutputTime*: Number of times that results are written. Default value is 5.
Obligatory attribute *FileChemSys*: Name of the input file for the chemical system.
Obligatory attribute *FileLocalChem*: Name of the input file for the local chemistries.
Obligatory attribute *ChemIni*: Name of the local chemistry that represents the initial condition.
Obligatory attribute *ChemBou*: Name of the local chemistry that represents chemistry of the inlet water.
Obligatory attribute *NCell*: Number of cells.

Example:
```
<?xml version="1.0"?>
<WMA1D Flux="6.342e-8"
Por="0.1"
PeInv="0.4"
Length="100.0"
NOutputTime="5"
NCell="20"
ChemIni="InitialWater"
ChemBou="BoundaryWater"
FileChemSys="Example. ChemSys.xml"
FileLocalChem="Example. LocalChem.xml"/>
```

3.6 DSA1D

The code DSA1D calculates reactive transport in a one-dimensional domain discretized by finite elements using the Direct Substitution Approach (DSA). Transport include advection, diffusion and dispersion in only the liquid phase. Liquid flow is assumed uniform in space, but it may vary in time.

The domain is divided into one or more zones, each zone having equal element size, diffusion coefficient, dispersivity and initial chemistry. Total time is divided into one or more periods, each period having equal time step, flow rate and boundary conditions. For the left boundary the following boundary condition is used

$$j^i = q(w_l^i)^0 + \beta \left((w_l^i)^0 - w_l^i \right) \quad q \geq 0 \text{ (inlet)}$$
$$j^i = qw_l^i + \beta \left((w_l^i)^0 - w_l^i \right) \qquad q < 0 \text{ (outlet)}$$

(25)

and for the right boundary

$$j^i = -q(w_l^i)^0 + \beta \left((w_l^i)^0 - w_l^i \right) \quad q \leq 0 \text{ (inlet)}$$
$$j^i = -qw_l^i + \beta \left((w_l^i)^0 - w_l^i \right) \qquad q > 0 \text{ (outlet)}$$

(26)

where superscript i refers to component, superscript ^{0} refers to external chemistry, w_l is the total aqueous concentration, q is Darcy's flux (positive means from left to right, negative from right to left) and β is a coefficient, which can be used to fixed concentrations or mixed boundary conditions. Large values of β tend to fix concentrations to those of the external chemistry.

DSA is an iterative method that may have convergence problems. Therefore, there is a timestep control, in which the size of a time step (Δt) is calculated from the number of iterations of the previous timestep (Δt_{old})

$$\Delta t = \min(f \Delta t_{old}, \Delta t_{max})$$
$$f = \left(\frac{N_{opt}}{N_{iter}} \right)^{1/4}$$

(27)

where f is the timestep increase/decrease factor, Δt_{max} is a maximum time increment, N_{opt} is a user given optimum number of iterations and N_{iter} is the number of actual iterations in the previous timestep. If the number of iterations reaches a maximum number (N_{max}), the calculations are repeated with a smaller timestep calculated according to timestep factor $f = \left(N_{opt}/N_{max} \right)^{1/4}$, in which $N_{opt} \leq N_{max}$.

DSA1D requires files for the chemical system and local chemistries (see 3.2) and an xml file with names of input and output files and transport parameter values.

Element <DSA1D>
Obligatory attribute *FileChemSys*: Name of the input file for the chemical system.
Obligatory attribute *FileLocalChem*: Name of the input file for the local chemistries.
Obligatorily contains one set of zones <ZoneS>
Obligatorily contains one set of periods <PeriodS>.
Optionally contains numerical options and parameters <Numerics>. If it is not specified default values are used.
Optionally contains output options <Output>.

Element <ZoneS> Set of one or more phases
No attributes.
Obligatorily contains one or more <Zone>.

Element <Zone>
Optional attribute *Name*: Name of the zone. Default is a number, which coincides with the order of appearance in the file.
Obligatory attribute *Length*: Length of the zone (m).
Optional attribute *nElem*: Number of elements of the zone. Default is 1.

Optional attribute *Dif*: Diffusion coefficient (m^2/s) of the zone. Default is 0.0.

