

Groundwater Flow and Solute Transport Modeling

Ye Zhang

Dept. of Geology & Geophysics

University of Wyoming

© *Draft date February 13, 2016*

Contents

Contents	i
0.1 Introduction	v
0.2 This Class	vi
0.2.1 Textbook	vii
0.2.2 Tools	viii
0.2.3 Questions and Answers	viii
0.2.4 Homework, Projects, Exams and Grades	viii
0.3 Basic Math Review	ix
0.4 Homework 1	ix
1 Groundwater Flow Equations (Review)	1
1.1 Introduction	1
1.2 Darcy's Law — Review	2
1.2.1 Darcy's Law of This Course	4
1.3 3D General Flow Equation (Confined Aquifer)	6
1.3.1 Less General Flow Equations	7
1.4 2D Planeview Flow Equation (Confined & Unconfined)	10
2 Modeling Overview for Groundwater Flow	15
2.1 General Methodology	15
2.2 Analytical versus Numerical Methods	19
2.3 Validity of Numerical Solution	23
2.4 Model Calibration and Parameter Estimation	24
2.5 Interpreting Model Results	27
2.6 Model Limitation and Uncertainty	28
2.7 References	29
3 Mathematical Notations, Definitions and Theorems	31
3.1 Differential Equations	31
3.2 Scalar, Vector, Tensor, Field	33
3.3 Taylor Series and Finite Difference Approximations	34
3.4 Numerical Error, Convergence, Stability	35
3.5 Matrix & Vector	36
3.6 Linear Algebra and Solution Techniques	39

3.6.1	Direct Approaches	40
3.6.2	Iterative Approaches	41
3.6.3	Comparison and Summary	44
3.7	Homework 2	46
3.8	Gauss-Jordan Direct Solver with Pivoting (Optional)	46
4	Lumped Parameter Models	53
4.1	Mass Balance Principle	53
4.2	Groundwater Flow Model	55
4.2.1	Analytical Solution	56
4.3	Solute Transport Model	57
4.3.1	Analytical Solution	58
4.4	Numerical Solution	59
4.4.1	Euler's Method	59
4.4.2	Huen's Method	60
4.4.3	Runga Kutta Methods	61
4.5	Homework 3	63
4.6	Project One — Flow & Transport Modeling	64
5	One-Dimensional Flow Modeling	67
5.1	1D Steady-State Flow	67
5.1.1	Homogeneous Porous Media	67
5.1.2	Heterogeneous Porous Media	69
5.2	Homework 4	71
5.3	FD Solution by Computer Codes	71
5.3.1	Matrix-Based Approaches	71
5.3.2	Matrix-Free Approaches	72
5.4	Irregular Grid	81
5.5	1D Transient Flow	81
5.5.1	FD Explicit Formulation	84
5.5.2	FD Implicit Formulation	85
5.5.3	FD Weighted Formulation	87
5.6	Homework 5	99
6	Two-Dimensional Steady-State Flow	103
6.1	Specified Head & No Flow boundaries	103
6.1.1	Finite Difference Discretization	107
6.2	Homework 6	110
6.3	Specified Head & Specified Flux boundaries	116
6.3.1	Variable Flux Boundary Condition	122
6.4	Project Two — 2D Steady State Flow Modeling	123
6.5	General Boundary Conditions	129
6.5.1	Boundary Fluxes	129
6.5.2	Internal Fluxes	131
6.6	Velocity Calculation	131
6.6.1	Darcy Flux	131

6.6.2	Mass Balance Calculation	135
6.6.3	Average Linear Velocity	137
6.6.4	Groundwater Pathline Generation	138
6.7	FD Formulations for Heterogeneous K Field	142
6.7.1	Project Two B (Optional) — 2D SS Modeling with Heterogeneous K	142
7	Two-Dimensional Transient Flow	145
7.1	Mathematical Formulation	146
7.2	Finite Difference Formulations	146
7.2.1	Explicit Formulation	147
7.2.2	Implicit Formulation	148
7.2.3	Weighted Formulation	148
7.3	Mass Balance Calculations	149
7.4	Project Three— 2D Transient Flow Modeling	149
8	Solute Transport Modeling	159
8.1	Introduction	159
8.1.1	Advection & Dispersion	159
8.1.2	Hydrodynamic Dispersion in Homogenous Medium	161
8.2	Derivation of the 1D Advection-Dispersion Equation (ADE)	163
8.3	Finite Difference Discretization	165
8.3.1	Explicit Formulation	166
8.3.2	Implicit Formulation	167
8.3.3	Weighted Formulation	167
8.4	Solute Transport Modeling	168
8.5	Homework 7	171
8.6	Extension of ADE to Higher Dimensions	173
8.7	Advanced Topics (Optional)	174
8.7.1	Groundwater Pathline & Particle Tracking	174
8.7.2	Solute Macrodispersion in Heterogeneous Media	174
8.7.3	Alternatives to ADE	174
9	3D Models of Flow and Transport	177
9.1	Available Software	177
9.2	Project Four—3D Flow Modeling with Groundwater Vista	178
9.3	Final Review	192
10	Parameter Estimation for Groundwater Models (Optional)	195
10.1	Introduction	195
10.2	Forward Estimation of Hydraulic Conductivity	197
10.2.1	Grid-Block Equivalent K^*	199
10.2.2	Issues with Numerical Upscaling	201
10.3	Upscaling for Irregular Domains	203
10.3.1	Numerical Flow Experiments	206
10.3.2	Tensor Analysis	208
10.4	Inverse Methods	213

10.4.1	Notation Conventions	214
10.4.2	Least Square Objective Function	214
10.4.3	Weight Matrix	216
10.4.4	Gauss-Newton Nonlinear Regression	217

This is a set of course notes written by Ye Zhang for a Groundwater Flow and Transport Modeling course.

Fall 2012

GEOL 5030

3 CREDITS

GRADING: S/U or A-F (It is up to each student.)

Lecture location: ESB1006

Lecture times: Tues + Thurs (9:35 am ~ 10:50 am)

Office hours: M(4:00~5:30 pm), F(3:00~4:30 pm), GE 220

Email: yzhang9@uwyo.edu

Phone: 307-766-2981

Prerequisite:

Calculus I & II;

Geohydrology;

Matlab Programming language*

*This course emphasizes the fundamental development of mathematical models of groundwater flow and solute transport and their applications using computer simulations. Students are expected to write small computer codes for (up to 2D) problems, thus rudimentary skills in programming with Matlab are necessary. If a student has not used Matlab before, please study an excellent tutorial before attending this class (it will take around 3 hours):

<http://faculty.gg.uwyo.edu/yzhang/teaching.html>

NOTE: The lecture notes will be periodically posted on the Wyoweb course website, usually 1 week before the relevant lectures. Please make a habit of regularly checking for notes or other announcements from the site. However, the lecture notes do NOT include: (1) solutions to the exercises and homework; (2) proofs to theories and equation derivations. These will be presented only during the lectures. So, do not rely on the notes for everything — attendance and in-class participation are key to doing well.

0.1 Introduction

Mathematical models of groundwater flow have been used since the late 1800s. A mathematical model consists of differential equations developed from analyzing groundwater flow (or solute transport in groundwater) and are known to govern the physics of flow (and transport). The reliability of model predictions depends on how well the model approximates the actual situation in the

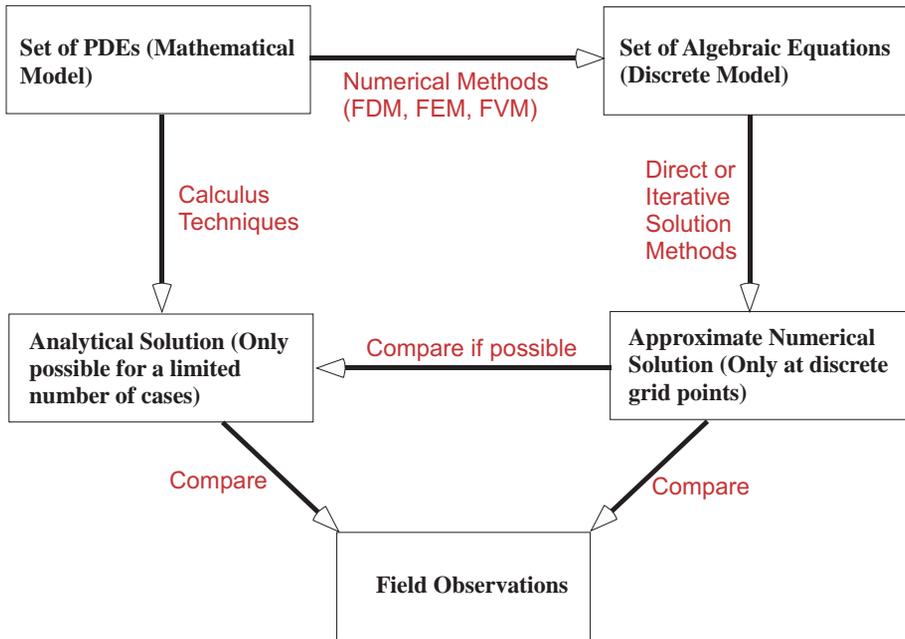


Figure 1: A flow chart showing the “big picture” of numerical modeling.

field. Inevitably, simplifying assumption must be made in order to construct a model, because the field situation is usually too complicated to be simulated exactly. In general, the assumptions necessary to solve a mathematical model *analytically* are very restrictive. For example, many analytical solutions are developed for homogeneous, isotropic, and/or infinite geological formations where flow is also steady-state (hydraulic head and groundwater velocity do not change with time). To deal with the more realistic situations (e.g., heterogeneous and anisotropic aquifer in which groundwater flow is transient), the mathematical model is commonly solved approximately using *numerical* techniques.

Ever since the 1960s, when computers first become widely available, numerical models have been the preferred type of model for analyzing groundwater flow and transport. Figure 1 illustrates the relationship between mathematical model, discrete algebraic model, analytical solution, approximate (numerical) solution, and field observations.

0.2 This Class

In this course, the goal is the use of numerical models to simulate groundwater flow and solute transport. It offers an introduction to the popular Finite

Difference Method¹, a widely used numerical method to solve the governing partial differential equations in hydrogeology. Fundamental understanding will be emphasized, as a fairly rigorous mathematical treatment is presented. This approach deviates from the typical introductory classes that may emphasize the more applied aspects, i.e., software application. I feel that mastering of such skills would come naturally after a student first develops a firm grasp of the fundamentals. This course is thus designed at the advanced upper undergraduate and graduate level, appropriate for the level of mathematical rigor contained herein. To comprehend the materials presented, a student should have sufficient knowledge of college math, e.g., Calculus I & II (required), Linear Algebra (optional), Differential Equations (optional). A prior class in Hydrogeology is required, e.g., Geohydrology (GEOL4444/5444).

Throughout the course, many formulations and equations are developed using mathematics. The emphasis is on understanding how these equations are obtained. However, you will rarely be tested on equation derivations in exams (those few that you may be tested on — I'll let you know), so it is unnecessary to memorize a large number of formulas or solutions. Typically, the exams will provide the necessary formulations so understanding what they mean and how to use them is key.

0.2.1 Textbook

The set of lecture notes I developed include the essential contents for this course. However, other books can serve as additional resources for self study:

- Introduction to groundwater modeling: finite difference and finite element methods, H. F. Wang and M. P. Anderson, 1995, Academic Press, 237 p.
- Applied Contaminant Transport Modeling, C. Zheng, G. D. Bennett, 2002, Wiley-Interscience, 656 p.
- Groundwater modeling by the finite element method, Jonathan Istok, 1989, American Geophysical Union Publication, 495 p.

Some materials are obtained by assembling course notes prepared by others (references will be given in the notes).

Due to time constraint, the current course cannot hope to cover every aspect of the subject as presented in these books. For example, variable-density flow is not presented, nor is flow in deforming porous media. Immiscible, multiphase fluids (e.g., flow in the soil zone) are not covered. In terms of numerical approaches, only the Finite Difference Method is presented. Some of these topics can be understood by independent study, others (e.g., Finite Element Method) should be the subjects of specialized classes. Thus, most of this class is devoted to the study of **single-phase** (water), **uniform-density** flow moving through non-deforming porous media (e.g., groundwater aquifers that are not

¹Due to time limitation, the Finite Element Method may be introduced in a separate future class.

going through compaction or subsidence). The solute transport equation and solutions are developed for a nonreactive solute with small concentrations so as not to affect the density of the groundwater in which it resides.

0.2.2 Tools

In this course, exercises, homework, labs, term project, exam problems can be solved using a variety of approaches, e.g., by hand, using Excel spreadsheet, writing Matlab codes, using software. The popular software packages (e.g., MODFLOW from USGS) will be introduced in the end, when we conduct 3D modeling of steady-state groundwater flow and advective transport.

0.2.3 Questions and Answers

Students can ask questions: (1) during office hour; (2) during lectures. As a rule, email is the last resort since I receive a lot of them every day and your message stands a chance of being overlooked by mistake.

0.2.4 Homework, Projects, Exams and Grades

When working on homework/lab/exam problems, read each assignment carefully. Read it twice if you have to. Do not skip anything or you might find that the later questions will not make sense. Of course, you should always point out to me if anything is unclear in the problem descriptions. A study guide will be provided before the exam.

Concerning the styles, 4 points must be emphasized: (1) Most problem sets involve writing equations, so if appropriate, provide a complete analysis rather than a single number. (2) Be professional in your presentations: write down the unit for your numerical results and round off the final number to 1 or 2 decimal point. (3) You can discuss the problem sets with fellow students, but complete your assignments by yourself. Copying other's work is considered cheating and no points will be given for that assignment. In particular, a student caught cheating may receive a "F" for the class. (4) Hand in your work on time. Unless otherwise stated, the general timeline is to hand in your homework in the beginning of the class a week from the day it is assigned (see syllabus for the detailed due dates). For larger lab assignments, you will be given 2 weeks to complete it. **If the assignment is not handed in on time, for every day it is delayed, 10 points will be taken out of the 100 points of this work until no points remain.**

Grades: In this course, emphasis is placed on the homework problems, labs and projects due to the time-consuming nature of most of these assignments. The final grade will be given based on your homework, labs, project and/or exams. The tentative percentage is shown:

- Homework 49% ($7\% \times 7$ homework)

Table 1: Letter versus numerical grade

		S	U		
		>60	< 60		
A	B	C	D	F	
90-100	80-89	70-79	60-69	< 60	

- Projects 36% (12% × 3 projects)
- Term Project (Or Final Exam) 15%

Note that each homework/lab/exam has a stand-alone grade of 100 points. When determining the final grade, these will be normalized reflecting the percentage distribution above. The final letter grade is shown in Table(1).

Finally, I set high expectation in this class. Be prepared to come to class, pay full attention, participate in exercises, work out the homework/lab by yourself, hand in your assignments on time, write professionally (clear, precise, concise). Please keep all course materials (notes, exercises, homework, exams, labs) to yourself and do not pass them on to future students. They must, as you have, work to earn the credit.

0.3 Basic Math Review

Please review the first handout: BasicMathReview.pdf, which is Chapter 1 **Basic Math** of the Geohydrology class (GEOL4444/5444) — for the mathematics you should already know before taking this class. If you find the math incomprehensible, please consider taking some other courses first, e.g., Calculus, Geohydrology. Please ignore 1.13 Test 1.

Gradient Tutorial

Please review the second handout: Gradient_123D.pdf, which is a tutorial of how to compute hydraulic head gradient in 1D, 2D and 3D for any coordinate system.

0.4 Homework 1

Exercise 1 and Exercise 2 of the first handout.

