

AN ANALYSIS OF THE STOCHASTIC
APPROACHES TO THE PROBLEMS OF FLOW
AND TRANSPORT IN POROUS MEDIA

by

DAVID W. DEAN

B. S., Illinois State University, 1967

M. S., University Of Illinois, 1969

A thesis submitted to the
University of Colorado at Denver in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Applied Mathematics

1997

This thesis for the Doctor of Philosophy

degree by

David W. Dean

has been approved

by

Thomas F. Russell

Leopoldo P. Franca

Karen Kafadar

Thomas A. Manteuffel

Steven F. McCormick

Date _____

Dean, David W. (Ph.D., Applied Mathematics)

An Analysis Of The Stochastic Approaches To The Problems
Of Flow And Transport In Porous Media

Thesis directed by Professor Thomas F. Russell

ABSTRACT

One need in the current theory of subsurface transport in porous media is an improved understanding of the basic transport physics in highly heterogeneous subsurface environments using models that are valid at multiple scales. The thesis addresses this problem by first developing a theoretical background for the spectral representation of stochastic processes which are then used to illustrate the more common aspects of the theoretical descriptions of dispersion. The analysis shows how the dispersion tensor in the homogeneous case must be modified in order to include mildly heterogeneous permeability fields and provides a transformation law for the conversion of the spectrum of velocity perturbations to the spectrum of log hydraulic conductivities. This theoretical connection is important because in Chapter II a Lagrangian approach is used to develop a description of dispersion in terms of the covariance of the hydraulic conductivities using a particle tracking algorithm. Chapter III describes the numerical methods used to implement the algorithm. Chapter IV treats the transport equation using stochastic calculus, specifically Itô's lemma, from which weak formulations of the mean and covariance equations can be derived. Chapter V considers the application of the theory of stochastic evolution equations to the problem of transport. By allowing both the dispersion and velocity to have random components, the evolution equation can be split into deterministic and stochastic parts. Using semigroup methods, the solution is given in terms of a Neumann expansion. Finally, Chapter VI uses the operator splitting method of Chapter V to illustrate a stochastic finite element method for solving the transport equation that uses the Karhunen-Loève expansion, the Galerkin method and the Homogeneous Chaos spaces of Wiener.

This abstract accurately represents the content of the candidate's thesis. I recommend its publication.

Signed _____
Thomas F. Russell

CONTENTS

Chapter	
1.	Introduction 1
1.1	Overview 1
1.2	Stochastic Processes 13
1.3	Stochastic Measures 16
1.4	Process With Orthogonal Increments 21
1.5	Spectral Representation 24
1.6	Space Correlations And Space Spectra 27
1.7	Ergodicity 29
1.8	Stochastic Solute Transport And Dispersion 31
1.9	Comments And Limitations 38
1.10	Velocity/Permeability Covariance Relationship 39
1.11	Summary 43
2.	Dispersivity Coefficients - Time And Distance Forms 45
2.1	Time Dependent Dispersivity Coefficients 45
2.1.1	Introduction 45
2.1.2	Transport Equation 46
2.1.3	Spatial Moments Of The Solute Concentration 46
2.2	Stochastic Differential Equations 53
2.2.1	Introduction 53
2.2.2	Integral Of A Stochastic Process 53
2.2.3	Wiener Process (Brownian Motion) 54
2.2.4	Stochastic Integration 55
2.2.5	Types Of Stochastic Integrals 56
2.3	The Concentration Equation 56
2.3.1	The Lagrangian Approach 57
2.3.2	Basic Form Of Transport Equation 61
2.3.3	Solution Of Basic Transport Equation 64
2.3.4	Dispersion As Velocity Covariances 69
2.3.5	Dagan's Approach 73
2.4	Distance Dependent Dispersivity Coefficients 74
2.4.1	Local Grid Block Dispersivity Coefficients 75
2.4.2	Global Dispersivity Coefficients 81
2.4.3	Molecular Diffusion 86
2.5	Random Variable Generation 86
2.5.1	Independent Random Variables 86

2.5.2	Correlated Random Variables	90
2.6	Summary	93
3.	Methods Of Solution	95
3.1	General	95
3.1.1	Simultaneous Solution	95
3.1.2	The Mixed Model	99
3.1.3	The Spaces \vec{Q}_h And V_h	102
3.2	2D Finite Element Solution	104
3.2.1	General	104
3.2.2	Finite Elements	104
3.2.3	Rectangular Elements	106
3.2.4	Numerical Integration	108
3.2.5	Groundwater Flow Equation	110
3.2.6	Matrix Assembly	116
3.2.7	Velocity Field	119
3.2.8	The Transport Equation	120
3.2.9	Imposition Of The Boundary Condition	122
3.2.10	The Boulder Experiments	124
3.3	Summary	128
4.	Moment Equations	131
4.1	Moments Derived From Distributed Parameters	132
4.2	An Itô Calculus Approach	138
4.2.1	System Definition	138
4.2.2	Types Of Problems	138
4.2.3	Existence Theory	139
4.2.4	Stochastic Integration	142
4.2.5	Itô's Lemma In Hilbert Space	147
4.2.6	Small o Notation	150
4.2.7	Hilbert Space Structures	151
4.2.8	Moment Equation Derivation	155
4.3	Summary	172
5.	Stochastic Evolution Equations	174
5.1	General Theoretical Foundations	174
5.2	Application To Transport And Scale-Up	185
5.3	Stochastic Parameters	187
5.4	Formal Solution	188
5.5	Convergence	203
5.6	Summary	214
6.	Future Research	215
6.1	General	215

6.2	Homogeneous Chaos	217
6.3	Stochastic Finite Elements	221
6.4	The Covariance Function	224
<u>References</u>	227

1. Introduction

1.1 Overview

This study is concerned with the basic forms of the equations of flow and transport of solutes or contaminants through porous media. Whenever two miscible fluids come together to form one phase, there is potential for either the density of the phase or the viscosity of the phase to change. This change can be brought about by changes in the concentration and/or changes in the pressure. In the case that the density is dependent on the concentration of the pollutant, a coupled system of PDE's is obtained that must be solved simultaneously, see Equation[3.3], page 99. Using Darcy's law and the Hubbert potential for a compressible fluid, the velocity can be related to the pressure via the permeability of the medium, see Equation[3.1], page 97. This equation can be solved for the velocity, pressure pair by the mixed finite element method.

In the case of constant density, the situation is less complicated. This is referred to as the *tracer* case. Here, the system can be written as an uncoupled system consisting of the flow equation, which can be solved for the piezometric head, and the transport equation. Once a velocity field has been derived from the piezometric head distribution, the transport equation is solved for the concentration plume that evolves over time. Section 3.2 discusses a 2D finite element implementation of the tracer case. This model is subsequently used to simulate horizontal tank experiments conducted in Boulder at the University of Colorado's Civil Engineering Department under the direction of Professor Tissa Illangsekare. The Boulder experiments are discussed in Section 3.2.10. Briefly, the tank environment is like that of a confined aquifer into which a tracer (pollutant) is injected. The tank has 45 port locations where tracer injections can be made and samples can be taken. Constant head conditions are assumed on the ends of the tank, and no flow boundary conditions are assumed on the sides of the tank.

Both homogeneous and heterogeneous experiments are run in the tank. In the homogeneous case, the tank is packed with a uniform sand, *i.e.*, of uniform hydraulic conductivity. Spectral methods can be used to study the homogeneous packing. As would be expected, Fourier series can be used to solve the 1D flow equation resulting in a solution of the form

$$\phi^*(x, t) = \psi(x, t) + \phi^*(0, t) + \frac{x}{L} (\phi^*(L, t) - \phi(0, t))$$

where

$$\psi(x, t) = \sum_{n=1}^{\infty} C_n e^{-\left(\frac{n\pi}{L}\right)^2 \frac{K}{S_0} t} \sin\left(\frac{n\pi}{L}\right) x$$

By letting $t \rightarrow \infty$, the steady state solution is then linear. In the two dimensional case, the constant head contours are nearly parallel, yielding a uniform velocity field. Both the model results and the experimental results show these characteristics.

The velocity field is determined from the flow equation or the velocity/pressure equation described earlier. However, prior to solving for the concentration distribution using the transport, advection-dispersion, equation, the dispersion parameter must be provided. The first two chapters of this study focus on the determination of the dispersion parameter. For homogeneous media, spectral methods can again be used to develop a theoretical formulation for the treatment of dispersion. These spectral methods can be extended to include mildly heterogeneous media by using stochastic process techniques. A stochastic process can be given a spectral representation defined in terms of special stochastic processes that have orthogonal increments. The first step of the procedure used to construct an integral spectral representation is to define orthogonal stochastic measures, Section 1.3. Once this is done, it can be shown that there is an isomorphism between processes with orthogonal increments, continuous from the right, and the orthogonal stochastic measures. It is this association that makes the integral spectral representation work, Sections 1.4 and 1.5.

Based on a review of the literature, it is our view that the theoretical descriptions of dispersion in porous media found in Gelhar and Axness[45], Neuman and Zhang[70] and Dagan[29, 30, 31, 32, 33, 34] are the most prominent. All of these theories link dispersion to the hydraulic conductivity properties of the media. However, there are fundamental differences which have generated a lot of debate in the literature. Section 1.8 seeks to select the common aspects of the Gelhar and Axness[45] and Neuman and Zhang[70] approaches which are based on the integral spectral representations of stochastic processes. The analysis shows how the dispersion tensor associated with the homogeneous case must be modified to include mildly heterogeneous cases, Equation[1.13], page

37. The modification is in terms of the integral of the spectrum of the velocity perturbations. Section 1.10 provides a transformation that allows the conversion of the spectrum of the velocity perturbations to the spectrum of the log hydraulic conductivity perturbations, Equation[1.21], page 43. This theoretical connection is important because in Section 2.4 a description of dispersion in terms of the covariance of the hydraulic conductivities using a particle tracking algorithm will be developed.

Chapter 2 discusses the time and distance forms of the dispersivity tensor. In this analysis, the solute body spatial moments are given by

$$M = \int_{\Omega} ncd\vec{x}$$

$$\vec{R} = \frac{1}{M} \int_{\Omega} n\vec{x}cd\vec{x}$$

$$\mathbf{S}_{ij} = \frac{1}{M} \int_{\Omega} n(x_i - R_i)(x_j - R_j)c(\vec{x}, t)d\vec{x} \quad i, j = 1, 2, 3$$

Here M is the mass, \vec{R} is the centroid coordinate, \mathbf{S}_{ij} is the second spatial moment which characterizes the spread around the centroid, n is the porosity and c is the contaminant concentration. Starting from the transport equation with \vec{V} representing fluid velocity,

$$\frac{\partial c}{\partial t} + \vec{V} \cdot \nabla c = \nabla \cdot (\mathbf{D}\nabla c)$$

and using integration by parts, it can be shown that

$$\text{Zero Moment} \quad \Rightarrow \quad \text{Conservation Of Mass}$$

$$\text{First Moment} \quad \Rightarrow \quad \text{Centroid Of The Mass Concentration Moves With Velocity } \vec{V}.$$

$$\text{Second Moment} \quad \Rightarrow \quad \mathbf{D}_{ij} = \frac{1}{2} \frac{d\mathbf{S}_{ij}}{dt}$$

Although the result that dispersion is related to the time rate of change of the second moment makes intuitive sense, in order to implement such a definition requires knowledge of the plume that is probably not available. For this reason, dispersion is dealt with as a stochastic entity. Dealing

with the problem from a stochastic point of view allows a certain level of uncertainty to be accounted for by the model. This means that instead of deterministic ordinary and partial differential equations, the parameters and variables in the differential equations will be allowed to have stochastic components. This switch to stochastic differential equations brings with it many difficulties that require a *stochastic* calculus to handle. For example, in discussing the movement of a fluid particle through a porous medium, stochastic differential equations of the form

$$dX(t) = a(t, X(t))dt + b(t, X(t))dW(t)$$

will arise where $X(t)$ is the trajectory of the particle through the medium and $W(t)$ is a special type of stochastic process called a *Wiener* process. This equation has the interpretation on the interval $[0, t]$ of

$$X(t) = X(0) + \int_0^t a(s, X(s))ds + \int_0^t b(s, X(s))dW(s)$$

The integrals in this equation cannot be interpreted in the usual sense. In the case of the first integral on the right hand side, the integrand, $a(s, X(s))$ is a random function, *i.e.*, for a given s , $a(s, X(s))$ is a random variable. The second integral is even more difficult to deal with since the measure part of the integral is a stochastic process which can be shown to have infinite variation. Hence, the usual Stieltjes interpretation is not applicable. How these integrals are dealt with is discussed in Section 2.2 and again in Section 4.2.4. As will be seen in those sections, the stochastic differential equation and stochastic integral have more than one interpretation.

Section 2.3 is the first section in which the *Lagrangian* approach is used to give the basic form of the transport equation and the basic form of the dispersion tensor that can be derived from it. In the Lagrangian framework, transport is characterized in terms of indivisible solute particles, *i.e.*, ensembles of molecules in a small volume, which are transported by the fluid. The total displacement of the fluid particle can be decomposed into a component due to convection and a component due to diffusion, Equation[2.12], page 58. The diffusion component is represented by a Brownian motion. The convective component is described by the fundamental kinematic equation

$$\vec{X}(t) = \int_0^t \vec{V}(\vec{X}_T(t'))dt'$$

where $\vec{V}(\vec{X}_T)$ is the Lagrangian velocity field associated with the fluid particle. Assuming the first order relationship between the displacement of \vec{X} , \vec{X}' , and the displacement of \vec{V} , \vec{V}' , is given by

$$\vec{X}'(t; \vec{x}_0, t_0) = \int_0^t \vec{V}'(\mathbf{E}[\vec{V}]t') dt'$$

it follows that the displacement covariance tensor is given by the time integration of the velocity covariances, Equation[2.14], page 59. This characterization of the displacement covariance will play a central role in the development of the dispersion tensor.

As will be seen in Section 3.1.1, the transport equation is usually derived from conservation of mass considerations. However, in Section 2.3.2 it will be demonstrated that the basic form of the transport equation can be derived by treating the fluid particles as obeying the following Itô stochastic differential equation

$$d\vec{X}_T = \mathbf{E}[\vec{V}]dt + d\vec{X}_d$$

Then, the transport equation follows from the Fokker-Planck or Kolmogorov forward equation. This derivation is restricted to the case where dispersion is created only by the Brownian motion, \vec{X}_d .

In Section 2.3.3, the solution of the transport equation is shown to be a multivariate Gaussian density function, and the dispersion tensor is shown to be one-half the time rate of change of the covariances of the trajectory displacements. Section 2.3.4 expands on this result to show that the components of the dispersion tensor should be related to the time rate of changes in the covariances of the fluid particle or the time integral of the velocity covariances which approximates the path integral of the velocity covariances along the fluid particle's trajectory. The integrand of this integral is a lagged covariance over the interval of integration. A numerical formulation of this integral is given which depends on a particle tracking algorithm which is suitable for implementation in a computer code. Section 2.3.5 is an extension of these results due to Kitanidis[58] and Dagan[32].

One of the objectives of this study is to be able to specify the dispersion coefficients on a numerical grid block basis so that they can be used in a finite element model. Section 2.4 contains the details of a proposed method of doing this. By dividing the domain of the problem into numerical grid blocks on which it can be assumed that gradients and hydraulic conductivity covariances are constant, the local displacement covariance tensor, which is represented in the literature as a bold-faced \mathbf{X} , can be expressed as

$$\mathbf{X} = \mathbf{E} \left[\left(\tilde{\mathbf{K}}\mathbf{E}[\mathbf{K}]^{-1}\mathbf{E}[\vec{r}] \right) \left(\tilde{\mathbf{K}}\mathbf{E}[\mathbf{K}]^{-1}\mathbf{E}[\vec{r}] \right)^\dagger \right]$$

where $\tilde{\mathbf{K}} = \mathbf{K} - \mathbf{E}[\mathbf{K}]$ and \mathbf{K} is hydraulic conductivity and the dagger, \dagger , represents the vector transpose. Since \vec{r} in this expression is the position vector of the fluid particle with reference to some injection point, this is really a distance dependent formula. One of the conclusions of experimental studies is that dispersivity varies with the distance from the input zone. The manner in which it varies depends on the degree and location of heterogeneities in the porous medium domain. The presence of heterogeneities in the porous medium will cause the velocity field to be non-uniform. In order to get an adequate representation of exactly where in the domain different magnitudes of dispersion are to be expected, ensembles of particles must be tracked. Figure 4 in Section 2.4 illustrates the tracking of 5 particles from each of two adjacent grid blocks located near the center of the domain. These particle paths are used to identify the numerical grid blocks that are most likely to be reached by a tracer plume that emanates from the grid blocks containing the origin of the plume. Once the grid blocks most likely to be reached have been identified, the previous formulation can be applied on a grid block by grid block basis to estimate how dispersion will develop over time. In this case, the displacement covariance matrix is given by

$$\mathbf{X} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \mathbf{E} \left[\left(\tilde{\mathbf{K}}^{(i+1)}\mathbf{E}[\mathbf{K}^{(i+1)}]^{-1}\mathbf{E}[\vec{r}^{(i+1)}] \right) \left(\tilde{\mathbf{K}}^{(j+1)}\mathbf{E}[\mathbf{K}^{(j+1)}]^{-1}\mathbf{E}[\vec{r}^{(j+1)}] \right)^\dagger \right]$$

where the superscripts represent individual block designations. Dispersivity estimates are derived by differentiating this expression with respect to time.

Since it may be difficult to identify exactly the numerical grid blocks where a local source is originating, or if more than one point source is involved, an entire column of grid blocks can be used to determine the particle paths. Figure 5, Section 2.4, shows the result of tracking a particle from each of the grid blocks in a column in the center of the domain. Of course, in a simulation, more than one particle from each grid block would be tracked.

Section 2.5 discusses some methods of random variable generation that are either currently being used or are planned to be used in the models of flow and transport being created. Since all of the methodology is based on exploiting randomness or uncertainty, it is necessary to have methods of simulating this randomness. In Section 2.5.1, it is shown that by using a Box-Muller

transformation, two independent uniform random variables can be converted into two independent Gaussian random variables. It is also shown, Section 2.5.2, how correlated random variables can be simply generated. However, for more sophisticated simulations in 3 dimensions, efficient computer codes are available that are capable of cogenerating pairs of 3 dimensional, cross-correlated random fields with different correlation scales, Robin, *et al*[80].

Chapter 3 discusses the numerical method being used to study the groundwater flow/transport problem. The particular form of the equations that are to be solved depends on the relationship that is assumed between the density, viscosity and concentration. For example, in an Enhanced Oil Recovery problem, a change in the viscosity of the single phase is brought about by the mixing of a surfactant with the oil in the reservoir. In problems that involve a pollutant entering an aquifer, many times it is the density that changes with the concentration of the pollutant. In these cases, it can be argued using standard definitions of the physical properties of a compressible fluid that the seepage velocity is related to hydraulic conductivity, pressure and density; or to the permeability, pressure, density and viscosity, Equation[3.1], page 97. The transport equation is usually derived by applying the Divergence theorem to the Conservation of Mass law. These equations result in a coupled system of equations, Equation[3.2], page 98. For this coupled system, the mixed finite element method can be used to solve for the pressure/velocity pair, followed by a solution of the transport equation by some method. Currently, the code that is available to solve the coupled system is the SEGMIX code. This code uses mixed finite elements to solve for the pressure/velocity pair and the modified method of characteristics (MMOC) to solve the transport equation. SEGMIX assumes a rectangular domain with no flow boundary conditions on the sides. In order to use this code to simulate the Boulder tank experiments, it would have to be modified to accept constant head conditions at the ends of the tank and no flow boundary conditions on the sides of the tank.

The types of experiments that are being conducted in the Boulder horizontal tank are tracer experiments. This means that a solute such as benzene or sodium chloride is injected at a selected port and samples are taken from a port downstream of the injection port. In this case, the effect on density is probably minimal, and the uncoupled flow/transport system of equations is adequate to study the experiments numerically.

The numerical method being used to solve the uncoupled system is finite elements. Because Gaussian integration formulas are used to evaluate the

integrals that arise in the system of equations resulting from the variational formulation, a reference element which is convenient for integration purposes is defined which is affinely equivalent to the elements in the domain. Affine equivalence can be defined in terms of special mappings called pull-backs and push-forwards. These concepts are explained in Sections 3.2.2 to 3.2.4.

Section 3.2.5 contains a derivation of the local system of equations that arise from the variational formulation of the *flow* equation, Equation[3.10], page 116. Once the systems of equations on the local rectangular elements have been established, they must be assembled into a global system of equations for the whole domain. This process is described in Section 3.2.6. The derivations of the velocity field from the piezometric head estimates is given in Section 3.2.7.

Section 3.2.8 shows the derivation of the system of equations that follow from the variational formulation of the *transport* equation. Since there are constant boundary conditions at the ends of the tank, and a pulsed-input is allowed to take place at an injection point, it is necessary to allow constant concentration conditions to exist at some grid points. The modification of the global system of equations to allow certain grid points to maintain a constant level of concentration is explained in Section 3.2.9. Section 3.2.10 describes in more detail the horizontal test tank used in the Boulder experiments. Two types of experiments are conducted in the tank. The *homogeneous* experiments are those in which the tank is packed with a single type of sand as rated by its hydraulic conductivity. In the *heterogeneous* experiments, the tank is packed in a block arrangement with 5 different types of sand. The hydraulic conductivities of the sands range from 3.618 m/day for Sand #1 to 1036.8 m/day for Sand #5. With this wide span of hydraulic conductivities, a significant amount of heterogeneity is represented in the tank. The block arrangement of the sands in the tank is represented graphically in Figures 13 and 14 in Section 3.2.10. Figure 15 provides a flowchart of the basic program components used and how they interact. Comparisons of computer simulation results shown in Figures 16 and 17 to actual tank measurements show very good agreement. Figure 18 illustrates a computed tracer plume.

Chapter 4 actually starts the second part of the thesis. The previous sections have investigated the components of the equations and the forms of the equations. However, only the expected or mean value of the concentration is predicted. Because of the uncertainties involved in specifying the physical characteristics of the porous medium, the concentration of a solute at a given

point in time is a random variable, and over a period of time it is a stochastic process. Consequently, in order to more accurately characterize the distribution of the solute concentration, higher order statistical moments such as the variance need to be estimated also. In theory, the more moments that can be predicted, the better this characterization will be. But, in practice, it is usually a difficult problem just to obtain information on the variance or covariance of variables in the system.

A much referenced paper in this area is the Graham and McGlaughlin[48] paper which specifies a set of three equations that are to be solved for the mean concentration, the velocity-concentration covariance and the concentration covariance. Section 4.1 derives and discusses these equations because they will be used as a basis of comparison for an approach to developing moment equations based on the Itô calculus.

Randomness can enter the boundary value problem in many different ways. Equation[4.12], page 138, is a statement of the stochastic boundary value problem, and the discussion following that equation specifies the various ways in which randomness can enter the picture. Existence theory for the stochastic boundary value problem is not unlike the nonstochastic case. A summary is included in Section 4.2.3.

Stochastic integration is again addressed in Section 4.2.4, this time from the more general perspective of a *martingale*. The Itô integral then follows from this more general definition as a special case. The use of the Itô integral requires that the rules of calculus have to be modified. The reason for this can be illustrated as follows: Suppose that $W(t)$ is a one dimensional Wiener process, then it is well known that it can be represented as the limiting form of a random walk. And, as part of this limiting process the step size of the random walk goes to zero as the square root of the time interval

$$\Delta W = \mathcal{O}(\Delta t^{\frac{1}{2}})$$

Gardiner[42] shows that for the Itô integral this property means that

$$[dW(t)]^{2+N} = \begin{cases} dt & \text{if } N = 0 \\ 0 & \text{if } N > 0 \end{cases}$$

and that $dt dW(t) = 0$. The most important new rule is that of Itô's lemma. It is a change of variable formula. The reason the change of variable formula

has to be modified is due to the above differential relationships. For example, if f is a smooth function and $X(t)$ satisfies the Itô differential equation

$$dX(t) = a dt + b dW(t)$$

Then, on expanding $df[X(t)]$

$$\begin{aligned} df[X(t)] &= f[X(t) + dX(t)] - f[X(t)] \\ &= f'[X(t)]dX(t) + \frac{f''[X(t)]}{2} (dX(t))^2 + \dots \\ &= f'[X(t)](a dt + b dW(t)) + \frac{f''[X(t)]}{2} (a dt + b dW(t))^2 + \dots \end{aligned}$$

And, using the differential relationships, it follows that

$$df[X(t)] = f'[X(t)](a dt + b dW(t)) + \frac{f''[X(t)]}{2} b^2 dt$$

which is different from the ordinary calculus rule. The Itô formula is a stochastic calculus chain-rule. It can also be extended to martingale type processes, Karatzas[57]. Curtain and Falb[26] have extended Itô's lemma to infinite dimensional Hilbert spaces. It is this form that is used to derive weak forms of the moment equations in Sections 4.2.5 to 4.2.8. For the purpose of illustrating this theory, the key equation is Equation[4.17], page 158, which is applied to two examples. The first example uses this theory to derive mean and covariance equations that in the weak form are identical to those used by Graham and McLaughlin[48]. The second example is cast in terms of accounting for the effects of measurement error that is assumed to enter the experiment as a random perturbation that takes the form of a Wiener process.

Chapter 5 considers the application of the theory of stochastic evolution equations to the problems of flow and transport. In the deterministic case, it is well known that the boundary value problem can be recast as an abstract evolution equation or Cauchy problem. Such a problem takes the form of

$$\begin{aligned} \frac{du}{dt} + A(t)u &= f(t) & 0 \leq t \leq T \\ u(0) &= u_0 \end{aligned}$$

The term *abstract* is attached to indicate the fact that the functions involved are mapping a time interval, $[0, T] \subset \mathfrak{R}^1$, into a Banach or Hilbert space. Hence, for a Hilbert space H , the function $u(t)$ is an H -valued function. It follows, then, that there must be an association between the abstract functions $u(t)$ and the real-valued functions $u(\vec{x}, t)$ of the boundary value problem. And, there must be an association between the differential operator of the boundary value problem and the operator $A(t)$ of the abstract evolution equation. Section 5.1 explores these connections and the forms of the solution to the abstract evolution equation for both the autonomous case, A independent of t , and the nonautonomous, $A(t)$ is dependent on t , cases. The solution, in general, is given as an integral equation.