Optional attribute *Dis*: Dispersivity (m) of the zone. Default is 0.0.

Obligatory attribute *ChemIni*: Name of the local chemistry that represents the initial condition.

Element <PeriodS> Set of one or more periods

No attributes.

Obligatorily contains one or more <Period>.

Element <Period>

Optional attribute *Name*: Name of the period. Default is a number, which coincides with the order of appearance in the file.

Obligatory attribute *Time*: Time length of the period (s).

Optional attribute *dTime*: Maximum time increment (s). Default is *Time*.

Optional attribute *Flux*: Darcy flux (m/s), q. Positive means flow from left to right, negative flow from right to left. Default is 0.0.

Optional attribute *CoefLeft*: Boundary coefficient (m/s) at the left boundary, β in equation (25). Default is 0.0.

Optional attribute *CoefRight*: Boundary coefficient (m/s) at the right boundary, β in equation (26). Default is 0.0.

Obligatory (if *Flux* > 0 or *CoefLeft* > 0) attribute *ChemLeft*: Name of the local chemistry that represents chemistry of the left boundary.

Obligatory (if *Flux* < 0 or *CoefRight* > 0) attribute *ChemRight*: Name of the local chemistry that represents chemistry of the right boundary.

Element <Numerics> Numerical options and parameters for reactive transport calculations.

Optional attribute *TimeWeight*: Interpolation time scheme for transport terms (0.0: explicit scheme, 1.0: implicit scheme, 0.5: Crank-Nicolson scheme). Default is 1.0

Optional attribute *TolVar*: Maximum relative error tolerance on variable (concentration of primary species or component, see section 3.2.1.4). Default is 1.0e-6.

Optional attribute *TolBal*: Maximum error tolerance on mass balance (mol/m^3 of medium). Default is 0.0.

Optional attribute *MaxIter*: Maximum number of iterations, N_{iter} in equation (27). MaxIter must be equal or larger than OptIter. Default is 20.

Optional attribute *OptIter*: Optimum number of iterations for timestep control, N_{opt} in equation (27). Default is 4.

Element <Output> Output options. Output filenames and chemical items to be written are specified in the <Output> element of the chemical system (see 3.2.1.5). Results of chemistry are always written for all nodes after each period to the file specified in the attribute *File* of section 3.2.1.5.

Optional attribute *Nodes*: Text containing node numbers separated by semicolons (;) for writing chemistry at each time step to the file specified in the attribute *FileTime* of section 3.2.1.5. If *Nodes* is not specified, chemistry is not written for any nodes.

4 Code structure

The code is written in Fortran 2003 using the technique of Object-Oriented Programming (Rouson, Xia and Xu 2011). Its structure or class diagram is based on that of Bea et al. (2009) and is shown in Figure 2. Basically a "Chem" consists of a LocalChem (local chemistry) and a ChemSys (chemical system). A LocalChem consists of the (state) variables, such as temperature, concentrations of all species and volume fractions of phases. A ChemSys contains all equations of state, such as mass action laws, kinetic rate laws, definitions of components, phase constraints and activity models. A ChemSys consists of zero or several Components, zero or several Reactions and one or more Phases. Each reaction, if it is kinetic, has a RateLaw, that can be of the type Monod or Lasaga. Each Phase has one or more Species and one PhaseModel. The various types of PhaseModels are defined according to the phase constraints of Table 1. They are further subdivided according to the way activities are calculated, e.g., assuming ideal mixtures or the Debye Hückel equation (5).
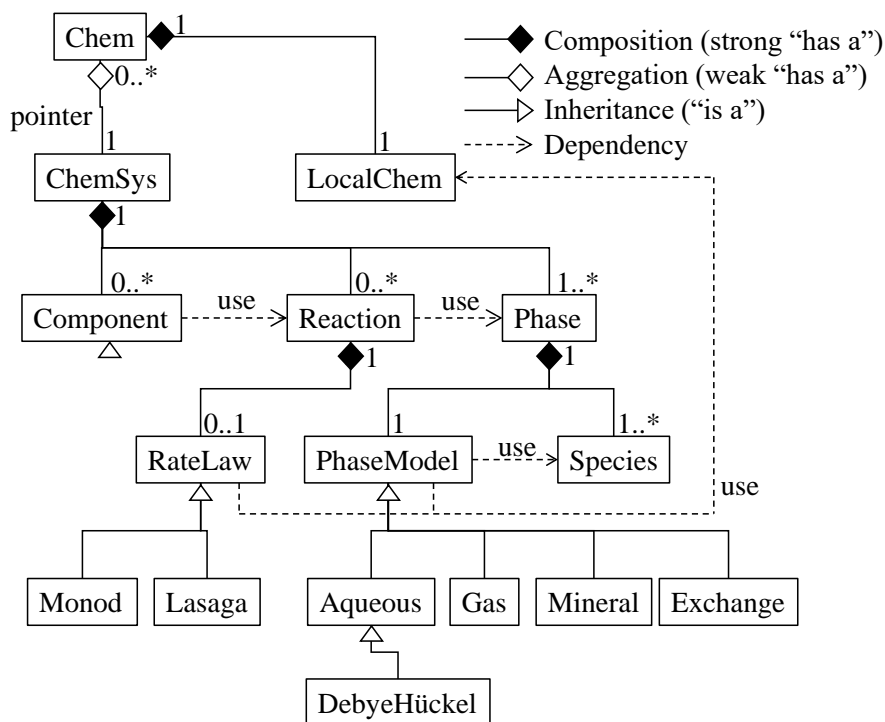