Chapter 1

Groundwater Flow Equations (Review)

1.1 Introduction

The first step in developing a mathematical model of almost any system is to formulate what are known as general equations. General equations are differential equations that are derived from the physical principles governing the process that is to be modeled. In the case of groundwater flow, the relevant physical principles are Darcy's law and mass balance. By combining the mathematical relation describing each principle, it is possible to come up with a general groundwater flow equation, which is a partial differential equation (PDE). Since in groundwater studies, the fluxes (Darcy flux, average linear velocity) are macroscopic quantities which are related to the head gradient and hydraulic properties of the aquifer (i.e., porosity, permeability, hydraulic conductivity) by the Darcy's law, the PDE developed for the general groundwater flow is thus established for macroscopic flow in porous media.¹

There are several different forms of the general flow equation depending on whether the flow is two-dimensional or three-dimensional, isotropic or anisotropic, and transient or steady state. In the following sections, we'll first review the most general flow equation (heterogenous, anisotropic and transient) and then look at the simplified equations under different simplifying assumptions (homogenous, isotropic, steady-state). To make the derivations as understandable as possible, a simple approach is adopted: first, the general equation is developed for the simple one-dimensional case; the results are then extended to the three-dimensional case.

¹There are even more fundamental equations developed for flow at the microscopic scale (i.e., pore scale), based on the theory of hydrodynamics. For example, with the Navier-Stokes equation, hydrologists have used various approaches to prove the macroscopic Darcy's law. These topics are however beyond this course. Interested students may find relevant discussions in Section 5.10 of Bear (1988).

1.2 Darcy's Law — Review

In 1856, engineer Henry Darcy was working on a project using sand to filter water supply for a French town. He performed laboratory experiments to examine the factors that govern the rate of water through sand. The results of his experiments defined the *empirical* principle of groundwater flow, in an equation known as Darcy's law. Darcy's experiment consisted of a sand-filled column with an inlet and an outlet for water (search Google image for Darcy's law will bring up many images of the Darcy's experiment). Two manometers measure the hydraulic head at two points within the column (h_1 and h_2). The sand is fully saturated, and a steady flow of water is forced through it at a volumetric rate of Q [L^3/T] (Q is sometimes called the flow rate or the discharge rate). Darcy found that $Q \sim$ head difference Δh between the manometers, inversely proportional to the distance between manometers Δs , and \sim cross sectional area of the sand column (A):

$$Q \propto \Delta h \quad Q \propto (1/\Delta s) \quad Q \propto A$$

Combining these observations and writing an equation in differential form gives Darcy's law for one-dimensional flow:

$$Q = -KA(dh/ds)$$

where Q is discharge rate in the s direction. The minus sign is necessary because head decreases in the direction of flow (i.e., water is always flowing from higher hydraulic head to lower hydraulic head). If there is flow in the positive s direction, Q is positive and dh/ds is negative². Conversely, when flow is in the negative s direction, Q is negative and dh/ds is positive. The constant of proportionality K is the hydraulic conductivity in the s direction, a property of the porous medium and the fluid (water) filling the pores. The common units for hydraulic conductivity are meters/year for regional studies, m/day for local aquifer-scale studies, and cm/sec for laboratory studies. Therefore, in some analysis, we often deviate from the rule of using the SI unit.

Another form of the Darcy's law is written for the **Darcy flux** (or the Darcy Velocity, or, the Specific Discharge) (q) which is the discharge rate per unit cross-sectional area:

$$q = Q/A = -K(dh/ds)$$

The Darcy flux q has unit of velocity [L/T]. Darcy flux is not the actual fluid *velocity* in the porous media, it is just discharge rate (Q) per unit (solid) cross-sectional area. Since porous rock does not have a solid cross-sectional area, Darcy flux is a not the fluid velocity.

In the real subsurface, groundwater flows in complex 3D patterns. Darcy's law in three dimensions is analogous to that of one dimension. In a Cartesian

²In this case, head is decreasing with s (since flow is towards $+s$), so by definition: $dh/ds \simeq \Delta h/\Delta s = \frac{h(s+\Delta s)-h(s)}{\Delta s} < 0$, since $h(s+\Delta s) < h(s)$.

x, y, z coordinate system, it is commonly expressed as:

$$\begin{aligned}q_x &= -K_x \partial h / \partial x \\q_y &= -K_y \partial h / \partial y \\q_z &= -K_z \partial h / \partial z\end{aligned}$$

where K_x , K_y , and K_z are the hydraulic conductivity in each of the coordinate direction, respectively. q_x , q_y , and q_z are 3 components of the Darcy flux \vec{q} .³ However, what exactly are the K_x , K_y , and K_z ? They are the directional hydraulic conductivity evaluated along each of the coordinate axis. To estimate these directional conductivities, Darcy test can be conducted along the x axis, in which case a horizontal hydraulic conductivity along the x direction can be determined: K_x . Same idea applies to estimating conductivities in the y and z directions.

In general, for a given porous medium, K_x , K_y and K_z do not need to be the same, in which case, the medium is called *anisotropic*. On the other hand, if $K_x = K_y = K_z$, the medium is called *isotropic*.

In general, x, y, z can have any orientation, but it is common to set z vertical and x and y horizontal. In three dimensions, all fluxes (\vec{Q} , \vec{q} , \vec{v}) are vector quantities (3 components), as is the hydraulic gradient $\vec{I} = [I_x, I_y, I_z]^T$ ($I_x = \partial h / \partial x$, $I_y = \partial h / \partial y$, $I_z = \partial h / \partial z$). The hydraulic head is still a scalar (1 component), while the hydraulic conductivity is a tensor quantity (9 components):

$$\mathbf{K} = \begin{bmatrix} K_{xx} & K_{xy} & K_{xz} \\ K_{yx} & K_{yy} & K_{yz} \\ K_{zx} & K_{zy} & K_{zz} \end{bmatrix}$$

Notice that in the above equation, \mathbf{K} contains more components than what is shown in the previous 3D Darcy's law with only K_x, K_y, K_z . Actually, the previous Darcy's law is established under a special requirement: **the 3 coordinate axes coincide with the principal axes of the hydraulic conductivity tensor (the direction of maximum, minimum, and intermediate hydraulic conductivity)**. When this requirement is satisfied, \mathbf{K} becomes a

³Note that K_x , K_y , K_z , q_x , q_y , and q_z are all specific to the given coordinate. When we change or rotate the coordinate, the magnitude of these component quantities will change with respect to the coordinate.

diagonal tensor⁴ and the Darcy's law thus becomes simplified:

$$\mathbf{K} = \begin{bmatrix} K_x & 0 & 0 \\ 0 & K_y & 0 \\ 0 & 0 & K_z \end{bmatrix}$$

We can write: $\mathbf{K} = \text{diag}[K_x, K_y, K_z]$ and K_x , K_y , and K_z are the three principal components (or eigenvalues) of \mathbf{K} .

When this coordinate alignment requirement is not satisfied, i.e., principal conductivity axes do not coincide with the coordinate axes, the most general form of 3D Darcy's law is written as:

$$\vec{q} = -\mathbf{K}\vec{I}$$

or, more explicitly:

$$\begin{Bmatrix} q_x \\ q_y \\ q_z \end{Bmatrix} = - \begin{bmatrix} K_{xx} & K_{xy} & K_{xz} \\ K_{yx} & K_{yy} & K_{yz} \\ K_{zx} & K_{zy} & K_{zz} \end{bmatrix} \begin{Bmatrix} \partial h / \partial x \\ \partial h / \partial y \\ \partial h / \partial z \end{Bmatrix}$$

we can easily verify that previous 3D Darcy's law (with 3 principal components: K_x, K_y, K_z) is just a reduced form of the above equation when \mathbf{K} is a diagonal tensor (under the condition that the coordinate axes are aligned with the principal axes of \mathbf{K}).

If we have time, I'll show you how to do tensor transformation with changing coordinate, i.e., relating the principal components of \mathbf{K} to its full tensor form.

1.2.1 Darcy's Law of This Course

In this course, to facilitate calculations, we make the assumption that the coordinate axes coincide with the principal axes of the hydraulic conductivity tensor, so the hydraulic conductivity is diagonal. This will considerably simplify the FD formulations we'll use. Thus, (in this course) we always use:

$$\begin{Bmatrix} q_x \\ q_y \\ q_z \end{Bmatrix} = - \begin{bmatrix} K_x & 0 & 0 \\ 0 & K_y & 0 \\ 0 & 0 & K_z \end{bmatrix} \begin{Bmatrix} \partial h / \partial x \\ \partial h / \partial y \\ \partial h / \partial z \end{Bmatrix} = - \begin{Bmatrix} K_x \partial h / \partial x \\ K_y \partial h / \partial y \\ K_z \partial h / \partial z \end{Bmatrix}$$

⁴If the time allows, I'll derive and explain the principle behind tensor rotation and eigen analysis. In turns out, a full tensor \mathbf{K} of any coordinate system can be projected, by coordinate transform, to a diagonal tensor $\mathbf{K} = \text{diag}[K_x, K_y, K_z]$ which are defined along 3 orthogonal principal axes (the direction of each can be defined by the unit eigen vector of \mathbf{K}). K_x , K_y , and K_z are the corresponding eigenvalues. In hydrology, the above K_x , K_y , and K_z are also called principal components, the corresponding eigen vectors are sometimes called principal axes. In Matlab, you can type "help eig" to find out how to do a eigen analysis for a (n x n) square symmetric matrix. Note that for tensors of n=2 or 3, we can visualize the relations in 2D and 3D geometric coordinate (thus tensor rotation can be proved by geometric means). But when n > 3, the eigen values and vectors cannot be visualized geometrically. A formal introduction on eigen analysis can be found in this online textbook: <http://tutorial.math.lamar.edu/Classes/LinAlg/LinAlg.aspx> (see the Chapter titled "Eigenvalues and Eigenvectors").

You can see that the above relation reduces to the previous 3D Darcy's law. **It is best if you memorize this equation**, but keep in mind the underlying assumptions.

Moreover, when analyzing 2D flow along the x-y plane (i.e., the velocity component in the z direction does not change: $q_z = 0$), the above relation becomes:

$$\begin{Bmatrix} q_x \\ q_y \end{Bmatrix} = - \begin{bmatrix} K_x & 0 \\ 0 & K_y \end{bmatrix} \begin{Bmatrix} \partial h / \partial x \\ \partial h / \partial y \end{Bmatrix} = - \begin{Bmatrix} K_x \partial h / \partial x \\ K_y \partial h / \partial y \end{Bmatrix}$$

This is because, according to Darcy's law of this course, $q_z = -K_z(\partial h / \partial z) = 0$, since conductivity is always greater than 0, this leads to $(\partial h / \partial z) = 0$. Substituting both into the above full 3D equation leads to the current 2D form (since the last equation becomes: $q_z = 0 = -K_z \times 0$ which no longer contains an unknown variable).

A similar form can be written for 2D flow in the x-z plane (what do you think it is?). Sometimes however the velocity is amenable to one-dimensional analysis (e.g., flow in the Darcy Tube). In this case, only the one-dimensional Darcy's law is needed:

$$q_s = -K_s(\partial h / \partial s)$$

Where K_s is hydraulic conductivity along the s direction (s can be aligned in the x , y , z , or any direction of interest), $\partial h / \partial s$ is the hydraulic head gradient along the s direction. For example, groundwater flow in consolidating clay tends to be largely vertical, in this case, if we decide to conduct 1D analysis, we write: $q_z = -K_z(\partial h / \partial z)$.

Finally, to facilitate analytical analysis by hand, many exercises in this class make the assumption of an *isotropic* conductivity in the aquifer of interest, which means: $K_x = K_y = K_z = K$, and the above general 3D equation can be simplified to become:

$$\begin{Bmatrix} q_x \\ q_y \\ q_z \end{Bmatrix} = -K \begin{Bmatrix} \partial h / \partial x \\ \partial h / \partial y \\ \partial h / \partial z \end{Bmatrix}$$

Similarly, depending on the dimension of the flow analysis, it can be reduced to a 2D or 1D form. For example, 2D flow in the x-y plane of an isotropic aquifer is written as:

$$\begin{Bmatrix} q_x \\ q_y \end{Bmatrix} = -K \begin{Bmatrix} \partial h / \partial x \\ \partial h / \partial y \end{Bmatrix}$$

Further, a 1D form is just the Darcy's law for an isotropic medium:

$$q_s = -K(\partial h / \partial s).$$

Compare this equation with the 1D Darcy's law written above: in this form, the conductivity without a subscript (K) implies isotropy; in the previous form (K_s or K_z), it implies anisotropy (though flow analysis may be one-dimensional, spatially the conductivity can be anisotropic, K_s is just the conductivity along the direction for which flow analysis is conducted).

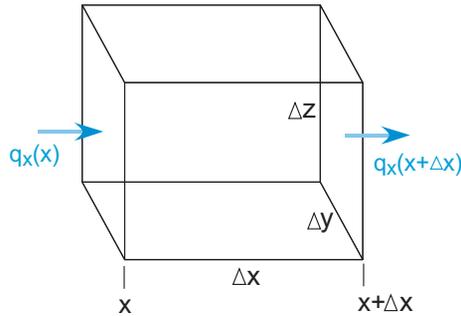


Figure 1.1: A control volume element within the saturated zone with fixed dimensions: Δx , Δy and Δz .

1.3 3D General Flow Equation (Confined Aquifer)

In a typical mass balance analysis, the net flux of *mass* through the boundary of a fixed control volume element is equated to the (time) rate of change of mass within the element:

(1.1)

$$\text{mass flux in} [M/T] - \text{mass flux out} [M/T] = \text{rate of change of mass} [M/T]$$

For example, if we pour water into a tub at 10 kg/min (mass flux in), but the tub is leaking at 4 kg/min (mass flux out), then the water in the tub experiences a rate of mass change at $10 - 4 = 6$ kg/min. This means at every minute, the water in the tub increases by 6 kg (so the rate of change of mass is + 6kg/min). This same mass balance principle can be applied to develop the groundwater flow equation. Instead of the stationary tub, we'll use the fixed and stationary control volume again.

We will consider the mass balance for a small rectangular control volume element within the saturated zone, as shown in Figure 1.1. The dimensions of the element are fixed in space, regardless of compression or dilation of the aquifer matrix or compression/expansion of pore water. For example, if the aquifer compresses, more aquifer solids will be squeezed into the element and some water will be displaced out of it. To make the derivation of the flow equations as clear as possible, we will assume that the macroscopic flow in the vicinity of this element is one-dimensional in the x direction: $q_x \neq 0$, $q_y = q_z = 0$. Using mass balance and Darcy's law, combined with the definition of specific storage (S_s) for a confined aquifer, we obtain for 1D flow (details given in class):

$$(1.2) \quad \frac{\partial}{\partial x} \left(K_x \frac{\partial h}{\partial x} \right) = S_s \frac{\partial h}{\partial t}$$

where S_s is the specific storage of a confined aquifer. Note that in developing equation (1.2), the spatial gradient of the water density is considered to be negligible (mostly fresh water).

If the previous analysis is carried out without the restriction of one-dimensional flow, there would be additional flux terms in the y and z directions (similar to the flux term in the x direction). For three-dimensional flow, the general equation is:

$$(1.3) \quad \frac{\partial}{\partial x} \left(K_x \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_y \frac{\partial h}{\partial y} \right) + \frac{\partial}{\partial z} \left(K_z \frac{\partial h}{\partial z} \right) = S_s \frac{\partial h}{\partial t}$$

Thus the hydraulic head ($h = h(x, y, z, t)$) must obey this PDE to be consistent with both the Darcy's law and mass balance. Note that in equation (1.3), the spatial gradient of the water density is considered to be negligible, and, the principal conductivity axes must be aligned with the coordinate axes (conductivity becomes a diagonal tensor: $K = \text{diag}[K_x, K_y, K_z]$).

Equation (1.3) is the most universal form of the saturated flow equation for a confined aquifer⁵, allowing flow in three dimensions, transient flow ($\partial h / \partial t \neq 0$), heterogenous conductivities (e.g., K_x, K_y, K_z are spatially variable), and anisotropic porous medium ($K_x \neq K_y \neq K_z$).