Curtain and Falb[27] are able to extend these results to the case where the forcing term of the abstract evolution equation contains an H -valued Wiener process, as defined in Section 4.2.4. The solution in this case is given in terms of an evolution operator $U(t, s)$ generated by $-A(t)$ as

$$u(t) = U(t, 0)u_0 + \int_0^t U(t, s)\Phi(s)dW(s)$$

where the integral is now a stochastic integral. Working from the Curtain and Falb solution, weak forms of the mean and covariance equations for the solution $u(t)$ are found that agree with the forms of these equations for the finite dimensional case as given in Åström[8].

In Section 5.2, the time stochastic process nature of the dispersion tensor is again considered. The fact that the dispersion tensor could be considered as a time dependent quantity was first shown in Section 2.1. But, since then, it has been treated in the models developed up to this section as an *effective* parameter by some type of averaging process. In the following sections, the time dependent dispersivity coefficient will be incorporated into the basic partial differential equation model. Both the dispersion coefficients and the velocity will be allowed to have random components. Section 5.3 gives the general form of the stochastic PDE that results in terms of the sum of deterministic and stochastic operator components.

Section 5.4 represents the solution of the stochastic PDE as an integral equation which has the form

$$u(t) = U(t, 0)u_0 + \int_0^t U(t, s)g(s)ds - \int_0^t U(t, s)R_{s,x}u(s)ds$$

where $g(s)$ is the forcing function which may be stochastic, $R_{s,x}$ is the stochastic operator component and $U(t, s)$ is the evolution operator that is the evolution operator that is derived from treating the deterministic part of the stochastic PDE as an abstract differential equation. As an example of this procedure, the following 1-D transport equation is solved

$$\frac{\partial u}{\partial t} - (\mathbf{E}[D] + D'(t, \omega)) \frac{\partial^2 u}{\partial x^2} + (\mathbf{E}[V] + V'(\omega)) \frac{\partial u}{\partial x} = 0$$

Since it is assumed that the $\mathbf{E}[D]$ is a constant in this example, the evolution operator $U(t, s)$ can be expressed as a strongly continuous semigroup generated by the operator

$$\mathcal{A} \equiv -\mathbf{E}[V] \frac{\partial}{\partial x} + \mathbf{E}[D] \frac{\partial^2}{\partial x^2}$$

Using a change of variables to a moving coordinate system, the semigroup is calculated and its properties verified. The solution is then given as an integral equation that has the form of a Volterra type two equation. Section 5.5 applies the classical integral equation methods to characterize the convergence property of the series solution to the integral equation, Hochstadt[53].

In order to provide some insight into the speed of convergence of this series, a test problem is constructed and three scenarios tested. The first case allows the dispersivity to have a random component, but not the velocity. The second case allows the velocity to have a random component, but not the dispersivity. Finally, the last case allows both the dispersivity and the velocity to have random components. The results are contained in the tables at the end of Section 5.5. Monte Carlo methods can be used to generate several sample paths for the random components, and from these, concentration means and variances can be calculated.

The Neumann expansion procedure discussed in Sections 5.4 and 5.5 can be extended to a stochastic finite element method by combining the Karhunen-Loève expansion, a Fourier type expansion of a stochastic process, with the Galerkin method. This approach is, however, subject to the same convergence criterion discussed in Section 5.5. Instead of pursuing the Neumann expansion any further, Chapter 6 on Future Research describes an alternative method which is based on the following three components:

- (1) The Karhunen-Loève Expansion
- (2) The Galerkin Method
- (3) Wiener's Homogeneous Chaos Spaces

As described in Section 5.4, the formal solution of the equation

$$\mathcal{L}_{t,x}u = g$$

is developed by splitting the operator $\mathcal{L}_{t,x}$ into a deterministic part and a zero mean stochastic part, *i.e.*,

$$\mathcal{L}_{t,x} \equiv L_{t,x} + R_{t,x}$$

and using the Karhunen-Loève expansion to express the random coefficients in the stochastic component, $R_{t,x}$, and the Homogeneous Chaoses of Wiener to represent the solution.

As described in Section 6.1, the Karhunen-Loève expansion requires a knowledge of the covariance function of the stochastic process. Specifically, the expansion requires that the eigenvalues and eigenfunctions associated with the covariance function of the process being represented be known. Although certain assumptions can be made regarding the covariance function associated with the stochastic coefficients of the stochastic operator, the covariance function of the solution process is not known. What is required is a way of representing the solution process that does not require a knowledge of its covariance function. The principal result in this approach is the Homogeneous Chaos decomposition of the L_2 -space of a Gaussian process. The fundamentals of this type of decomposition are provided in Section 6.2. The stochastic finite element method using Homogeneous Chaoses is outlined in Section 6.3. Since the success of this finite element procedure depends on obtaining the eigenvalues and eigenfunctions of the covariance function, some details of this problem are covered in Section 6.4.

1.2 Stochastic Processes

Some of the more famous and fundamental results on contaminant transport through porous media involve the characterization of dispersion in terms of the spectral representation of stochastic processes, Gelhar and Axness[45], Neuman, *et al*[69]. This introduction will trace the development of these spectral representations from their beginnings in classical functional analysis to their use in deriving fundamental theoretical formulations of the dispersion tensor. In doing so, the key role that the dispersion tensor plays in the process of scaling-up from small homogeneous laboratory experiments to larger heterogeneous field problems will become clear. The following definitions can be found in many references, for example Burrill[17], Doob[36] and Todorovic[96]:

Definition: Let $(\Omega, \mathcal{B}, \mathcal{P})$ be a probability space. A real (complex) valued measurable function with domain Ω is a real (complex) *random variable*.

Definition: A real (complex) *stochastic process* is a family of real (complex) random variables $\{X(t) : t \in T\}$ defined on a common probability space $(\Omega, \mathcal{B}, \mathcal{P})$ with $T \subset \mathfrak{R}^1$.

Definition: A stochastic process is said to be *strictly stationary* if its distributions do not change with time, *i.e.*, if for any $t_1, t_2, \dots, t_n \in T$ and for any $h \in T$, the multivariate distribution function of the random variable $(X_{t_1+h}, \dots, X_{t_n+h})$ does not depend on h .

Ex:

$$f(X_1, X_2, \dots, X_n; t_1, \dots, t_n) = f(X_1, X_2, \dots, X_n; t_1 + h, \dots, t_n + h)$$

Definition: A stochastic process $\{X_t : t \in T\}$ is *wide sense stationary* if

- $\mathbf{E}[|X_t|^2] < \infty$

- $\mathbf{E}[X_{t+\tau}, X_t] = \mathbf{E}[X_\tau, X_0] \quad \forall t \in T$

where \mathbf{E} represents the expected value. Doob[36], shows that for wide sense stationary processes, the following holds:

Proposition If $\{X(t); t \in \mathfrak{R}^1\}$ is a process which is wide sense stationary, then there is a group of unitary transformations $\{U_t; t \in \mathfrak{R}^1\}$ such that for each t ,

$$X(t) = U_t X(0)$$

$$U_{s+t} = U_s U_t$$

$$U_0 \equiv \text{Identity Element}$$

$$U_{-t} \equiv \text{Inverse Of } U_t$$

Using the notation (\cdot, \cdot) to represent the Hilbert space inner product, the following spectral representation theorem exists for unitary operators, Riesz and Sz.-Nagy[79], Section 137,

Stone's Theorem Every one-parameter group $\{U_t; t \in \mathfrak{R}^1\}$ of unitary transformations for which $(U_t f, g)$ is a continuous function of t , for all elements f and g , admits the spectral representation

$$U_t = \int_{-\infty}^{+\infty} e^{i\lambda t} dE_\lambda$$

where $\{E_\lambda\}$ is a spectral family. Furthermore, E_λ is uniquely determined.

E_λ is an orthogonal projection with the properties

- (1) $E_\lambda \leq E_\mu$ for $\lambda < \mu$
- (2) $E_{\lambda+0} = E_\lambda$
- (3) $E_\lambda \rightarrow 0$ for $\lambda \rightarrow -\infty$
- (4) $E_\lambda \rightarrow I$ for $\lambda \rightarrow \infty$

A second theorem related to this is, Riesz and Sz.-Nagy[79], Section 138,

Bochner's Theorem In order for the function $p(t)$ ($-\infty < t < \infty$) to admit the representation

$$p(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dV(\lambda)$$

with a nondecreasing and bounded real function $V(\lambda)$, it is necessary and sufficient that $p(t)$ be continuous and nonnegative definite in the sense that

$$\sum_{\mu, \nu=1}^m p(t_\mu - t_\nu) \rho_\mu \bar{\rho}_\nu \geq 0$$

whatever the positive integer m , the real numbers t_1, t_2, \dots, t_m and the complex numbers $\rho_1, \rho_2, \dots, \rho_m$.

The reason that these two theorems are connected is that the function

$$p(t) = (U_t f, f)$$

is nonnegative definite, and from this it follows that, Riesz and Sz.-Nagy[79], Section 138,

$$(U_t f, g) = \int_{-\infty}^{\infty} e^{i\lambda t} d(E_\lambda f, g) \tag{1.1}$$

where E_λ is a spectral family.

From the *polarization identity*, the following holds in a complex Hilbert space

$$\begin{aligned}
(E_\lambda f, g) &= \Re(E_\lambda f, g) + i\Im(E_\lambda f, g) \\
&= \Re(E_\lambda^2 f, g) + i\Im(E_\lambda^2 f, g) \\
&= \Re(E_\lambda f, E_\lambda g) + i\Im(E_\lambda f, E_\lambda g) \\
&= \|E_\lambda \frac{f+g}{2}\|^2 - \|E_\lambda \frac{f-g}{2}\|^2 + i \left(\|E_\lambda \frac{f+ig}{2}\|^2 - \|E_\lambda \frac{f-ig}{2}\|^2 \right)
\end{aligned}$$

In particular, if $g = f$, then

$$(E_\lambda f, f) = \|E_\lambda f\|^2$$

and Equation[1.1] then gives

$$(U_t f, f) = \int_{-\infty}^{\infty} e^{i\lambda t} d\|E_\lambda f\|^2 \quad (1.2)$$

From the properties of the spectral family,

$$\begin{aligned}
0 \leq \|E_\lambda f\|^2 &= (E_\lambda f, E_\lambda f) = (E_\lambda^2 f, f) \\
&= (E_\lambda f, f)
\end{aligned}$$

and, if $\lambda \geq \mu$ then $E_\lambda \geq E_\mu \Rightarrow (E_\lambda f, f) \geq (E_\mu f, f)$, so that $\|E_\lambda f\|^2$ is *real* and *nondecreasing*. Furthermore,

$$\|E_\lambda f\|^2 \leq \|E_\infty f\|^2 = \|f\|^2$$

so that $\|E_\lambda f\|^2$ is also *bounded*. Equation[1.2] will be useful in the discussion of the Bochner-Khinchin theorem in Section 1.4.

1.3 Stochastic Measures

In the preceding section, it was shown that a stochastic process $X(t)$ that is wide sense stationary has the representation

$$X(t) = U_t X(0) = \int_{-\infty}^{\infty} e^{i\lambda t} dE_\lambda X(0)$$

The process $\{E_\lambda X(0) : \lambda \in \mathfrak{R}^1\}$ turns out to be a special type of stochastic process. Also, in order to be able to interpret this integral in the usual sense, this process must be associated in some way with a type of measure. In this section a stochastic measure and a stochastic integral are defined that allow the representation as a stochastic process in terms of a unique orthogonal stochastic measure which corresponds to a stochastic process with orthogonal increments, Section 1.4. The details of this section on *Stochastic Measures* are found in Todorovic[96].

Definition: A complex-valued random variable Z on $(\Omega, \mathcal{B}, \mathcal{P})$ is called *second order* if

$$\mathbf{E}[|Z|^2] < \infty$$

The family of all such random variables is denoted by

$$L_2\{\Omega, \mathcal{B}, \mathcal{P}\}$$

Todorovic[96] shows that given the definition of equality that

$$Z_1 = Z_2 \quad \text{iff} \quad Z_1 = Z_2 \quad (\text{a.s.})$$

ie, they differ on at most a set of measure zero, so that the space $L_2\{\Omega, \mathcal{B}, \mathcal{P}\}$ consists of equivalence classes, and using the inner product definition

$$(Z_1, Z_2) = \mathbf{E}[Z_1 Z_2^*] = \int_{\Omega} Z_1 Z_2^* d\mathcal{P}(\omega)$$

where the symbol * represents complex conjugation, then $L_2(\Omega, \mathcal{B}, \mathcal{P})$ is an inner product space. The norm for this space is defined as

$$\|Z\| = +(Z, Z)^{\frac{1}{2}} = \mathbf{E}[|Z|^2]^{\frac{1}{2}}$$

Using this norm to define distance, $L_2(\Omega, \mathcal{B}, \mathcal{P})$ is a metric space. The Cauchy-Schwarz inequality holds, Todorovic[96], and the inner product is uniformly continuous, for if $Z_0, Z_1, Z_2 \in L_2\{\Omega, \mathcal{B}, \mathcal{P}\}$

$$|(Z_1, Z_0) - (Z_2, Z_0)| \leq |(Z_1 - Z_2, Z_0)| \leq \|Z_0\| \|Z_1 - Z_2\|$$

so that for $\|Z_1 - Z_2\| < \frac{\epsilon}{\|Z_0\|}$, $|(Z_1, Z_0) - (Z_2, Z_0)| < \epsilon$.

The mode of convergence on this metric space is that of mean squared convergence and defined as

$$(m^2) \lim_{n \rightarrow \infty} Z_n = Z \quad \text{iff} \quad \|Z_n - Z\| \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

And, this definition of convergence leads to the following form of the

Riesz-Fisher Theorem If $\{Z_n\}_{n=1}^{\infty}$ is a Cauchy sequence, then there exists a $Z \in L_2\{\Omega, \mathcal{B}, P\}$ such that

$$(m^2) \lim_{n \rightarrow \infty} Z_n = Z \quad \square$$

Let $\{S, \mathcal{S}\}$ be an arbitrary measurable space, and let \mathcal{S}_0 be the algebra of sets that generates the σ -algebra \mathcal{S} .

Definition: A mapping

$$\eta : \mathcal{S}_0 \rightarrow L_2\{\Omega, \mathcal{B}, \mathcal{P}\}$$

such that

$$\begin{aligned} \eta(\emptyset) &= 0 \\ \eta(A \cup B) &= \eta(A) + \eta(B) \quad (a.s.) \end{aligned}$$

for any disjoint sets $A, B \in \mathcal{S}_0$ is called an *elementary random or stochastic measure*.

Definition: Let $m(A) = \|\eta(A)\|^2 < \infty$ for any $A \in \mathcal{S}_0$ where $\|\cdot\|^2 = \mathbf{E}[\cdot^2]$

Definition: An elementary random measure is said to be *orthogonal* if

$$(\eta(A), \eta(B)) = 0 \quad \forall \text{ disjoint } A, B \in \mathcal{S}_0$$

It is clear from the orthogonality property of the random measure $\eta(\cdot)$ that the set function $m(\cdot)$ is finitely additive. And, if it is assumed that $m(\cdot)$ is

countably subadditive, then it can be extended to a measure on \mathcal{S}_0 . According to Proposition 12.3.9 of Royden[81], all that is required is a semialgebra of sets \mathcal{C} that generates \mathcal{S}_0 and that on \mathcal{C} , $m(\emptyset) = 0$, $m(\cdot)$ is finitely additive on \mathcal{C} and countable subadditive on \mathcal{C} . Furthermore, the measure $m(\cdot)$ so defined on \mathcal{S}_0 can then be extended to the measurable space $\{S, \mathcal{S}\}$.

Definition: $m(\cdot)$ is called the *measure associated with $\eta(\cdot)$* . Let

$$L_2(m) = L_2\{S, \mathcal{S}, m\}$$

the Hilbert space of complex-valued functions on S which are square integrable with respect to m .

Definition: For step functions

$$h(s) = \sum_{k=1}^n c_k I_{B_k}(s) \quad \text{disjoint } B_k \in \mathcal{S}_0$$

define

$$\psi(h) = \int_S h(s) \eta(ds) = \sum_{k=1}^n c_k \eta(B_k)$$

It is clear that the mapping ψ takes step functions in $L_2(m)$ and maps them into random variables in $L_2(\Omega, \mathcal{B}, \mathcal{P})$. Furthermore, given two step functions $h(s)$ and $f(s)$

$$\begin{aligned} (\psi(h), \psi(f))_{L_2(\Omega, \mathcal{B}, \mathcal{P})} &= \left(\sum_{i=1}^n x_i \eta(A_i), \sum_{j=1}^n y_j \eta(B_j) \right)_{L_2(\Omega, \mathcal{B}, \mathcal{P})} \\ &= \sum_{i=1}^n \sum_{j=1}^n x_i \bar{y}_j (\eta(A_i), \eta(B_j))_{L_2(\Omega, \mathcal{B}, \mathcal{P})} \end{aligned}$$

and, since the sets A_i and B_j can be written in terms of disjoint sets as

$$A_i = (A_i \setminus B_j) \cup (A_i \cap B_j)$$

$$B_j = (B_j \setminus A_i) \cup (A_i \cap B_j)$$

it follows from the orthogonality of $\eta(\cdot)$ that the inner product $(\eta(A_i), \eta(B_j))_{L_2(\Omega, \mathcal{B}, \mathcal{P})}$ can be written as

$$\begin{aligned} (\eta(A_i), \eta(B_j))_{L_2(\Omega, \mathcal{B}, \mathcal{P})} &= (\eta(A_i \cap B_j), \eta(A_i \cap B_j))_{L_2(\Omega, \mathcal{B}, \mathcal{P})} \\ &= m(A_i \cap B_j) \end{aligned}$$

consequently, the inner product $(\psi(h), \psi(f))_{L_2(\Omega, \mathcal{B}, \mathcal{P})}$ can be written as

$$\begin{aligned} (\psi(h), \psi(f))_{L_2(\Omega, \mathcal{B}, \mathcal{P})} &= \sum_{i=1}^n \sum_{j=1}^n x_i \bar{y}_j m(A_i \cap B_j) \\ &= \int_S h(s) \overline{f(s)} m(ds) \\ &= (h, f)_{L_2(m)} \end{aligned}$$

Therefore, the mapping ψ preserves the inner product of the step functions from $L_2(\Omega, \mathcal{B}, \mathcal{P})$ to $L_2(m)$. If $g \in L_2(m)$, then let $\{h_n\}_{n=1}^{\infty}$ be a sequence of step functions such that

$$\|g - h_n\| \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

then from the orthogonality of $\eta(\cdot)$

$$\|\psi(h_n) - \psi(h_m)\|^2 = \|h_n - h_m\|^2 \rightarrow 0 \quad \text{as } m, n \rightarrow \infty$$

So, $\{\psi(h_n)\}_{n=1}^{\infty}$ is a Cauchy sequence, and by the Riesz-Fisher Theorem, $\exists \psi(g) \in L_2\{\Omega, \mathcal{B}, \mathcal{P}\}$ such that

$$\|\psi(g) - \psi(h_n)\| \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

Definition: The random variable $\psi(g)$ is called the stochastic integral of $g \in L_2(m)$ with respect to the elementary orthogonal random measure η . It is denoted by

$$\psi(g) = \int_S g(s) \eta(ds) \tag{1.3}$$

The elementary stochastic measure η which is defined on \mathcal{S}_0 can be extended to the σ -algebra \mathcal{S} , Todorovic[96]. Hence, Equation[1.3] is taken to be the integral of a $g \in L_2(m)$ over the set S with respect to the orthogonal stochastic measure η .

1.4 Process With Orthogonal Increments

Let $\{S, \mathcal{S}, m\} = \{\mathfrak{R}^1, \mathcal{R}, m\}$ where $\mathfrak{R}^1 \equiv$ Real Line, and $\mathcal{R} \equiv \sigma$ -algebra of Borel sets of \mathfrak{R}^1 . Let $\{Z(t); t \in \mathfrak{R}^1\} \subset L_2\{\Omega, \mathcal{B}, P\}$ with $\mathbf{E}[Z(t)] = 0, \forall t \in \mathfrak{R}^1$. Then, using Stone's theorem, these stochastic processes can be written as

$$Z(t) = U_t Z(0) = \int_{-\infty}^{\infty} e^{i\lambda t} dE_{\lambda} Z(0)$$

From the properties of the spectral family that the E_{λ} are symmetric and for $\lambda_1 \leq \lambda_2$,

$$E_{\lambda_1} E_{\lambda_2} = E_{\lambda_1}$$

it follows that, for $\lambda_0 \leq \lambda_1 \leq \lambda_2$

$$\mathbf{E}[(E_{\lambda_1} Z(0) - E_{\lambda_0} Z(0))(E_{\lambda_2} Z(0) - E_{\lambda_1} Z(0))^*] =$$

$$((E_{\lambda_1} - E_{\lambda_0})Z(0), (E_{\lambda_2} - E_{\lambda_1})Z(0)) = (Z(0), (E_{\lambda_1} - E_{\lambda_0})(E_{\lambda_2} - E_{\lambda_1})Z(0))$$

The operator product in the last term can be expanded as

$$\begin{aligned} (E_{\lambda_1} - E_{\lambda_0})(E_{\lambda_2} - E_{\lambda_1}) &= E_{\lambda_1} E_{\lambda_2} - E_{\lambda_1} E_{\lambda_1} - E_{\lambda_0} E_{\lambda_2} + E_{\lambda_0} E_{\lambda_1} \\ &= E_{\lambda_1} - E_{\lambda_1} - E_{\lambda_0} + E_{\lambda_0} \\ &= 0 \end{aligned}$$

Consequently, the stochastic processes $E_{\lambda} Z(0), -\infty < \lambda < \infty$ that appear in the Stone representation have the special property

$$\mathbf{E}[(E_{\lambda_1} Z(0) - E_{\lambda_0} Z(0))(E_{\lambda_2} Z(0) - E_{\lambda_1} Z(0))^*] = 0$$

Todorovic[96] makes the following

Definition: The stochastic process $\{Z(t); t \in \mathfrak{R}^1\}$ is said to have *orthogonal increments* if, for any $t_0 < t_1 < t_2$,

$$\mathbf{E}[Z(t_1) - Z(t_0)][Z(t_2) - Z(t_1)]^* = 0$$

where the symbol $*$ represents complex conjugation.

Definition: The process $Z(t)$ is *right continuous* in the mean squared sense if $\forall t \in \mathfrak{R}^1, t$ fixed,

$$\|Z(t) - Z(t_k)\| \rightarrow 0 \text{ as } t_k \downarrow t$$

Todorovic[96] shows that there exists an isomorphism between processes $\{Z(t) : t \in \mathfrak{R}^1\} \subset L_2\{\Omega, \mathcal{B}, P\}$ with orthogonal increments continuous from the right in the mean squared sense and the orthogonal stochastic measures η with $m(\cdot)$ the measure associated with it. This correspondence is given by

$$Z(t) = \eta((-\infty, t])$$

From this, for $t > s$,

$$\begin{aligned} Z(t) &= \eta((-\infty, t]) = \eta((-\infty, s]) \cup (s, t]) \\ &= \eta((-\infty, s]) + \eta((s, t]) \\ &= Z(s) + \eta((s, t]) \\ \Rightarrow \eta((s, t]) &= Z(t) - Z(s) \end{aligned}$$

Ash[6] characterizes the Borel sets in various ways. In particular, they can be defined as the smallest σ -algebra of subsets of \mathfrak{R}^1 that contains the intervals $(-\infty, t], t \in \mathfrak{R}^1$. Furthermore, from this definition the function $F(t)$ can be defined as

$$\begin{aligned} F(t) &= m((-\infty, t]) = \|\eta((-\infty, t])\|^2 \\ &= \mathbf{E}[|Z(t)|^2] = \mathbf{E}[Z(t)Z(t)^*] \end{aligned}$$

where $*$ represents the complex conjugate. Since $m((-\infty, t]) = m((-\infty, s] \cup (s, t])$, then

$$F(t) = F(s) + m((s, t])$$

So that,

$$m((s, t]) = F(t) - F(s)$$

It can be shown, Todorovic[96], Ash[6], that the function

$$F(t) = \mathbf{E}[Z(t)Z(t)^*]$$

is bounded, right continuous and non-decreasing.

Letting $\eta(dt) = dZ(t)$, then

$$\begin{aligned} dF(t) &= m(dt) = \|\eta(dt)\|^2 \\ &= \|dZ(t)\|^2 \\ &= \mathbf{E}[|dZ(t)|^2] \\ &= \mathbf{E}[dZ(t)dZ(t)^*] \end{aligned}$$

Hence, the following holds

$$dF(t) = \mathbf{E}[dZ(t)dZ(t)^*]$$

And, if the process $Z(t)$ has orthogonal increments, then

$$\mathbf{E}[dZ(t)dZ(t')^*] = \begin{cases} 0 & \text{if } t \neq t' \\ dF(t) & \text{if } t = t' \end{cases} \quad (1.4)$$

Definition: The stochastic integral

$$\int_{-\infty}^{+\infty} h(t)dZ(t)$$

is to be interpreted as

$$\int_{-\infty}^{+\infty} h(t)\eta(dt)$$

From the definition of the stochastic integral, Equation[1.3], page 20, $h(t)$ is a complex-valued function. If

$$h(t) = e^{i\lambda t}$$

then

$$\begin{aligned} U_t Z(0) &= \int_{-\infty}^{\infty} e^{i\lambda t} dE_\lambda Z(0) \\ &= \int_{-\infty}^{\infty} e^{i\lambda t} d\eta(d\lambda) \end{aligned}$$

where η is the orthogonal stochastic measure associated with the stochastic process with orthogonal increments, $E_\lambda Z(0)$.