Figure 2. Class diagram.

The main functions and subroutines to be used by the user code (e.g., reactive transport module or Codebright) are:

Subroutine *ChemSysFin*(*This*)
    Description: Deallocates a chemical system.
    This: type(ChemSys_t): The chemical system
Subroutine *ChemSysIni*(*This,fname*)
    Description: Allocates and reads from an xml file a chemical system.
    This: type(ChemSys_t): The chemical system
    fname: character*(*): Name of the xml input file
Subroutine *ChemSysWrite*(*This,fname*)
    Description: Writes the complete chemical system to an xml output file.

This: type(ChemSys_t): The chemical system

fname: character*(*): Name of the xml output file

Function *ChemSysGetNpri*(*This*)

Description: Gets the number of components that are not homogeneous immobile components. It is the number of components necessary for transport. It is also the number of primary species.

This: type(ChemSys_t): The chemical system

Subroutine *ChemFin*(*This,n*)

Description: Deallocates local chemistries.

This: type(Chem_t): An array of local chemistries of dimension *n*.

n: integer: dimension of This

Subroutine *ChemIni*(*This,ChemSys,n,fname*)

Description: Allocates and reads from an xml file local chemistries.

This: type(Chem_t): An array of local chemistries of dimension *n*.

ChemSys: type(ChemSys_t): The chemical system

n: integer: dimension of This

fname: character*(*): Name of the xml input file

Subroutine ChemSpeciate(*This,ini,fin,PrimVar,nPri,dTime*)

Description: Speciates or updates local chemistries.

This: type(Chem_t): An array of local chemistries of dimension *ini:fin*.

ini, *fin*: integer: First and last local chemistry

PrimVar: real*8: An array of primary variables of dimension (1:*npri,ini:fin*).

npri: integer: Number of primary variables.

dTime. Time increment (s), required for kinetic terms in some speciation types. This is an optional argument. If not used no kinetic terms are considered.

Subroutine *ChemUpdateOld*(*This,ini,fin*)

Description: Copies the content of the new chemistry (LocalChem) to the old chemistry (LocalOldChem)

This: type(Chem_t): An array of local chemistries of dimension *ini:fin*.

ini, *fin*: integer: First and last local chemistry

Subroutine *ChemWrite*(*This,ini,fin,fname*)

Description: Writes complete local chemistries to an xml output file.

This: type(Chem_t): An array of local chemistries of dimension ini:fin.

ini, *fin*: integer: First and last local chemistry

fname: character*(*): Name of the xml output file

Subroutine *ChemWriteGid*(*This,ini,fin,Time,iou*)

Description: Writes local chemistries in GiD format to a post.res file according to output options given in the chemical system.

This: type(Chem_t): An array of local chemistries of dimension *ini:fin*.

ini, *fin*: integer: First and last local chemistry

Time: real*8: Time.

iou: integer: Unit number of the post.res file.