Equation (1.3) can be alternatively written in more condensed forms, e.g., Figure 1.2 shows two ways this equation is sometimes presented. On the other hand, from the last of this equation:

$$-\nabla \cdot \vec{q} = S_s \frac{\partial h}{\partial t}$$

we can obtain the “most” general form of the 3D flow equation, see Figure 1.3. Note that the most general 3D flow equation can also be written with the Einstein summation (see Geohydrology, Advanced material, for explanation):

$$\frac{\partial}{\partial x_i} [K_{ij} \frac{\partial h}{\partial x_j}] = S_s \frac{\partial h}{\partial t}$$

Finally, equation (1.3) is just a reduced form of the most general equation under the condition that conductivity principal directions are aligned with the axes of the working coordinate.

1.3.1 Less General Flow Equations

Other, less general, forms of the flow equations can be derived from Equation(1.3) by making various simplifying assumptions. If the hydraulic conductivities are assumed to be homogenous (K_x, K_y, K_z are independent of x, y, z),

⁵Actually, it is not. The most general form is when the coordinate axes are not aligned with the conductivity principal axes, thus K has 9 components. This most general form of groundwater flow equation will be presented in class, if time allows. However, many existing groundwater programs and softwares solves equation (1.3), assuming that during the model building stage, the coordinate system had been designed to align with the principal conductivity axes. This is a reasonable approach for many groundwater applications, however, for complex geometry with multiple undulating layers or with intersecting structures, the most general form may be necessary, allowing the more accurate calculations of the cross flows.

3D General Flow Equation in a Confined Aquifer

$$\frac{\partial}{\partial x} \left(K_x \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_y \frac{\partial h}{\partial y} \right) + \frac{\partial}{\partial z} \left(K_z \frac{\partial h}{\partial z} \right) = S_s \frac{\partial h}{\partial t} \quad (6.10)$$



$$\begin{aligned} q_x &= -K_x \frac{\partial h}{\partial x} \\ q_y &= -K_y \frac{\partial h}{\partial y} \\ q_z &= -K_z \frac{\partial h}{\partial z} \end{aligned} \quad \begin{array}{l} \text{3D Darcy's Law of this} \\ \text{course (assuming K prin-} \\ \text{cipal directions aligned} \\ \text{with coordinate axes)} \end{array}$$

$$-\left[\frac{\partial}{\partial x} (q_x) + \frac{\partial}{\partial y} (q_y) + \frac{\partial}{\partial z} (q_z) \right] = S_s \frac{\partial h}{\partial t}$$

$$-\nabla \cdot \vec{q} = S_s \frac{\partial h}{\partial t}$$

Figure 1.2: Alternative forms of equation (1.3) by substituting the Darcy's law into the equation.

the general equation can be written as:

$$(1.4) \quad K_x \frac{\partial^2 h}{\partial x^2} + K_y \frac{\partial^2 h}{\partial y^2} + K_z \frac{\partial^2 h}{\partial z^2} = S_s \frac{\partial h}{\partial t}$$

This simplifies further when conductivity is also isotropic ($K_x = K_y = K_z = K$):

$$(1.5) \quad \frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} + \frac{\partial^2 h}{\partial z^2} = \nabla^2 h = \frac{S_s}{K} \frac{\partial h}{\partial t}$$

where ∇^2 is called the Laplacian operator— a shorthand for the sum of the second derivatives: $\nabla^2() = \frac{\partial^2()}{\partial x^2} + \frac{\partial^2()}{\partial y^2} + \frac{\partial^2()}{\partial z^2}$.

In this section, all the previous equations have included the storage term ($S_s \frac{\partial h}{\partial t}$) which occurs only with transient flow ($\frac{\partial h}{\partial t} \neq 0$). If however the flow is steady state ($\frac{\partial h}{\partial t} = 0$), the right-hand-side (RHS) of all the above equations becomes zero. For example, under steady state, equation (1.5) becomes:

$$(1.6) \quad \nabla^2 h = 0$$

This is a well-known PDE called the Laplace equation. Equation (1.6) has many applications in fluid flow, heat conduction, electronics and elasticity.

The above equations in three-dimensional forms can be reduced to two or one dimensions by simply dropping y and/or z terms from the equation. For example, equation (1.3) can be reduced to solve a two-dimensional flow problem in the horizontal plane (x,y axes only):

$$\frac{\partial}{\partial x} \left(K_x \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_y \frac{\partial h}{\partial y} \right) = S_s \frac{\partial h}{\partial t}$$

3D Most General Flow Equation in a Confined Aquifer

$$-\nabla \cdot \vec{q} = S_s \frac{\partial h}{\partial t}$$

$$\vec{q} = -\overline{\overline{K}} \nabla h \quad \overline{\overline{K}} = \begin{bmatrix} K_{xx} & K_{xy} & K_{xz} \\ K_{yx} & K_{yy} & K_{yz} \\ K_{zx} & K_{zy} & K_{zz} \end{bmatrix}$$

Most general Form of 3D Darcy's Law (K principal directions NOT aligned with coordinate axes)

$$\nabla \cdot (\overline{\overline{K}} \nabla h) = S_s \frac{\partial h}{\partial t} \rightarrow$$

$$\begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \cdot \begin{bmatrix} K_{xx} & K_{xy} & K_{xz} \\ K_{yx} & K_{yy} & K_{yz} \\ K_{zx} & K_{zy} & K_{zz} \end{bmatrix} \begin{pmatrix} \frac{\partial h}{\partial x} \\ \frac{\partial h}{\partial y} \\ \frac{\partial h}{\partial z} \end{pmatrix} = S_s \frac{\partial h}{\partial t} \rightarrow$$

When conductivity principle directions are aligned with the coordinate axes, K becomes a diagonal tensor (Kx, Ky, Kz are the principal components)

3D Most General Flow Equation Reduced to (6.10)

$$\begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \cdot \begin{bmatrix} K_x & 0 & 0 \\ 0 & K_y & 0 \\ 0 & 0 & K_z \end{bmatrix} \begin{pmatrix} \frac{\partial h}{\partial x} \\ \frac{\partial h}{\partial y} \\ \frac{\partial h}{\partial z} \end{pmatrix} = S_s \frac{\partial h}{\partial t} \quad (6.10)$$

Figure 1.3: The most general form of the flow equation. Also shown is the condition under which this equation is reduced to equation (1.3).

This implies that $\frac{\partial}{\partial z} \left(K_z \frac{\partial h}{\partial z} \right) = 0$ which becomes true if $q_z = -K_z \frac{\partial h}{\partial z} = 0$ or $\partial h / \partial z = 0$ (envison a 3D flow field where velocity (\vec{q}) varies along the horizontal plane, but its vertical component is zero).

Equation (1.3) can also be reduced to solve a two-dimensional flow problem in the vertical plane (e.g., x,z axes only):

$$\frac{\partial}{\partial x} \left(K_x \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial z} \left(K_z \frac{\partial h}{\partial z} \right) = S_s \frac{\partial h}{\partial t}$$

Similarly, this implies that $\frac{\partial}{\partial y} \left(K_y \frac{\partial h}{\partial y} \right) = 0$ which becomes true if $q_y = -K_y \frac{\partial h}{\partial y} = 0$ or $\partial h / \partial y = 0$. A schematic diagram shows the conditions under which the groundwater flow in aquifers can be modeled as two-dimensional or fully three-dimensional (Figure 1.4).

1.4 2D Planeview Flow Equation (Confined & Unconfined)

Groundwater flow in aquifers is often modeled as two-dimensional in the *horizontal* plane. This is because most aquifers have large aspect ratio like a thin pancake, with horizontal dimensions hundreds times greater than the vertical thickness. In such a setting, groundwater more or less flows along the horizontal plane, i.e., the z component of the velocity is comparatively small. Therefore, a two-dimensional analysis is carried out in conjuncture with the use of transmissivity, by assuming that $h=h(x,y,t)$ only (h does not vary with z, thus $\partial h / \partial z = 0$). This simplification of modeling 3D aquifer flow as horizontal two-dimensional flow is called the **Dupuit-Forchheimer approximation**.

To picture what this simplification represents in a physical sense, imagine an aquifer that is perforated by numerous tiny vertical tubes, each possessing an infinitely large hydraulic conductivity (thus no resistance to flow in the vertical direction). This leads to a constant hydraulic head long each vertical tube, but laterally the head is still varying. Also, under this approximation, along any vertical profile, the pressure is hydrostatic (we'll prove this when working on the homework). Figure 1.5 illustrates the difference between actual three-dimensional flow and flow modeled with the Dupuit-Forchheimer approximation.

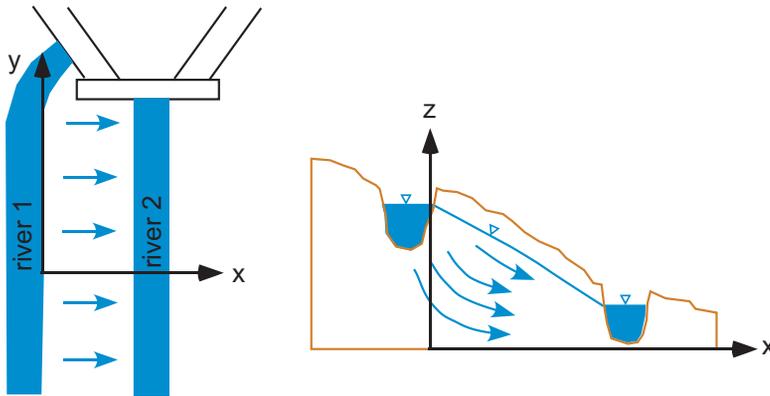
The general equation for the two-dimensional (horizontal) flow is developed first for one-dimensional flow ($q_x \neq 0, q_y = 0$), and then, it is extended to two-dimensions. By assuming that the spatial gradient of water density is small (i.e., ρ_w is mostly constant as in fresh water), a volume balance is done:

(1.7)

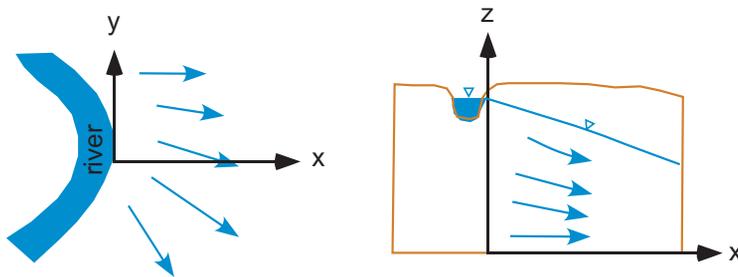
$$\text{volume flux in} [L^3/T] - \text{volume flux out} [L^3/T] = \text{rate of change of volume} [L^3/T]$$

Consider a fixed control volume that is a vertical prism with cross-sectional area $\Delta x \times \Delta y$, extending the full saturated thickness of the aquifer (b) (Figure 1.6). By computing the discharge fluxes across the two sides of the prism *as well as* computing another influx from recharge or leakage (N ([L/T]) is the net

2D (x-z) Flow ($q_y=0$)



3D Flow



3D Flow To a Well

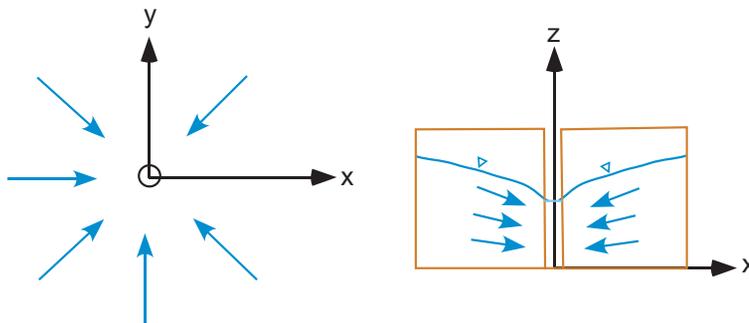


Figure 1.4: Planeview (x-y plane) of groundwater flow in aquifers with corresponding cross-sectional view (x-z plane). The top panel illustrates the configuration where flow can be modeled as 2D (in this case, $q_y = 0$); the bottom two panels illustrates the configurations where 3D flow must be modeled (in these cases, $q_x \neq 0$, $q_y \neq 0$, $q_z \neq 0$).

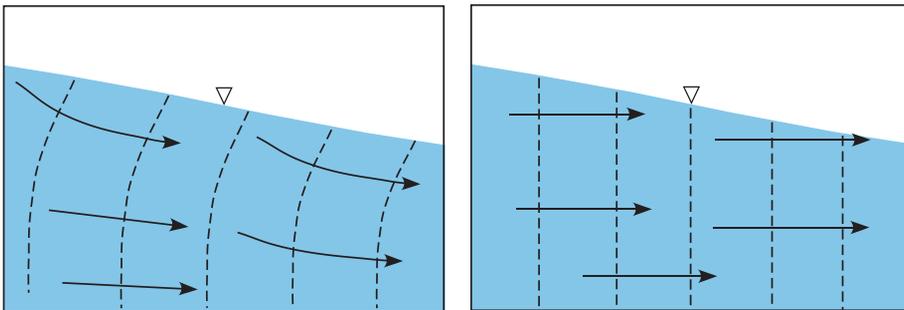


Figure 1.5: Cross-section of actual unconfined flow (left) and the same situation as modeled with the Dupuit-Forchheimer approximation (right). Hydraulic head contours are dashed lines. In the Dupuit-Forchheimer model, there is no resistance to vertical flow, resulting in constant head along the vertical lines ($\partial h/\partial z = 0$).

flux of water entering the prism from the top and bottom), the net volume flux across the prism boundaries is equated to the rate of change in volume stored in the prism (which is related to the storativity (S) for a confined aquifer or the specific yield (S_y) for an unconfined aquifer). After some equation manipulations and extending 1D flow to 2D (details given in class), we find:

$$(1.8) \quad \frac{\partial}{\partial x} \left(T_x \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(T_y \frac{\partial h}{\partial y} \right) + N = S \frac{\partial h}{\partial t}$$

This equation is the general equation for two-dimensional horizontal flow in an aquifer (under the Dupuit-Forchheimer approximation), allowing for heterogeneity and anisotropy in transmissivity.

Similar to the 3D flow analysis of the last section, the above two-dimensional general equation can be simplified under various assumptions. For example, if transmissivity is homogeneous and isotropic ($T_x = T_y = T = \text{constant}$):

$$(1.9) \quad T \left[\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} \right] + N = S \frac{\partial h}{\partial t}$$

This can be simplified further to $\nabla^2 h + N/T = (S/T)\partial h/\partial t$. If there is zero net vertical recharge or leakage ($N=0$), this becomes:

$$(1.10) \quad \nabla^2 h = \frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} = \frac{S}{T} \frac{\partial h}{\partial t}$$

If flow is steady-state but net vertical recharge/leakage is nonzero, this becomes:

$$(1.11) \quad \nabla^2 h = -N/T$$

Equation (1.11) is also called the Poisson equation.

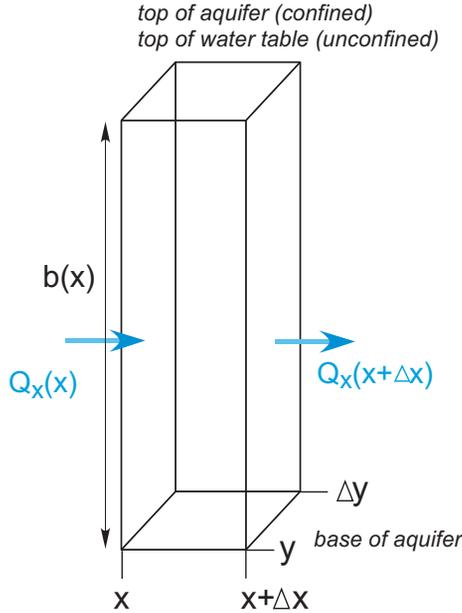


Figure 1.6: Prism control volume element of a two-dimensional aquifer.