1.5 Spectral Representation

Let $\{Z(t); t \in \mathfrak{R}^1\} \subset L_2\{\Omega, \mathcal{B}, P\}$ be a wide sense stationary process with

$$\mathbf{E}[Z(t)] = 0 \quad C(t) = \mathbf{E}[Z(s+t)Z^*(s)]$$

The following proposition follows from the non-negative definiteness of the covariance function, Todorovic[96].

Proposition The covariance function $C(t)$ is continuous on \mathfrak{R}^1 if it is continuous at zero.

From the definition of the inner product,

$$C(t) = (Z(s+t), Z(s))$$

Let $f = Z(0)$, then from the discussion in Section 1.2, the covariance can be expressed in terms of the unitary operators as

$$\begin{aligned} C(t) &= (U_{s+t}f, U_s f) \\ &= (U_t U_s f, U_s f) \end{aligned}$$

Letting $g = U_s f = Z(s)$ and using Equation[1.2], page 16, it follows that

$$C(t) = (U_t g, g) = \int_{-\infty}^{\infty} e^{i\lambda t} d\|E_\lambda g\|^2$$

Furthermore, since $\|E_\lambda g\|^2$ is real, bounded and nondecreasing and E_λ is uniquely determined, then

$$F(\lambda) = \|E_\lambda g\|^2 = \mathbf{E}[E_\lambda Z(s)(E_\lambda Z(s))^*] = \mathbf{E}[Z(\lambda)Z(\lambda)^*] \quad (1.5)$$

and obtain

$$C(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dF(\lambda)$$

In particular, letting $t = 0$ and using the spectral family properties, the covariance can be written as

$$\begin{aligned} C(0) &= \int_{-\infty}^{\infty} d\|E_\lambda g\|^2 = \|E_\infty g\|^2 \\ &= \|g\|^2 \\ &= \mathbf{E}[Z(s)Z^*(s)] \end{aligned}$$

The forgoing discussion illustrates to the famous

Bochner-Khinchin Theorem A complex-valued function $C(t)$ defined on R continuous at zero is the covariance function of a wide sense stationary stochastic process iff it can be written in the form

$$C(t) = \int_{-\infty}^{+\infty} e^{it\lambda} dF(\lambda)$$

where $F(\cdot)$ is a real nondecreasing bounded function on \mathfrak{R}^1 called the *spectral distribution* of the process $\xi(t)$.

If the spectral distribution function $F(\cdot)$ is absolutely continuous, then the *spectral density* of the process is given by

$$f(\lambda) = F'(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i\lambda t} C(t) dt$$

So, the covariance function and the spectral density form a Fourier Transform pair.

Comparing this with Equation[1.4], page 23, and Equation[1.5], page 25, the spectral density can be written as

$$f(\lambda)d\lambda = dF(\lambda) = \mathbf{E}[dZ(\lambda)dZ(\lambda)^*]$$

So that the covariance function takes the form

$$C(t) = \int_{-\infty}^{\infty} f(\lambda)e^{it\lambda}d\lambda$$

$f(\lambda)$ is called the continuous spectrum of $C(t)$ and the integral expresses $C(t)$ as a continuous bundle of waves having amplitudes $f(\lambda)$.

Similarly, the cross-spectrum $S_{XY}(\omega)$ of two processes $X(t)$ and $Y(t)$ is the Fourier transform of their cross-correlation $R_{XY}(\tau)$

$$\mathbf{E}[X(t + \tau)Y^*(t)] = R_{XY}(\tau) = \int_{-\infty}^{+\infty} S_{XY}(\omega)e^{i\omega\tau}d\omega$$

and

$$S_{XY}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} R_{XY}(\tau)e^{-i\omega\tau}d\tau$$

Let $\{\xi(t); t \in \mathfrak{R}^1\} \subset L_2\{\Omega, \mathcal{B}, P\}$ be a wide sense stationary process such that

$$\mathbf{E}[\xi(t)] = 0 \quad \text{and} \quad C(t) = \mathbf{E}[\xi(t + s)\xi^*(s)]$$

Then from Todorovic[96] there is the following existence

Proposition Let the covariance function $C(t)$ be continuous at zero. Then, there exists a unique orthogonal stochastic measure η with values in $L_2\{\Omega, \mathcal{B}, P\}$ such that

$$\xi(t) = \int_{-\infty}^{+\infty} e^{i\lambda t}\eta(d\lambda) \quad (a.s.)$$

and

$$\|\eta(A)\|^2 = m(A) = \int_A dF \quad \forall A \in \mathcal{R}$$

where $m(\cdot)$ is the Lebesgue-Stieltjes measure associated with η and generated by the spectral distribution F .

Hence, if $\{Z_\xi(\lambda); \lambda \in \mathfrak{R}^1\}$ is a stochastic process with orthogonal increments corresponding to the orthogonal stochastic measure $\eta(\cdot)$, the process $\{\xi(t); t \in \mathfrak{R}^1\}$ has the spectral representation

$$\xi(t) = \int_{-\infty}^{+\infty} e^{i\lambda t} \eta(d\lambda) \equiv \int_{-\infty}^{+\infty} e^{i\lambda t} dZ_\xi(\lambda)$$

1.6 Space Correlations And Space Spectra

The interpretation of the independent variable as a space variable instead of a time variable does not change things, Lumley and Panofsky[64].

Homogeneity is the property that for the space variable corresponds to *stationarity* for the time variable.

$$\xi(x) = \int_{-\infty}^{+\infty} e^{i\lambda x} dZ_\xi(\lambda)$$

In 3-dimensions, the corresponding spectral representation is given as

$$\xi(\vec{x}) = \int_{\mathfrak{R}^3} e^{i\vec{k}\cdot\vec{x}} dZ_\xi(\vec{k})$$

where \vec{k} is the wave number and $Z_\xi(\vec{k})$ is a stochastic process with orthogonal increments.

In the case that time is a contributing factor, the spectral representation can be written as

$$\xi(\vec{x}, t) = \int_{\mathfrak{R}^3} e^{i\vec{k}\cdot\vec{x}} dZ_\xi(\vec{k}, t)$$

As in the one-dimensional case, if $Z(\vec{k})$ is a stochastic process with orthogonal increments, then

$$\mathbf{E}[dZ(\vec{k})dZ(\vec{k}')^*] = \begin{cases} 0 & \text{if } \vec{k} \neq \vec{k}' \\ dF(\vec{k}) & \text{if } \vec{k} = \vec{k}' \end{cases} \quad (1.6)$$

where $dF(\vec{k}) = f(\vec{k}) d\vec{k}$ and $f(\vec{k})$ is the spectral density of the process.

The cross-spectrum of two processes X and Y is given by

$$S_{XY}(\vec{k}) = \frac{1}{(2\pi)^3} \int_{\mathfrak{R}^3} R_{XY}(\vec{x}) e^{-i\vec{k}\cdot\vec{x}} d\vec{x}$$

and, its Fourier inverse is the cross-correlation of the processes X and Y , given by

$$R_{XY}(\vec{x}) = \int_{\mathfrak{R}^3} S_{XY}(\vec{k}) e^{i\vec{k}\cdot\vec{x}} d\vec{k}$$

For a *homogeneous* process, the following is true

$$R_{XY}(-\vec{\theta}) = \mathbf{E}[X(\vec{x} - \vec{\theta})Y^*(\vec{x})] = R_{YX}^*(\vec{\theta})$$

In particular, if X and Y are *real* stochastic processes,

$$R_{XY}(-\vec{\theta}) = R_{YX}(\vec{\theta})$$

and, if $X = Y$,

$$R_{XX}(-\vec{\theta}) = R_{XX}(\vec{\theta})$$

so that the covariance function of a *real* stochastic process is an *even* function. Also,

$$R_{XY}(-\vec{\theta}) + R_{YX}(-\vec{\theta}) = R_{YX}(\vec{\theta}) + R_{XY}(\vec{\theta})$$

So, clearly, $R_{XY}(\vec{\theta}) + R_{YX}(\vec{\theta})$ is an even function. Furthermore, if R_{XX} is an even function, then using the mapping $T(\vec{x}) = -\vec{x} = \vec{\tau}$ and the change of variable formula it follows that

$$\begin{aligned} S_{XX}(-\vec{k}) &= \frac{1}{(2\pi)^3} \int_{\mathfrak{R}^3} R_{XX}(\vec{\tau}) e^{-i(-\vec{k})\cdot\vec{\tau}} d\vec{\tau} \\ &= \frac{1}{(2\pi)^3} \int_{\mathfrak{R}^3} R_{XX}(\vec{x}) e^{-i\vec{k}\cdot(-\vec{\tau})} d\vec{\tau} \\ &= \frac{1}{(2\pi)^3} \int_{\mathfrak{R}^3} R_{XX}(\vec{x}) e^{-i\vec{k}\cdot\vec{x}} (-1) d\vec{x} \\ &= S_{XX}(\vec{k}) \end{aligned}$$

The last equality follows from the change of variable formula and the fact that -1 is the Jacobian of the transformation T . Hence, $S_{XX}(\vec{k})$ is an *even* function. Similarly, $S_{XY}(\vec{k}) + S_{YX}(\vec{k})$ is an even function.

1.7 Ergodicity

For many properties that are to be measured, it is easier to obtain one observation at each point in time of a random sequence over a long period of time than to obtain several observations at the same point in time.

The former method is called *averaging along the process* (time averages) and the latter method is called *averaging across the process* (ensemble averages).

A statement that these two averages are the same is called an *ergodic theorem*. Burrill[17] states the following:

Definition: Let M be a 1-1 mapping of Ω onto Ω . M is a *measure-preserving* transformation whenever the following condition holds: E is a measurable set iff $M(E)$ is and $P(E) = P(M(E))$ for each measurable set E . For every integer t , the function M^t defined inductively by $M^t = M^{(t-1)} \circ M$ is a measure preserving transformation.

Weak Ergodic Theorem: If M is a measuring-preserving transformation and if $X \in L_2$, then

$$(m^2) \lim_{n-m \rightarrow \infty} \frac{1}{n-m} \sum_{t=m}^{n-1} M^t X = \tilde{X}$$

exists.

The important corollary is

Corollary:

$$\mathbf{E}[\tilde{X}] = \mathbf{E}[X]$$

The transformation M is called *ergodic* if \tilde{X} is a.s. constant. If this is true, then from the corollary

$$(m^2) \lim_{n-m \rightarrow \infty} \frac{1}{n-m} \sum_{t=m}^{n-1} X(t) = \mathbf{E}[X]$$

where $X(t) = M^t X$. So, it is clear from this that the mean-squared limit taken over time is equal to the expected value of the random variable X . The

term ergodic can have several interpretations. If a stochastic process has the property that the mean taken along the process is equal to the mean taken across the process, then the process is said to be *mean-ergodic*. In general, a stochastic process is ergodic if the ensemble averages are equal to the time averages.

In the Hilbert space context, F. Riesz proved the following form of the

Mean Ergodic Theorem Let H be a Hilbert space and $U_t : H \rightarrow H$ be a strongly continuous one parameter unitary group. Let the closed subspace H_0 be defined by

$$H_0 = \{x \in H : U_t x = x \quad \forall t \in \mathfrak{R}^1\}$$

and let P be the orthogonal projection onto H_0 . Then for any $x \in H$

$$\lim_{t \rightarrow \infty} \frac{1}{T} \int_0^T U_s x ds = Px \quad \square$$

The statement and proof are found in Abraham, *et al*[2]. Using the Hilbert space of second order random variables, $L_2(\Omega, \mathcal{B}, \mathcal{P})$ and the Proposition from Section 1.2, there is a group of unitary operators $\{U_t; t \in \mathfrak{R}^1\}$ such that

$$X(t) = U_t X(0)$$

The theorem then says that the limit of the time averages is a random variable that is invariant under the group of operators, $\{U_t : t \in \mathfrak{R}^1\}$.

If $X(t, \hat{\omega})$ is a sample path of a stochastic process $X(t, \omega)$, then the integral

$$\frac{1}{T} \int_0^T X(t, \hat{\omega}) dt$$

represents a time average of $X(t, \hat{\omega})$ over the interval $[0, T]$. From this, the random variable

$$Y_T(\omega) = \frac{1}{T} \int_0^T X(t, \omega) dt$$

can be formed. Clearly, if $X(t, \omega)$ is a stationary process, then $\mathbf{E}[X] = \mathbf{E}[X(t)] \quad \forall t$ and

$$\mathbf{E}[Y_T] = \frac{1}{T} \int_0^T \mathbf{E}[X(t)] dt = \mathbf{E}[X]$$

The variance of $Y_T(\omega)$ is given by

$$\sigma_{Y_T}^2 = \mathbf{E} \left[(Y_T - \mathbf{E}[Y_T])^2 \right] = \mathbf{E} \left[(Y_T - \mathbf{E}[X])^2 \right]$$

Clearly, if $(m^2) \lim_{T \rightarrow \infty} Y_T(\omega) = \mathbf{E}[X]$, then $\sigma_{Y_T}^2 \rightarrow 0$ as $T \rightarrow \infty$. Similarly, if $\sigma_{Y_T}^2 \rightarrow 0$ as $T \rightarrow \infty$, then $(m^2) \lim_{T \rightarrow \infty} Y_T(\omega) = \mathbf{E}[X]$.

Assuming homogeneity, if the independent variable is a spatial variable instead of a time variable, the interpretation is that as more spatial points are included in the average, the variance of the spatial average tends to zero.

1.8 Stochastic Solute Transport And Dispersion

A key component in the scale-up problem is the correct formulation of the dispersion tensor used in the transport equation. Using arguments similar to those found in Gelhar and Axness[45] and Neuman[69], the reason for the focus on dispersion can be illustrated as follows:

Starting from the transport equation,

$$\frac{\partial nc}{\partial t} = -\nabla \cdot n[c\vec{V} - \mathbf{D}\nabla c]$$

and letting the porosity be constant, it follows that

$$\frac{\partial c}{\partial t} = -\nabla \cdot [c\vec{V} - \mathbf{D}\nabla c]$$

where $c \equiv$ concentration, $\vec{V} \equiv$ seepage velocity, and $\mathbf{D} \equiv$ molecular diffusion.

From the conservation of mass equation

$$\frac{\partial(n\rho)}{\partial t} = -\nabla \cdot (\rho\vec{q})$$

with constant porosity and density, it follows that

$$\nabla \cdot \vec{V} = 0 \quad \text{and} \quad \nabla \cdot \mathbf{E}[\vec{V}] = 0$$

Writing the transport equation as

$$\frac{\partial c}{\partial t} + \nabla \cdot (c\vec{V}) = \nabla \cdot (\mathbf{D}\nabla c) \quad (1.7)$$

and letting the concentration and velocity be stochastic processes, which are distributed as follows:

$$\begin{aligned} c &= \mathbf{E}[c] + c' & \mathbf{E}[c'] &= 0 \\ \vec{V} &= \mathbf{E}[\vec{V}] + \vec{V}' & \mathbf{E}[\vec{V}'] &= 0 \end{aligned}$$

the substitution of these distributed parameters into Equation[1.7] yields

$$\frac{\partial(\mathbf{E}[c] + c')}{\partial t} + \nabla \cdot [(\mathbf{E}[c] + c')(\mathbf{E}[\vec{V}] + \vec{V}')] = \nabla \cdot (\mathbf{D}\nabla(\mathbf{E}[c] + c'))$$

Expanding,

$$\begin{aligned} \frac{\partial(\mathbf{E}[c] + c')}{\partial t} + \nabla \cdot \{ \mathbf{E}[c]\mathbf{E}[\vec{V}] + \mathbf{E}[c]\vec{V}' + c'\mathbf{E}[\vec{V}] + c'\vec{V}' \} = & \quad (1.8) \\ \nabla \cdot [\mathbf{D}\nabla(\mathbf{E}[c] + c')] & \end{aligned}$$

Taking expectations and using $\mathbf{E}[c'] = \mathbf{E}[\vec{V}'] = 0$,

$$\frac{\partial\mathbf{E}[c]}{\partial t} + \nabla \cdot (\mathbf{E}[c]\mathbf{E}[\vec{V}]) + \nabla \cdot \mathbf{E}[c'\vec{V}'] = \nabla \cdot [\mathbf{D}\nabla(\mathbf{E}[c])] \quad (1.9)$$

Subtracting Equation[1.9] from Equation[1.8],

$$\frac{\partial c'}{\partial t} + \nabla \cdot [\mathbf{E}[c]\vec{V}' + c'\mathbf{E}[\vec{V}] + c'\vec{V}' - \mathbf{E}[c'\vec{V}']] = \nabla \cdot [\mathbf{D}\nabla(c')]$$

Important assumption: Assuming that the perturbations from their means in c and \vec{V} are *small*, the second order perturbation term

$$c'\vec{V}' - \mathbf{E}[c'\vec{V}']$$

is eliminated because for small perturbations this difference would be close to zero leaving the first order approximation

$$\frac{\partial c'}{\partial t} + \nabla \cdot [\mathbf{E}[c]\vec{V}' + c'\mathbf{E}[\vec{V}]] \approx \nabla \cdot [\mathbf{D}\nabla(c')]$$

Expanding and using the zero divergence of \vec{V}' and $\mathbf{E}[\vec{V}]$,

$$\frac{\partial c'}{\partial t} + \nabla\mathbf{E}[c] \cdot \vec{V}' + \nabla c' \cdot \mathbf{E}[\vec{V}] = \nabla \cdot (\mathbf{D}\nabla c') \quad (1.10)$$

Assuming that c' depends on both \vec{x} and t and that there is spatial homogeneity, the following spectral representations can be made:

$$c'(\vec{x}, t) = \int_{\mathfrak{R}^3} e^{i(\vec{k} \cdot \vec{x})} dZ_{c'}(\vec{k}, t)$$

$$V_i'(\vec{x}) = \int_{\mathfrak{R}^3} e^{i(\vec{k} \cdot \vec{x})} dZ_{V_i'}(\vec{k})$$

$$\vec{V}' = (V_1', V_2', V_3')^\dagger$$

where $\vec{k} = (k_1, k_2, k_3)$ is the wave number and $\vec{x} = (x_1, x_2, x_3)$ is the position vector.

Defining

$$dZ_{\vec{V}'}(\vec{k}) = (dZ_{V_1'}(\vec{k}), dZ_{V_2'}(\vec{k}), dZ_{V_3'}(\vec{k}))^\dagger$$

and making the change of variables

$$y_i = x_i - \mathbf{E}[V_i]t, \quad i = 1, 2, 3$$

and letting

$$c(\vec{x}, t) = \nu(\vec{y}, t)$$

$$\Rightarrow \quad \mathbf{E}[c(\vec{x}, t)] = \mathbf{E}[\nu(\vec{y}, t)] \quad \Rightarrow \quad c'(\vec{x}, t) = \nu'(\vec{y}, t)$$

so that $\nu'(\vec{y}, t)$ is $c'(\vec{x}, t)$ in the moving coordinate system. In other words, c' and ν' are the same stochastic process represented in different coordinate systems. This allows us to write in terms of the spectral representations,

$$\begin{aligned}
\int_{\mathfrak{R}^3} e^{i\vec{k}\cdot\vec{y}} dZ_{\nu'}(\vec{k}, t) &= \nu'(\vec{y}, t) = c'(\vec{x}, t) = \int_{\mathfrak{R}^3} e^{i\vec{k}\cdot\vec{x}} dZ_{c'}(\vec{k}, t) \\
&= \int_{\mathfrak{R}^3} e^{i\vec{k}\cdot(\vec{y}+\mathbf{E}[\vec{V}]t)} dZ_{c'}(\vec{k}, t) \\
&= \int_{\mathfrak{R}^3} e^{i\vec{k}\cdot\vec{y}} e^{i\vec{k}\cdot\mathbf{E}[\vec{V}]t} dZ_{c'}(\vec{k}, t)
\end{aligned}$$

The stochastic process $\nu'(\vec{y}, t)$ has a unique orthogonal stochastic measure associated with it, and from the above equality and the isomorphism between the orthogonal stochastic measures and the processes with orthogonal increments, continuous from the right, it follows that

$$e^{i\vec{k}\cdot\mathbf{E}[\vec{V}]t} dZ_{c'}(\vec{k}, t) = dZ_{\nu'}(\vec{k}, t)$$

$$\Rightarrow \left[i\vec{k} \cdot \mathbf{E}[\vec{V}] dZ_{c'}(\vec{k}, t) + \frac{\partial}{\partial t} (dZ_{c'}(\vec{k}, t)) \right] e^{i\vec{k}\cdot\mathbf{E}[\vec{V}]t} = \frac{\partial}{\partial t} (dZ_{\nu'}(\vec{k}, t)) \quad (1.11)$$

Substituting $\nu'(\vec{y}, t)$ into Equation[1.10], page 33, and simplifying yields

$$\begin{aligned}
\frac{\partial \nu'(\vec{y}, t)}{\partial t} - \nabla_y \nu'(\vec{y}, t) \cdot \mathbf{E}[\vec{V}] + \nabla_y \mathbf{E}[\nu(\vec{y}, t)] \cdot \vec{V}' + \nabla_y \nu'(\vec{y}, t) \cdot \mathbf{E}[\vec{V}] \\
= \nabla_y \cdot [\mathbf{D} \nabla_y \nu'(\vec{y}, t)]
\end{aligned}$$

Then letting $\nabla = \nabla_y$, it follows that

$$\frac{\partial \nu'(\vec{y}, t)}{\partial t} + \nabla \mathbf{E}[\nu(\vec{y}, t)] \cdot \vec{V}' = \nabla \cdot [\mathbf{D} \nabla \nu'(\vec{y}, t)]$$

Substituting the spectral representations for ν' and \vec{V}' in this equation gives

$$\begin{aligned}
\frac{\partial}{\partial t} \left(\int_{\mathfrak{R}^3} e^{i\vec{k}\cdot(\vec{y}+\mathbf{E}[\vec{V}]t)} dZ_{c'}(\vec{k}, t) \right) + \nabla \mathbf{E}[\nu(\vec{y}, t)] \cdot \int_{\mathfrak{R}^3} e^{i\vec{k}\cdot\vec{x}} dZ_{\vec{V}'}(\vec{k}, t) \\
= \nabla \cdot \left(\mathbf{D} \nabla \int_{\mathfrak{R}^3} e^{i\vec{k}\cdot\vec{x}} dZ_{c'}(\vec{k}, t) \right)
\end{aligned}$$

Hence, using Equation[1.11]

$$\begin{aligned} \int_{\mathfrak{R}^3} e^{i\vec{k}\cdot\vec{x}} \left[i\vec{k} \cdot \mathbf{E}[\vec{V}] dZ_{c'}(\vec{k}, t) + \frac{\partial}{\partial t} (dZ_{c'}(\vec{k}, t)) \right] + \int_{\mathfrak{R}^3} e^{i\vec{k}\cdot\vec{x}} \nabla \mathbf{E}[\nu(\vec{y}, t)] \cdot dZ_{\vec{v}'}(\vec{k}, t) \\ = \nabla \cdot \left(\mathbf{D} \nabla \int_{\mathfrak{R}^3} e^{i\vec{k}\cdot\vec{x}} dZ_{c'}(\vec{k}, t) \right) \end{aligned}$$

Taking derivatives and simplifying yields

$$\begin{aligned} \int_{\mathfrak{R}^3} e^{i\vec{k}\cdot\vec{x}} \frac{\partial}{\partial t} (dZ_{c'}(\vec{k}, t)) + \int_{\mathfrak{R}^3} e^{i\vec{k}\cdot\vec{x}} [i\vec{k} \cdot \mathbf{E}[\vec{V}] + \mathbf{D}\vec{k} \cdot \vec{k}] dZ_{c'}(\vec{k}, t) \\ = - \int_{\mathfrak{R}^3} e^{i\vec{k}\cdot\vec{x}} \nabla \mathbf{E}[\nu(\vec{y}, t)] \cdot dZ_{\vec{v}'}(\vec{k}, t) \end{aligned}$$

And, using the uniqueness of the spectral representation, the following first order ordinary differential equation is obtained

$$\begin{aligned} \frac{\partial}{\partial t} (dZ_{c'}(\vec{k}, t)) + [i\vec{k} \cdot \mathbf{E}[\vec{V}] + \mathbf{D}\vec{k} \cdot \vec{k}] dZ_{c'}(\vec{k}, t) \\ = -\nabla \mathbf{E}[\nu(\vec{y}, t)] \cdot dZ_{\vec{v}'}(\vec{k}, t) \end{aligned}$$

Letting $a = [i\vec{k} \cdot \mathbf{E}[\vec{V}] + \mathbf{D}\vec{k} \cdot \vec{k}]$, this equation has the solution, assuming that $dZ_{c'}(\vec{k}, 0) = 0$,

$$dZ_{c'}(\vec{k}, t) = -e^{-at} \int_0^t e^{a\tau} \nabla \mathbf{E}[\nu] \cdot dZ_{\vec{v}'}(\vec{k}) d\tau$$

Furthermore, if it is assumed that the gradient $\nabla \mathbf{E}[\nu]$ is constant, then $\nabla \mathbf{E}[\nu] = \nabla \mathbf{E}[c]$, and it follows by evaluating the integral that

$$a dZ_{c'}(\vec{k}, t) = -b(t) \nabla \mathbf{E}[c] \cdot dZ_{\vec{v}'}(\vec{k})$$

where $b(t) = 1 - e^{-at}$. The assumption that the gradient of the average concentration is constant means that the concentration is spatially linear. This is another assumption that requires only mild heterogeneities. Multiplying both sides of this by $\bar{a} = \overline{[i\vec{k} \cdot \mathbf{E}[\vec{V}] + \mathbf{D}\vec{k} \cdot \vec{k}]}$, then gives

$$\frac{-\nabla \mathbf{E}[c] \cdot dZ_{\vec{v}'}(\vec{k}) [-i\vec{k} \cdot \mathbf{E}[\vec{V}] + \mathbf{D}\vec{k} \cdot \vec{k}] b(t)}{(\vec{k} \cdot \mathbf{E}[\vec{V}])^2 + (\mathbf{D}\vec{k} \cdot \vec{k})^2} = dZ_{c'}(\vec{k})$$

Multiplying both sides by $dZ_{\vec{v}'}(\vec{k})^*$, the complex conjugate transpose, taking expectations and using the fact that

$$\begin{aligned} \mathbf{E} \left[\left(\nabla \mathbf{E}[c] \cdot dZ_{\vec{v}'}(\vec{k}) \right) dZ_{\vec{v}'}(\vec{k})^* \right] &= \mathbf{E} \left[dZ_{\vec{v}'}(\vec{k}) dZ_{\vec{v}'}(\vec{k})^* \right]^\dagger \nabla \mathbf{E}[c] \\ &= \left(S_{\vec{v}'\vec{v}'}^\dagger(\vec{k}) \nabla \mathbf{E}[c] \right)^\dagger d\vec{k} \end{aligned}$$

yields

$$\frac{b(t) \left[i\vec{k} \cdot \mathbf{E}[\vec{V}] - \mathbf{D}\vec{k} \cdot \vec{k} \right] \left(S_{\vec{v}'\vec{v}'}^\dagger(\vec{k}) \nabla \mathbf{E}[c] \right)^\dagger}{(\vec{k} \cdot \mathbf{E}[\vec{V}])^2 + (\mathbf{D}\vec{k} \cdot \vec{k})^2} = S_{c'\vec{v}'}^\dagger(\vec{k})$$

where $S_{\vec{v}'\vec{v}'}(\vec{k})$ is the cross-spectrum matrix and $S_{c'\vec{v}'}(\vec{k})$ is the cross-spectrum vector.