Subroutine *ChemWriteTxt*(*This,ini,fin,Title*)

Description: Writes local chemistries to an txt output file according to output options given in the chemical system.

This: type(Chem_t): An array of local chemistries of dimension *ini:fin*.

ini, *fin*: integer: First and last local chemistry

Title: character*(*): A title that is written to the txt output file.

Function *ChemGetCom*(*This,ini,fin,Mob*)

Description: Gets the component concentrations in a group of phases according to their mobility or component kinetic reaction rates. The dimension of ChemGetCom is (*ini:fin*).

This: type(Chem_t): A local chemistry.

ini, *fin*: integer: First and last component.

Mob: character*(*): Mobility. It can be 'liq': phases with liquid mobility (in mol/m3 of PHASE), 'gas': phases with gas mobility (in mol/m3 of PHASE), 'imm': immobile phases (in mol/m3 of MEDIUM), 'kin': kinetic reaction rate terms (in mol/s/m3 of MEDIUM)

Function *ChemGetOldCom*(*This,ini,fin,Mob*)

Description: Same as *ChemGetCom*, but it gets the component concentrations of the old (or previous) local chemistry.

Function *ChemGetDerCom*(*This,ini,fin,inip,finp,Mob*)

Description: Gets the derivatives wrt primary variables of the component concentrations in a group of phases according to their mobility or component kinetic reaction rates. The dimension of ChemGetDerCom is (*ini:fin,inip:finp*).

This: type(Chem_t): A local chemistry.

ini, *fin*: integer: First and last component.

inip, *finp*: integer: First and last primary variable.

Mob: character*(*): Mobility, see above.

Function *ChemGetPrim*(*This,ini,fin,dTime*)

Description: Gets values of the primary variables. The dimension of ChemGetPrim is (*ini:fin*).

This: type(Chem_t): A local chemistry.

ini, *fin*: integer: First and last primary variable.

dTime. Time increment (s), required for kinetic terms in some speciation types. This is an optional argument. If not used no kinetic terms are considered.

Function *ChemGetVolF(This,Mob,Change)*

Description: Gets the volume fractions (m^3_{phase}/m^3_{medium}) or volume fraction change of phases according to their mobility.

This: type(Chem_t): A local chemistry.

Mob: character*(*): Mobility, see above.

Change: logical: Mobility. If false, it gets the volume fraction; if true, it gets the volume fraction change.

References

Bea, S. A., J. Carrera, C. Ayora, F. Batlle, and M. W. Saaltink. 2009. "Cheproo: a Fortran 90 object oriented module to solve chemical processes in earth science models." *Comput. Geosci.* 35 (6): 1098-1112. http://dx.doi.org/10.1016/j.cageo.2008.08.010.

Debye, P., and E. Hückel. 1923. "The theory of electrolytes, I. Lowering of freezing point and related phenomena." *Phys. Z.* 24: 185-206.

Kräutle, S. 2011. "The semismooth Newton method for multicomponent reactive transport with minerals." *Adv. Water Resour.* 34: 137-151. https://www.doi.org/10.1016/j.advwatres.2010.10.004.

Lasaga, A. C., J. M. Soler, J. Ganor, T. E. Burch, and K. L. Nagy. 1994. "Chemical weathering rate laws and global geochemical cycles." *Geochim. Cosmochim. Acta* 58: 2361-2386.

Meier, C. D. 2000. *Matrix analysis and Applied Linear Algebra.* Philadelphia: SIAM.

Molins, S., J. Carrera, C. Ayora, y M. W. Saaltink. 2004. «A formulation for decoupling components in reactive transport problems.» *Water Resources Research* W10301. http://doi.org/10.1029/2003WR002970.

Olivella, S., A. Gens, J. Carrera, y E.E. Alonso. 1996. «Numerical formulation for a simulator (CODE_BRIGHT) for the coupled analysis of saline media.» *Engineering Computations* 13 (7): 87-112. https://doi.org/10.1108/02644409610151575.