When the recharge/leakage is zero, the Poisson equation further simplifies to the Laplace equation:

$$(1.12) \quad \nabla^2 h = 0$$

Unconfined Aquifer with Large Change in Saturated Thickness

For unconfined aquifer where its change in saturated thickness is small, the previous equations derived for the confined aquifer is generally applicable (4.2 3D General Flow Equation). However, there are situations where the saturated thickness of an unconfined aquifer experiences large variations. In these cases, sometimes special formulations are obtained.

Flow in a *homogeneous* and *isotropic unconfined* aquifer underlain by a horizontal impermeable base is a special case of the previous 2D analysis (4.3 2D Planeview Flow Equation). If the base of the aquifer is considered the head datum, then $h = b$, thus $T = Kh$, Equation (1.8) in this case becomes (details given in class):

$$(1.13) \quad \frac{\partial}{\partial x} \left(h \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(h \frac{\partial h}{\partial y} \right) + N/K = \frac{S_y}{K} \frac{\partial h}{\partial t}$$

If recharge and leakage can be ignored, the above relation becomes:

$$\frac{\partial}{\partial x} \left(h \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(h \frac{\partial h}{\partial y} \right) = \frac{S_y}{K} \frac{\partial h}{\partial t}$$

which is also called the *Boussinesq* equation.

The above two equations are nonlinear PDEs, since the terms involve h multiplied by its derivative. Nonlinear equations are more difficult to solve than linear ones. However, the nonlinearity disappears when writing these equations for h^2 . For example, equation (1.13) becomes (details given in class):

$$\nabla^2(h^2) + \frac{2N}{K} = \frac{2S_y}{K} \frac{\partial h}{\partial t}$$

For steady flow this becomes the linear Poisson equation:

$$(1.14) \quad \nabla^2(h^2) = -\frac{2N}{K}$$

Further, if the recharge/leakage is zero, equation (1.14) becomes the linear Laplace equation:

$$(1.15) \quad \nabla^2(h^2) = 0$$

Note that in these two last equations, the unknown is the head squared h^2 .

Chapter 2

Modeling Overview for Groundwater Flow

Models of groundwater flow are widely used for a variety of purposes ranging from water supply studies to designing contaminant cleanup. In general, groundwater flow system can be divided into steady-state and transient. In a steady-state flow system, the RHS of the flow equation is zero, while it is nonzero in transient systems (as reviewed in the last chapter).¹

There are dozens of free or commercial computer programs available to do the calculations involved with these models. In this course, by working on our own codes, we'll understand the fundamental mathematical principles and numerical methods behind these computer programs.

2.1 General Methodology

In general, mathematical modeling of groundwater flow and transport involves the following 3 steps (Fitts, 2002):

- Review all the available data about the material properties (porosity, permeability, storativity), heads, and fluxes in the vicinity of the region to be modeled.
- Develop a conceptual system that is simpler than the real flow system, but captures the important overall features of the real system.
- Simulate the conceptual system developed in step (2) using a mathematical model which is solved numerically with a computer code.

Usually developing a model requires several iterations through this process, revisiting steps (1) and (2) in light of the results of the simulations developed

¹Solute movement in groundwater is generally always transient. A steady-state concentration field (the RHS of the transport equation is zero), though possible, is rarely solved. We will defer the solution of the transient transport equations to Chapter 8.

in step (3). During this iteration, heads and fluxes predicted by the computer model are compared to actual measurements of heads and fluxes observed in the natural groundwater system. If significant discrepancies exist, the model parameters (e.g., porosity, permeability, storativity) and boundary conditions (specified head, specified flux) will be adjusted. The simulation is carried out again until a suitable (and usually rather subjective) fit between predictions and observations is achieved.

In the third step outlined above, the conceptual system should be examined to determine which general flow equation applies. There are several different general equations, each valid under certain conditions (as reviewed in the last Chapter). Which one applies depends on circumstances such as:

1. Should the flow be represented as one-, two-, or three-dimensional?
2. Can the flow be approximated as steady state or should it be represented as transient?
3. Is the flow in a confined or an unconfined aquifer?

For a given general equation, there is an infinite number of possible solutions. For steady-state flow, the unique and appropriate solution is one that matches the particular boundary conditions (BC) of the conceptual model. (For transient flow system and for solute transport, both initial condition and BC are required to obtain the unique solutions of head and concentrations.)

Boundary conditions include things like heads at surface waters in contact with the aquifer, the location and discharge rate of a pumping well or a leaching irrigation field. For the no-flow boundary (which is a special type of flux boundary with specified flux of zero), we need to investigate the involvement of aquitard for the spatial/temporal scale of interest: no-flow boundary (small scale engineering problems) versus low-K zone as part of model domain (larger regional flow problems)?

The two most common types of BC are specified head and specified flux conditions, as illustrated in Figure 2.1. At specified head boundaries, the head is known. These are usually where the groundwater is in direct contact with a surface water like a lake or a river. There are several types of specified flux conditions, including impermeable boundaries that allow zero flux, recharge boundaries at the top of the saturated zone, and wells that are pumped at a known rate. The less common, so-called mixed boundary condition involves some combination of head and flux specification².

Along a particular stretch of boundary, only one condition can be specified. For example, it is not possible to specify both the head and the flux along the same portion of a boundary (in class, we'll show how one type of BC

²For example, leakage through a silty river bed to an underlying aquifer is represented by a flux that is proportional to the vertical conductivity of the silt layer and proportional to the head difference from the river to the underlying aquifer (hint: write a vertical Darcy's Law for the silt layer)

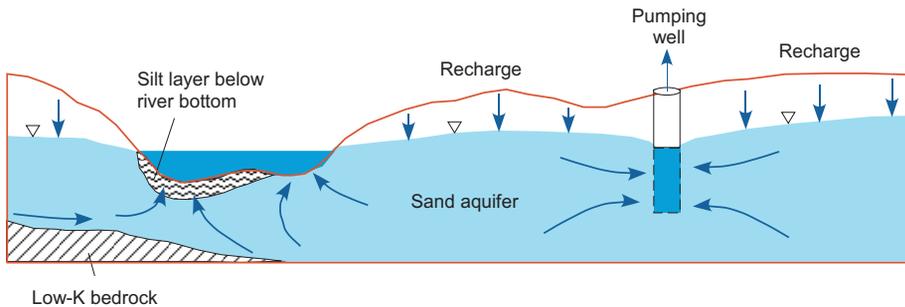


Figure 2.1: Examples of flow domain boundary conditions. The pumping well is a specified flux condition at the permeable section of the well. The recharge is a specified flux applied at the water table. The low-K bedrock is considered a special specified flux boundary (no-flow boundary). The leaky silt layer below the river is a mixed condition where the flux through the layer is proportional to the difference between the head in the river and the head in the aquifer beneath the silt layer. Where the river is in direct contact with the aquifer, there is a specified head condition.

can be translated to another type). You may specify either, but not both. Once the general equation is known and the boundary conditions are assessed, review the available modeling techniques and select one that can simulate the general equation and boundary conditions of the problem at hand. Construct the model and adjust its parameters and/or boundary conditions as necessary to fit the observations of the real system. The parameters that are input (conductivity, storativity, etc.) should be within the range of measured values, or if lacking measurements, within the range of expected values for the geologic materials present (Figure 2.2). Also the discharge fluxes that are modeled should be reasonable. For example, the discharge to a river segment should be similar to the measured or expected baseflow for that river segment.

Computer programs are now widely used to develop complex two- and three-dimensional models. These programs implement numerical (approximate) or analytic (exact) solutions to the general flow equations, and allow solutions with diverse and irregular boundary conditions. This class focuses on the numerical solutions, i.e., the Finite Difference Method (FDM), which solve the flow or transport equations at *discrete* grid points, for the unknown hydraulic head, groundwater fluxes or solute concentrations. (For transient problems, both initial and boundary conditions are needed, the flow and transport equations are solved at *both* discrete grid points and discrete times.)

On the other hand, if exists, the analytical solutions predict these values exactly at every *continuous* point in space (if steady-state flow) and continuous point in time (if transient flow). For the same problem, the numerical solution will generally converge to the analytical solution (if exists) when the numerical model grid employed is dense (i.e., the spacing between model grid cells is small).

Hydraulic Conductivity (K) and Permeability (k) Range for Unconsolidated Deposits and Rocks

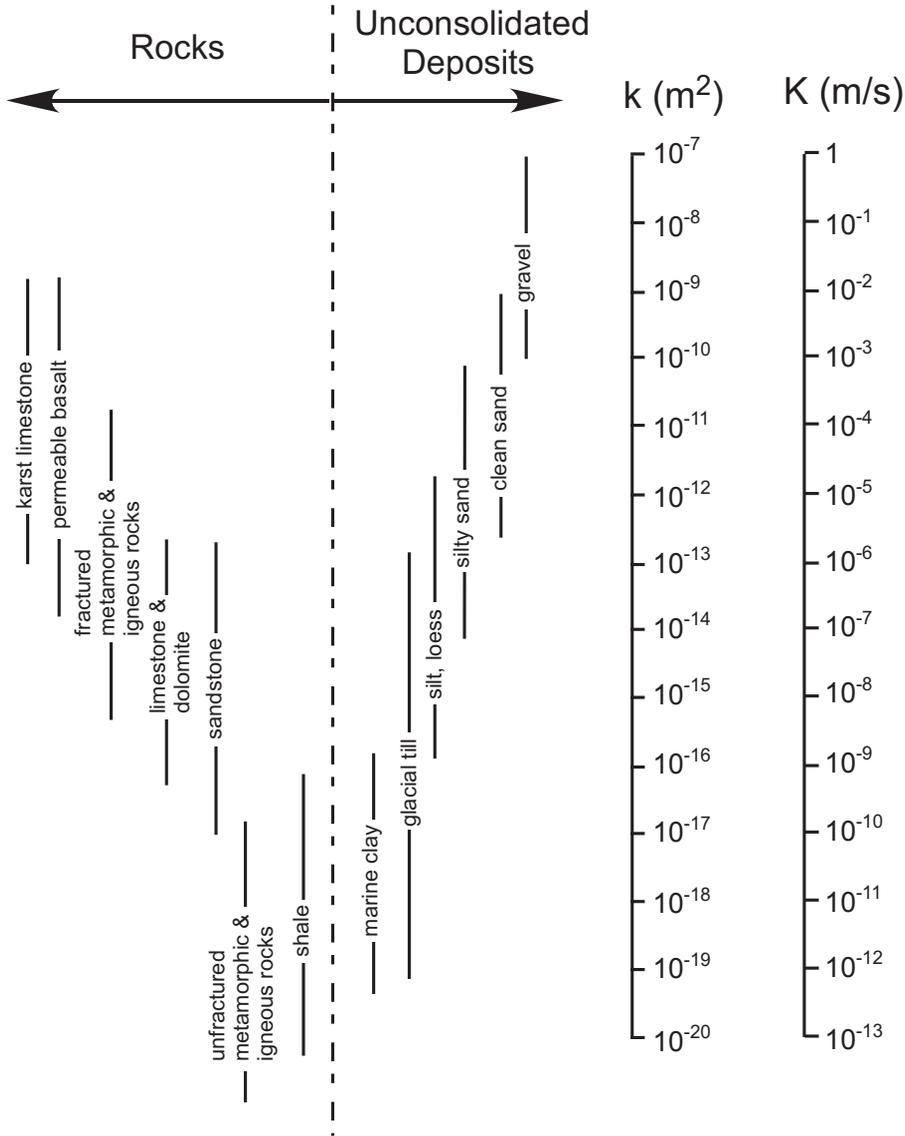


Figure 2.2: General range of hydraulic conductivity [L/T] and permeability [L^2] for different consolidated and unconsolidated geological materials (Freeze & Cherry, 1979).

As discussed by Fitts (2002), some groundwater models are very complicated that require months of labor to create and adjust. At the other end of the spectrum, simple models can be created in a matter of minutes. There still is a role for hand calculations though (particularly those we've done in the Geohydrology class), as it is often reasonable to neglect components of flow in one or two directions, and use a 1D or 2D analysis of a simple conceptual model. In these cases, a modeler can perform simple calculations, quickly. Hand calculations can also teach the modeler useful insights that can apply in general ways to other, more complex situations. In Geohydrology, I introduced several practical methods for analyzing 1D and 2D *steady-state* flow (Chapter 7) as well as the commonly used methods for analyzing *transient* flow and pumping tests (Chapter 8).

2.2 Analytical versus Numerical Methods

Mathematical models of groundwater flow and solute transport can be solved generally with two broad approaches:

1. Analytical solution of the mathematical equation gives *exact* solution to the problem, i.e., the unknown variable is solved continuously for every point in space (steady-state flow) and time (transient flow).
2. Numerical solution of the mathematical equation gives *approximate* solution to the problem, i.e., the unknown variable is solved at discrete points in space (steady-state flow) and time (transient flow).

For example, shown in Figure 2.3a is a regional confined aquifer: along the top boundary of this aquifer, the hydraulic head is described by the potentiometric surface which is higher than the aquifer top. Within the aquifer, groundwater flow forms several regional flow cells. In particular, the one near the middle is the model domain of interest. To understand the hydraulic head and streamline distribution in this region, mathematical model needs to be built and solved.

Toth (1962) derived an analytical solution for the hydraulic head in the solution domain (Figure 2.3b):

$$h(x, z) = h_0 + \frac{cL}{2} - \frac{4cL}{\pi^2} \sum_{m=0}^{\infty} \frac{\cos[(2m+1)\pi x/L] \cosh[(2m+1)\pi z/L]}{(2m+1)^2 \cosh[(2m+1)\pi h_0/L]}$$

To obtain this solution, Toth (1962) made several simplifying assumptions. For example, the aquifer is assumed homogeneous and isotropic for which the hydraulic conductivity is a scalar constant value. The flow is steady-state. Under these assumptions, the flow equation becomes the Laplace equation (see Figure 2.3b). The potentiometric surface of the confined aquifer is assumed linear. For this 2D problem, the hydraulic head of the top aquifer boundary is thus linear: $h(x, H) = cx + h_0$. Finally, the aquifer domain is assumed to be a rectangle (in reality, real aquifer thickness usually varies). Note that the advantage

A Regional Confined Aquifer

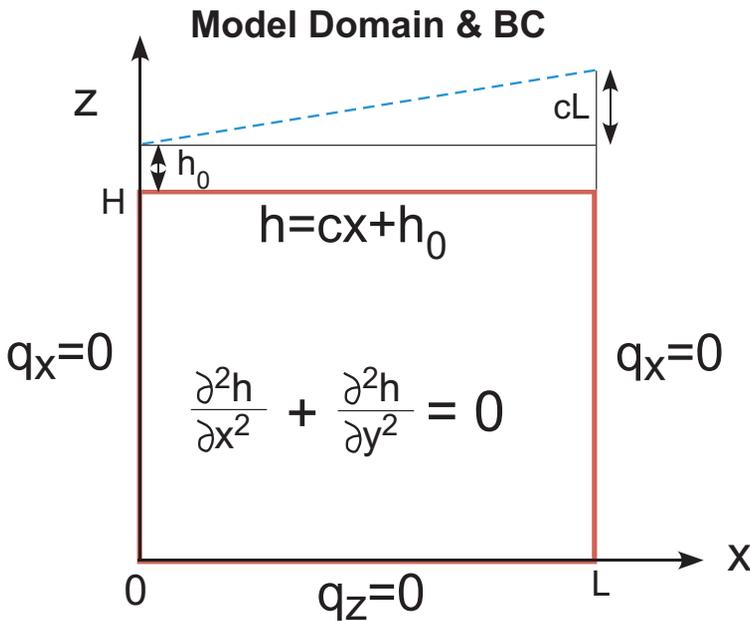
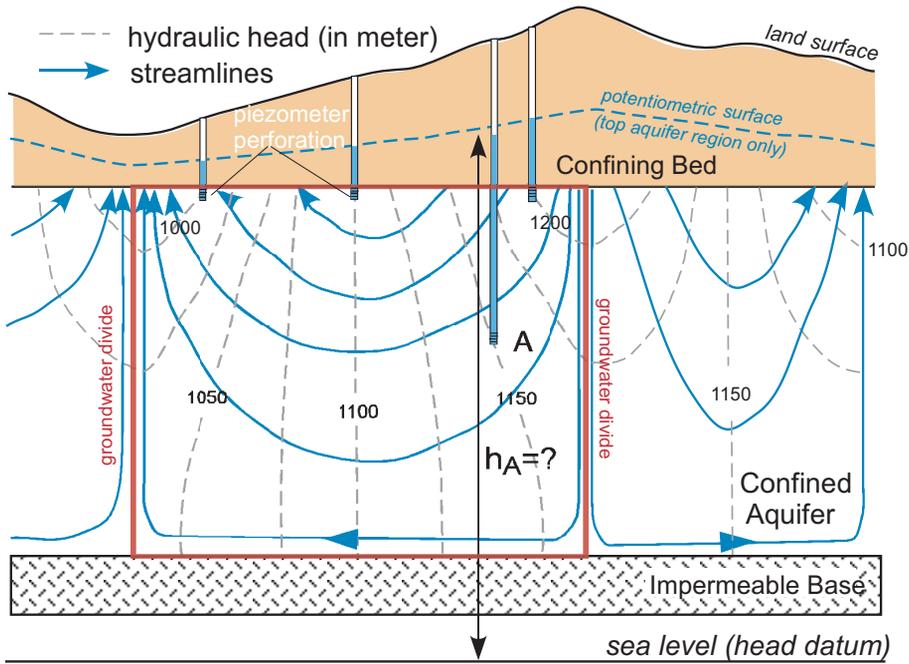


Figure 2.3: A regional confined aquifer: (a) a geological transect of the aquifer and its confining units; (b) the model domain and boundary conditions. Note that the region to be modeled has a length of L and height of H , corresponding

of the analytical solution is that it is an exact solution for the above problem under the various assumptions. A solution also exists for every point (x, z) in space. The disadvantage is the various assumptions as required to obtain the exact analytical solution. None of these assumptions is necessary to obtain a numerical (approximate) solution.