Now, taking the inverse Fourier Transform

$$R_{c'\vec{v}'}(\vec{x}) = \int_{\mathfrak{R}^3} e^{i(\vec{k} \cdot \vec{x})} S_{c'\vec{v}'}(\vec{k}) d\vec{k}$$

and letting $\vec{x} = 0$, it follows that

$$\begin{aligned} \mathbf{E}[c'\vec{V}'] = R_{c'\vec{v}'}(\vec{0}) &= \int_{\mathfrak{R}^3} S_{c'\vec{v}'}(\vec{k}) d\vec{k} \\ &= - \int_{\mathfrak{R}^3} \frac{b(t) \left[-i\vec{k} \cdot \mathbf{E}[\vec{V}] + \mathbf{D}\vec{k} \cdot \vec{k} \right] S_{\vec{v}'\vec{v}'}^\dagger(\vec{k}) \nabla \mathbf{E}[c] d\vec{k}}{(\vec{k} \cdot \mathbf{E}[\vec{V}])^2 + (\mathbf{D}\vec{k} \cdot \vec{k})^2} \end{aligned} \quad (1.12)$$

Take the k_j^{th} term of the matrix $S_{\vec{v}'\vec{v}'}^\dagger(\vec{k})$ which, as shown in Section 1.6, has the property that $S_{V'_k V'_j}(\vec{k}) + S_{V'_j V'_k}(\vec{k})$ is an even function, and assume that $S_{V'_k V'_j}(\vec{k}) = S_{V'_j V'_k}(\vec{k})$ and that \mathbf{D} is positive definite, then using the mapping $T(\vec{k}) = -\vec{k} = \vec{\omega}$ and the change of variable formula, it follows that

$$\begin{aligned} 2 \int_{\mathfrak{R}^3} \frac{b(t) \left[-i\vec{k} \cdot \mathbf{E}[\vec{V}] + \mathbf{D}\vec{k} \cdot \vec{k} \right] S_{V'_k V'_j}(\vec{k}) d\vec{k}}{(\vec{k} \cdot \mathbf{E}[\vec{V}])^2 + (\mathbf{D}\vec{k} \cdot \vec{k})^2} = \\ \int_{\mathfrak{R}^3} \frac{b(t) \left[-i\vec{k} \cdot \mathbf{E}[\vec{V}] + \mathbf{D}\vec{k} \cdot \vec{k} \right] S_{V'_k V'_j}(\vec{k}) d\vec{k}}{(\vec{k} \cdot \mathbf{E}[\vec{V}])^2 + (\mathbf{D}\vec{k} \cdot \vec{k})^2} \end{aligned}$$

$$\begin{aligned}
& + \int_{\mathfrak{R}^3} \frac{b(t) \left[-i(\vec{w}) \cdot \mathbf{E}[\vec{V}] + \mathbf{D}(\vec{w}) \cdot (\vec{w}) \right] S_{V'_k V'_j}(\vec{w})}{(\vec{w} \cdot \mathbf{E}[\vec{V}])^2 + (\mathbf{D}(\vec{w}) \cdot (\vec{w}))^2} d\vec{w} \\
= & \int_{\mathfrak{R}^3} \frac{b(t) \left[-i\vec{k} \cdot \mathbf{E}[\vec{V}] + \mathbf{D}\vec{k} \cdot \vec{k} \right] S_{V'_k V'_j}(\vec{k})}{(\vec{k} \cdot \mathbf{E}[\vec{V}])^2 + (\mathbf{D}\vec{k} \cdot \vec{k})^2} d\vec{k} \\
& + \int_{\mathfrak{R}^3} \frac{b(t) \left[i\vec{k} \cdot \mathbf{E}[\vec{V}] + \mathbf{D}\vec{k} \cdot \vec{k} \right] S_{V'_k V'_j}(\vec{k})}{(\vec{k} \cdot \mathbf{E}[\vec{V}])^2 + (\mathbf{D}\vec{k} \cdot \vec{k})^2} d\vec{k} \\
= & 2 \int_{\mathfrak{R}^3} \frac{b(t) (\mathbf{D}\vec{k} \cdot \vec{k}) S_{V'_k V'_j}(\vec{k})}{(\vec{k} \cdot \mathbf{E}[\vec{V}])^2 + (\mathbf{D}\vec{k} \cdot \vec{k})^2} d\vec{k}
\end{aligned}$$

Without the symmetry assumption, this result would contain the term $S_{V'_k V'_j}(\vec{k}) + S_{V'_j V'_k}(\vec{k})$. Returning to the transport equation

$$\frac{\partial c}{\partial t} = -\nabla \cdot [c\vec{V} - \mathbf{D}\nabla c]$$

we know that by writing the concentration and velocity in terms of distributed parameters,

$$c = \mathbf{E}[c] + c' \quad \mathbf{E}[c'] = 0$$

$$\vec{V} = \mathbf{E}[\vec{V}] + \vec{V}' \quad \mathbf{E}[\vec{V}'] = 0$$

we obtain Equation[1.9], page 32, namely,

$$\frac{\partial \mathbf{E}[c]}{\partial t} + \nabla \cdot (\mathbf{E}[c]\mathbf{E}[\vec{V}]) - \nabla \cdot (\mathbf{D}\nabla \mathbf{E}[c]) + \nabla \cdot \mathbf{E}[c'\vec{V}'] = 0$$

The term $\mathbf{E}[c'\vec{V}']$ is an additional dispersive flux term that has already been expressed in terms of $\nabla \mathbf{E}[c]$ via Equation[1.12], page 36. So, the transport equation for $\mathbf{E}[c]$ can be stated in terms of a new dispersion tensor, $\tilde{\mathbf{D}}$, as

$$\frac{\partial \mathbf{E}[c]}{\partial t} + \nabla \cdot (\mathbf{E}[c]\mathbf{E}[\vec{V}]) - \nabla \cdot (\tilde{\mathbf{D}}\nabla \mathbf{E}[c]) = 0$$

where

$$\tilde{\mathbf{D}} = \mathbf{D} + \int_{\mathfrak{R}^3} \frac{b(t) (\mathbf{D}\vec{k} \cdot \vec{k}) S_{\vec{V}'\vec{V}'}(\vec{k})}{(\vec{k} \cdot \mathbf{E}[\vec{V}])^2 + (\mathbf{D}\vec{k} \cdot \vec{k})^2} d\vec{k} \quad (1.13)$$

It is clear from this formulation that in general, $\tilde{\mathbf{D}}$ is a time dependent quantity. If $\vec{k} \in \mathfrak{R}^3$ and if $\mathbf{D}\vec{k} \cdot \vec{k}$ is positive definite, then since $a = [i\vec{k} \cdot \mathbf{E}[V] + \mathbf{D}\vec{k} \cdot \vec{k}]t$,

$$\begin{aligned} \lim_{t \rightarrow \infty} b(t) &= \lim_{t \rightarrow \infty} 1 - e^{-at} \\ &= 1 - \lim_{t \rightarrow \infty} e^{-[\mathbf{D}\vec{k} \cdot \vec{k} + i\vec{k} \cdot \mathbf{E}[V]]t} \\ &= 1 \quad \text{since } \mathbf{D}\vec{k} \cdot \vec{k} \geq 0 \end{aligned}$$

And, assuming that the interchange of limit and integral makes sense, the asymptotic or steady state limit as $t \rightarrow \infty$ for $\tilde{\mathbf{D}}$ is

$$\tilde{\mathbf{D}} = \mathbf{D} + \int_{\mathfrak{R}^3} \frac{(\mathbf{D}\vec{k} \cdot \vec{k}) S_{\vec{V}'\vec{V}'}(\vec{k})}{(\vec{k} \cdot \mathbf{E}[\vec{V}])^2 + (\mathbf{D}\vec{k} \cdot \vec{k})^2} d\vec{k}$$

1.9 Comments And Limitations

The assumption of ergodicity is implicit in the study, *i.e.*, the solute transport in an ensemble of aquifers approximates the real field situation. This means that if the independent variable is a spatial variable, then the variance around the average must go to zero as the size of the domain gets arbitrarily large. Hence, the scale of the system must be large in comparison to the correlation scale, the length scale over which variables remain correlated. So, the estimates of macroscopic dispersivity and effective hydraulic conductivity are meaningful only if the scale of the problem is large in comparison to the correlation scale. Consequently, Equation[1.12], page 36, is valid only after a large displacement distance has been reached, perhaps tens or hundreds of meters.

The adequacy of the first order approximation of the solute transport equation

$$\frac{\partial c'}{\partial t} + \nabla \cdot [\mathbf{E}[c]\vec{V}' + c'\mathbf{E}[\vec{V}]] = \nabla \cdot [\mathbf{D}\nabla(c')]$$

depends on small perturbations in c and \vec{V} and is not certain for large variance of hydraulic conductivity. In this same vein, the assumption that the gradient of the expected concentration is constant also depends on mild heterogeneities.

Existing linear theories predict that transverse dispersivities tend asymptotically to zero as *Fickian* conditions are reached. The assumption of mildly fluctuating hydraulic conductivities has been used to justify eliminating nonlinear terms in establishing the linear theories. However, Rubin[82] found that higher order terms may cause some reduction in longitudinal mixing and a significant enhancement in the transverse spread. Hence, the importance of non-linearities should not be disregarded in studying dispersion in geologic media. Although it is the point of view of this study that significant heterogeneities should be modelled using numerical models, attempts at incorporating heterogeneity and therefore nonlinearities in the analytical models have been made. Neuman and Zhang[70] recover part of the nonlinearity due to the deviation of the plume particles from their mean trajectories using what has become known as *Corrsin's Conjecture*, which is a statement relating the Lagrangian covariance of the velocity to the Eulerian covariance of the velocity through the probability density of the particle's position. From this, nonlinear analytical expressions for the time dependent dispersivity have been developed. The relationship between the Lagrangian and Eulerian velocity fields is key to the developments of dispersion that are to follow.

1.10 Velocity/Permeability Covariance Relationship

Spectral arguments similar to those used in the preceding subsection can also be used to develop a fundamental relationship between the Fourier Transforms of the velocity covariances and the log-hydraulic conductivity covariances, Equation[1.21], page 43. This relationship is important because it shows the connection between velocity covariances and log-hydraulic conductivity covariances, and hence permeability covariances. This means that estimates of dispersion can be based on either velocity covariances or permeability covariances. In essence, the characterization of dispersion as a phenomenon created by permeability says that dispersion is a result of deviations in local permeabilities from a global average permeability.

By Darcy's law

$$\vec{q} = -\mathbf{K}\nabla\phi$$

where \mathbf{K} is the hydraulic conductivity, in general for \mathfrak{R}^2 a second rank symmetric tensor, and $-\nabla\phi$ is the hydraulic gradient. By assuming that the medium is *isotropic*, the second rank tensor can be replaced by a scalar hydraulic conductivity, K .

Let $Y = \ln(K)$ and suppose that

$$Y = \mathbf{E}[Y] + Y'; \quad \phi = \mathbf{E}[\phi] + \phi'; \quad \mathbf{E}[\phi'] = 0; \quad \mathbf{E}[Y'] = 0$$

Then,

$$\vec{q} = -e^Y \nabla(\mathbf{E}[\phi] + \phi')$$

And, since $K = e^Y = e^{\mathbf{E}[Y]} e^{Y'}$, it follows by expanding $e^{Y'}$ that

$$\vec{q} = -e^{\mathbf{E}[Y]} \left(1 + Y' + \frac{(Y')^2}{2} + \dots\right) (\nabla \mathbf{E}[\phi] + \nabla \phi')$$

Letting $\vec{J} = -\nabla \mathbf{E}[\phi]$, then

$$\vec{q} = e^{\mathbf{E}[Y]} \left[(\vec{J} - \nabla \phi') + Y' (\vec{J} - \nabla \phi') + \frac{(Y')^2}{2} (\vec{J} - \nabla \phi') + \dots \right]$$

Take expectations and dropping terms higher than first order, yields *the first order approximation*

$$\mathbf{E}[\vec{q}] \approx e^{\mathbf{E}[Y]} \vec{J}$$

From Darcy's law and the incompressibility condition, respectively,

$$\vec{q} = -K \nabla \phi, \quad \nabla \cdot \vec{q} = 0$$

it follows that since $K = e^Y$, then

$$\nabla \cdot (e^Y \nabla \phi) = 0$$

Expanding this it follows that

$$e^Y \nabla Y \cdot \nabla \phi + e^Y \nabla^2 \phi = 0$$

which yields

$$\nabla Y \cdot \nabla \phi + \nabla^2 \phi = 0 \tag{1.14}$$

Then using the distributed parameters

$$Y = \mathbf{E}[Y] + Y' \quad \phi = \mathbf{E}[\phi] + \phi' \tag{1.15}$$

with

$$\mathbf{E}[\phi'] = 0; \quad \mathbf{E}[Y'] = 0$$

and substituting Equation[1.15] into Equation[1.14], it follows that

$$\nabla(\mathbf{E}[Y] + Y') \cdot \nabla(\mathbf{E}[\phi] + \phi') + \nabla^2(\mathbf{E}[\phi] + \phi') = 0$$

Expanding and treating $\mathbf{E}[Y]$ as a constant, gives

$$\nabla Y' \cdot \nabla \mathbf{E}[\phi] + \nabla Y' \cdot \nabla \phi' + \nabla^2 \mathbf{E}[\phi] + \nabla^2 \phi' = 0$$

And, by retaining only first order perturbation terms, *i.e.* dropping the $\nabla Y' \cdot \nabla \phi'$ term, it follows that

$$\nabla Y' \cdot \nabla \mathbf{E}[\phi] + \nabla^2 \mathbf{E}[\phi] + \nabla^2 \phi' \approx 0$$

Taking expectations and using $\mathbf{E}[Y'] = 0$ and $\mathbf{E}[\phi'] = 0$, results in

$$0 = -\nabla^2 \mathbf{E}[\phi] \approx \nabla Y' \cdot \nabla \mathbf{E}[\phi] + \nabla^2 \phi' \quad (1.16)$$

Substituting in Equation[1.16] the spectral representations

$$Y' = \int_{\mathfrak{R}^3} e^{i\vec{k} \cdot \vec{x}} dZ_{Y'}(\vec{k}) \quad \phi' = \int_{\mathfrak{R}^3} e^{i\vec{k} \cdot \vec{x}} dZ_{\phi'}(\vec{k})$$

it follows that

$$\int_{\mathfrak{R}^3} i\vec{k} \cdot \nabla \mathbf{E}[\phi] e^{i\vec{k} \cdot \vec{x}} dZ_{Y'}(\vec{k}) - \int_{\mathfrak{R}^3} \|\vec{k}\|^2 e^{i\vec{k} \cdot \vec{x}} dZ_{\phi'}(\vec{k}) \approx 0$$

And, using the uniqueness of the spectral representation theorem,

$$i\vec{k} \cdot \nabla \mathbf{E}[\phi] dZ_{Y'}(\vec{k}) \approx \|\vec{k}\|^2 dZ_{\phi'}(\vec{k}) \quad (1.17)$$

Starting from the expansion

$$\vec{q} = -e^{\mathbf{E}[Y]} \left(1 + Y' + \frac{(Y')^2}{2} + \dots \right) (\nabla \mathbf{E}[\phi] + \nabla \phi')$$

and disregarding terms involving the products of perturbed quantities yields

$$\vec{q} \approx -e^{\mathbf{E}[Y]} (\nabla \mathbf{E}[\phi] + \nabla \phi' + Y' \nabla \mathbf{E}[\phi]) \quad (1.18)$$

And, taking expectations and using $\mathbf{E}[Y'] = \mathbf{E}[\phi'] = 0$,

$$\mathbf{E}[\vec{q}] \approx -e^{\mathbf{E}[Y]} \nabla \mathbf{E}[\phi]$$

Then since $\vec{q} = \mathbf{E}[\vec{q}] + \vec{q}'$, it follows from Equation[1.18] that

$$\vec{q}' \approx -e^{\mathbf{E}[Y]} (\nabla \phi' + Y' \nabla \mathbf{E}[\phi])$$

Using the spectral representations

$$\vec{q}' = \int_{\mathfrak{R}^3} e^{i\vec{k} \cdot \vec{x}} dZ_{\vec{q}'}(\vec{k}); \quad Y' = \int_{\mathfrak{R}^3} e^{i\vec{k} \cdot \vec{x}} dZ_{Y'}(\vec{k}); \quad \phi' = \int_{\mathfrak{R}^3} e^{i\vec{k} \cdot \vec{x}} dZ_{\phi'}(\vec{k})$$

it follows that

$$\int_{\mathfrak{R}^3} e^{i\vec{k} \cdot \vec{x}} dZ_{\vec{q}'}(\vec{k}) \approx -e^{\mathbf{E}[Y]} \left(\nabla \int_{\mathfrak{R}^3} e^{i\vec{k} \cdot \vec{x}} dZ_{\phi'}(\vec{k}) + \int_{\mathfrak{R}^3} e^{i\vec{k} \cdot \vec{x}} dZ_{Y'}(\vec{k}) \nabla \mathbf{E}[\phi] \right)$$

By the uniqueness of the spectral representation theorem

$$dZ_{\vec{q}'}(\vec{k}) \approx -e^{\mathbf{E}[Y]} i\vec{k} dZ_{\phi'}(\vec{k}) - e^{\mathbf{E}[Y]} \nabla \mathbf{E}[\phi] dZ_{Y'}(\vec{k})$$

Using the *first order approximation*, $\mathbf{E}[\vec{q}] \approx e^{\mathbf{E}[Y]} \vec{J} = e^{\mathbf{E}[Y]} \nabla \mathbf{E}[\phi]$,

$$dZ_{\vec{q}'}(\vec{k}) \approx \mathbf{E}[\vec{q}] dZ_{Y'}(\vec{k}) - ie^{\mathbf{E}[Y]} \vec{k} dZ_{\phi'}(\vec{k}) \quad (1.19)$$

Then from Equation[1.17]

$$dZ_{\phi'}(\vec{k}) \approx i\|\vec{k}\|^{-2} \vec{k} \cdot \nabla \mathbf{E}[\phi] dZ_{Y'}(\vec{k}) \quad (1.20)$$

Equations [1.19] and [1.20] along with the *first order approximation*, $\mathbf{E}[\vec{q}] \approx e^{\mathbf{E}[Y]} \vec{J}$, and the fact that $\vec{k}(\vec{k} \cdot \mathbf{E}[\vec{q}]) = \vec{k}\vec{k}^\dagger \mathbf{E}[\vec{q}]$ yield

$$dZ_{\vec{q}'}(\vec{k}) \approx \left(\mathbf{I} - \|\vec{k}\|^{-2} \vec{k}\vec{k}^\dagger \right) \mathbf{E}[\vec{q}] dZ_{Y'}(\vec{k})$$

where \mathbf{I} is the identity matrix. Multiplying this expression by its conjugate transpose yields

$$\mathbf{E}[dZ_{\vec{q}'}(\vec{k}) dZ_{\vec{q}'}(\vec{k})^*] \approx \left(\mathbf{I} - \|\vec{k}\|^{-2} \vec{k}\vec{k}^\dagger \right) \mathbf{E}[\vec{q}] \mathbf{E}[\vec{q}]^\dagger \left(\mathbf{I} - \|\vec{k}\|^{-2} \vec{k}\vec{k}^\dagger \right) \mathbf{E}[dZ_{Y'}(\vec{k}) dZ_{Y'}(\vec{k})^*]$$

And, using the formulas, Equation[1.6], page 27,

$$\mathbf{E}[dZ_{\vec{q}}(\vec{k})dZ_{\vec{q}}(\vec{k})^*] = d\mathbf{F}(\vec{k}) = S_{\vec{q}\vec{q}}(\vec{k}) d\vec{k}$$

$$\mathbf{E}[dZ_{Y'}(\vec{k})dZ_{Y'}(\vec{k})^*] = dF(\vec{k}) = S_{Y'Y'}(\vec{k}) d\vec{k}$$

it follows that

$$S_{\vec{q}\vec{q}}(\vec{k}) \approx (\mathbf{I} - \|\vec{k}\|^{-2}\vec{k}\vec{k}^\dagger) \mathbf{E}[\vec{q}]\mathbf{E}[\vec{q}]^\dagger (\mathbf{I} - \|\vec{k}\|^{-2}\vec{k}\vec{k}^\dagger) S_{Y'Y'}(\vec{k}) \quad (1.21)$$

This result agrees with Dagan[31, Equation 4.10].

1.11 Summary

This section introduced some important concepts of stochastic processes that will be used and expanded on in subsequent sections to study the general problem of *scale-up*. Some basic conclusions can be drawn from the presentation in this section. First, from Equation[1.13], page 37, it is clear that in the absence of variations in the velocity field no adjustment to the local dispersion tensor \mathbf{D} is necessary. Secondly, from Equation[1.21], page 43, it is clear that the variations in the velocities are due to the variations in the hydraulic conductivities. Consequently, the conclusion is that the dispersion tensor requires *scaling-up* in the presence of variations in the hydraulic conductivities. Another way of saying this is that the effects of heterogeneities in the porous medium cannot be adequately modeled using a local dispersion tensor only. This means that although a local dispersion tensor may be adequate for describing plume development in small homogeneous laboratory experiments, it must be modified in a way that takes into consideration either variations in the velocities or variations in the hydraulic conductivities if it is to be used to adequately describe plume development in highly heterogeneous field problems.

The linear theories that have been characterized in this section depend on the assumption of only mild heterogeneities being present in the porous medium. However, researchers have found that this restriction may lead to erroneous conclusions regarding both the longitudinal and transverse spread of the plume. For this reason, attempts have been made to modify the linear theories. To this end, Corrsin's conjecture has been borrowed from the field of plasma diffusion to aid in the modification of the linear theories. Corrsin's conjecture relates the Lagrangian velocity covariance to the Eulerian velocity covariance through the probability density of the particles's position.

$$\mathbf{E}[\vec{V}(0)\vec{V}(t)^\dagger] = \int_{\Omega} \mathbf{E}[\vec{V}(\vec{0}, 0)\vec{V}(\vec{y}, t)^\dagger]p(\vec{y}, t)d\vec{y}$$

As will be seen in the next section, the idea of Lagrangian velocity covariance is key to the characterization of dispersion in a highly heterogeneous environment.

Field studies made by various researchers to determine dispersivity values have concluded that:

- Field dispersivity values are larger than laboratory dispersivity values by a few orders of magnitude.
- Dispersivity varies with the distance from the solute input zone.

Consequently, the upscaled dispersion tensor can be represented either as a time dependent quantity, as in the case of Equation[1.13], page 37, or as a distance dependent quantity. Chapter 2 of this study will investigate further the origins of time and distance characterizations of the dispersion tensor. In addition, the next chapter will provide information on some methods that will be used to study the scaling up of dispersion from a stochastic point of view.

2. Dispersivity Coefficients - Time And Distance Forms

2.1 Time Dependent Dispersivity Coefficients

2.1.1 Introduction

In the previous section, it was shown that when heterogenieties are introduced into the porous medium, the component of the transport equation that is significantly affected is the dispersion tensor. And, it was concluded that the dispersion tensor could be characterized as either a time dependent or a distance dependent quantity.