Palandri, James L., and Yousif K. Kharaka. 2004. *A compilation of rate parameters of water-mineral interaction kinetics for application to geochemical modeling. Open file report 2004-1068.* Menlo Park: U.S. Geological Survey.

Parkhurst, D.L., and C.A.J. Appelo. 2013. *Description of input and examples for PHREEQC version 3-A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations.* U.S. Geological Survey Techniques and Methods, book 6, chap. A43, Denver: U.S. Geological Survey, 497 p. http://pubs.usgs.gov/tm/06/a43/.

Redlich, O., and S. J. N. Kwong. 1949. "On the thermodynamics of solutions. V. An equation of state. Fugacities of gaseous solutions." *Chem. Rev.* 44: 233-244.

Rouson, Damian, Jim Xia, and Xiaofeng Xu. 2011. *Scientific Software Design, The Object-Oriented Way.* Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi, Tokyo, Mexico City: Cambridge University Press.

Soler-Sagarra, J., M. W. Saaltink, A. Nardi, F. De Gaspari, and J. Carrera. 2022. "Water Mixing Approach (WMA) for reactive transport modeling." *Advances in Water Resources Research* 161: 104131. https://doi.org/10.1016/j.advwatres.2022.104131.

Steefel, C. I., y K. T. B. MacQuarrie. 1996. «Approaches to modeling reactive transport.» *Rev. Mineral* 34: 83-129.

Van Cappellen, P., and J. F. Gaillard. 1996. *Biogeochemical dynamics in aquatic systems.* Vol. 34, in *Reactive transport in porous media, Reviews in mineralogy*, edited by P. C. Lichtner, C. I. Steefel and E. H. Oelkers, 335-376. Washington DC: Mineralogical Society of America.

Appendix A. Calculation of kernel matrices through RREF

One way of constructing the kernel matrices is by transforming the stoichiometric matrices to their Reduced Row Echelon Form (RREF), which can be seen as a generalization for rectangular matrices of the Gauss-Jordan method (Meier 2000). A matrix is said to be in a row echelon form if:

1. All nonzero rows are above any rows of all zeroes, i.e., all zero rows are at the bottom.
2. The first nonzero number from the left, or pivot (p), of a nonzero row is always to the right of the pivot of the row above it.

This means it has a jagged triangular form, such as:

$$\begin{array}{ccccccc} p_1 & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} \\ 0 & 0 & p_2 & a_{24} & a_{25} & a_{26} & a_{27} \\ 0 & 0 & 0 & p_3 & a_{35} & a_{36} & a_{37} \\ 0 & 0 & 0 & 0 & 0 & 0 & p_4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}$$

where a can be a zero or nonzero entry. A matrix is in a <u>reduced</u> row echelon form if:

1. It is in a row echelon form
2. Each pivot equals one
3. All entries above the pivot equal zero.

This means it has a form, such as:

$$\begin{array}{ccccccc} 1 & a_{12} & 0 & 0 & a_{15} & a_{16} & 0 \\ 0 & 0 & 1 & 0 & a_{25} & a_{26} & 0 \\ 0 & 0 & 0 & 1 & a_{35} & a_{36} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}$$

Every matrix can be transformed, by linear combinations of its rows, to a reduced row echelon form. The rank of a matrix equals the number of pivots or number of nonzero rows of its reduced row echelon form (for instance, the above matrix has a rank of 4). A kernel of a matrix in a reduced row echelon form can be constructed simply by setting the columns without pivots (columns 2, 5 and 6) to the identity matrix. The rest of the columns (columns 1, 3, 4 and 7) are set to the transpose of the columns without pivots of the matrix in its row reduced echelon form of which the bottom zero rows have been removed. This gives the following kernel matrix for the matrix of above:

$$\begin{array}{ccccccc} -a_{12} & 1 & 0 & 0 & 0 & 0 & 0 \\ -a_{15} & 0 & -a_{25} & -a_{35} & 1 & 0 & 0 \\ -a_{16} & 0 & -a_{26} & -a_{36} & 0 & 1 & 0 \end{array}$$