For many problems, the assumptions that must be made to obtain an analytical solution will not be realistic. In these cases, to solve the mathematical models, we must resort to approximate methods using numerical solution techniques. In general, three main numerical methods are used in hydrogeology (listed in the order of popularity): (1) Finite Difference Method (FDM); (2) Finite Element Method (FEM); (3) Boundary Element Method (AEM). The first two methods are the most popular. They transform the differential equation that governs flow (or transport) into a set of algebraic equations.

Before computers became widely available, only hand calculations (developing analytical solutions for highly simplified and idealized problems) or analogue studies were possible³, and the numerical techniques (FDM, FEM, AEM) were of limited value. Using computers, we can solve large number of algebraic equations by iterative methods or direct matrix-based methods (next chapter will introduce these methods). We can test the numerical solution by comparing that head (or solute concentration) distribution computed by the computer with that determined from an analytical solution, if one is available. We can also check the simulated head (or solute concentration) against the values observed in the field. This procedure is summarized in Figure 1.

The mathematical formulations for the FDM and FEM are quite different: the former is based on the Taylor's Series Expansion; the later is based on the Variational Principle which involves more advanced mathematics. In this class, due to time limitation and the popular demand to use a software package (e.g., MODFLOW is based on FDM), only the FDM is introduced. The FDM has certain advantage over FEM, e.g., it provides close approximation of mass continuity across grid cell boundaries (I'll draw a plot on board to demonstrate what this means). The traditional FEM's approximation of mass continuity is weak at the local cell boundaries. However, new approaches are continuously implemented to address the FEM local mass balance problems. The hybrid Finite Volume Method is becoming increasingly popular in the research community since it combines the flexibility of FEM in grid cell shape (Figure 2.5) with the local mass conservativeness of the FDM. However, FDM and FEM have similarities: they are both grid-based. Compared to the analytical solution which exists at any point in the solution domain, the numerical methods yield values

³Before computers were commonplace and inexpensive, complex groundwater flow problems were modeled using physical models or analogs. These models would typically be a miniature scaled model of the flow domain in a tank. The most common analog method uses the flow of electricity through a network of wires and resistors, where voltage is analogous to head, resistance is analogous to $1/K$ or $1/T$, current is analogous to discharge, and capacitance is analogous to storage. However, computer simulation methods have essentially replaced physical and analog modeling, so these methods are not discussed in this course. Interested students can find more coverage of analog methods in Walton (1970) *Groundwater Resource Evaluation*, McGraw-Hill, New York.

for only a fixed finite number of grid points (N) in the solution domain. Using either FDM or FEM, we convert a partial differential equations to a set of N algebraic equations involving N unknown values at these particular points. However, it is important to note that there are advantages of using FEM to solve the flow/transport problems. The most important one is its flexibility to simulate complex internal geometry and irregular domain boundary since the grid cells of the FEM can be of any shape. Depending on student demands, the FEM can be taught in future, in an advanced class or seminar in hydrogeology.

For example, in Figure 2.5a, an aquifer is shown in map view, with the locations of a well field, observation wells, and aquifer boundary. In Figure 2.5b, c, d, three different grids are used to represent the model domain to solve for either head (if only solving the flow equation) and concentration (if both flow and solute equations are solved). We can immediately see that the FDM can only give an approximate representation of the irregular boundary, while the FEM can honor irregular boundary more realistically. Also, two different grids are used in FDM: block-centered and mesh-centered. **In this class, to simplify the derivations, we use the block-centered grid to derive the FD approximations of the flow and transport equations.** Note that in hydrogeology, MODFLOW, the most popular groundwater/transport code, is based on block-centered FDM. Also, block-centered FDM is the preferred method for most general purpose petroleum reservoir simulators (Lee et al., 1998; Ertekin et al., 2001). The derivations for the mesh-centered grid are very similar (only minor details vary): with the understanding for the former method, the later one is easy to learn. Moreover, the grid spacing ($\Delta x, \Delta y$) are also of interest. The smaller we make Δx and Δy , the numerical approximate solution converges to the analytical solution. The trade-off is that the number of unknowns (N) increases. For very large problems, e.g., radionuclide transport at regional scale, oil reservoir simulations, the limitation on grid size is often set by the computer code and memory. We generally select a N value that will not be overwhelmingly large (exceeding the computation limit), but large enough that most of the main hydrogeological features are represented.

Finally, a variety of solution techniques can be used to solve a set of linear algebraic equations (separate classes on Numerical Linear Algebra exist, possibly from the Math Department, which should cover these techniques in a more comprehensive fashion). Based on my own past experiences, I've summarized the ones often used in hydrogeology (Figure 2.4). In this class, due to time limitation, we'll use only some of the solution techniques. For example, though we introduce the matrix-free iterative methods, due to their computation inefficiency, they're no longer used very much for solving large systems of equations which derive from large models. So, we'll not reply on them too much. On the other hand, the matrix-based iterative solver will only be mentioned in passing, as setting up these solvers will generally require access to a Unix cluster on which the various libraries are usually installed (currently our PC labs do not have these libraries). Working on Unix systems require knowledge on Unix (e.g., compile and link to external libraries), which is beyond the purpose of this course. Therefore, the most often used solution technique **in this class** is the

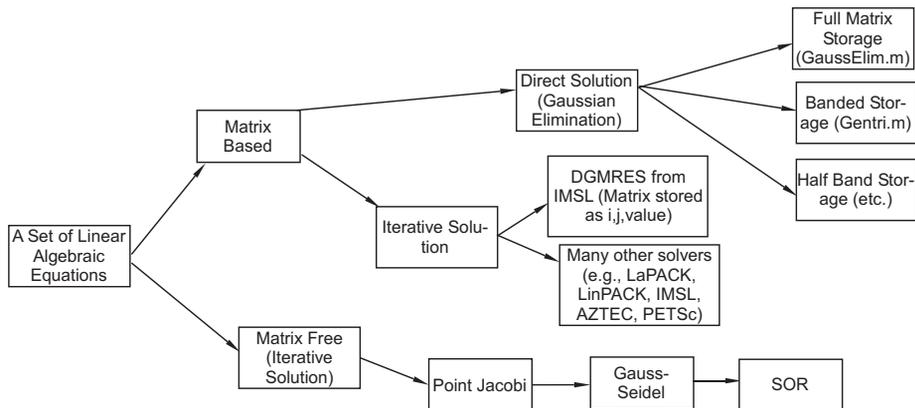


Figure 2.4: Commonly used linear solution techniques in hydrogeology and their relations.

matrix-based direct solver (the various forms of Gaussian Elimination solvers, tailed to the specific storage styles of the global coefficient matrix, more on this later).

2.3 Validity of Numerical Solution

In solving simple examples, we can check the validity of the numerical solution by comparing the numbers generated using the numerical solution with those calculated from an analytical solution. However, analytical solutions are not available for many problems of practical interest. Numerical methods allow us to solve the governing equation in more than one dimension, for complex boundary conditions and for heterogeneous and anisotropic aquifers, whereas most analytical solutions are restricted to consideration of homogeneous, isotropic aquifers.

An important point which is sometimes overlooked is the necessity to verify the validity of every numerical solution. Therefore, several checks on the correctness of the solution should be made routinely when an analytical solution is not available. Specifically, it is a good practice to check the sensitivity of the solution to (1) the choice of error tolerance when matrix-free numerical techniques are used to solve the governing equations, e.g., Point-Jacobi, Gauss-Seidel and SOR (these are most suitable for simple domain geometry and small to medium-sized grids); (2) nodal spacing. Note that both matrix-free and matrix-based ⁴ numerical techniques are based on a grid. So this sensitivity test applies to both groups. Moreover, it is desirable to perform a mass balance calculation.

⁴In the matrix-based numerical techniques, the coefficient matrix and right hand side vector are assembled; their values sent to a direct Gaussian-elimination-based solver or an iterative solver.

If a matrix-free technique is employed to solve the flow equation, as a general rule, several computer runs should be made using successively smaller values for the error tolerance until a small enough value is selected and the solution does not change with the further decrease in error tolerance. In all numerical techniques, the solution may also be improved through the use of smaller nodal spacing. Several computer runs should be made with successively smaller nodal spacing until the solution does not change with further decrease in nodal spacing. However, it is generally a time-consuming process to modify the finite difference grid and redefine the input. Instead, most modelers rely on the third type of validity check—the mass balance calculation. A mass balance calculation is an expression of the fact that **at steady state, the amount of water entering the system equals the amount leaving the system**. If inflow does not equal outflow for a steady-state computer simulation, there may be something wrong with the numerical solution. However, a large error in the mass balance may also indicate a lack of precision in the formula of mass balance calculation itself.

2.4 Model Calibration and Parameter Estimation

Model calibration is the process of adjusting the input parameters (hydraulic conductivity, porosity, specific storage, well or natural recharge/discharge rate) and boundary conditions (head or Darcy flux along sections of the boundary) of a model to achieve a close fit to observed data (e.g., hydraulic head, flow rate) in a real groundwater system. In flow model calibration, simulated heads and Darcy fluxes are typically compared to their observed counterparts. If a model is well calibrated, there will be some *random* deviations between simulated and observed data, but there will not be *systematic* deviations. If there are systematic deviations such as if most of the simulated heads exceed the observed heads, the calibration is poor and adjustments should be made.

In thinking about how to calibrate a flow model, it helps to go back to the basics — the Darcy’s law (here it is written in its 1D form along the x axis):

$$q_x = -K_x \frac{\partial h}{\partial x}$$

If the head gradients predicted by a model are too large, the relations in Darcy’s law indicate that either the modeled fluxes are too large, or the modeled conductivities are too low, or both.

Consider an example where the simulated heads in a regional flow model are systematically higher than the observed heads, as shown in Figure 2.6 (Fitts, 2002). The aquifer is unconfined (so its heads correspond to the water table), and water enters the aquifer as recharge and leaves as discharge to local streams. Assuming that the constant heads assigned at the streams are correct, there are two adjustments to the model that could be made to eliminate the systematic error in the heads:

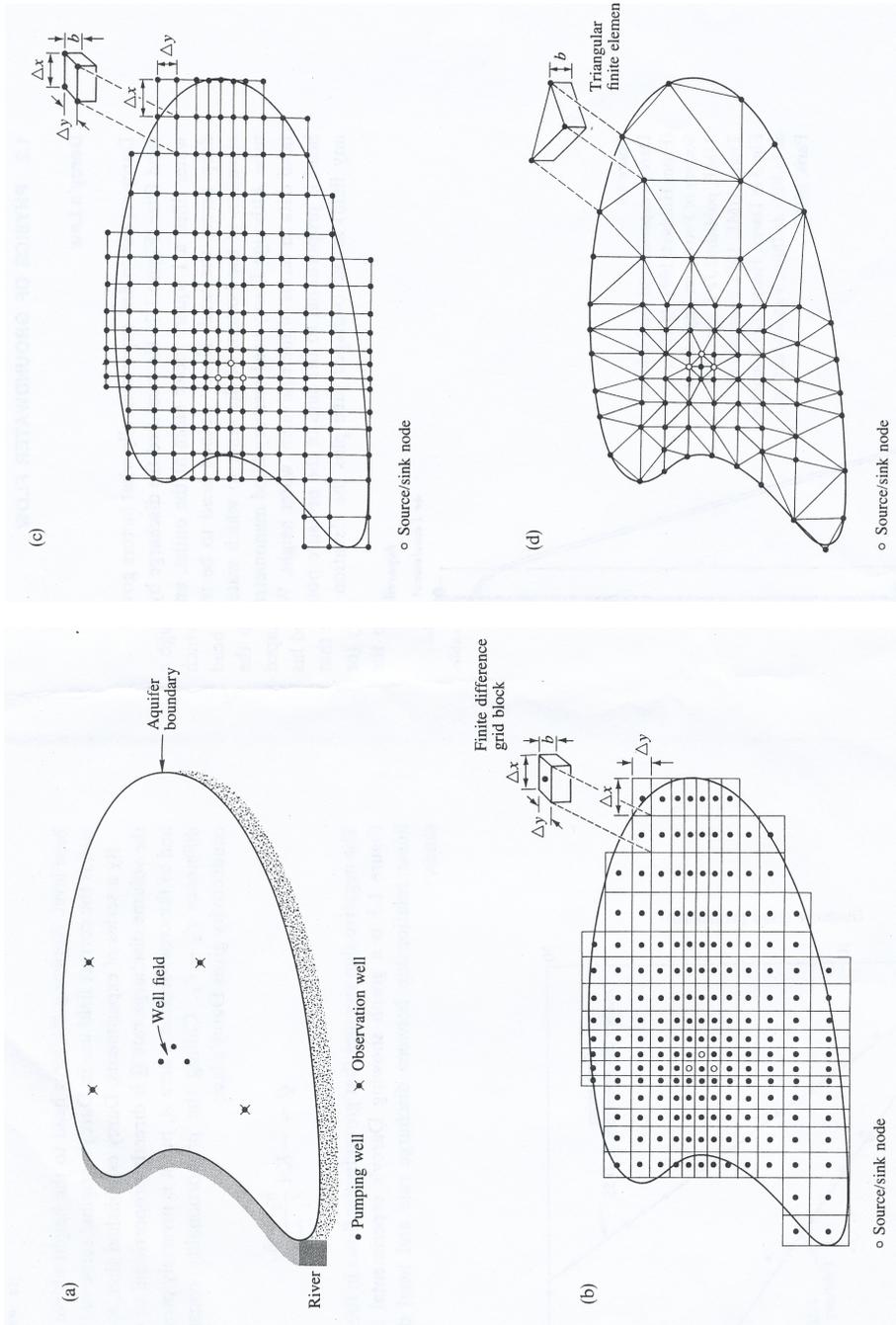


Figure 2.5: Finite difference and finite element representations of an aquifer region (from Wang & Anderson, 1995). (a) Map view of aquifer showing a well field, observation wells, and aquifer boundary. (b) Finite difference grid with block-centered nodes, where Δx is the spacing in x direction, Δy is the spacing in y direction, and b is the aquifer thickness at a particular block (this thickness can vary in space). (c) Finite difference grid with mesh-centered nodes. (d) Finite element mesh with triangular elements.