In this section, these characterizations will be developed further, albeit with some simplifying assumptions. The intent is to provide motivation for the formulations of dispersion that will be used in subsequent sections. In Section 2.1.3, a characterization of the dispersion tensor is derived from the transport equation and shown to be equal to half the time rate of change of the second spatial moment tensor. This makes physical sense in that the second spatial moment represents the spread around the centroid of the mass plume. Dispersion, characterized in this manner, then represents how the spread of the plume with respect to the centroid is changing over time. It is also shown that the centroid of the plume moves with the velocity of the flow. So, by using the centroid of the plume as a reference point, this characterization eliminates any changes in the plume due to convective influences. The following is required:

Let Ω be a bounded, open connected domain in \Re^n with a Lipschitz-continuous boundary, $\partial\Omega$. Then the *Fundamental Green's Formula*, integration by parts, is given by

$$\int_{\Omega} \vec{q} \cdot \nabla p d\Omega + \int_{\Omega} p \nabla \cdot \vec{q} d\Omega = \int_{\partial\Omega} p \vec{q} \cdot \vec{\nu} d\gamma \quad (2.1)$$

and follows from the Divergence Theorem

$$\int_{\Omega} \nabla \cdot \vec{\alpha} d\Omega = \int_{\partial\Omega} \vec{\alpha} \cdot \vec{\nu} d\gamma$$

by letting $\vec{\alpha} = p\vec{q}$ and using the expansion

$$\nabla \cdot (p\vec{q}) = \nabla p \cdot \vec{q} + p \nabla \cdot \vec{q} \quad (2.2)$$

2.1.2 Transport Equation

Let $c(\vec{x}, t)$ represent the concentration distribution of a solute, then the mass transport equation of this solute is given by

$$\frac{\partial c}{\partial t} + \vec{V} \cdot \nabla c = \nabla \cdot (\mathbf{D} \nabla c) \quad (2.3)$$

Here \vec{V} is the seepage velocity

$$\vec{V} = - \left(\frac{\mathbf{K}}{n} \right) \nabla \phi$$

where \mathbf{K} is the hydraulic conductivity, n is the porosity and ϕ is the hydraulic head.

Initially, a *divergence free* velocity field is assumed,

$$\nabla \cdot \vec{V} = 0$$

2.1.3 Spatial Moments Of The Solute Concentration

In this section expressions are developed for the first three spatial moments of the solute concentration. In these calculations the porosity, n , is assumed to be constant and included in the definition of $c(\vec{x}, t)$.

Zero Moment

$$M_0 = \int_{\Omega} c(\vec{x}, t) d\Omega$$

Taking the time derivative of this integral, gives

$$\frac{d}{dt} M_0 = \frac{d}{dt} \int_{\Omega} c(\vec{x}, t) d\Omega = \int_{\Omega} \frac{\partial c}{\partial t} d\Omega$$

And, from the mass transport equation and the Divergence theorem

$$\begin{aligned} \frac{d}{dt} M_0 &= - \int_{\Omega} \vec{V} \cdot \nabla c d\Omega + \int_{\Omega} \nabla \cdot (\mathbf{D} \nabla c) d\Omega \\ &= - \left[- \int_{\Omega} c \nabla \cdot \vec{V} d\Omega + \int_{\partial\Omega} c \vec{V} \cdot \vec{\nu} d\gamma \right] \\ &\quad + \int_{\Omega} \nabla \cdot (\mathbf{D} \nabla c) d\Omega \\ &= - \int_{\partial\Omega} c \vec{V} \cdot \vec{\nu} d\gamma + \int_{\partial\Omega} (\mathbf{D} \nabla c) \cdot \vec{\nu} d\gamma \end{aligned}$$

So, the time rate of change of M_0 depends on the boundary conditions. And, if $c = 0$ and $\nabla c = 0$ on $\partial\Omega$, then $\frac{d}{dt}M_0 = 0$ implies M_0 is a constant. Since M_0 is the total mass, this is a statement about the conservation of mass of the solute.

First Moment

$$M_1 = \int_{\Omega} \vec{x}c(\vec{x}, t)d\Omega$$

Taking the time derivative,

$$\frac{d}{dt}M_1 = \frac{d}{dt} \int_{\Omega} \vec{x}c(\vec{x}, t)d\Omega = \int_{\Omega} \vec{x} \frac{\partial c}{\partial t} d\Omega$$

And, from the Transport Equation[2.3], page 46,

$$\frac{d}{dt}M_1 = - \int_{\Omega} \vec{x}(\vec{V} \cdot \nabla c)d\Omega + \int_{\Omega} \vec{x}(\nabla \cdot (\mathbf{D}\nabla c))d\Omega$$

From Equation[2.2], page 45,

$$\vec{V} \cdot \nabla c = \nabla \cdot (c\vec{V}) - c\nabla \cdot \vec{V}$$

By letting $\{\vec{e}_i\}$, $i = 1, \dots, n$ be a standard basis of \mathfrak{R}^n , it follows that

$$\begin{aligned} \frac{d}{dt}M_1 &= - \sum_{i=1}^n \vec{e}_i \int_{\Omega} x_i \nabla \cdot (c\vec{V})d\Omega \\ &+ \int_{\Omega} c\vec{x} \nabla \cdot \vec{V}d\Omega + \sum_{i=1}^n \vec{e}_i \int_{\Omega} x_i (\nabla \cdot (\mathbf{D}\nabla c))d\Omega \end{aligned}$$

Using the assumption of a divergence free velocity field, the second integral on the right hand side of the previous equation vanishes so that

$$\frac{d}{dt}M_1 = - \sum_{i=1}^n \vec{e}_i \int_{\Omega} x_i \nabla \cdot (c\vec{V})d\Omega + \sum_{i=1}^n \vec{e}_i \int_{\Omega} x_i (\nabla \cdot (\mathbf{D}\nabla c))d\Omega \quad (2.4)$$

Let \vec{V} be independent of \vec{x} , then from Green's formula[2.1] the two integrals on the right are evaluated as

$$\begin{aligned} \int_{\Omega} x_i \nabla \cdot (c\vec{V})d\Omega &= - \int_{\Omega} (c\vec{V}) \cdot \nabla x_i d\Omega + \int_{\partial\Omega} x_i (c\vec{V}) \cdot \vec{\nu} d\gamma \\ &= - \int_{\Omega} cV_i d\Omega + \int_{\partial\Omega} x_i (c\vec{V}) \cdot \vec{\nu} d\gamma \\ &= -V_i \int_{\Omega} cd\Omega \end{aligned} \quad (2.5)$$

if $c = 0$ on $\partial\Omega$.

$$\begin{aligned}\int_{\Omega} x_i (\nabla \cdot (\mathbf{D}\nabla c)) d\Omega &= - \int_{\Omega} \mathbf{D}\nabla c \cdot \nabla x_i d\Omega + \int_{\partial\Omega} x_i (\mathbf{D}\nabla c) \cdot \vec{\nu} d\Omega \\ &= - \int_{\Omega} (\mathbf{D}\nabla c)_i d\Omega\end{aligned}$$

if $\nabla c = 0$ on $\partial\Omega$.

Combining these results it follows that

$$\frac{d}{dt} M_1 = \sum_{i=1}^n V_i \vec{e}_i \int_{\Omega} c d\Omega - \sum_{i=1}^n \vec{e}_i \int_{\Omega} (\mathbf{D}\nabla c)_i d\Omega = \vec{V} \int_{\Omega} c d\Omega - \int_{\Omega} \mathbf{D}\nabla c d\Omega$$

Now, if \mathbf{D} is dependent on only t , and letting

$$\mathbf{D} = \begin{bmatrix} \vec{D}_1 \\ \vec{D}_2 \\ \vdots \\ \vec{D}_n \end{bmatrix} \implies \mathbf{D}\nabla c = \begin{bmatrix} \vec{D}_1 \cdot \nabla c \\ \vec{D}_2 \cdot \nabla c \\ \vdots \\ \vec{D}_n \cdot \nabla c \end{bmatrix}$$

it follows that

$$\int_{\Omega} \mathbf{D}\nabla c d\Omega = \int_{\Omega} \begin{bmatrix} \vec{D}_1 \cdot \nabla c \\ \vec{D}_2 \cdot \nabla c \\ \vdots \\ \vec{D}_n \cdot \nabla c \end{bmatrix} d\Omega$$

But, by integrating by parts, each component integral is equal to zero,

$$\int_{\Omega} \vec{D}_i \cdot \nabla c d\Omega = - \int_{\Omega} c \nabla \cdot \vec{D}_i d\Omega + \int_{\partial\Omega} c \vec{D}_i \cdot \vec{\nu} d\gamma = 0 \quad (2.6)$$

since \vec{D}_i is independent of \vec{x} and $c = 0$ on $\partial\Omega$.

Then, from Equations [2.4], [2.5] and [2.6] it follows that

$$\frac{d}{dt} M_1 = \vec{V} \int_{\Omega} c d\Omega$$

And, since $M_0 = \int_{\Omega} c d\Omega$ is a constant, this yields

$$\frac{d}{dt} \left(\frac{M_1}{M_0} \right) = \vec{V} \quad (2.7)$$

But since $\frac{M_1}{M_0}$ is the centroid of the concentration mass, Equation[2.7] says that the centroid of the concentration mass moves with velocity \vec{V} .

Second Moment

The first moment, M_1 , is a vector quantity and the zero moment is a constant. The quantity

$$\vec{R}(t) = \frac{M_1}{M_0}$$

is the *centroid*(*center of mass*) and is also a vector quantity. The second moment about the centroid is given by

$$M_2 = \int_{\Omega} (\vec{x} - \vec{R}(t))(\vec{x} - \vec{R}(t))^{\dagger} c(\vec{x}, t) d\Omega$$

M_2 is a matrix since the quantity

$$\mathbf{S} = (\vec{x} - \vec{R}(t))(\vec{x} - \vec{R}(t))^{\dagger}$$

is a matrix. So, M_2 can be written as

$$M_2 = \int_{\Omega} \mathbf{S} c(\vec{x}, t) d\Omega$$

In order to evaluate the time rate of change of M_2 , first consider the

Diagonal Element: $i = j$

First, from the definition of the centroid and Equation[2.7], page 49,

$$\begin{aligned} \frac{d}{dt} \int_{\Omega} (x_i - R_i)^2 c(\vec{x}, t) d\Omega &= -2V_i \int_{\Omega} (x_i - R_i) c d\Omega + \int_{\Omega} (x_i - R_i)^2 \frac{\partial c}{\partial t} d\Omega \\ &= -2V_i \int_{\Omega} (x_i - R_i) c d\Omega + \int_{\Omega} (x_i - R_i)^2 [-\vec{V} \cdot \nabla c + \nabla \cdot (\mathbf{D} \nabla c)] d\Omega \\ &= -2V_i \int_{\Omega} (x_i - R_i) c d\Omega - \left\{ - \int_{\Omega} c \nabla \cdot (x_i - R_i)^2 \vec{V} d\Omega \right. \\ &\quad \left. + \int_{\partial\Omega} c (x_i - R_i)^2 \vec{V} \cdot \vec{\nu} d\gamma \right\} + \int_{\Omega} (x_i - R_i)^2 \nabla \cdot (\mathbf{D} \nabla c) d\Omega \end{aligned}$$

Now,

$$\int_{\Omega} c \nabla \cdot (x_i - R_i)^2 \vec{V} d\Omega = 2V_i \int_{\Omega} (x_i - R_i) c d\Omega$$

so that the first two terms of the last equality cancel, and using the boundary condition $c = 0$ on $\partial\Omega$, the boundary integral vanishes so that

$$\frac{d}{dt} \int_{\Omega} (x_i - R_i)^2 c(\vec{x}, t) d\Omega = \int_{\Omega} (x_i - R_i)^2 \nabla \cdot (\mathbf{D} \nabla c) d\Omega$$

Integrating by parts and using the boundary condition that $\nabla c = 0$ on $\partial\Omega$,

$$\begin{aligned} \int_{\Omega} (x_i - R_i)^2 \nabla \cdot (\mathbf{D} \nabla c) d\Omega &= - \int_{\Omega} \mathbf{D} \nabla c \cdot \nabla (x_i - R_i)^2 d\Omega + \int_{\partial\Omega} (x_i - R_i)^2 \mathbf{D} \nabla c \cdot \vec{\nu} d\gamma \\ &= -2 \int_{\Omega} \left[\sum_{j=1}^n \mathbf{D}_{1j} \frac{\partial c}{\partial x_j}, \dots, \sum_{j=1}^n \mathbf{D}_{nj} \frac{\partial c}{\partial x_j} \right] \begin{bmatrix} 0 \\ \vdots \\ (x_i - R_i) \\ \vdots \\ 0 \end{bmatrix} d\Omega \\ &= -2 \int_{\Omega} \sum_{j=1}^n \mathbf{D}_{ij} \frac{\partial c}{\partial x_j} (x_i - R_i) d\Omega \\ &= -2 \int_{\Omega} \nabla c \cdot \vec{\alpha} d\Omega \quad \vec{\alpha} = \begin{bmatrix} \mathbf{D}_{i1}(x_i - R_i) \\ \mathbf{D}_{i2}(x_i - R_i) \\ \vdots \\ \mathbf{D}_{in}(x_i - R_i) \end{bmatrix} \\ &= -2 \left[- \int_{\Omega} c \nabla \cdot \vec{\alpha} d\Omega + \int_{\partial\Omega} c \vec{\alpha} \cdot \vec{\nu} d\gamma \right] \\ &= 2 \int_{\Omega} c \mathbf{D}_{ii} d\Omega \quad c(\vec{x}, t) = 0 \quad \vec{x} \in \partial\Omega \end{aligned}$$

Consequently, a diagonal element can be written as

$$\mathbf{D}_{ii} = \frac{1}{2} \frac{d}{dt} \left[\frac{\int_{\Omega} (x_i - R_i)^2 c(\vec{x}, t) d\Omega}{\int_{\Omega} c(\vec{x}, t) d\Omega} \right]$$

Off-Diagonal Element: $i \neq j$

In this case, take the time derivative of the ij^{th} component of M_2 ,

$$\frac{d}{dt} \int_{\Omega} (x_i - R_i)(x_j - R_j) c(\vec{x}, t) d\Omega = \int_{\Omega} \left(-\frac{dR_i}{dt} \right) (x_j - R_j) c(\vec{x}, t) d\Omega$$

$$\begin{aligned}
& + \int_{\Omega} \left(-\frac{dR_j}{dt} \right) (x_i - R_i) c(\vec{x}, t) d\Omega \\
& + \int_{\Omega} (x_i - R_i)(x_j - R_j) \frac{\partial c}{\partial t} d\Omega \\
= & -V_i \int_{\Omega} (x_j - R_j) c d\Omega - V_j \int_{\Omega} (x_i - R_i) c d\Omega \\
& + \int_{\Omega} (x_i - R_i)(x_j - R_j) [-\vec{V} \cdot \nabla c + \nabla \cdot (\mathbf{D} \nabla c)] d\Omega \\
= & -V_i \int_{\Omega} (x_j - R_j) c d\Omega - V_j \int_{\Omega} (x_i - R_i) c d\Omega \tag{2.8} \\
& + \int_{\Omega} (x_i - R_i)(x_j - R_j) \vec{V} \cdot \nabla c d\Omega \\
& + \int_{\Omega} (x_i - R_i)(x_j - R_j) \nabla \cdot (\mathbf{D} \nabla c) d\Omega
\end{aligned}$$

Integrating by parts the integral

$$\begin{aligned}
\int_{\Omega} (x_i - R_i)(x_j - R_j) \vec{V} \cdot \nabla c d\Omega & = - \int_{\Omega} c \nabla \cdot (x_i - R_i)(x_j - R_j) \vec{V} d\Omega \\
& + \int_{\partial\Omega} c (x_i - R_i)(x_j - R_j) \vec{V} \cdot \vec{\nu} d\gamma \\
& = - \int_{\Omega} c [(x_i - R_i)V_j + (x_j - R_j)V_i] d\Omega \tag{2.9}
\end{aligned}$$

And, integrating by parts the last integral in Equation[2.8], it follows from Equation[2.1], page 45,

$$\begin{aligned}
\int_{\Omega} (x_i - R_i)(x_j - R_j) \nabla \cdot (\mathbf{D} \nabla c) d\Omega & = - \int_{\Omega} \mathbf{D} \nabla c \cdot \nabla (x_i - R_i)(x_j - R_j) d\Omega \\
& + \int_{\partial\Omega} (x_i - R_i)(x_j - R_j) \mathbf{D} \nabla c \cdot \vec{\nu} d\gamma \\
= & - \int_{\Omega} \mathbf{D} \nabla c \cdot \vec{\alpha} d\Omega \quad \vec{\alpha} = \begin{bmatrix} 0 \\ \vdots \\ (x_j - R_j) \\ \vdots \\ (x_i - R_i) \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \leftarrow i^{th} \text{ element} \\ \leftarrow j^{th} \text{ element} \end{array}
\end{aligned}$$

$$\begin{aligned}
&= - \int_{\Omega} \left[\sum_{k=1}^n \mathbf{D}_{1k} \frac{\partial c}{\partial x_k}, \dots, \sum_{k=1}^n \mathbf{D}_{nk} \frac{\partial c}{\partial x_k} \right] \begin{bmatrix} 0 \\ \vdots \\ (x_j - R_j) \\ \vdots \\ (x_i - R_i) \\ \vdots \\ 0 \end{bmatrix} d\Omega \\
&= - \int_{\Omega} \left[\sum_{k=1}^n \mathbf{D}_{ik} \frac{\partial c}{\partial x_k} (x_j - R_j) + \sum_{k=1}^n \mathbf{D}_{jk} \frac{\partial c}{\partial x_k} (x_i - R_i) \right] d\Omega
\end{aligned} \tag{2.10}$$

Consider the integral

$$\begin{aligned}
\int_{\Omega} \sum_{k=1}^n \mathbf{D}_{ik} \frac{\partial c}{\partial x_k} (x_j - R_j) d\Omega &= \int_{\Omega} \sum_{k=1}^n \frac{\partial c}{\partial x_k} \mathbf{D}_{ik} (x_j - R_j) d\Omega \\
&= \int_{\Omega} \nabla c \cdot \vec{\beta} d\Omega \quad \vec{\beta} = \begin{bmatrix} \mathbf{D}_{i1} (x_j - R_j) \\ \mathbf{D}_{i2} (x_j - R_j) \\ \vdots \\ \mathbf{D}_{in} (x_j - R_j) \end{bmatrix} \\
&= - \int_{\Omega} c \nabla \cdot \vec{\beta} d\Omega + \int_{\partial\Omega} c \vec{\beta} \cdot \vec{\nu} d\gamma \\
&= - \int_{\Omega} c \mathbf{D}_{ij} d\Omega \quad c(\vec{x}, t) = 0 \quad \vec{x} \in \partial\Omega \\
&= -\mathbf{D}_{ij} \int_{\Omega} c d\Omega
\end{aligned}$$

So, Equation[2.10], page 52, and the fact that $\mathbf{D}_{ij} = \mathbf{D}_{ji}$, yield

$$\int_{\Omega} (x_i - R_i)(x_j - R_j) \nabla \cdot (\mathbf{D} \nabla c) d\Omega = 2\mathbf{D}_{ij} \int_{\Omega} c d\Omega$$

And, from Equations [2.8], [2.9] and [2.10] it follows that

$$\mathbf{D}_{ij}(t) = \frac{1}{2} \frac{d}{dt} \left[\frac{\int_{\Omega} (x_i - R_i)(x_j - R_j) c(\vec{x}, t) d\Omega}{\int_{\Omega} c(\vec{x}, t) d\Omega} \right]$$

Since the mass of the solute is given by

$$M = \int_{\Omega} c(\vec{x}, t) d\Omega$$

this result can be written as

$$\mathbf{D}_{ij}(t) = \frac{1}{2} \frac{d}{dt} \left[\frac{1}{M} \int_{\Omega} (x_i - R_i)(x_j - R_j) c(\vec{x}, t) d\Omega \right]$$

which is in agreement with the definition given by Dagan for the *Actual Dispersion Coefficients*.

2.2 Stochastic Differential Equations

2.2.1 Introduction

In order to proceed further with the investigation of the dispersion tensor in terms of a stochastic analysis, it is necessary to define the type of integration that will be required to analyze certain stochastic differential equations that arise naturally. This will be the subject of Subsections 2.2.2 to 2.2.5. Although the present discussion is very superficial, it will be adequate for the immediate need. However, this material will be expanded later when more details are required.

2.2.2 Integral Of A Stochastic Process

According to Jazwinski[54] a stochastic process $X(t)$ is *mean square Riemann integrable over $[a, b]$* if for

$$a = t_0 < t_1 < \cdots < t_n = b$$

and

$$\rho = \max_i (t_{i+1} - t_i), \quad t_i \leq t'_i < t_{i+1}$$

the following mean squared limit exists

$$(m^2) \lim_{\rho \rightarrow 0} \sum_{i=0}^{n-1} X(t'_i)(t_{i+1} - t_i) = \int_a^b X(t) dt$$

Or,

$$\lim_{\rho \rightarrow 0} \mathbf{E} \left[\left| \sum_{i=0}^{n-1} X(t'_i)(t_{i+1} - t_i) - \int_a^b X(t) dt \right|^2 \right] = 0$$

As an existence theorem for the mean square Riemann integral, the following can be shown to be true, Jazwinski[54].

Theorem: $X(t)$ is mean square Riemann integrable over $[a, b]$ if and only if $\mathbf{E}[X(t)X(\tau)]$ is Riemann integrable over $[a, b] \times [a, b]$.

This notion of integrating a stochastic process is different from *stochastic integration* which will be defined next.

2.2.3 Wiener Process (Brownian Motion)

In Section 1.2, the concepts of a random variable and a stochastic process were defined. In short, a stochastic process can be thought of as a mathematical model that describes the occurrence of a random phenomenon at each point in time subsequent to some initial time. Karatzas and Shreve[57] make the following

Definition: If $(\Omega, \mathcal{B}, \mathcal{P})$ is the probability space on which the stochastic process $X(t, \omega)$ is defined, then $X(t, \omega)$ is *measurable* if the mapping

$$X(t, \omega) : ([0, \infty) \times \Omega, \mathcal{B}([0, \infty)) \otimes \mathcal{B}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$$

is measurable. Here $\mathcal{B}([0, \infty))$ and $\mathcal{B}(\mathbb{R}^d)$ represent the Borel sets of $[0, \infty)$ and \mathbb{R}^d , respectively.

Definition: Given a probability space $(\Omega, \mathcal{B}, \mathcal{P})$, a nondecreasing family $\{\mathcal{B}_t : t \geq 0\}$ of sub- σ -algebras of \mathcal{B} , $\mathcal{B}_s \subset \mathcal{B}_t \subset \mathcal{B}$ for $0 \leq s < t < \infty$ is called a *filtration*.

Definition: The stochastic process $X(t, \omega)$ is *adapted* to the filtration $\{\mathcal{B}_t\}$ if for each $t \geq 0$, $X(t, \omega)$ is a \mathcal{B}_t -measurable random variable.

With these definitions, the one-dimensional Brownian motion can be defined as follows:

Definition: A *one-dimensional Brownian Motion* is a continuous, adapted process $W = \{W_t, \mathcal{B}_t; 0 \leq t < \infty\}$ defined on the probability space $(\Omega, \mathcal{B}, \mathcal{P})$, with the properties:

- (1) $W_0 = 0$ (a.s.)
- (2) For $0 \leq s < t$, the increment $W_t - W_s$ is independent of \mathcal{B}_s and is normally distributed with mean 0 and variance $t - s$.

Definition:

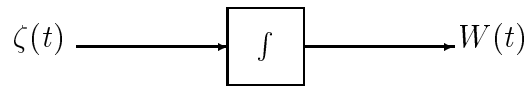
A process $\zeta(t)$ is a *white noise* process if its values $\zeta(t_i)$ and $\zeta(t_j)$ are uncorrelated for every t_i and t_j such that $t_i \neq t_j$. For a white noise process with zero mean,

$$\mathbf{E}[\zeta(t)\zeta(\tau)] = q(t)\delta(t - \tau), \quad q(t) \geq 0$$

The Wiener process can be defined as the limit of a *random walk*, or as the integral of a Gaussian white noise process with zero mean.

$$W(t) = \int_0^t \zeta(s)ds$$

The following is a block diagram representation of this equation:



2.2.4 Stochastic Integration

Because the solutions of the stochastic evolution equations come from spaces whose members are random functions, the solution process will require integrating these functions. For reasons given below, this will require a new type of integration called *stochastic integration*.

Consider an integral of the form

$$\int_0^t B(s)dW(s) \tag{2.11}$$

where $W(s)$ is a *Brownian motion* process.

To illustrate the difficulty involved with interpreting this type of integral in the usual Riemann-Stieltjes sense, it is necessary to define a *sample path* of a Brownian motion process.

Definition: The mapping $t \rightarrow W(t, \omega)$ (ω fixed) is called a *sample path*.

The following facts regarding sample paths of a Brownian motion process are proved in Friedman[40].

- Almost all sample paths of a Brownian motion are nowhere differentiable

- Almost all sample paths of a Brownian motion have infinite variation on any finite interval

A non-rigorous reason for the first bullet is that since

$$\Delta W = \mathcal{O}(\Delta t^{\frac{1}{2}})$$

then

$$\frac{\Delta W}{\Delta t} \approx \frac{1}{\Delta t^{\frac{1}{2}}} \rightarrow \infty \quad \text{as } \Delta t \rightarrow 0$$

Consequently, given the preceding discussion of the interpretation of $W(t)$ as the integral of a Gaussian white noise, the derivative $\frac{dW}{dt}$ must be interpreted as a δ -correlated Gaussian white noise which is a purely mathematical idealization.

So, the integral [2.11] cannot be defined as a Stieltjes integral in the usual sense, for in order to do so, the sample paths would have to have bounded variation.

2.2.5 Types Of Stochastic Integrals

Stochastic integrals have been defined in different ways. Two of these methods of defining the stochastic integral are:

- Itô Integral
 - * Applicable to a larger class of functions
 - * Does not follow the formal rules of calculus
- Stratonovich Integral
 - * Applicable to a restricted class of functions
 - * Follows the formal rules of calculus

Stochastic integrals are defined in the sense of convergence in measure or mean squared convergence. Although the definition of stochastic integrals will be expanded in more detail below, the present discussion is adequate for the immediate purpose of exploring the dispersion tensor.

2.3 The Concentration Equation

Traditionally, the concentration or solute transport equation is developed from the consideration of conservation of mass in terms of variables averaged over a Representative Elementary Volume (REV), Bear[12], Gray[49]. In this section, the solute transport equation will be derived using stochastic considerations. In addition, the all important dispersion tensor will be further scrutinized with the objective in mind of finding expressions for this tensor that can be used in our computational work.