This method has a tendency for the unity matrix to be found more at the right part of the kernel matrix. Normally, we place the most important species (or species with highest concentrations) at the left, and, prefer as much as possible the identity matrix of the kernel to be at the left as well. We can do this by inverting the order of the columns. This is illustrated by the construction of the kernel of the aqueous stoichiometric matrix (S_a) of the example of section 2.4:

$$\boldsymbol{S}_a = \begin{array}{c} \\ \\ \\ \\ \\ \end{array} \begin{array}{|cccccccc|c} \text{H}_2\text{O} & \text{O}_2 & \text{H}^+ & \text{OH}^- & \text{Fe}^{2+} & \text{Fe}^{3+} & \text{SO}_4^{2-} & \text{FeSO}_4 & \\ \hline 0 & 0 & 0 & 0 & 1 & 0 & 1 & -1 & \text{R1} \\ -1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & \text{R2} \\ 0 & 0 & 0 & 3 & 0 & 1 & 0 & 0 & \text{R3} \\ \hdashline 0 & 0.25 & 0.5 & 0.5 & 1 & -1 & 0 & 0 & \text{R4} \\ 0 & -3.5 & 1 & 1 & 1 & 0 & 2 & 0 & \text{R5} \\ \hline \end{array}$$

Inverting the order of the columns gives:

$$\boldsymbol{S}_a = \begin{array}{|cccccccc|c} \text{FeSO}_4 & \text{SO}_4^{2-} & \text{Fe}^{3+} & \text{Fe}^{2+} & \text{OH}^- & \text{H}^+ & \text{O}_2 & \text{H}_2\text{O} & \\ \hline -1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & \text{R1} \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & -1 & \text{R2} \\ 0 & 0 & 1 & 0 & 3 & 0 & 0 & 0 & \text{R3} \\ \hdashline 0 & 0 & -1 & 1 & 0.5 & 0.5 & 0.25 & 0 & \text{R4} \\ 0 & 2 & 0 & 1 & 1 & 1 & -3.5 & 0 & \text{R5} \\ \hline \end{array}$$

The reduced row echelon form (RREF) of this matrix is:

$$\text{RREF}(\boldsymbol{S}_a) = \begin{array}{|cccccccc|c} \text{FeSO}_4 & \text{SO}_4^{2-} & \text{Fe}^{3+} & \text{Fe}^{2+} & \text{OH}^- & \text{H}^+ & \text{O}_2 & \text{H}_2\text{O} & \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & -1.625 & 0.75 & \text{R1} \\ 0 & 1 & 0 & 0 & 0 & 2 & -1.875 & -0.75 & \text{R2} \\ 0 & 0 & 1 & 0 & 0 & -3 & 0 & 3 & \text{R3} \\ \hdashline 0 & 0 & 0 & 1 & 0 & -2 & 0.25 & 2.5 & \text{R4} \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & -1 & \text{R5} \\ \hline \end{array}$$

The kernel matrix ($\boldsymbol{U}_{holc}$) of $\boldsymbol{S}_a$ is:

$$\boldsymbol{U}_{holc} = \begin{array}{|cccccccc|} \text{FeSO}_4 & \text{SO}_4^{2-} & \text{Fe}^{3+} & \text{Fe}^{2+} & \text{OH}^- & \text{H}^+ & \text{O}_2 & \text{H}_2\text{O} \\ \hline 0 & -2 & 3 & 2 & -1 & 1 & 0 & 0 \\ 1.625 & 1.875 & 0 & -0.25 & 0 & 0 & 1 & 0 \\ -0.75 & 1.75 & -3 & -2.5 & 1 & 0 & 0 & 1 \\ \hline \end{array}$$

Inverting back the order of the columns and rows gives:

$$\boldsymbol{U}_{holc} = \begin{array}{|cccccccc|} \text{H}_2\text{O} & \text{O}_2 & \text{H}^+ & \text{OH}^- & \text{Fe}^{2+} & \text{Fe}^{3+} & \text{SO}_4^{2-} & \text{FeSO}_4 \\ \hline 1 & 0 & 0 & 1 & -2.5 & -3 & 1.75 & -0.75 \\ 0 & 1 & 0 & 0 & -0.25 & 0 & 1.625 & 1.625 \\ 0 & 0 & 1 & -1 & 2 & 3 & 0 & 0 \\ \hline \end{array}$$