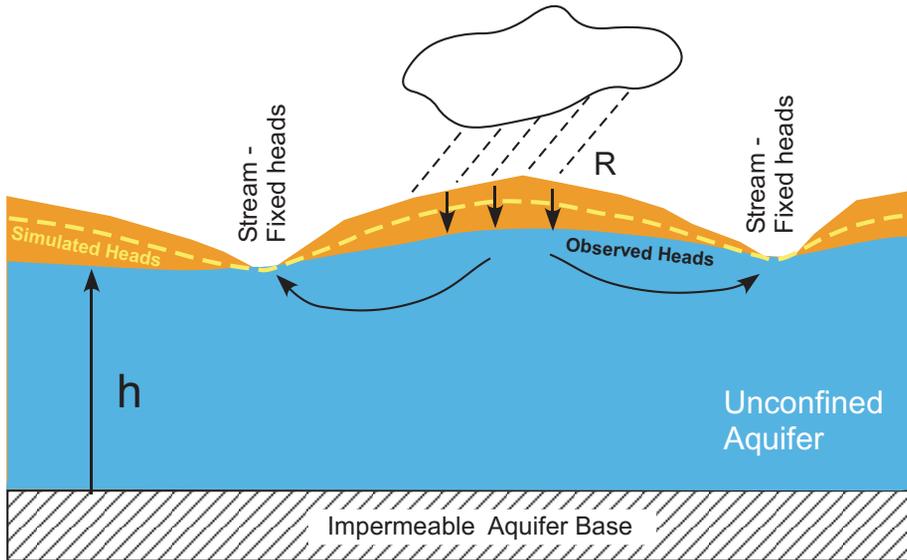


Figure 2.6: Vertical cross section comparing modeled heads with observed heads.

1. Increase hydraulic conductivities in the aquifer, or
2. Decrease the rate of recharge applied.

If the fluxes to streams are known to be correct (the model predicted fluxes to streams are equal to the observed fluxes within a small random error), then the conductivities should be increased. Why this is so can be seen from Darcy's law; increasing K allows the same flux of water to be transmitted with smaller head gradients. Thus, if we increase the value of K , to provide the same (correct) Darcy flux to the streams, the head gradient will drop, thus the simulated head will be lower than the current estimation and the simulated head gradient will become smaller.

On the other hand, if the conductivities specified to the model aquifer are known to be correct (for example, we have lots of detailed measurements of K in the aquifer to back up this claim) and the predicted fluxes to streams are too high compared to observed fluxes, the recharge rate should be lowered. Note that smaller recharge rate specified to the model will allow lesser amount of water to enter the aquifer for a fixed time period, so the predicted head mound in between the streams will be lower.

The above calibration procedure involves a trial-and-error approach by the modeler involving both the hard hydrogeological principles (relationship between head gradient, fluxes and conductivity) and soft intuition (which parameter might be the most uncertain thus requiring extensive adjustment). However, a recent development in flow modeling is the automated estimation of parameters by special computer algorithms that will optimize the calibration of models

(Hill, 1992; Doherty, 2000; Hill & Tiedeman, 2007). These techniques are based on minimizing an objective function which is defined to be a measure of the fit between model results and actual observations. The larger the computed objective function is, the greater the discrepancy between the model results and corresponding real observations. One simple and common definition of the objective function (F) is the sum of the squared residuals (differences between modeled and observed heads):

$$F = \sum_{i=1}^n (h_i^{obs} - h_i^{sim})^2$$

where n is the number of observed heads, h_i^{obs} is the i th observed head and h_i^{sim} is the modeled head corresponding to the i th observed head. The above equation can also be modified by multiplying weighting factors times each term in the sum. The weighting factor depends on the importance of an observation to the overall model validity and the confidence in that observation (“True Error” defined in Chapter 9). Automated calibration techniques will find the optimal set of parameter values that result in a minimal value of the objective function. Such techniques can save a modeler some time in the calibration process, but they are no substitute for careful thinking. The automated techniques can yield unreasonable results if insufficient constraints are supplied. Chapter 9 will discuss the methods of parameter estimation, based on both the forward upscaling theory and the inverse regression theory.

2.5 Interpreting Model Results

When interpreting flow model results, it is important to bear in mind the simplifying assumptions that went into creating the conceptual model. Although a well-calibrated mathematical model may simulate the *conceptual* model with great accuracy and aesthetic appeal, it is likely to be only a crude representation of the real flow system. In some cases, there is no unique solution to the calibration problem; several different combinations of input parameters will result in models that all equally fit the observations. Problems of non-uniqueness are more prevalent when the observation data are limited. For example, in a model where there are no known fluxes such as well rates or recharge rates, and the only known observations are heads, there will be a range of recharge and hydraulic conductivity combinations that can result in a similar calibration. As long as the hydraulic conductivities (or transmissivities) and fluxes are increased or decreased by the same proportion, the pattern of heads will remain the same. The general equation for steady two-dimensional flow with recharge is shown below to show why this is so: (please review GEOL4444/5444, Geohydrology)

$$\nabla^2 h = -N/T$$

where N is the recharge rate, T is aquifer transmissivity. It is clear from the above that the same pattern of h could be achieved with any number

of models that maintain a fixed ratio N/T . Therefore, for many modeling projects, it is important to have some flux measurements, e.g., well rates, aquifer recharge/discharge rates to lakes or streams, or precipitation-induced recharge rate to the aquifer.

After a model is calibrated to observed existing conditions in the flow system, it is then used to predict future conditions with a different set of boundary conditions. For example, models are often calibrated to the known flow conditions at a contaminated site prior to clean-up and then used to simulate the response of the system to various proposed remediation designs including the addition of pumping wells, barriers and/or drains. In the geological community, there are also scientific studies to evaluate the vulnerability of regional aquifers under future change in climate. For example, the Global Climate Models predict that under Global Warming, the western US may become dryer (Levi, 2008). This would mean less recharge water will reach the regional aquifers in this region which are sources of freshwater to many communities for both domestic, industrial and agricultural uses. Then, if the west were to become dryer, there will be less recharge water to the aquifer to balance the (possibly increasing) water demands due to the increased population. We need a computer model to predict the head decline under these hypothesized scenarios.

2.6 Model Limitation and Uncertainty

However, actual groundwater flow systems are much more complex than the conceptual models can typically represent. The accuracy of predictive simulations is thus difficult to assess, so it's wise to assume a fair amount of uncertainty when using models to make predictions. In particular, the subsurface has complex distributions of materials with transient groundwater fluxes. No matter how much effort is spent drilling, sampling, and testing the subsurface, only a small fraction of it is ever sampled or tested hydraulically. The available data will provide only an *incomplete* picture of the actual subsurface system. Because of the inherent difficulty of characterizing subsurface regions, substantial uncertainty is always introduced in the conceptual model created.

Typically, the complex distribution of subsurface materials is represented in the conceptual model as regions with locally homogeneous and isotropic hydraulic conductivity (anisotropic representation is also common, see Zhang et al (2006), Zhang & Gable (2008)). The parameters assigned to these regions are chosen to represent the large-scale average (or *effective*) hydraulic behaviors of flow and transport. Complex transient fluxes like recharge or pumping rates are represented in the conceptual model as either steady-state average values or as transient rates that change in some simplified manner.

In a well-constructed mathematical model, most of the uncertainty in the results stems from discrepancies between the real system and the conceptual system. Most mathematical models provide a fairly accurate simulation of the *conceptual* system. Therein lies the danger. Accurate simulation of the conceptual system is often taken to mean accurate simulation of the real system.

With sophisticated model-generated graphics, there is a tendency to forget the unavoidable uncertainties in representing the real system with a simpler conceptual model. Regardless of these limitations, models are usually the best way to develop judgment when solving quantitative groundwater flow problems.

2.7 References

Toth J, 1962, A theory of groundwater motion in small drainage basins in Center Alberta, Canada, Journal of Geophysical Research, Vol. 67, p. 4375-4387.

Lee S H, L J Durlofsky, M F Lough and W H Chen (1998) Finite difference simulation of geologically complex reservoirs with tensor permeabilities, SPE Reservoir Evaluation and Engineering, December, p. 567-574.

Ertekin T, J H Abou-Kassem, G R King (2001) Basic Applied Reservoir Simulation, SPE Textbook Series, 406 p.

Hill M C (1992) A computer program (MODFLOWP) for estimating parameters of a transient, three-dimensional ground-water flow model using nonlinear regression, USGS Water Resources Investigation Report, 91-484.

Doherty J (2000) Manual for PEST2000, Watermark Numerical Computing, Brisbane, Australia.

Levi B G (2008) Trends in the hydrology of the western US bear the imprint of manmade climate change, Physics Today, p. 16-18, April 2008.

Ye Zhang, Carl W Gable, Mark Person, 2006, Equivalent Hydraulic Conductivity of an Experimental Stratigraphy - Implications for Basin-Scale Flow Simulations, Vol. 42, W05404, Water Resources Research, doi:10.1029/2005WR004720.

Ye Zhang, Carl W Gable, 2008, Two-Scale Modeling of Solute Transport in an Experimental Stratigraphy, Journal of Hydrology, Vol. 348, p. 395-411, doi: 10.1016/j.jhydrol.2007.10.017.

Chapter 3

Mathematical Notations, Definitions and Theorems

This chapter introduces the basic mathematical notations, definitions and theorems. It supplements the handout of *Basic Math Review* given earlier but focuses more specifically on the type of problems we'll encounter in this course.

3.1 Differential Equations

Differential Equations The mathematical formulations of most flow and transport problems involve the rates of change of a dependent variable (e.g., hydraulic head, temperature, or concentration) with respect to one or more independent variables (e.g., time, distance, angle). A mathematical equation that contains a dependent variable and one or more of its derivatives is referred to as a Differential Equation. The *order* of a differential equation is defined by its highest derivative.

Ordinary Differential Equation (ODE): a type of differential equation that contains only 1 independent variable (e.g., time or one spatial axis). For example, a first-order ODE is:

$$\frac{dc}{dt} = -kc$$

The solution is $c = c(t)$, the one independent variable is time t . Note that for ODE, the differential symbol used is d . By definition, the ordinary derivative: $dc/dt = \lim_{\Delta t \rightarrow 0} [c(t + \Delta t) - c(t)]/\Delta t$.

Partial Differential Equation (PDE): a type of differential equation that contains more than 1 independent variables (e.g., several spatial axes, time and several spatial axes). For example, a second-order PDE is:

$$(\rho c) \frac{\partial T}{\partial t} = \lambda \frac{\partial^2 T}{\partial z^2}$$

The solution is $T = T(t, z)$, the two independent variables are time t and the z axis. Note that for PDE, the differential symbol used is ∂ (partial derivative symbol). By definition, the partial derivative: $\partial T / \partial t = \lim_{\Delta t \rightarrow 0} [T(t + \Delta t, z) - T(t, z)] / \Delta t$.

Initial and Boundary Conditions:

- Initial Condition: Constrains that are specified at the initial point, generally *time* point, are called initial conditions. Problems with specified initial conditions are called initial value problems.
- Boundary Condition: Constrains that are specified at the boundary points, generally *space* points, are called boundary conditions. Problems with specified boundary conditions are called boundary value problems.

ODE: Separation of Variables:

Given a first-order ordinary differential equation

$$(3.1) \quad \frac{dy}{dx} = f(x, y)$$

If $f(x, y)$ can be expressed using separation of variables as

$$f(x, y) = p(x)q(y)$$

then the equation can be expressed as

$$\frac{dy}{q(y)} = p(x)dx$$

and equation (3.1) can be solved by integrating both sides to obtain:

$$\int \frac{dy}{q(y)} = \int p(x)dx$$

Any first-order ODE of the form

$$(3.2) \quad \frac{dy}{dx} + p(x)y = q(x)$$

has a solution (details given in class):

$$(3.3) \quad y(x) = \frac{\int e^{\int p(x)dx} q(x)dx + C}{e^{\int p(x)dx}}$$

where C is a constant of integration.

3.2 Scalar, Vector, Tensor, Field

Scalar: a numerical quantity which is defined by its magnitude alone. Examples include concentration, temperature, pressure, hydraulic head.

Vector: a numerical quantity having both magnitude and direction. Examples include the Darcy flux vector, the average linear velocity vector. In 3D, the Darcy flux has 3 components, e.g., qx is the normal projection of \vec{q} along the x axis.

$$\vec{q} = \begin{pmatrix} qx \\ qy \\ qz \end{pmatrix}$$

Magnitude of the vector above is:

$$|\vec{q}| = \sqrt{qx^2 + qy^2 + qz^2}$$

Tensor: a matrix quantity which relates one vector to another. One example is the hydraulic conductivity tensor (\mathbf{K}) which relates the hydraulic head gradient (\vec{I}) with Darcy flux (\vec{q}): $\vec{q} = -\mathbf{K}\vec{I}$, or written out in full:

$$\begin{pmatrix} qx \\ qy \\ qz \end{pmatrix} = - \begin{bmatrix} K_{xx} & K_{xy} & K_{xz} \\ K_{yx} & K_{yy} & K_{yz} \\ K_{zx} & K_{zy} & K_{zz} \end{bmatrix} \begin{pmatrix} \partial h / \partial x \\ \partial h / \partial y \\ \partial h / \partial z \end{pmatrix}$$

This equation is also known as the general 3D Darcy's law (where the coordinate axes do not aligned with the principle components of the hydraulic conductivity tensor).

The above equation can also be written out explicitly using the rule of matrix-vector multiplication, one for each Darcy flux component,

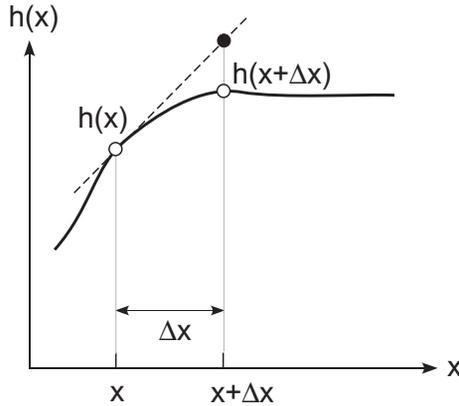
$$qx = -K_{xx} \frac{\partial h}{\partial x} - K_{xy} \frac{\partial h}{\partial y} - K_{xz} \frac{\partial h}{\partial z}$$

Field

A *scalar* field is created by simply assigning scalar quantities (numbers) to each point in space. Think of temperature – each point in the room has a different temperature.

A *vector* field is created by assigning vectors to each point in space. A flow field is an example – groundwater at each point in space has a certain speed and direction as represented by the Darcy flux vector at that point.

A *tensor* field has a tensor corresponding to each point in space. An example is the hydraulic conductivity of an aquifer: at each point in space there is a \mathbf{K} that relates the head gradient with Darcy flux at that point.

Figure 3.1: 1D Taylor series expansion along the x axis.