2.3.1 The Lagrangian Approach

By the Dupuit-Forcheimer equation, Bear[12], the filtration velocity, \vec{V} is given in terms of the specific discharge, \vec{q} , and the porosity, n , as

$$\vec{V} = \frac{\vec{q}}{n}$$

In a heterogeneous porous medium, the properties of the medium cannot be precisely known. Hence they are considered to be composed of an average value plus some type of random component that describes the uncertainty associated with the medium. Because of this, both \vec{q} and n are random and so is the velocity \vec{V} . The uncertainty associated with the velocity field of a fluid particle can be illustrated by the situation depicted in Figure 1. The black dot in this figure represents a fluid particle about to begin its journey through a porous medium represented by the open circles. As shown, there are several paths that the particle can ultimately take. And, there is no way of knowing which path will be the actual path.

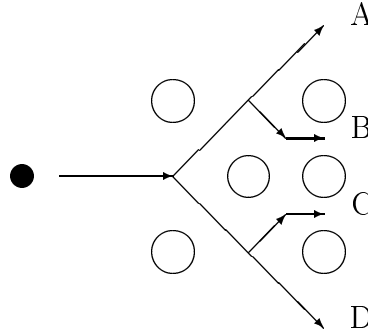


Figure 1

Figure 1 depicts the mechanical mixing component of dispersion. If the velocity field could be perfectly described, *i.e.*, if at each juncture it was known which way the particle was going to go, then there would be no mechanical dispersion to account for, only molecular diffusion. As shown in Chapter 1 and as will be discussed further below, the mechanical dispersion can be accounted for through the velocity covariances, or equivalently, through the permeability covariances.

The velocity covariances are defined as

$$\begin{aligned}\vec{V}' &= \vec{V} - \mathbf{E}[\vec{V}] \\ \rho_{jk}(\vec{x}, \vec{y}) &= \mathbf{E}[\vec{V}'_j(\vec{x})\vec{V}'_k(\vec{y})]\end{aligned}$$

In the Lagrangian framework, transport is developed in terms of indivisible solute particles which are transported by the fluid. As discussed in Bear and Verruijt[11], the solute particles can be thought of as ensembles of molecules in a small volume. If the vector \vec{X}_T represents the total displacement of the particle which started its motion at $\vec{x} = \vec{x}_0$, $t = t_0$, then the vector \vec{X}_T can be decomposed into

$$\underbrace{\vec{X}_T(t; \vec{x}_0, t_0)}_{\text{Total Displacement}} = \underbrace{\vec{X}(t; \vec{x}_0, t_0)}_{\text{Convection}} + \underbrace{\vec{X}_d(t; t_0)}_{\text{Diffusion}} \quad (2.12)$$

In this expression, the first term on the right hand side is due to a mechanical mixing of the fluid and requires movement of the fluid in order to exist. The second term on the right hand side is due to molecular diffusion which can take place without motion of the fluid. Hence, it does not contribute to the velocity field of the fluid particle. It does, however, contribute to the total displacement of the fluid particle. The vector \vec{X}_d represents a *Normalized Brownian motion or Brownian motion with zero drift* type of displacement. Hence, \vec{X}_d can be defined as the integral of a *white noise* process

$$\vec{X}_d(t) = \int_0^t \nu(s) ds \quad \mathbf{E}[\nu] = 0$$

where the autocovariance of the white noise, assuming a constant spectral density $f(\lambda) = K$ and using the integral representation for the delta function, $\delta(\tau) = \frac{1}{2\pi} \int_{\mathbb{R}^1} e^{i\tau\lambda} d\lambda$, is given by

$$\mathbf{E}[\nu(t + \tau)\nu(t)^*] = \int_{\mathbb{R}^1} K e^{i\tau\lambda} d\lambda = 2\pi K \delta(\tau) \equiv \alpha \delta(\tau)$$

Even though $\delta(\tau)$ is not a function in the mathematical sense, the idea is that if $\tau \neq 0$, then the autocovariance is zero.

The vector \vec{X} comes from convective transport and is related to the velocity field by the kinematic relationship

$$\vec{X}(t; \vec{x}_0, t_0) = \int_0^t \vec{V}(\vec{X}_T) dt'$$

$$\begin{aligned}
&= \int_0^t [\mathbf{E}[\vec{V}] + \vec{V}'(\vec{X}_T)] dt' & (2.13) \\
&= \mathbf{E}[\vec{V}]t + \int_0^t \vec{V}'(\vec{X}_T) dt' \\
&\equiv \mathbf{E}[\vec{V}]t + \vec{X}'(t; \vec{x}_0, t_0)
\end{aligned}$$

By taking expectations, it follows that $\mathbf{E}[\vec{X}_T] = \mathbf{E}[\vec{X}] = \mathbf{E}[\vec{V}]t$, and if \vec{X}_T in the integrand of the above integral is replaced by its average, a *first order approximation* of \vec{X}_T in terms of \vec{V}' is obtained, Dagan[33]

$$\vec{X}'(t; \vec{x}_0, t_0) = \int_0^t \vec{V}'(\mathbf{E}[\vec{V}]t') dt'$$

Replacing the vector $\vec{X}_T(t')$ with the vector $\mathbf{E}[\vec{V}]t'$ in essence assumes that the displacements of the velocity about the expected velocity path are not unlike those about the actual trajectory. This type of assumption is similar to that used by Taylor who assumed that since the expected velocity was much greater than the velocity fluctuations, any disturbances or eddies in a wind tunnel were transported with the expected velocity without significant distortion, Monin and Yaglom[68].

The displacement covariance tensor, \mathbf{X}_{jk} , is then expressed in terms of the velocity covariances as

$$\begin{aligned}
\mathbf{X}_{jk}(t; \vec{x}_0, t_0) &= \mathbf{E}[X'_j(t; \vec{x}_0, t_0)X'_k(t; \vec{x}_0, t_0)] \\
&= \int_0^t \int_0^t \rho_{jk}(\mathbf{E}[\vec{V}]t', \mathbf{E}[\vec{V}]t'') dt' dt'' & (2.14)
\end{aligned}$$

The mass of the indivisible particle is denoted m_{fp} , and the solute is assumed to be inert, *i.e.*, it does not react with the fluid that transports it nor with the solid matrix. A solute particle follows a path through the porous medium according to Equation[2.12]. The velocity of the particle is given by

$$\frac{d\vec{X}_T}{dt} = \vec{V} \quad t > 0 \quad (2.15)$$

The concentration field associated with the particle is given by

$$C_{fp}(\vec{x}, t; \vec{x}_0, t_0) = \frac{m_{fp}}{n} \delta(\vec{x} - \vec{X}_T(t; \vec{x}_0, t_0)) \quad (2.16)$$

where n is the effective porosity. The concentration is defined as mass per unit volume, but, only a portion of the unit volume is available for flow, and that portion is given by the porosity. Hence, if the concentration is multiplied by the porosity and integrated over \mathfrak{R}^n , the result is the mass of the fluid particle. From Equation[2.16], this is equal to m_{fp} at each point of the particle's trajectory. This illustrates the indivisible nature of the fluid particle.

The concentration field C_{fp} is a stochastic function, *i. e.*, at each time t , the function $C_{fp}(\vec{x}, t)$ is a random variable of the spatial variable, \vec{x} . This follows from the fact that the trajectory \vec{X}_T is a stochastic function. Hence, if $p(\vec{X}_T, t)$ represents the probability density function of \vec{X}_T , then

$$\begin{aligned} \mathbf{E}[C_{fp}(\vec{x}, t; \vec{x}_0, t_0)] &= \int_{\mathfrak{R}^n} C_{fp} p(\vec{X}_T; t, \vec{x}_0, t_0) d\vec{X}_T \\ &= \int_{\mathfrak{R}^n} \frac{m_{fp}}{n} \delta(\vec{x} - \vec{X}_T) p(\vec{X}_T; t, \vec{x}_0, t_0) d\vec{X}_T \\ &= \frac{m_{fp}}{n} p(\vec{x}; t, \vec{x}_0, t_0) \end{aligned}$$

It is clear from this equation that the probability density function of the concentration field is the same as that of the trajectory \vec{X}_T . So, the concentration field is random because the trajectory is random. As mentioned before, the velocity field is also a stochastic function. From Equation[2.12], page 58, the following stochastic differential equation can be formed

$$\frac{d\vec{X}_T}{dt} = \frac{d\vec{V}}{dt} + \frac{d\vec{X}_d}{dt}$$

And, since \vec{X}_d is a Wiener process, the second term on the right hand side is formally a Gaussian white noise. Furthermore, Equation[2.13], page 59, yields,

$$\frac{d\vec{X}_T}{dt} = \mathbf{E}[\vec{V}] + \vec{V}'(\vec{X}_T) + \frac{d\vec{X}_d}{dt}$$

The resulting stochastic differential equation then takes the form

$$d\vec{X}_T = \left(\mathbf{E}[\vec{V}] + \vec{V}'(\vec{X}_T) \right) dt + d\vec{X}_d \quad (2.17)$$

which is to be interpreted in terms of the stochastic integral equation as

$$\vec{X}_T(t) = \vec{X}_T(0) + \int_0^t \left(\mathbf{E}[\vec{V}] + \vec{V}'(\vec{X}_T) \right) ds + \int_0^t d\vec{X}_d(s)$$

The uncertainty of the position of the fluid particle, \vec{X}_T , can be demonstrated by solving the finite-difference form of this stochastic differential equation, *viz.*, for $t_0 < t_1 < \dots < t_n$,

$$\vec{X}_T(t_{i+1}) = \vec{X}_T(t_i) + \left(\mathbf{E}[\vec{V}] + \vec{V}'(\vec{X}_T(t_i)) \right) \Delta t_i + \Delta \vec{X}_d(i)$$

where

$$\begin{aligned} \Delta t_i &= t_{i+1} - t_i \\ \Delta \vec{X}_d(i) &= \vec{X}_d(t_{i+1}) - \vec{X}_d(t_i) \end{aligned}$$

Figure 2 shows various realizations of the path of a fluid particle. The average velocity vector is along the line $x = y$. What is immediately clear from this figure is that the final position of a fluid particle can vary greatly depending on the exact path taken through the porous medium. And, this is the essence of dispersion. If we knew exactly which path each fluid particle takes, there would be no mechanical dispersion. Since this is impossible to know, the transport equation must include a term to compensate for this uncertainty. This was the case in Chapter 1 where the dispersion tensor was modified to include a component in part described by the spectrum of the velocity covariances.

2.3.2 Basic Form Of Transport Equation

It can be shown, Jazwinski[54], that under certain circumstances the solution process of an Itô stochastic differential equation (SDE) is a Markov process and that the probability density or transition probability density function of the solution process solves the **Fokker-Planck** or **Kolmogorov forward equation**, *i.e.*, given the Itô SDE

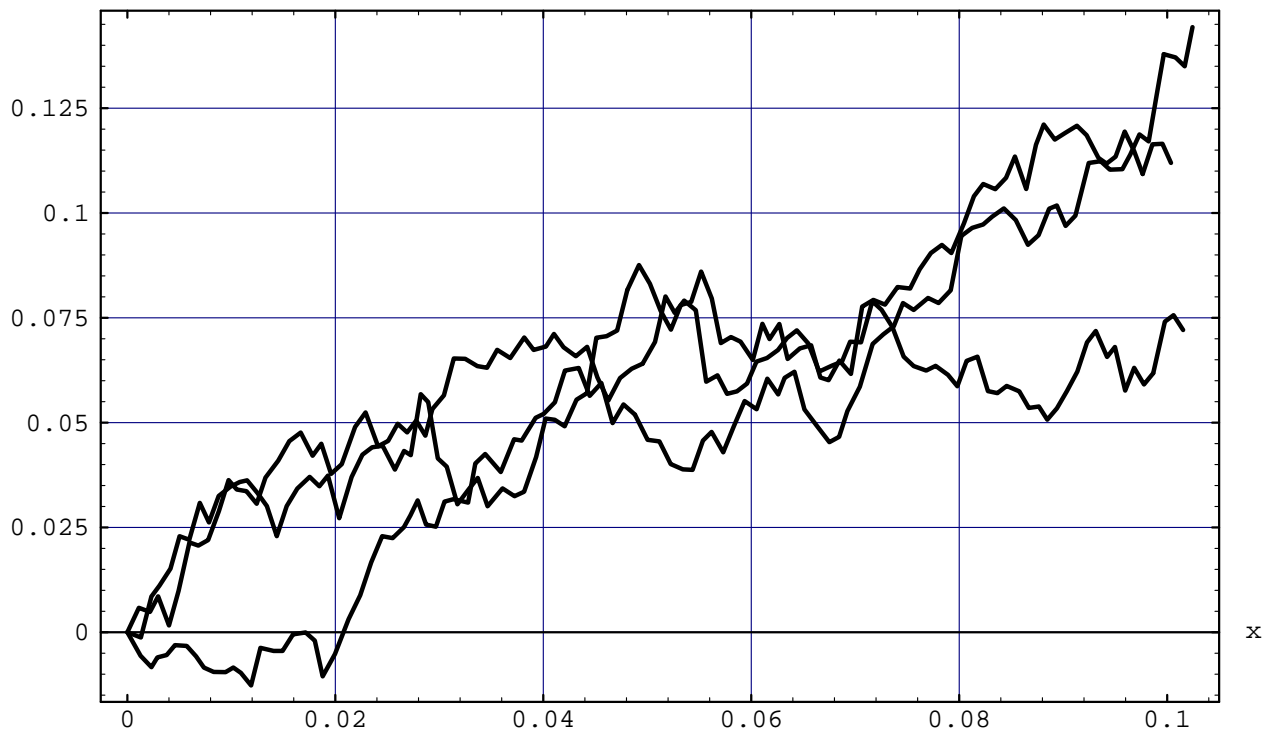
$$d\vec{x} = \vec{a}(\vec{x}, t)dt + \mathbf{B}(\vec{x}, t)d\vec{X}_d$$

where \vec{x} and \vec{a} are n -vectors, \mathbf{B} is an $n \times m$ matrix and \vec{X}_d is an m -vector Wiener process with $\mathbf{Q}(t)dt = \mathbf{E} \left[d\vec{X}_d d\vec{X}_d^\dagger \right]$, then the density function, p , satisfies

$$\frac{\partial p}{\partial t} + \nabla \cdot (p\vec{a}) = \frac{1}{2} \nabla \cdot \nabla (p\mathbf{B}\mathbf{Q}\mathbf{B}^\dagger)$$

It is important to note that the coefficient of the $d\vec{X}_d$ term in Equation[2.17], page 60, which in this case is the identity, is not allowed to be spatially

Figure 2 - Sample Particle Paths



varying. If this term had a spatially varying coefficient, then the form of the Fokker-Planck equation would depend on whether the Itô theory or the Stratonovich theory is applied. For example, Gardiner[42], if the stochastic differential equation (SDE)

$$d\vec{x} = \vec{a}(\vec{x}, t)dt + \mathbf{B}(\vec{x}, t)d\vec{X}_d$$

is an Itô SDE, then the equivalent Stratonovich SDE is given by

$$d\vec{x} = \vec{a}^S(\vec{x}, t)dt + \mathbf{B}^S(\vec{x}, t)d\vec{X}_d$$

where

$$\vec{a}_i^S = \vec{a}_i - \frac{1}{2} \sum_{j,k} \mathbf{B}_{kj} \partial_k \mathbf{B}_{ij}$$

$$\mathbf{B}_{ij}^S = \mathbf{B}_{ij}$$

And, the form of the Fokker-Planck equation depends on which approach is used.

If, for simplicity, it is assumed that $\vec{V}'(\vec{X}_T) = 0$ in Equation[2.17], page 60, then the resulting stochastic differential equation is

$$d\vec{X}_T = \mathbf{E}[\vec{V}]dt + d\vec{X}_d$$

Given that the solution process \vec{X}_T is a Markov process, then the probability density function, p , satisfies

$$\frac{\partial p}{\partial t} + \nabla \cdot (p\mathbf{E}[\vec{V}]) = \nabla \cdot \left(\frac{1}{2} \mathbf{Q}(t) \nabla p \right) \quad (2.18)$$

So, if we let $\mathbf{D} = \frac{1}{2} \mathbf{Q}(t)$, then the tensor \mathbf{D} is either a constant or a time dependent quantity. To be accurate, Equation[2.18] represents a diffusion process with *drift* coefficient $\mathbf{E}[\vec{V}]$ and a *diffusion* coefficient \mathbf{D} .

Obviously, this equation has the form of the transport equation. And, the expected value of the concentration field solves the same equation, *e.g.*,

$$\frac{\partial \mathbf{E}[C]}{\partial t} + \nabla \cdot (\mathbf{E}[C] \mathbf{E}[\vec{V}]) = \nabla \cdot (\mathbf{D} \nabla \mathbf{E}[C])$$

Hence, the basic form of the concentration equation follows from the fundamental displacement Equation[2.12], page 58 and the associated stochastic differential equation.

Clearly, this equation accomodates only the dispersion created by the Brownian motion term, X_d . Whereas, according to Equation[2.12], page 58, the total dispersion is going to come from that associated with the convective transport and that associated with the Brownian motion.

2.3.3 Solution Of Basic Transport Equation

In general terms, the multivariate Gaussian probability density function (pdf) for n dependent random variables is given by

$$f(z_1, z_2, \dots, z_n) = \frac{|\mathbf{V}|^{-\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \exp \left\{ -\frac{1}{2} (\vec{z} - \vec{\mu})^\dagger (\mathbf{V}^{-1}) (\vec{z} - \vec{\mu}) \right\} \quad (2.19)$$

where

$$\vec{z} = (z_1, z_2, \dots, z_n)^\dagger$$

$$\vec{\mu} = (\mathbf{E}[z_1], \mathbf{E}[z_2], \dots, \mathbf{E}[z_n])^\dagger$$

$$\mathbf{V} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ & & \vdots & \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_N^2 \end{pmatrix}$$

so that \mathbf{V} is the variance-covariance matrix. And, the multivariate characteristic function for the Gaussian probability density is given by

$$\phi(\zeta_1, \zeta_2, \dots, \zeta_n) = \exp \left(-\frac{1}{2} \vec{\zeta}^\dagger \mathbf{V} \vec{\zeta} \right) \exp \left(i \vec{\zeta}^\dagger \vec{\mu} \right) \quad (2.20)$$

In theory, the characteristic function of a random variable is given by

$$\hat{f}(\vec{\omega}) = \int_{\mathfrak{R}^n} f(\vec{x}) e^{i\vec{\omega} \cdot \vec{x}} d\vec{x}$$

Clearly, \hat{f} is the Fourier Transform of the function $f(\vec{x})$ so that

$$f(\vec{\omega}) = \frac{1}{(2\pi)^n} \int_{\mathfrak{R}^n} \hat{f}(\vec{\omega}) e^{-i\vec{\omega} \cdot \vec{x}} d\vec{\omega}$$

Given that the trajectory \vec{X}_T is given by Equation[2.12], page 58, and assuming the the probability density function is Gaussian and has the form

$$p(\vec{X}_T, t) = \frac{|\mathbf{V}|^{-\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \exp \left\{ -\frac{1}{2} (\vec{X}_T - \mathbf{E}[\vec{V}]t)^\dagger (\mathbf{V}^{-1}) (\vec{X}_T - \mathbf{E}[\vec{V}]t) \right\}$$

then this pdf solves a non-divergence form of Equation[2.18], page 63, *e.g.*,

$$\frac{\partial p}{\partial t} + \mathbf{E}[\vec{V}] \cdot \nabla p = \nabla \cdot \mathbf{D} \nabla p$$

To see this, first from Ortega[73]:

If \mathbf{D} is a real symmetric matrix, then there is an orthogonal matrix \mathbf{P} whose columns are the eigenvectors of \mathbf{D} , such that

$$\mathbf{P}^\dagger \mathbf{D} \mathbf{P} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = \tilde{\mathbf{D}}$$

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of \mathbf{D} . The change of variables

$$\vec{x} = \mathbf{P}^\dagger \vec{y}$$

has the effect of aligning the principle axes of the matrix \mathbf{D} with the coordinate axes and \mathbf{D} becomes a diagonal matrix.

In this case, the non-divergence form of the transport equation in the original \vec{y} system, *viz.*,

$$\frac{\partial p}{\partial t} + \mathbf{E}[\vec{V}] \cdot \nabla p = \nabla \cdot \mathbf{D} \nabla p$$

becomes in the rotated \vec{x} system

$$\frac{\partial p}{\partial t} + \widetilde{\mathbf{E}}[\vec{V}] \cdot \nabla p = \sum_{i=1}^n \tilde{\mathbf{D}}_{ii} \frac{\partial^2 p}{\partial x_i^2} \quad \widetilde{\mathbf{E}}[\vec{V}] = \mathbf{P}^\dagger \mathbf{E}[\vec{V}]$$

And, by making the change of variables to a moving coordinate system,

$$X_1 = x_1 - \widetilde{\mathbf{E}}[\vec{V}_1]t; X_2 = x_2 - \widetilde{\mathbf{E}}[\vec{V}_2]t; \dots; X_n = x_n - \widetilde{\mathbf{E}}[\vec{V}_n]t; T = t$$

and,

$$\nu(\vec{X}, T) = p(\vec{x}, t)$$

then it follows that

$$\begin{aligned} \frac{\partial p}{\partial t} &= \frac{\partial \nu}{\partial t} = \frac{\partial \nu}{\partial T} - \mathbf{E}[\widetilde{\vec{V}}] \cdot \nabla_{\vec{X}} \nu \\ \nabla_{\vec{x}} p &= \nabla_{\vec{x}} \nu = \nabla_{\vec{X}} \nu \\ \frac{\partial^2 p}{\partial x_i^2} &= \frac{\partial^2 \nu}{\partial X_i^2} \end{aligned}$$

So that the problem becomes

$$\begin{aligned} \frac{\partial \nu}{\partial T} &= \sum_{i=1}^n \mathbf{D}_{ii} \frac{\partial^2 \nu}{\partial X_i \partial X_i} = \tilde{\mathbf{D}} \Delta \nu \quad \vec{X} \in \mathfrak{R}^n, \quad T > 0 \\ \nu(\vec{X}, 0) &= f(\vec{X}) \end{aligned}$$

This equation is transformed by forming

$$0 = \int_{\mathfrak{R}^n} e^{i\vec{\omega} \cdot \vec{X}} \left[\frac{\partial \nu}{\partial T} - \tilde{\mathbf{D}} \Delta \nu \right] d\vec{X}$$

Assuming that $\nu(\vec{X}, T)$ and $\nabla \nu(\vec{X}, T) \rightarrow 0$ as $|\vec{X}| \rightarrow \infty$ and integrating by parts, it follows that

$$\int_{\mathfrak{R}^n} \nu_{X_j X_j} e^{i\vec{X} \cdot \vec{\omega}} d\vec{X} = -\omega_j^2 \int_{\mathfrak{R}^n} e^{i\vec{X} \cdot \vec{\omega}} \nu(\vec{X}, T) d\vec{X}$$

So, letting $\nu_T = \frac{\partial \nu}{\partial T}$

$$0 = \hat{\nu}_T(\vec{\omega}, T) + \vec{\omega}^\dagger \tilde{\mathbf{D}} \vec{\omega} \hat{\nu}(\vec{\omega}, T)$$

And, by transforming the initial condition, the following ordinary differential equation results

$$\begin{cases} \hat{\nu}_T(\vec{\omega}, T) &= -\vec{\omega}^\dagger \tilde{\mathbf{D}} \vec{\omega} \hat{\nu}(\vec{\omega}, T) \\ \hat{\nu}(\vec{\omega}, 0) &= \hat{f}(\vec{\omega}) \end{cases} \quad (2.21)$$

which has the solution

$$\hat{\nu}(\vec{\omega}, T) = \hat{f}(\vec{\omega})e^{-\vec{\omega}^\dagger \tilde{\mathbf{D}}\vec{\omega}T}$$

$\nu(\vec{X}, T)$ can be retrieved by writing

$$\begin{aligned}\nu(\vec{X}, T) &= \frac{1}{(2\pi)^n} \int_{\mathfrak{R}^n} \hat{\nu}(\vec{\omega}, T) e^{-i\vec{X} \cdot \vec{\omega}} d\vec{\omega} \\ &= \frac{1}{(2\pi)^n} \int_{\mathfrak{R}^n} e^{-i\vec{X} \cdot \vec{\omega} - \vec{\omega}^\dagger \tilde{\mathbf{D}}\vec{\omega}T} \hat{f}(\vec{\omega}) d\vec{\omega}\end{aligned}$$

And, using the transform

$$\hat{f}(\vec{\omega}) = \int_{\mathfrak{R}^n} e^{i\vec{\omega} \cdot \vec{Y}} f(\vec{Y}) d\vec{Y}$$

it follows that

$$\nu(\vec{X}, T) = \frac{1}{(2\pi)^n} \int_{\mathfrak{R}^n} \int_{\mathfrak{R}^n} e^{-i\vec{\omega} \cdot (\vec{X} - \vec{Y}) - \vec{\omega}^\dagger \tilde{\mathbf{D}}\vec{\omega}T} d\vec{\omega} f(\vec{Y}) d\vec{Y} \quad (2.22)$$

Next, let $\vec{\beta} = \vec{X} - \vec{Y}$ and write the inside integral as

$$\int_{\mathfrak{R}^n} e^{-i\vec{\omega} \cdot \vec{\beta} - \vec{\omega}^\dagger \tilde{\mathbf{D}}\vec{\omega}T} d\vec{\omega}$$

and let

$$\hat{g}(\vec{\omega}) = e^{-\vec{\omega}^\dagger \tilde{\mathbf{D}}\vec{\omega}T}$$

then

$$\begin{aligned}g(\vec{\beta}) &= \frac{1}{(2\pi)^n} \int_{\mathfrak{R}^n} e^{-i\vec{\omega} \cdot \vec{\beta} - \vec{\omega}^\dagger \tilde{\mathbf{D}}\vec{\omega}T} d\vec{\omega} \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega_1 \beta_1 - \tilde{\mathbf{D}}_{11} \omega_1^2 T} d\omega_1 \cdots \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega_n \beta_n - \tilde{\mathbf{D}}_{nn} \omega_n^2 T} d\omega_n\end{aligned}$$

Then, using the integral from Guenther and Lee[51, p. 167]

$$\int_{-\infty}^{\infty} e^{-i\omega\beta - D\omega^2 T} d\omega = \left(\frac{\pi}{DT}\right)^{\frac{1}{2}} e^{-\frac{\beta^2}{4DT}}$$

it follows that

$$\begin{aligned}
g(\vec{\beta}) &= \frac{1}{(2\pi)^n} \frac{\pi^{\frac{n}{2}}}{(\tilde{\mathbf{D}}_{11} \cdots \tilde{\mathbf{D}}_{nn})^{\frac{1}{2}} T^{\frac{n}{2}}} e^{-\sum_{i=1}^n \left(\frac{\beta_i^2}{4\tilde{\mathbf{D}}_{ii}T} \right)} \\
&= \frac{1}{(4\pi T)^{\frac{n}{2}} (\tilde{\mathbf{D}}_{11} \cdots \tilde{\mathbf{D}}_{nn})^{\frac{1}{2}}} e^{-\sum_{i=1}^n \left(\frac{\beta_i^2}{4\tilde{\mathbf{D}}_{ii}T} \right)}
\end{aligned}$$

Consequently, Equation[2.22] can be written as

$$\nu(\vec{X}, T) = \int_{\mathfrak{R}^n} K(\vec{X} - \vec{Y}, \tilde{\mathbf{D}}_{11} \cdots \tilde{\mathbf{D}}_{nn}T) f(\vec{Y}) d\vec{Y}$$

where

$$K(\vec{\beta}, \tilde{\mathbf{D}}_{11} \cdots \tilde{\mathbf{D}}_{nn}T) = \frac{1}{(4\pi T)^{\frac{n}{2}} (\tilde{\mathbf{D}}_{11} \cdots \tilde{\mathbf{D}}_{nn})^{\frac{1}{2}}} e^{-\sum_{i=1}^n \left(\frac{\beta_i^2}{4\tilde{\mathbf{D}}_{ii}T} \right)}$$

Hence, if $f(\vec{Y}) = \delta(\vec{Y})$, then

$$\begin{aligned}
\nu(\vec{X}, T) &= \int_{\mathfrak{R}^n} K(\vec{X} - \vec{Y}, \tilde{\mathbf{D}}_{11} \cdots \tilde{\mathbf{D}}_{nn}T) \delta(\vec{Y}) d\vec{Y} \\
&= K(\vec{X}, \tilde{\mathbf{D}}_{11} \cdots \tilde{\mathbf{D}}_{nn}T) \\
&= \frac{1}{(4\pi T)^{\frac{n}{2}} (\tilde{\mathbf{D}}_{11} \cdots \tilde{\mathbf{D}}_{nn})^{\frac{1}{2}}} e^{-\sum_{i=1}^n \left(\frac{X_i^2}{4\tilde{\mathbf{D}}_{ii}T} \right)} \\
&= \frac{1}{(2\pi)^{\frac{n}{2}} (2\tilde{\mathbf{D}}_{11}T \cdots 2\tilde{\mathbf{D}}_{nn}T)^{\frac{1}{2}}} e^{-\frac{1}{2} \left(\frac{1}{2T} \vec{X}^\dagger \tilde{\mathbf{D}}^{-1} \vec{X} \right)} \\
&= \frac{1}{(2\pi)^{\frac{n}{2}} (2\tilde{\mathbf{D}}_{11}t \cdots 2\tilde{\mathbf{D}}_{nn}t)^{\frac{1}{2}}} e^{-\frac{1}{2} \left[\frac{1}{2t} (\vec{x} - \mathbf{E}[\vec{V}]t)^\dagger \tilde{\mathbf{D}}^{-1} (\vec{x} - \mathbf{E}[\vec{V}]t) \right]} \\
&= \frac{1}{(2\pi)^{\frac{n}{2}} (2\tilde{\mathbf{D}}_{11}t \cdots 2\tilde{\mathbf{D}}_{nn}t)^{\frac{1}{2}}} e^{-\frac{1}{2} \left[(\mathbf{P}^\dagger (\vec{y} - \mathbf{E}[\vec{V}]t))^\dagger \frac{1}{2t} \tilde{\mathbf{D}}^{-1} (\mathbf{P}^\dagger (\vec{y} - \mathbf{E}[\vec{V}]t)) \right]} \\
&= \frac{1}{(2\pi)^{\frac{n}{2}} (2\tilde{\mathbf{D}}_{11}t \cdots 2\tilde{\mathbf{D}}_{nn}t)^{\frac{1}{2}}} e^{-\frac{1}{2} \left[(\vec{y} - \mathbf{E}[\vec{V}]t)^\dagger \mathbf{P} \frac{1}{2t} \tilde{\mathbf{D}}^{-1} \mathbf{P}^\dagger (\vec{y} - \mathbf{E}[\vec{V}]t) \right]}
\end{aligned}$$

Since, $\mathbf{P} \frac{1}{2t} \tilde{\mathbf{D}}^{-1} \mathbf{P}^\dagger = \frac{1}{2t} \mathbf{D}^{-1}$ and since $\det(\mathbf{P}^\dagger \mathbf{P}) = \det(\mathbf{I}) = 1$, then

$$p(\vec{y}, t) = \frac{|2t\mathbf{D}|^{-\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}[(\vec{y} - \mathbf{E}[\vec{V}]t)^\dagger \frac{1}{2t} \mathbf{D}^{-1} (\vec{y} - \mathbf{E}[\vec{V}]t)]}$$

For given \vec{y} , t , and $\mathbf{E}[\vec{V}]$, the probability $p(\vec{y}, t)$ is then linked to the size of the components of the \mathbf{D} tensor; which indicates spreading due to dispersion. Comparing this equation with Equation[2.19], page 64, it is seen that $p(\vec{x}, t)$ is a Gaussian density function with mean $\mathbf{E}[\vec{V}]t$ and covariance matrix $\mathbf{V} = 2t\mathbf{D}$, from which it follows immediately that

$$\mathbf{D}_{ij} = \frac{1}{2} \dot{\mathbf{V}}_{ij} = \left[\frac{1}{2} \frac{d}{dt} \text{Cov}(\vec{X}_{Ti}, \vec{X}_{Tj}) \right] = \frac{1}{2} \frac{d\mathbf{X}_{Tij}}{dt} \quad (2.23)$$

This method does not work if the matrix \mathbf{D} is allowed to be a function of t , for then $\mathbf{E}[\vec{V}]$ would depend on t also by virtue of the dependence of \mathbf{P}^\dagger on t . However, in the case that \mathbf{D} does depend on t , if it is assumed that p is a Gaussian density function, then it can be shown by direct differentiation that in order for p to satisfy the transport equation it is necessary that Equation[2.23] hold.

2.3.4 Dispersion As Velocity Covariances

If \vec{X}' and \vec{X}_d are not correlated, then it follows from Equation[2.14], page 59, and Equation[2.23] that

$$\begin{aligned} \mathbf{D}_{ij} &= \frac{1}{2} \frac{d\mathbf{X}_{Tij}}{dt} = \frac{1}{2} \frac{d\mathbf{X}_{ij}}{dt} + \frac{1}{2} \frac{d\mathbf{X}_{dij}}{dt} \\ &= \frac{1}{2} \frac{d}{dt} \int_0^t \int_0^t \rho_{ij}(\mathbf{E}[\vec{V}]t', \mathbf{E}[\vec{V}]t'') dt' dt'' + \frac{1}{2} \frac{d\mathbf{X}_{dij}}{dt} \end{aligned}$$

Letting

$$f(t'', t) = \int_0^t \rho_{ij}(\mathbf{E}[\vec{V}]t', \mathbf{E}[\vec{V}]t'') dt'$$

then $\mathbf{X}_{ij} = \int_0^t f(t'', t) dt''$ and by the Liebnitz formula

$$\frac{d\mathbf{X}_{ij}}{dt} = f(t, t) + \int_0^t \frac{\partial}{\partial t} f(t'', t) dt''$$

$$\begin{aligned}
&= f(t, t) + \int_0^t \rho_{ij}(\mathbf{E}[\vec{V}]t, \mathbf{E}[\vec{V}]t'') dt'' \\
&= \int_0^t \rho_{ij}(\mathbf{E}[\vec{V}]t', \mathbf{E}[\vec{V}]t) dt' + \int_0^t \rho_{ij}(\mathbf{E}[\vec{V}]t, \mathbf{E}[\vec{V}]t'') dt''
\end{aligned}$$

Furthermore, it is assumed that a type of stationarity or homogeneity (Section 1.2 or Section 1.6) in the sense that $\mathbf{E}[\vec{V}]$ is constant and

$$\begin{aligned}
\rho_{ij}(\mathbf{E}[\vec{V}]t', \mathbf{E}[\vec{V}]t) &= \mathbf{E} [\vec{V}'_i(\mathbf{E}[\vec{V}]t') \vec{V}'_j(\mathbf{E}[\vec{V}]t)] \\
&= \mathbf{E} [\vec{V}'_i(\mathbf{E}[\vec{V}]t) \vec{V}'_j(\mathbf{E}[\vec{V}]t)] \\
&= \rho_{ij}(\mathbf{E}[\vec{V}]t, \mathbf{E}[\vec{V}]t)
\end{aligned}$$

This means that ρ_{ij} depends only on the separation vector $\mathbf{E}[\vec{V}](t' - t)$. Also, since the stochastic processes considered are all real,

$$\begin{aligned}
\rho_{ij}(\mathbf{E}[\vec{V}]t', \mathbf{E}[\vec{V}]t) &= \mathbf{E} [\vec{V}'_i(\mathbf{E}[\vec{V}]t') \vec{V}'_j(\mathbf{E}[\vec{V}]t)] \\
&= \rho_{ij}(\mathbf{E}[\vec{V}]t, \mathbf{E}[\vec{V}]t')
\end{aligned}$$

which leads to the conclusion that

$$\frac{1}{2} \frac{d\mathbf{X}_{ij}}{dt} = \int_0^t \rho_{ij}(\mathbf{E}[\vec{V}](t' - t)) dt'$$

And, if the mapping $\vec{\sigma}$ is defined as

$$\begin{aligned}
\vec{\sigma} : [0, t] &\rightarrow \mathfrak{R}^n \ni \\
\vec{\sigma}(t') &= \mathbf{E}[\vec{V}]t'
\end{aligned}$$

then it follows from the definition of a path integral that

$$\begin{aligned}
\int_{\vec{\sigma}} \rho_{ij} ds &= \int_0^t \rho_{ij}(\vec{\sigma}(t'), \mathbf{E}[\vec{V}]t) \|\vec{\sigma}'(t')\| dt' \\
&= \|\mathbf{E}[\vec{V}]\| \int_0^t \rho_{ij}(\mathbf{E}[\vec{V}]t', \mathbf{E}[\vec{V}]t) dt'
\end{aligned}$$

Hence, $\int_0^t \rho_{ij}(\mathbf{E}[\vec{V}]t', \mathbf{E}[\vec{V}]t)dt'$ is a path integral. And, consequently,

$$\frac{1}{2} \frac{d\mathbf{X}_{ij}}{dt} = \int_0^t \rho_{ij}(\mathbf{E}[\vec{V}]t', \mathbf{E}[\vec{V}]t)dt' = \frac{1}{\|\mathbf{E}[\vec{V}]\|} \int_{\vec{\sigma}} \rho_{ij}ds$$

If the true Lagrangian path integral is given by

$$\int_{\vec{\sigma}_T} \rho_{ij}ds$$

then the path σ is an approximation to the path σ_T so that for the true Lagrangian path

$$\begin{aligned} \int_{\vec{\sigma}_T} \rho_{ij}ds &= \int_0^t \rho_{ij}(\vec{\sigma}_T(t'), \vec{\sigma}_T(t)) \|\vec{\sigma}'_T(t')\| dt' \\ &\approx \int_0^t \rho_{ij}(\mathbf{E}[\vec{V}]t', \mathbf{E}[\vec{V}]t) \|\mathbf{E}[\vec{V}]\| dt' \end{aligned}$$

so that

$$\mathbf{D}_{ij} = \frac{1}{2} \frac{d\mathbf{X}_{ij}}{dt} = \int_0^t \rho_{ij}(\mathbf{E}[\vec{V}]t', \mathbf{E}[\vec{V}]t)dt' \approx \frac{1}{\|\mathbf{E}[\vec{V}]\|} \int_{\vec{\sigma}_T} \rho_{ij}ds \quad (2.24)$$

This result says that the components of the dispersion tensor should be related to time rate of changes in the covariances of the position of a fluid particle or the time integral of the velocity covariances. Furthermore, it also says that the calculation of the components of the dispersion tensor should take into consideration the orientation of the path of the fluid particle.

In the pure Lagrangian sense, it is not assumed that the expected velocity is constant along the fluid particle's path. If \vec{x} and \vec{y} represent two points on the particle path, then

$$\rho_{jk}(\vec{x}, \vec{y}) = \mathbf{E} \left[\left(\vec{V}_j(\vec{x}) - \mathbf{E}[\vec{V}_j(\vec{x})] \right) \left(\vec{V}_k(\vec{y}) - \mathbf{E}[\vec{V}_k(\vec{y})] \right) \right]$$

Using the kinematic relationship

$$\vec{X}(t; \vec{x}_0, t_0) = \int_0^t \vec{V}(\vec{X}_T(t'))dt'$$

it follows that $\mathbf{E}[\vec{X}(t; \vec{x}_0, t_0)] = \int_0^t \mathbf{E}[\vec{V}(\vec{X}_T(t'))]dt'$. The convective displacement is given by

$$\begin{aligned}
\vec{X}'(t; \vec{x}_0, t_0) &= \vec{X}(t; \vec{x}_0, t_0) - \mathbf{E}[\vec{X}(t; \vec{x}_0, t_0)] \\
&= \int_0^t \left(\vec{V}(\vec{X}_T(t')) - \mathbf{E}[\vec{V}(\vec{X}_T(t'))] \right) dt'
\end{aligned}$$

And, the displacement covariance is given by

$$\mathbf{X}_{jk}(t) = \int_0^t \int_0^t \rho_{jk}(\vec{X}_T(t'), \vec{X}_T(t'')) dt' dt''$$

And, by differentiation,

$$\mathbf{D}_{jk} = \frac{1}{2} \frac{d}{dt} \mathbf{X}_{jk} = \int_0^t \rho_{jk}(\vec{X}_T(t'), \vec{X}_T(t)) dt'$$

The integrand represents a *lagged* covariance. If $s = t - t'$, then $t' = t - s$ and

$$\begin{aligned}
\rho_{jk}(\vec{X}_T(t'), \vec{X}_T(t)) &= \rho_{jk}(\vec{X}_T(t-s), \vec{X}_T(t)) \\
&= \mathbf{E} \left[\left(\vec{V}_j(\vec{X}_T(t-s)) - \mathbf{E}[\vec{V}_j(\vec{X}_T(t-s))] \right) \right. \\
&\quad \left. \times \left(\vec{V}_k(\vec{X}_T(t)) - \mathbf{E}[\vec{V}_k(\vec{X}_T(t))] \right) \right]
\end{aligned}$$

In order to implement this type of dispersion estimate, the paths that the fluid particles take have to be identified. This can be done by generating velocity fields and computing the paths using a particle tracking algorithm such as

$$\vec{X}_T((k+1)\Delta t) = \vec{X}_T(k\Delta t) + \vec{V}(\vec{X}_T(k\Delta t))\Delta t \quad k = 0, 1, \dots, n-1$$

Then by choosing an appropriate Δt so that $t = n\Delta t$ where n is an integer and letting $s = m\Delta t$, it follows that $t' = t - s = (n-m)\Delta t$. The formula for \mathbf{D}_{jk} can be approximated by

$$\begin{aligned}
\mathbf{D}_{jk} = \frac{1}{2} \frac{d}{dt} \mathbf{X}_{jk} &\approx \sum_{m=0}^n \mathbf{E} \left[\left(\vec{V}_j(\vec{X}_T((n-m)\Delta t)) - \mathbf{E}[\vec{V}_j(\vec{X}_T((n-m)\Delta t))] \right) \right. \\
&\quad \left. \times \left(\vec{V}_k(\vec{X}_T(n\Delta t)) - \mathbf{E}[\vec{V}_k(\vec{X}_T(n\Delta t))] \right) \right] \Delta t
\end{aligned}$$

The velocities in this representation are elements of the Lagrangian velocity field. A distinction must be made between an Eulerian velocity field and the

Lagrangian velocity field. It is true that at any given point the fluid particle will move with the velocity of the fluid at that point. But, since the fluid particle is following a changing path through the porous medium, it will encounter only a subset of velocities that comprise the entire velocity field of the domain. This subset is the *Lagrangian* velocity field. The velocity field of the entire domain is the *Eulerian* velocity field. In Section 1.11 the relationship between the Lagrangian velocity covariance and the Eulerian velocity covariance was described in terms of the probability density function of the fluid particles's position. Consequently, the statistical properties of the Lagrangian velocity field may well be different from those of the Eulerian velocity field.

2.3.5 Dagan's Approach

Formulations of the dispersion tensor in terms of the velocity covariances appear quite often in the literature. Section 1.7 describes a version based on arguments in Gelhar and Axness[45] and Neuman[69]. Dagan[32] offers an approach that allows the specification of the dispersion tensor on a numerical grid block by numerical grid block basis.

As discussed in Section 2.1.3, the second spatial moment, S_{ij} , which characterizes the spread of a mass around its centroid, is given by

$$\mathbf{S}_{ij} = \frac{1}{M} \int_{\Omega} n(X_i - R_i)(X_j - R_j)c(\vec{X}, t)d\vec{X} \quad i, j = 1, 2, 3$$

where M is the mass, \vec{R} is the centroid coordinate, $c(\vec{X}, t)$ is the concentration and n is the porosity.

Since concentration is mass per unit volume, the second spatial moment of the plume with respect to the centroid can be written as

$$\mathbf{S}_{ij}(t) = \frac{1}{V} \int_V [X_i(t, \vec{a}) - R_i(t)][X_j(t, \vec{a}) - R_j(t)]d\vec{a}$$

The *Actual Dispersion Coefficients* are defined as half the rate of change of the plume's second spatial moment with respect to the centroid in the given realization

$$\frac{1}{2} \frac{d\mathbf{S}_{ij}}{dt}$$

Since \mathbf{S}_{ij} is a random variable, Dagan[30, 32] defines *effective dispersion coefficients* as:

$$\mathbf{D}_{ij} = \frac{1}{2} \frac{d\mathbf{E}[\mathbf{S}_{ij}]}{dt}$$

The key to Dagan's final result is the fundamental relationship, Kitanidis[58], Dagan[29],

$$\mathbf{E}[\mathbf{S}_{ij}(t)] = \mathbf{S}_{ij}(0) + \mathbf{X}_{ij}(t, 0) - \mathbf{R}_{ij}(t)$$

From which it follows that

$$\mathbf{D}_{ij} = \frac{1}{2} \frac{d\mathbf{E}[\mathbf{S}_{ij}]}{dt} = \frac{1}{2} \frac{d\mathbf{X}_{ij}}{dt} - \frac{1}{2} \frac{d\mathbf{R}_{ij}}{dt}$$

If the rectangle V_0 is of dimension l_1 in the direction of the mean flow, x_1 , and l_2 be transverse to it, then the dispersion tensor components $\mathbf{D}_{ij}(t, \omega)$ where ω is the $l_1 \times l_2$ rectangle are given by Dagan[30, Equation 15]

$$\begin{aligned} \mathbf{D}_{ij}(t, l_1, l_2) &= \frac{4}{l_1^2 l_2^2} \int_0^{l_1} \int_0^{l_2} \int_0^t (l_1 - b_1)(l_2 - b_2) \\ &\quad \times [\rho_{ij}(\mathbf{E}[\vec{V}]t', 0) - \frac{1}{2}\rho_{ij}(\mathbf{E}[\vec{V}]t' + b_1, b_2) \\ &\quad - \frac{1}{2}\rho_{ij}(\mathbf{E}[\vec{V}]t' - b_1, b_2)] dt' db_2 db_1 \end{aligned}$$

So, by assuming a particular autocovariance function (exponential or Gaussian) the $\mathbf{D}_{ij}(t, l_2)$ can be solved for, and applied using the following steps:

- Determine the log-transmissivity variogram
- Calculate \mathbf{D}_{ij} for the assumed transverse dimension l_2 of the numerical blocks
- Attach the resulting \mathbf{D}_{ij} to rows of blocks at a distance $x = \|V\|t$ from the input zone

2.4 Distance Dependent Dispersivity Coefficients

In Gelhar[44] it is demonstrated that for a stratified aquifer to which an hydraulic gradient which varies only in the z direction is applied parallel to the layers, the variance of the displacement of a fluid particle is given in terms of the hydraulic conductivity, K , by

$$\sigma_x^2 = \frac{\sigma_K^2}{\mathbf{E}[K]^2} \mathbf{E}[x]^2$$

And, using the definition of dispersivity this means that

$$\begin{aligned}\mathbf{D} &= \frac{1}{2} \frac{d\sigma_x^2}{dt} = \frac{1}{2} \frac{d\sigma_x^2}{d\mathbf{E}[x]} \frac{d\mathbf{E}[x]}{dt} \\ &= \frac{\sigma_K^2}{\mathbf{E}[K]^2} \mathbf{E}[x] \mathbf{E}[v]\end{aligned}$$

This result shows the dependence of the dispersivity on the mean distance. The following seeks to extend this result to higher dimensions. In order to determine the effect of one variable on a second variable, it is desirable to have all *other* variables held constant during the experiment. In economics this is referred to as *ceteris paribus*. If the object is to determine the effect of permeability or hydraulic conductivity on velocity, the experimenter would take a material of known permeability, put it in a test tank, establish a hydraulic gradient in the tank and measure the velocity. Next, a second material with a different permeability would be placed in the tank, a hydraulic gradient of the *same* magnitude as before would be established and the resulting velocity measured. By maintaining a constant hydraulic gradient in both measurements, the effect of permeability on velocity can be determined. Because we want to know only the effect of permeability on velocity, the hydraulic gradient and conductivity covariance will be assumed to be constant in the following calculations.