3.3 Taylor Series and Finite Difference Approximations

Taylor Series Expansion: approximation of a function h at a point $h(x + \Delta x)$ using derivatives of the function at $h(x)$ (Figure 3.1):

- Forward Approximation:

$$h(x + \Delta x) = h(x) + \frac{\Delta x}{1!} \frac{\partial h(x)}{\partial x} + \frac{\Delta x^2}{2!} \frac{\partial^2 h(x)}{\partial x^2} + \frac{\Delta x^3}{3!} \frac{\partial^3 h(x)}{\partial x^3} + \dots$$

If we truncate the higher order terms (i.e., $\frac{\Delta x^n}{n!} \frac{\partial^n h(x)}{\partial x^n}$, $n \geq 2$), we get:

$$h(x + \Delta x) \approx h(x) + \Delta x \frac{\partial h(x)}{\partial x} \Rightarrow$$

$$(3.4) \quad \frac{\partial h(x)}{\partial x} \approx \frac{h(x + \Delta x) - h(x)}{\Delta x}$$

- Backward Approximation:

$$h(x - \Delta x) = h(x) - \frac{\Delta x}{1!} \frac{\partial h(x)}{\partial x} + \frac{\Delta x^2}{2!} \frac{\partial^2 h(x)}{\partial x^2} - \frac{\Delta x^3}{3!} \frac{\partial^3 h(x)}{\partial x^3} + \dots$$

If we truncate the higher order terms ($n = 2$), we get:

$$h(x - \Delta x) \approx h(x) - \Delta x \frac{\partial h(x)}{\partial x} \Rightarrow$$

$$(3.5) \quad \frac{\partial h(x)}{\partial x} \approx \frac{h(x) - h(x - \Delta x)}{\Delta x}$$

If the function contains more than one independent variables, we will need to fix one independent variable and evaluate the function in relation to the other independent variable. For example, $h = h(x, t)$, we can write the forward approximation as:

- Forward Approximation:

$$h(x+\Delta x, t) = h(x, t) + \frac{\Delta x}{1!} \frac{\partial h(x, t)}{\partial x} + \frac{\Delta x^2}{2!} \frac{\partial^2 h(x, t)}{\partial x^2} + \frac{\Delta x^3}{3!} \frac{\partial^3 h(x, t)}{\partial x^3} + \dots \Big|_{fixt}$$

If we truncate the higher order terms (i.e., $\frac{\Delta x^n}{n!} \frac{\partial^n h(x, t)}{\partial x^n}$, $n \geq 2$), we get:

$$h(x + \Delta x, t) \approx h(x, t) + \Delta x \frac{\partial h(x, t)}{\partial x} \Rightarrow$$

$$(3.6) \quad \frac{\partial h(x, t)}{\partial x} \approx \frac{h(x + \Delta x, t) - h(x, t)}{\Delta x} \Big|_{fixt}$$

The Taylor Series is key to understanding the finite difference method. We'll see that by truncating the higher order terms in the Taylor Series, we can approximate any ordinary or partial derivative in a ODE or PDE.

Integration by Parts (1D)¹:

$$\int_a^b u dv = uv \Big|_a^b - \int_a^b v du$$

Exercise 1 Use integration by parts, evaluate $\int_a^b x e^x dx$.

Exercise 2 Evaluate the Taylor Series expansion for $h(x) = 3x + 2 + 9x^2$ at $x = 1$ for truncation error of order Δx^2 and Δx^3 , with $\Delta x = 1$. Hint: $x_0 = 0$, use Taylor Series expansion for $h(x_0 + \Delta x) = h(0 + 1) = h(1)$. Repeat the analysis for $\Delta x = 0.1$: now h is evaluated at $x = 0 + 0.1 = 0.1$.

3.4 Numerical Error, Convergence, Stability

Numerical Error Two types of errors occur in computational science:

- **Truncation Error**— due to truncation of higher order terms in the Taylor Series when approximating a derivative. In the previous forward or backward approximation, the truncation error $\rightarrow 0$ when $\Delta x \rightarrow 0$. For the above approximation where $\frac{\Delta x^n}{n!} \frac{\partial^n h(x)}{\partial x^n}$ ($n \geq 2$) terms are truncated, we say that the truncation error is of order Δx^2 .

¹ $d(uv) = udv + vdu \rightarrow \int_a^b d(uv) = \int_a^b udv + \int_a^b vdu \rightarrow uv \Big|_a^b = \int_a^b udv + \int_a^b vdu$

- **Roundoff Error** — due to the finite arithmetics of conducting mathematical operations using a computer. For example, when computing $x = 1/2 = 0.5$, the memory of the computer where it is designated to store a “real” number is 4 bytes (=4x8=32 bits). This size is sufficient to hold the value 0.5 without loss of accuracy. However, when computing $y = 1/3 = 0.333\dots$, the memory of the computer where it is designated to store a “real” number is also 4 bytes. This size is insufficient to hold the value $0.33\dots$: instead 0.33333333 is stored, so all digits after the 8th get rounded off. Thus, when we do an operation like $z = y \times 3$, we get $z = 0.99999999$ instead of the mathematically correct value of 1.0. This type of error is called roundoff error and can either cancel out or accumulate during numerical operations done by a computer.

The above numerical errors influence whether a numerical approximation of a differential equation converges and is stable:

- **Convergence** — a numerical approximation is said to converge if the numerical solution approaches the analytical solution as Δx approaches zero.
- **Stability** (Only applicable to initial-value problems) — if the numerical error, $e(\mathbf{x}) = h(\mathbf{x}) - \widehat{h}(\mathbf{x})$; $h(\mathbf{x})$ is the exact analytical value of h at \mathbf{x} , $\widehat{h}(\mathbf{x})$ is the numerical approximation of $h(\mathbf{x})$ at \mathbf{x} (\mathbf{x} is a position vector), does not grow through sequential calculations of the solution through time, then the numerical approximation is said to be stable. If the error grows without bound, then the numerical approximation is unstable.

3.5 Matrix & Vector

Consider the following matrix A consisting of n rows and m columns:

$$A_{n \times m} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}$$

Each entry in the matrix is called an element.

If $n = m$, the matrix of $n \times n$ is considered a square matrix:

$$A_{n \times n} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

This course is mostly concerned with the solution of square matrices.

A square matrix is symmetric if:

$$a_{ij} = a_{ji} \quad (i = 1, \dots, n; j = 1, \dots, n)$$

Note that the diagonal elements always satisfy $a_{ii} = a_{ii}$ when $j = i$, so we only look at the off-diagonal elements. For example, the matrix A_1 is symmetric:

$$A_1 = \begin{bmatrix} 10.0 & 2.0 & 1.0 \\ 2.0 & 23.0 & 5.0 \\ 1.0 & 5.0 & 1.5 \end{bmatrix}$$

A matrix is diagonally dominant if:

$$a_{ii} \geq |a_{ij}| \quad (i = 1, \dots, n; j \neq i)$$

This means, across any row in the matrix, the diagonal element is positive and greater than the *magnitude* of the off-diagonal elements. For example, the matrix A_2 is diagonally dominant:

$$A_2 = \begin{bmatrix} 10.0 & 0.0 & 1.0 \\ -5.0 & 23.0 & 20.0 \\ -0.1 & 0.0 & 1.5 \end{bmatrix}$$

In A_2 , for example, across the first row ($i = 1$), $a_{11} > |a_{12}|$, $a_{11} > |a_{13}|$. In the second row ($i = 2$), $a_{22} > |a_{21}|$, $a_{22} > |a_{23}|$, etc.

A sparse matrix is one where *most* off-diagonal elements are equal to zero (the diagonal elements are non-zeros), for example, the matrix A_3 is considered sparse:

$$A_3 = \begin{bmatrix} 10.0 & 0.0 & 0.0 & 0.0 \\ 3.0 & 23.0 & 0.0 & 0.0 \\ 0.0 & 2.0 & 1.5 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 \end{bmatrix}$$

A special type of sparse matrix is the Banded Sparse matrix which arises in finite difference (or finite element) discretization of PDEs describing groundwater flow and transport. Such banded matrices contain zeros except along the diagonal where there are a band of elements, e.g., a 8×8 banded sparse matrix can look like this:

$$\begin{bmatrix} a_{11} & a_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} & a_{34} & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{43} & a_{44} & a_{45} & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{54} & a_{55} & a_{56} & 0 & 0 \\ 0 & 0 & 0 & 0 & a_{65} & a_{66} & a_{67} & 0 \\ 0 & 0 & 0 & 0 & 0 & a_{76} & a_{77} & a_{78} \\ 0 & 0 & 0 & 0 & 0 & 0 & a_{87} & a_{88} \end{bmatrix}$$

The bandwidth m refers to the *maximum* number of non-zero columns occupied across a single row of elements. In the above example, the bandwidth $m = 3$.

As we will show, the discretization of the 1D flow (or transport) equation gives rise to a matrix with a bandwidth of 3. Such a matrix is also called a

tridiagonal matrix. It is the left-hand-side coefficient matrix of a linear set of equations — results of turning the PDE into a set of discrete algebraic equations (more on this later). In solving 2D to 3D problems, the coefficient matrices are still Banded Sparse, but they have higher bandwidth.²

The determinant of a square matrix is defined as (only square matrices can have determinants):

$$\begin{aligned} \text{Det} \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix} &= a_{11} \text{Det} \begin{bmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix} \\ &- a_{12} \text{Det} \begin{bmatrix} a_{21} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n3} & \dots & a_{nn} \end{bmatrix} + \dots \\ &\pm a_{1n} \text{Det} \begin{bmatrix} a_{21} & a_{22} & \dots & a_{2,n-1} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{n,n-1} \end{bmatrix} \end{aligned}$$

For example, $\text{Det}[a_{11}] = a_{11}$, $\text{Det} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}$. If the determinant of a matrix is 0, the matrix is said to be singular.

A square matrix is positive definite if the determinants of itself and all its submatrices are positive:

$$\begin{aligned} &\text{Det}[a_{11}] > 0 \\ &\text{Det} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} > 0 \\ &\text{Det} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} > 0 \\ &\dots \\ &\text{Det} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} > 0 \end{aligned}$$

²The characteristics of the matrix and bandwidth varies for groundwater flow and solute transport problems. In solving the groundwater flow equation, a symmetric coefficient matrix is produced, thus in a computer, commonly only half of the bandwidth is stored (this is a practice common in the early days of numerical simulation where a computer has very limited memory, e.g., 64KB). In using older codes, we'll still see the definition of a reduced bandwidth: $m_r = m/2$, which defines the half band of a matrix. The same situation occurs in solving the equation describing solute diffusion. However, whenever solving the typical solute transport problem described by advection, dispersion, and diffusion, the coefficient matrix can be non-symmetrical and sometimes off-diagonally dominated. Special solution technique are needed to deal with such matrices, e.g., iterative solutions that are not restricted to any special matrix characteristics, e.g., the IMSL DGMRES works for any general sparse non-symmetric matrix.

Diagonally dominant positive definite square matrices have properties that make them amenable to special solutions if they constitute the left-hand-side coefficient matrix of a linear system of equations.

As shown in the handout in the beginning of the class, there is an important relation between matrix and vector product, e.g., we can write the vector outcome for a multiplication of a 3×3 matrix and 3×1 vector:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{Bmatrix} b_1 \\ b_2 \\ b_3 \end{Bmatrix} = \begin{Bmatrix} a_{11}b_1 + a_{12}b_2 + a_{13}b_3 \\ a_{21}b_1 + a_{22}b_2 + a_{23}b_3 \\ a_{31}b_1 + a_{32}b_2 + a_{33}b_3 \end{Bmatrix}$$

The resulting vector has 3 components or a 3×1 vector. This can be extended for the multiplication of $n \times n$ matrix and $n \times 1$ vector: the resulting vector is $n \times 1$.

3.6 Linear Algebra and Solution Techniques

The previous properties of the matrix and matrix-vector product is the cornerstone of linear algebra. The mathematics of linear algebra is broad and continuously expanding. To study it systematically requires a separate class which we can ill afford. So, we select the important concepts and solution procedures from linear algebra which are widely used in modeling hydrogeological problems.

What is a linear algebra problem?

A linear algebra problem typically consists of a set of n linear equations with n unknowns, e.g.,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ a_{31}x_1 + a_{32}x_2 + \dots + a_{3n}x_n = b_3 \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases}$$

This equation can be written alternatively using matrix vector format as:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{Bmatrix} = \begin{Bmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{Bmatrix}$$

To write the above equation in a more condensed fashion, we use:

$$A\vec{x} = \vec{b}$$

where the left-hand-side (LHS) matrix $A_{n \times n}$ is the coefficient matrix of the set of linear equations, the solution vector $\vec{x} = \{x_1, x_2, \dots, x_n\}^T$ contains the unknown

values.³ The right-hand-side (RHS) vector \vec{b} contains the known values.

The key in solving a set of linear equations is thus to find the solution vector \vec{x} . In general, two broad solution techniques exist:⁴

- Direct Approach (Gaussian Elimination Schemes)
- Indirect Approach (Various Iterative Methods)

3.6.1 Direct Approaches

The direct approach have many variants. The classic Gaussian Elimination is based on a two-step procedure:

1. Triangulation of Matrix: Eliminate all terms that occur below the diagonal elements (i.e., $a_{11}, a_{22}, \dots, a_{nn}$). For example, after triangulation of a 3×3 matrix, we get:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} = \begin{Bmatrix} b_1 \\ b_2 \\ b_3 \end{Bmatrix}$$

Note that compared to the original coefficient matrix (e.g., it can be a full matrix), the new system will have different coefficient matrix and RHS vector.

2. Back Substitution: After triangulation, the bottom equation (see above) has only one unknown x_n (i.e., x_3). This can be solved for and the value of x_n substituted into the equation *immediately above it* to find x_{n-1} (i.e., x_2). This process is repeated until we solve for the first equation to find x_1 .

Exercise 3 Consider the following system of linear equations — 3 equations to solve for 3 unknowns. Use the two-step Gaussian Elimination to solve for the unknowns.

$$\begin{cases} x_1 + 2x_2 - x_3 = 3 & (1) \\ 4x_1 + x_2 + 6x_3 = 10 & (2) \\ 2x_1 - 3x_2 + 6x_3 = 1 & (3) \end{cases}$$

Exercise 4 Consider the following system of linear equations — 3 equations to solve for 3 unknowns. Use the two-step Gaussian Elimination to solve for

³In this class, a vector is considered a column vector: a vector of size n is thus of $n \times 1$. To allow easy writing, we write the column vector as the transpose of a row vector, as shown above.

⁴In hydrologically modeling, typically the Direct Approach is efficient in solving smaller problems, e.g., 1D or 2D grid with less than 10,000 nodes. The Indirect Approach is commonly used to solve much larger problems, e.g., 3D grid with more than 10,000 nodes.

the unknowns.

$$\begin{cases} x_1 + x_2 + x_3 = 1 & (1) \\ x_1 + x_2 + 2x_3 = 2 & (2) \\ x_1 + 2x_2 + 2x_3 = 3 & (3) \end{cases}$$

Note that in **Exercise 4**, the Gaussian Elimination is interrupted when we found $a_{22} = 0$ after the first elimination. The procedure to deal with this is to exchange this zero-diagonal equation to the one below, after which triangulation continues to arrive at the final set which is then solved via back substitution. This exchange (to allow triangulation) is also called pivoting.

3.6.2 Iterative Approaches

Iterative approaches also attempt to solve the system of linear equations, though the methodology is quite different. Consider the following example of a 4×4 system of equations:

$$\begin{cases} a_{11} \underbrace{x_1} + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 = b_1 \\ a_{21}x_1 + a_{22} \underbrace{x_2} + a_{23}x_3 + a_{24}x_4 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33} \underbrace{x_3} + a_{34}x_4 = b_3 \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + a_{44} \underbrace{x_4} = b_4 \end{cases}$$

We can make an initial guess for *all* of the unknowns, i.e., for the 4×4 system, they are $\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_3, \widetilde{x}_4$, and re-arrange the equations: in sequence, we solve for the diagonal terms with the underbraces $\underbrace{x_i}$:

$$(3.7) \quad \begin{cases} x_1 = \frac{1}{a_{11}} [b_1 - a_{12}\widetilde{x}_2 - a_{13}\widetilde{x}_3 - a_{14}\widetilde{x}_4] \\ x_2 = \frac{1}{a_{22}} [b_2 - a_{21}\widetilde{x}_1 - a_{23}\widetilde{x}_3 - a_{24}\widetilde{x}_4] \\ x_3 = \frac{1}{a_{33}} [b_3 - a_{31}\widetilde{x}_1 - a_{32}\widetilde{x}_2 - a_{34}\widetilde{x}_4] \\ x_4 = \frac{1}{a_{44}} [b_4 - a_{41}\widetilde{x}_1 - a_{42}\widetilde{x}_2 - a_{43}\widetilde{x}_3] \end{cases}$$

After a set of values are obtained using the above equations (x_1, x_2, x_3, x_4) , they become the new “guesses” $(\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_3, \widetilde{x}_4)$, and are substituted back into the same above equation to find the next iteration values. We repeat this recursive process until there is almost no difference between the successive approximation of the unknown vector \vec{x} . That is, another round of iteration brings a negligible change in the values of $\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_3, \widetilde{x}_4$ compared to the previous round of values. The final converged values are considered the solution set for the linear equations.

Note that *before* implementing the iterative method, the diagonal elements must be all non-zero. This can be accomplished via pivoting. We have n equations for n unknowns (x_1, x_2, \dots, x_n) , thus the coefficient for each unknown (a_{ij}) must be non-zero at least once. Therefore we can always shift the order of the equations to make the *diagonal* elements non-zero.

However, the recursive relation can be fine-tuned. There are several types of Iterative Methods which differ in implementation details:

1. Point Jacobi
2. Gauss-Seidel
3. Successive Over-Relaxation

Point Jacobi is exactly what is described by equation (3.7) which can be written concisely by the following recursive relationship:

$$(3.8) \quad x_i^{m+1} = \frac{1}{a_{ii}} [b_i - \sum_{j=1}^n a_{ij} x_j^m] \quad j \neq i$$

where

x^m — guess at beginning of an iteration step;

x^{m+1} — updated value at the end of an iteration step.

We continue the recursion ($m \rightarrow \infty$) until some small tolerance (δ) has been achieved:

$$(3.9) \quad \max | \frac{h_i^{m+1} - h_i^m}{h_i^m} | \leq \delta$$

For example, equation (3.7) is just this condensed equation written out for $n = 4$. For $i = 1$, we write: ($j = 2, 3, 4$)

$$x_1^{m+1} = \frac{1}{a_{11}} [b_1 - a_{12}x_2^m - a_{13}x_3^m - a_{14}x_4^m]$$

For $i = 2$, we write: ($j = 1, 3, 4$)

$$x_2^{m+1} = \frac{1}{a_{22}} [b_2 - a_{21}x_1^m - a_{23}x_3^m - a_{24}x_4^m]$$

For $i = 3$, we write: ($j = 1, 2, 4$)

$$x_3^{m+1} = \frac{1}{a_{33}} [b_3 - a_{31}x_1^m - a_{32}x_2^m - a_{34}x_4^m]$$

For $i = 4$, we write: ($j = 1, 2, 3$)

$$x_4^{m+1} = \frac{1}{a_{44}} [b_4 - a_{41}x_1^m - a_{42}x_2^m - a_{43}x_3^m]$$

The above equations boil down to equation(3.7). Note that during a given recursive round ($m + 1$), only the results from the last round ($x_1^m, x_2^m, x_3^m, x_4^m$) are used.

At the end of each update, we found $x_1^{m+1}, x_2^{m+1}, x_3^{m+1}, x_4^{m+1}$. These values are compared to the previous results $x_1^m, x_2^m, x_3^m, x_4^m$, for which we can find a *relative* difference for each unknown variable: $(x_1^{m+1} - x_1^m)/x_1^m, (x_2^{m+1} - x_2^m)/x_2^m, (x_3^{m+1} - x_3^m)/x_3^m, (x_4^{m+1} - x_4^m)/x_4^m$. From these 4 relative differences, we can find a maximum absolute value (the highest magnitude of relative difference

between iteration m and $m + 1$), if this maximum is still greater than δ , the recursion continues until equation (3.9) is satisfied.

Gauss-Seidel Iteration Method is the same as Point Jacobi Method except it uses the most recent updated value of x^{m+1} . It can be written concisely by the following recursive relationship:

$$x_i^{m+1} = \frac{1}{a_{ii}} \left[b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{m+1} - \sum_{j=i+1}^n a_{ij} x_j^m \right]$$

where

When $j = 1, \dots, i - 1$, x_j^{m+1} are the *updated* value found during the current ($m + 1$) iteration step;

When $j = i + 1 \dots, n$, x_j^m are the old values found during the last iteration (m) step.

We continue the recursion ($m \rightarrow \infty$) until the same small tolerance (δ) has been achieved (note for each unknown variable, it is comparing its updated value x^{m+1} with its previous old value x^m , just like that for the Point Jacobi Method):

$$\max_i \left| \frac{h_i^{m+1} - h_i^m}{h_i^m} \right| \leq \delta$$

For example, equation (3.7) can be rewritten for Gauss-Seidel Iteration. For $i = 1$, we write: ($j = 2, 3, 4$)

$$x_1^{m+1} = \frac{1}{a_{11}} [b_1 - a_{12}x_2^m - a_{13}x_3^m - a_{14}x_4^m]$$

For $i = 2$, we write ($j = 1, 3, 4$), now x_1^{m+1} is known:

$$x_2^{m+1} = \frac{1}{a_{22}} [b_2 - a_{21}x_1^{m+1} - a_{23}x_3^m - a_{24}x_4^m]$$

For $i = 3$, we write ($j = 1, 2, 4$), now x_1^{m+1} and x_2^{m+1} is known:

$$x_3^{m+1} = \frac{1}{a_{33}} [b_3 - a_{31}x_1^{m+1} - a_{32}x_2^{m+1} - a_{34}x_4^m]$$

For $i = 4$, we write ($j = 1, 2, 3$), now x_1^{m+1} , x_2^{m+1} and x_3^{m+1} is known:

$$x_4^{m+1} = \frac{1}{a_{44}} [b_4 - a_{41}x_1^{m+1} - a_{42}x_2^{m+1} - a_{43}x_3^{m+1}]$$

Note that in the Gauss-Seidel Method, during a given recursive round ($m + 1$), for x_1 , only results from the last round ($x_1^m, x_2^m, x_3^m, x_4^m$) are used; however, for x_2, x_3, \dots , increasingly, results from the current round are used.

Successive Over-Relaxation (SOR) — In this iterative method, we accelerate the step change in x_i from iteration m to $m + 1$. A two-step approach is used. In step (1):

$$c_i = x_i^{m+1} (\text{guess}) - x_i^m$$

x_i^{m+1} (*guess*) is the initial guess for x_i^{m+1} , estimated using the previous Gauss-Seidel Iteration approach. In step (2), we use:

$$x_i^{m+1} = x_i^m + wc_i$$

where w is the Relaxation Factor ($1 \leq w \leq 2$). Note that when $w = 1$, SOR collapses to Gauss-Seidel Iteration.

Combining step (1) and (2), we can also write a single SOR formulation:

$$x_i^{m+1} = x_i^m + w \left\{ \frac{1}{a_{ii}} \left[b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{m+1} - \sum_{j=i+1}^n a_{ij}x_j^m \right] - x_i^m \right\}$$

3.6.3 Comparison and Summary

1. Direct methods (Gaussian Elimination) provide the most reliable means of solving a set of linear algebra equations. However, the chief shortcoming of using direct methods is that it can be very slow in solving large matrix problems, compared to the speed with which the same problems can be solved by iterative methods. Typically, direct methods are used to solve for a problem with 10,000 or less unknowns. Since numerical approximations of equations of flow/transport are based on grids, this translates to a grid with 10,000 or less grid cells (FD block centered) or mesh points (FD mesh centered) or nodes (FEM).
2. Numerical round-off error occurs in both direct and iterative methods since the finite arithmetics of using a limited number of bytes to store real numbers is unavoidable as long as computers are used to solve the problem.
3. In the matrix-free iterative methods (Point-Jacobi, Gauss-Seidel, SOR), there is no matrix storage since the coefficient matrix does not need to be explicitly formed (this will become apparent when you work on the Second Program of Homework 2). In general, Gauss-Seidel converges twice as fast as the Point Jacobi method. So, for the same set of linear equations, using Gauss-Seidel iteration method to solve for the solution vector will need only half of the computation time as the Point Jacobi method. The convergence speed of the SOR is sensitive to the choice of w used. For a given set of equations, methods for selecting an optimum value of w can be found by trial-and-error or through theoretical means (see *Remson, Hornberger and Molz (1971), Numerical Methods in Subsurface Hydrology*).
4. Matrix-free iterative solver is nowadays rarely used, since they are much slower compared to matrix-based iterative methods (IMSL, SPLIB, LAPACK, LINPACK, etc). In the matrix-based iterative methods, matrix needs to be explicitly formed, and according to the requirement of each solver, specific format is required in assembling the global coefficient matrix. This can be time consuming at first but very useful in solving large problems. Though matrix-based iterative solver is not covered in this

class, you're welcome to talk to me after class where I can show you some examples of calling these solvers (mostly done on Unix clusters).

5. Compared to the Direct method, iterative methods can have trouble converging or can develop instability (solution vector fluctuates greatly from m to $m + 1$ without convergence as $m \rightarrow \infty$), for highly deformed grids or problems with large contrast in material properties. For example, in modeling a multilayered aquifer-aquitard system, if the contrast in hydraulic conductivity between aquifer and aquitard is large, iterative methods may have trouble converging. In this case, investigate the condition number of the matrix to see if it is approaching singular (e.g., calling a subroutine in IMSL which computes the matrix condition number). If so, consider improving the aspect ratio of the grid (so it is less deformed) and consider adding buffer zones between very-high-K and very-low-K zones. Moreover, in solving transient problems (i.e., hydraulic head, temperature (T), or velocity change with time), the convergence speed of iterative methods can be sensitive to the choice of initial conditions.⁵ On the other hand, direct methods seem generally less affected by these issues though there is no guarantee.
6. For large problems where we're not sure if iterative methods can provide sufficiently accurate solutions, the same problem should be solved first by a direct method (as long as the computer is sufficiently fast). By comparing the solutions using both approaches, the accuracy of the iterative method will be ascertained. Since the iterative method can generally solve bigger problems faster, it is still the best choice for large simulation models. However, if computer time is not excessive, verification with direct methods is a prudent procedure to take.
7. The process of ensuring that an iterative method is giving the correct numerical results (as compared to those predicted by the direct method for the same problem) is called code verification. The process of ensuring that an iterative method (or direct method) gives the results that are close to the analytical (exact) solutions (when they exist) is called code validation. However, for complicated problems with complex geometry, parameter heterogeneity, and spatially-varying and/or time-dependent boundary conditions, analytical solutions to such problems generally do not exist. In such cases, though numerical simulations are used to model such problems, analytical solutions usually do not exist, thus the numerical codes can be verified but not validated. In these cases, we conduct (1) grid resolution study (for transient problems both grid and temporal resolution can be changed); (2) local and global mass balance analysis (for transient problems, mass balance can be checked for every time step).

⁵For example, to solve the transient responses of a geotherm field to groundwater infiltration and discharge, if the initial T field is chosen to reflect the final expected field more closely, the iterative methods will converge to find the final solution faster than choosing an initial T field that is far from reality (e.g., giving T a uniform value of 0 degree C).

3.7 Homework 2

1. Solve the below equations by hand to find the unknowns (x_1, x_2, x_3) using Gaussian Elimination. Hint: pivoting may be needed.

$$\begin{cases} x_1 + x_2 + x_3 = 1.0 & (1) \\ x_1 + 1.0001x_2 + 2x_3 = 2.0 & (2) \\ x_1 + 2x_2 + 2x_3 = 1.0 & (3) \end{cases}$$

2. Solve the below equations first by hand using Gaussian Elimination. Then, write Matlab (alternatively you can write in Fortran or C) scripts to solve for the unknowns by the following iterative methods: (a) Point-Jacobi; (b) Gauss-Seidel; (c) SOR. Note that your earlier hand-calculated values (via Direct Gaussian Elimination) help you determine if your codes are giving you correct results. Due to numerical rounding errors, your numerical results (via each Iterative Method coded by you) should be extremely close to the hand-calculated values, but not necessarily the same.

For each iterative method, use a convergence tolerance $\delta=0.005$ and initial guesses of $\{0.0, 0.0, 0.0, 0.0\}$. Keep an account of how many iteration steps it will take for each method to reach convergence. Which of the three methods appear to converge the fastest? Which is the slowest? For SOR, try different w values to see how it affects the speed of convergence, i.e., how many steps it takes to converge. For example, you can try: $w = 1.1, 1.3, 1.5, 1.7, 1.9$. Which value is optimal (giving the least number of steps towards convergence)? If the above values do not determine an optimal value, you can further fine-tune w , e.g., try $w = 1.05$.

$$\begin{cases} 4x_1 - x_2 - x_3 = 1.0 & (1) \\ -x_1 + 4x_2 - x_4 = 2.0 & (2) \\ -x_1 + 4x_3 - x_4 = 0.0 & (3) \\ -x_2 - x_3 + 4x_4 = 1.0 & (4) \end{cases}$$

For these exercises, you must hand in your codes as well. For hand calculations, present the full steps.

(Optional) You can also find out what happens to the speed of convergence when you increase or decrease the tolerance.

3.8 Gauss-Jordan Direct Solver with Pivoting (Optional)

Solving linear algebra problems for which the matrix is not well-behaved (e.g., pivoting is required as one of the class hand-exercises illustrates) is normally not an issue with solving the groundwater flow and transport problems. Since after

3.8. GAUSS-JORDAN DIRECT SOLVER WITH PIVOTING (OPTIONAL)47

FD (or FE) discretization of the flow and transport equations, the matrix is generally well-behaved: for flow problem, the matrix is symmetric, positive-definite (diagonally dominant), and sparse; for transport, the matrix is symmetric and sparse, but can be off-diagonally dominant for advection-dominated problems (in which case special solution technique is implemented, e.g., Modified Method of Characteristics). In all these cases, we do not need to worry about switching the orders of the unknowns as required by pivoting.

However, sometimes linear matrix problems do arise for which pivoting is required, or the matrix is not symmetric, positive-definite, or possesses any special characteristics (for example, in solving the 1D flow equation, the coefficient matrix is always tridiagonal for which a special direct Gaussian-Elimination solver `Gentri.m` can be used which specifically works for this type of matrix). A case in mind is solving the kriging or co-kriging equations (please take my Geostatistics class to learn these methods) which have zero diagonal elements! Whenever such cases happen (you'll be surprised how often they do once you start doing research), a more general solver can be used which has no special requirements on matrix characteristics. A popular **direct solver** is the Gauss-Jordan method which is capable of automatic pivoting (parallel to our hand-exercises when our brain was doing the pivoting). For example, a full code solving a 5×5 problem is shown below. If we write out the $A\vec{x} = b$ and try to do this by hand, we soon realize that we have to do pivoting. However, this code calls the Gauss-Jordan method to do the pivoting and Gaussian-Elimination directly.

```
% Gaussian-Jordan method of solving linear equations with full
% pivoting: this method is capable of solving non-symmetrical,
% non-positive-definite, non-diagonal-dominated matrix. IN general,
% any linear set of equations that are NOT singular (If the matrix is
% singular, there exist infinite sets of solution vectors for any one
% solve of Ax=b. That is, x cannot be uniquely determined).
```

```
PARAMETER(n=5, np=5, m=1, mp=1)
REAL*8 a(np,np), b(np,mp)

INTEGER i,j,k
% OPEN(UNIT=8, FILE='debug_jordan.data')
OPEN(UNIT=9, FILE='output_jordan.data')

a(1,1)=1
a(1,2)=1
a(1,3)=1
a(1,4)=1
a(1,5)=0

a(2,1)=0
a(2,2)=95.75
```