2.4.1 Local Grid Block Dispersivity Coefficients

The seepage velocity is given by

$$\vec{V} = -\frac{\mathbf{K}}{n} \nabla \phi$$

And, the mean seepage velocity is given by

$$\mathbf{E}[\vec{V}] = -\frac{\mathbf{E}[\mathbf{K}]}{n} \nabla \phi$$

Let \vec{r} be the position vector of a fluid particle at time t . This vector is given by the kinematic relationship

$$\vec{r} = \int_0^t \vec{V}(t') dt' = -\int_0^t \frac{\mathbf{K}}{n} \nabla \phi dt'$$

The deviation of the position vector \vec{r} from its expected value is given by $\vec{r} - E[\vec{r}]$, and the matrix of covariances of these deviations is given by

$$\begin{aligned}
\mathbf{X} &= \mathbf{E} \left[(\vec{r} - \mathbf{E}[\vec{r}]) (\vec{r} - \mathbf{E}[\vec{r}])^\dagger \right] \\
&= \mathbf{E} \left[\left(\int_0^t (\mathbf{K} - \mathbf{E}[\mathbf{K}]) \frac{\nabla \phi}{n} dt'' \right) \left(\int_0^t (\mathbf{K} - \mathbf{E}[\mathbf{K}]) \frac{\nabla \phi}{n} dt' \right)^\dagger \right] \\
&= \mathbf{E} \left[\int_0^t \int_0^t \left(\mathbf{K}' \frac{\nabla \phi}{n} \right) \left(\mathbf{K}' \frac{\nabla \phi}{n} \right)^\dagger dt'' dt' \right] \\
&= \int_0^t \int_0^t \mathbf{E} \left[\vec{V}' \vec{V}'^\dagger \right] dt'' dt'
\end{aligned}$$

But, since $\left(\mathbf{E}[\vec{V}'(\vec{x}) \vec{V}'^\dagger(\vec{y})] \right)_{ij} = \rho_{ij}(\vec{x}, \vec{y})$, this is basically the same form as Equation[2.14], page 59. In order to generalize the one dimensional result, we can argue as follows:

$$\begin{aligned}
\mathbf{X} &= \mathbf{E} \left[(\vec{r} - \mathbf{E}[\vec{r}]) (\vec{r} - \mathbf{E}[\vec{r}])^\dagger \right] \\
&= \mathbf{E} \left[\left(\int_0^t \frac{\mathbf{K}}{n} \nabla \phi dt' - \mathbf{E} \left[\int_0^t \frac{\mathbf{K}}{n} \nabla \phi dt' \right] \right) \left(\int_0^t \frac{\mathbf{K}}{n} \nabla \phi dt' - \mathbf{E} \left[\int_0^t \frac{\mathbf{K}}{n} \nabla \phi dt' \right] \right)^\dagger \right] \\
&= \mathbf{E} \left[\left\{ (\mathbf{K} - \mathbf{E}[\mathbf{K}]) \frac{t \nabla \phi}{n} \right\} \left\{ (\mathbf{K} - \mathbf{E}[\mathbf{K}]) \frac{t \nabla \phi}{n} \right\}^\dagger \right]
\end{aligned}$$

Letting $\tilde{\mathbf{K}} = \mathbf{K} - \mathbf{E}[\mathbf{K}]$, and noting that

$$\mathbf{E}[\vec{r}] = t \mathbf{E}[\vec{V}] = \frac{t}{n} \mathbf{E}[\mathbf{K}] \nabla \phi$$

which implies that

$$\frac{t}{n} \nabla \phi = \mathbf{E}[\mathbf{K}]^{-1} \mathbf{E}[\vec{r}]$$

So

$$\mathbf{X} = \mathbf{E} \left[\left(\tilde{\mathbf{K}} \mathbf{E}[\mathbf{K}]^{-1} \mathbf{E}[\vec{r}] \right) \left(\tilde{\mathbf{K}} \mathbf{E}[\mathbf{K}]^{-1} \mathbf{E}[\vec{r}] \right)^\dagger \right]$$

Let \mathbf{A} be an $r \times s$ matrix and \mathbf{B} be a $t \times u$ matrix, then the Kronecker product is the $rt \times su$ matrix

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1s}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2s}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1}\mathbf{B} & a_{r2}\mathbf{B} & \cdots & a_{rs}\mathbf{B} \end{bmatrix}$$

In particular, if $\vec{y} \in \mathfrak{R}^r$, $z \in \mathfrak{R}^s$, then \vec{y} is $r \times 1$ and \vec{z} is $s \times 1$ so that $\vec{y} \otimes \mathbf{I}_s$ is $rs \times s$ and given by

$$\vec{y} \otimes \mathbf{I}_s = \begin{bmatrix} y_1 \mathbf{I}_s \\ y_2 \mathbf{I}_s \\ \vdots \\ y_r \mathbf{I}_s \end{bmatrix}$$

and $\mathbf{I}_r \otimes \vec{z}$ is $rs \times s$ and given by

$$\mathbf{I}_r \otimes \vec{z} = \begin{bmatrix} \vec{z} & \vec{0} & \vec{0} & \cdots & \vec{0} \\ \vec{0} & \vec{z} & \vec{0} & \cdots & \vec{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vec{0} & \vec{0} & \vec{0} & \cdots & \vec{z} \end{bmatrix}$$

Given that

$$\mathbf{X} = \mathbf{E} \left[(\vec{r} - \mathbf{E}[\vec{r}])(\vec{r} - \mathbf{E}[\vec{r}])^\dagger \right]$$

then using the differentiation formula from Marlow[66] that if $\vec{y} \in \mathfrak{R}^r$, $\vec{z} \in \mathfrak{R}^s$, $\vec{x} \in \mathfrak{R}^n$, then

$$\frac{d}{d\vec{x}} (\vec{y}\vec{z}^\dagger) = (\vec{y} \otimes \mathbf{I}_s) \frac{d\vec{z}}{d\vec{x}} + (\mathbf{I}_r \otimes \vec{z}) \frac{d\vec{y}}{d\vec{x}}$$

where

$$\frac{d\vec{y}}{d\vec{x}} = \begin{bmatrix} \partial_1 y_1 & \partial_2 y_1 & \cdots & \partial_n y_1 \\ \partial_1 y_2 & \partial_2 y_2 & \cdots & \partial_n y_2 \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 y_r & \partial_2 y_r & \cdots & \partial_n y_r \end{bmatrix}$$

and

$$\partial_i y_j = \frac{\partial y_j}{\partial x_i}$$

In our case, we let $r = s = n$ and

$$\vec{y} = \vec{z} = \tilde{\mathbf{K}}\mathbf{E}[\mathbf{K}]^{-1}\mathbf{E}[\vec{r}] = \hat{\mathbf{K}}\mathbf{E}[\vec{r}] = \begin{bmatrix} \sum_{j=1}^n \hat{\mathbf{K}}_{1j}\mathbf{E}[r_j] \\ \sum_{j=1}^n \hat{\mathbf{K}}_{2j}\mathbf{E}[r_j] \\ \vdots \\ \sum_{j=1}^n \hat{\mathbf{K}}_{nj}\mathbf{E}[r_j] \end{bmatrix}$$

which is an $n \times 1$ vector. It has been shown that the dispersivity tensor contains the term

$$\frac{1}{2} \frac{d}{dt} \mathbf{X} = \frac{1}{2} \frac{d\mathbf{X}}{d\mathbf{E}[\vec{r}]} \frac{d\mathbf{E}[\vec{r}]}{dt} = \frac{1}{2} \frac{d\mathbf{X}}{d\mathbf{E}[\vec{r}]} \mathbf{E}[\vec{V}]$$

According to Marlow[66], the matrix $\frac{d\mathbf{X}}{dt}$ can be written as

$$\frac{d\mathbf{X}}{dt} = \begin{bmatrix} \frac{dX_{11}}{dt} \\ \frac{dX_{12}}{dt} \\ \vdots \\ \frac{dX_{1n}}{dt} \\ \vdots \\ \frac{dX_{n1}}{dt} \\ \frac{dX_{n2}}{dt} \\ \vdots \\ \frac{dX_{nn}}{dt} \end{bmatrix}$$

which is $n^2 \times 1$. Also, $\frac{d\mathbf{X}}{d\mathbf{E}[\vec{r}]}$ is an $n^2 \times n$ matrix and $\mathbf{E}[\vec{V}]$ is an $n \times 1$ vector so that $\frac{d\mathbf{X}}{d\mathbf{E}[\vec{r}]} \mathbf{E}[\vec{V}]$ is also $n^2 \times 1$.

Now,

$$\begin{aligned} \frac{d\mathbf{X}}{d\mathbf{E}[\vec{r}]} &= \mathbf{E} \left[\frac{d}{d\mathbf{E}[\vec{r}]} (\vec{y}\vec{z}^\dagger) \right] \\ &= \mathbf{E} \left[(\vec{y} \otimes \mathbf{I}_n) \frac{d\vec{z}}{d\mathbf{E}[\vec{r}]} + (\mathbf{I}_n \otimes \vec{z}) \frac{d\vec{y}}{d\mathbf{E}[\vec{r}]} \right] \end{aligned}$$

Since $\vec{y} \otimes \mathbf{I}_n$ has dimension $n^2 \times n$ and $\mathbf{I}_n \otimes \vec{z}$ has dimension $n^2 \times n$ and $\frac{d\vec{z}}{d\mathbf{E}[\vec{r}]}$ and $\frac{d\vec{y}}{d\mathbf{E}[\vec{r}]}$ have dimension $n \times n$, the matrix $\frac{d\mathbf{X}}{d\mathbf{E}[\vec{r}]}$ has dimension $n^2 \times n$.

The components of this equation are evaluated as follows:

$$\vec{y} \otimes \mathbf{I}_n = \begin{bmatrix} \sum_{j=1}^n \hat{\mathbf{K}}_{1j} \mathbf{E}[r_j] \mathbf{I}_n \\ \vdots \\ \sum_{j=1}^n \hat{\mathbf{K}}_{nj} \mathbf{E}[r_j] \mathbf{I}_n \end{bmatrix}$$

And,

$$\mathbf{I}_n \otimes \vec{z} = \begin{bmatrix} \vec{z} & \vec{0} & \cdots & \vec{0} \\ \vec{0} & \vec{z} & \cdots & \vec{0} \\ \vdots & \vdots & \ddots & \vdots \\ \vec{0} & \vec{0} & \cdots & \vec{z} \end{bmatrix}$$

where

$$\vec{z} = \begin{bmatrix} \sum_{j=1}^n \hat{\mathbf{K}}_{1j} \mathbf{E}[r_j] \\ \vdots \\ \sum_{j=1}^n \hat{\mathbf{K}}_{nj} \mathbf{E}[r_j] \end{bmatrix} \quad \text{and} \quad \vec{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Furthermore,

$$\frac{d\vec{z}}{d\mathbf{E}[\vec{r}]} = \frac{d\vec{y}}{d\mathbf{E}[\vec{r}]} = \begin{bmatrix} \hat{\mathbf{K}}_{11} & \hat{\mathbf{K}}_{12} & \hat{\mathbf{K}}_{13} & \cdots & \hat{\mathbf{K}}_{1n} \\ \hat{\mathbf{K}}_{21} & \hat{\mathbf{K}}_{22} & \hat{\mathbf{K}}_{23} & \cdots & \hat{\mathbf{K}}_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{K}}_{n1} & \hat{\mathbf{K}}_{n2} & \hat{\mathbf{K}}_{n3} & \cdots & \hat{\mathbf{K}}_{nn} \end{bmatrix}$$

So, we can write

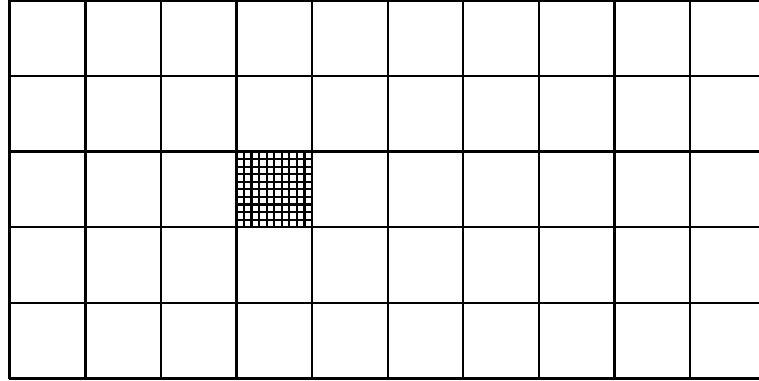
$$(\vec{y} \otimes \mathbf{I}_n) \frac{d\vec{z}}{d\mathbf{E}[\vec{r}]} = \begin{bmatrix} \sum_{j=1}^n \hat{\mathbf{K}}_{1j} \mathbf{E}[r_j] \mathbf{I}_n \\ \vdots \\ \sum_{j=1}^n \hat{\mathbf{K}}_{nj} \mathbf{E}[r_j] \mathbf{I}_n \end{bmatrix} \begin{bmatrix} \hat{\mathbf{K}}_{11} & \hat{\mathbf{K}}_{12} & \hat{\mathbf{K}}_{13} & \cdots & \hat{\mathbf{K}}_{1n} \\ \hat{\mathbf{K}}_{21} & \hat{\mathbf{K}}_{22} & \hat{\mathbf{K}}_{23} & \cdots & \hat{\mathbf{K}}_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{K}}_{n1} & \hat{\mathbf{K}}_{n2} & \hat{\mathbf{K}}_{n3} & \cdots & \hat{\mathbf{K}}_{nn} \end{bmatrix}$$

which is an $n^2 \times n$ matrix. And,

$$(\mathbf{I}_n \otimes \vec{z}) \frac{d\vec{y}}{d\mathbf{E}[\vec{r}]} =$$

$$\begin{bmatrix}
\sum_{j=1}^n \hat{\mathbf{K}}_{1j} \mathbf{E}[r_j] & 0 & \cdots & 0 \\
\vdots & \vdots & & \vdots \\
\sum_{j=1}^n \hat{\mathbf{K}}_{nj} \mathbf{E}[r_j] & 0 & \cdots & 0 \\
\text{---} & \text{---} & \text{---} & \text{---} \\
0 & \sum_{j=1}^n \hat{\mathbf{K}}_{1j} \mathbf{E}[r_j] & \cdots & 0 \\
\vdots & \vdots & & \vdots \\
0 & \sum_{j=1}^n \hat{\mathbf{K}}_{nj} \mathbf{E}[r_j] & \cdots & 0 \\
\text{---} & \text{---} & \text{---} & \text{---} \\
& \vdots & & \\
\text{---} & \text{---} & \text{---} & \text{---} \\
0 & 0 & \cdots & \sum_{j=1}^n \hat{\mathbf{K}}_{1j} \mathbf{E}[r_j] \\
\vdots & \vdots & & \vdots \\
0 & 0 & \cdots & \sum_{j=1}^n \hat{\mathbf{K}}_{nj} \mathbf{E}[r_j]
\end{bmatrix}
\begin{bmatrix}
\hat{\mathbf{K}}_{11} & \hat{\mathbf{K}}_{12} & \hat{\mathbf{K}}_{13} & \cdots & \hat{\mathbf{K}}_{1n} \\
\hat{\mathbf{K}}_{21} & \hat{\mathbf{K}}_{22} & \hat{\mathbf{K}}_{23} & \cdots & \hat{\mathbf{K}}_{2n} \\
& & \vdots & & \\
\hat{\mathbf{K}}_{n1} & \hat{\mathbf{K}}_{n2} & \hat{\mathbf{K}}_{n3} & \cdots & \hat{\mathbf{K}}_{nn}
\end{bmatrix}$$

which is also an $n^2 \times n$ matrix. The result will be demonstrated using a two dimensional example:



Local Grid Block

Example: $n = 2$

In this example we can write

$$\frac{1}{2} \frac{d}{dt} \mathbf{X} = \begin{bmatrix} \frac{1}{2} \frac{d\mathbf{X}_{11}}{dt} \\ \frac{1}{2} \frac{d\mathbf{X}_{12}}{dt} \\ \frac{1}{2} \frac{d\mathbf{X}_{21}}{dt} \\ \frac{1}{2} \frac{d\mathbf{X}_{22}}{dt} \end{bmatrix} = \frac{1}{2} \frac{d\mathbf{X}}{d\mathbf{E}[\vec{r}]} \mathbf{E}[\vec{V}] = \mathbf{E} \left[(\vec{y} \otimes \mathbf{I}_2) \frac{d\vec{z}}{d\mathbf{E}[\vec{r}]} + (\mathbf{I}_2 \otimes \vec{z}) \frac{d\vec{y}}{d\mathbf{E}[\vec{r}]} \right] \mathbf{E}[\vec{V}]$$

Expanding this yields

$$\frac{1}{2} \frac{d}{dt} \mathbf{X} = \frac{1}{2} \mathbf{E} \left\{ \left[\begin{array}{cc} \sum_{j=1}^2 \hat{\mathbf{K}}_{11} \hat{\mathbf{K}}_{1j} \mathbf{E}[r_j] & \sum_{j=1}^2 \hat{\mathbf{K}}_{12} \hat{\mathbf{K}}_{1j} \mathbf{E}[r_j] \\ \sum_{j=1}^2 \hat{\mathbf{K}}_{21} \hat{\mathbf{K}}_{1j} \mathbf{E}[r_j] & \sum_{j=1}^2 \hat{\mathbf{K}}_{22} \hat{\mathbf{K}}_{1j} \mathbf{E}[r_j] \\ \sum_{j=1}^2 \hat{\mathbf{K}}_{11} \hat{\mathbf{K}}_{2j} \mathbf{E}[r_j] & \sum_{j=1}^2 \hat{\mathbf{K}}_{12} \hat{\mathbf{K}}_{2j} \mathbf{E}[r_j] \\ \sum_{j=1}^2 \hat{\mathbf{K}}_{21} \hat{\mathbf{K}}_{2j} \mathbf{E}[r_j] & \sum_{j=1}^2 \hat{\mathbf{K}}_{22} \hat{\mathbf{K}}_{2j} \mathbf{E}[r_j] \end{array} \right] + \left[\begin{array}{cc} \sum_{j=1}^2 \hat{\mathbf{K}}_{11} \hat{\mathbf{K}}_{1j} \mathbf{E}[r_j] & \sum_{j=1}^2 \hat{\mathbf{K}}_{12} \hat{\mathbf{K}}_{1j} \mathbf{E}[r_j] \\ \sum_{j=1}^2 \hat{\mathbf{K}}_{11} \hat{\mathbf{K}}_{2j} \mathbf{E}[r_j] & \sum_{j=1}^2 \hat{\mathbf{K}}_{12} \hat{\mathbf{K}}_{2j} \mathbf{E}[r_j] \\ \sum_{j=1}^2 \hat{\mathbf{K}}_{21} \hat{\mathbf{K}}_{1j} \mathbf{E}[r_j] & \sum_{j=1}^2 \hat{\mathbf{K}}_{22} \hat{\mathbf{K}}_{1j} \mathbf{E}[r_j] \\ \sum_{j=1}^2 \hat{\mathbf{K}}_{21} \hat{\mathbf{K}}_{2j} \mathbf{E}[r_j] & \sum_{j=1}^2 \hat{\mathbf{K}}_{22} \hat{\mathbf{K}}_{2j} \mathbf{E}[r_j] \end{array} \right] \right\} \\ \times \left[\begin{array}{c} \mathbf{E}[V_1] \\ \mathbf{E}[V_2] \end{array} \right]$$

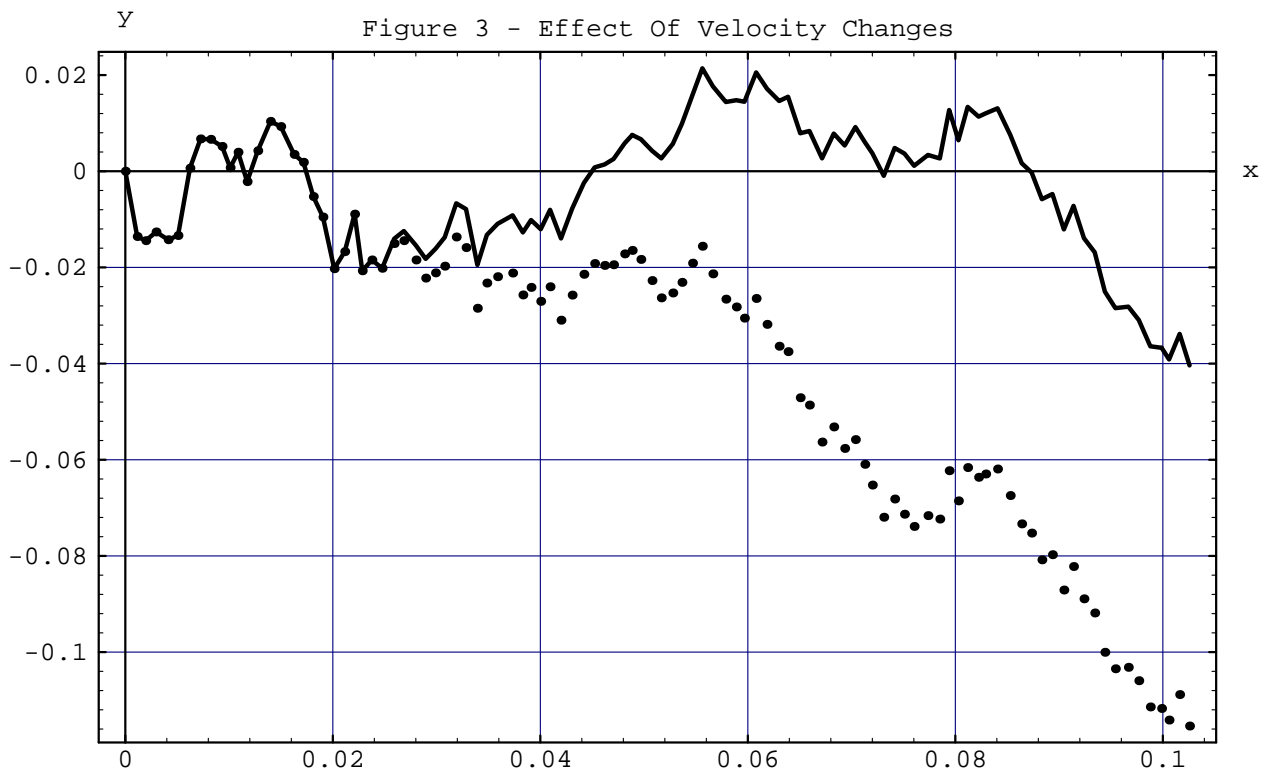
Using the first order approximation of the expected value of the fluid particle's position, *viz.*,

$$\mathbf{E}[\vec{r}] = \mathbf{E}[\vec{V}]t$$

and assuming that $\hat{\mathbf{K}}_{ij} = 0$ for $i \neq j$, it follows that

$$\frac{1}{2} \frac{d}{dt} \mathbf{X} = \left[\begin{array}{c} \mathbf{E}[\hat{\mathbf{K}}_{11}^2] \mathbf{E}[V_1]^2 t \\ \mathbf{E}[\hat{\mathbf{K}}_{11} \hat{\mathbf{K}}_{22}] \mathbf{E}[V_1] \mathbf{E}[V_2] t \\ \mathbf{E}[\hat{\mathbf{K}}_{11} \hat{\mathbf{K}}_{22}] \mathbf{E}[V_1] \mathbf{E}[V_2] t \\ \mathbf{E}[\hat{\mathbf{K}}_{22}^2] \mathbf{E}[V_2]^2 t \end{array} \right]$$

The presence of the expected velocity vector in this expression is key since it can be changing from numerical grid block to numerical grid block. Figure 3 shows the effect of a non-constant expected velocity on the path of a fluid particle. This figure shows the result of solving Equation[2.17] first with $\mathbf{E}[\vec{V}] = (0, -0.1)$, the dotted line, and then allowing $\mathbf{E}[\vec{V}] = (0, -0.1)$ for the first 25 time steps, $\mathbf{E}[\vec{V}] = (0, 0.0)$ for the next 25 time steps, $\mathbf{E}[\vec{V}] = (0, 0.1)$ for the next 25 time steps, and finally, $\mathbf{E}[\vec{V}] = (0, -0.1)$ for the last 25 time steps, the solid line in Figure 3.



2.4.2 Global Dispersivity Coefficients

The presence of heterogeneities in the porous medium will cause the velocity field to be non-uniform. To maintain dispersive symmetry, the dispersion tensor should be recalculated on a grid block by grid block basis, taking into consideration the expected velocity on the grid block. The preceding formulation can be extended to this case in the following manner: Let \vec{X}_T be the trajectory of a fluid particle. Then, the position vector of the particle is given by

$$\vec{r} = \int_0^t \vec{V}(\vec{X}_T(t')) dt' = - \int_0^t \frac{\mathbf{K}}{n} \nabla \phi dt'$$

It is assumed that $\nabla \phi$ is steady state so that it does not depend on t' . Furthermore, it is assumed that the porous medium is locally homogeneous, *i.e.*, in the sense of Section 1.1 and Section 1.5, and that this local homogeneity applies to numerical grid blocks.

Suppose that \vec{X}_T spends from t_0 to t_1 on grid block 1, from t_1 to t_2 on grid block 2, \dots , and from t_{n-1} to t_n on grid block n . Then since $\nabla \phi$ does not depend on time and the statistics of \mathbf{K} do not depend on time on individual grid blocks, we can write

$$\begin{aligned} \vec{r} &= - \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \frac{\mathbf{K}^{(i+1)}}{n} \nabla \phi^{(i+1)} dt' = - \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbf{K}^{(i+1)} dt' \frac{\nabla \phi^{(i+1)}}{n} \\ &\Rightarrow \mathbf{E}[\vec{r}] = - \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbf{E}[\mathbf{K}^{(i+1)}] dt' \frac{\nabla \phi^{(i+1)}}{n} \end{aligned}$$

then,

$$\vec{r} - \mathbf{E}[\vec{r}] = - \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} (\mathbf{K}^{(i+1)} - \mathbf{E}[\mathbf{K}^{(i+1)}]) dt' \frac{\nabla \phi^{(i+1)}}{n}$$

Using the formulas

$$\vec{b} \otimes \vec{a}^\dagger = \vec{b} \vec{a}^\dagger$$

$$(\vec{a} + \vec{b}) \otimes (\vec{c} + \vec{d}) = (\vec{a} \otimes \vec{c}) + (\vec{a} \otimes \vec{d}) + (\vec{b} \otimes \vec{c}) + (\vec{b} \otimes \vec{d})$$

it follows that

$$\begin{aligned}
\mathbf{X} &= \mathbf{E} \left[(\vec{r} - \mathbf{E}[\vec{r}]) (\vec{r} - \mathbf{E}[\vec{r}])^\dagger \right] \\
&= \mathbf{E} \left[\left\{ - \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} (\mathbf{K}^{(i+1)} - \mathbf{E}[\mathbf{K}^{(i+1)}]) dt' \frac{\nabla \phi^{(i+1)}}{n} \right\} \right. \\
&\quad \left. \left\{ - \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} (\mathbf{K}^{(j+1)} - \mathbf{E}[\mathbf{K}^{(j+1)}]) dt'' \frac{\nabla \phi^{(j+1)}}{n} \right\}^\dagger \right] \\
&= \mathbf{E} \left[\left\{ - \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \tilde{\mathbf{K}}^{(i+1)} dt' \frac{\nabla \phi^{(i+1)}}{n} \right\} \left\{ - \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \tilde{\mathbf{K}}^{(j+1)} dt'' \frac{\nabla \phi^{(j+1)}}{n} \right\}^\dagger \right] \\
&= \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \mathbf{E} \left[\left\{ \int_{t_i}^{t_{i+1}} \tilde{\mathbf{K}}^{(i+1)} dt' \frac{\nabla \phi^{(i+1)}}{n} \right\} \otimes \left\{ \int_{t_j}^{t_{j+1}} \tilde{\mathbf{K}}^{(j+1)} dt'' \frac{\nabla \phi^{(j+1)}}{n} \right\}^\dagger \right] \\
&= \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \mathbf{E} \left[\left\{ \int_{t_i}^{t_{i+1}} \tilde{\mathbf{K}}^{(i+1)} dt' \frac{\nabla \phi^{(i+1)}}{n} \right\} \left\{ \int_{t_j}^{t_{j+1}} \tilde{\mathbf{K}}^{(j+1)} dt'' \frac{\nabla \phi^{(j+1)}}{n} \right\}^\dagger \right] \\
&= \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \mathbf{E} \left[\left\{ \sum_{l=1}^2 \int_{t_i}^{t_{i+1}} \tilde{\mathbf{K}}_{ul}^{(i+1)} dt' \frac{\nabla \phi_l^{(i+1)}}{n} \right\} \left\{ \sum_{m=1}^2 \int_{t_j}^{t_{j+1}} \tilde{\mathbf{K}}_{vm}^{(j+1)} dt'' \frac{\nabla \phi_m^{(j+1)}}{n} \right\}^\dagger \right]
\end{aligned}$$

Define the vectors \vec{a} and \vec{b} such that for $u, v = 1, 2$

$$a_u = \sum_{l=1}^2 \int_{t_i}^{t_{i+1}} \tilde{\mathbf{K}}_{ul}^{(i+1)} dt' \frac{\nabla \phi_l^{(i+1)}}{n}$$

and

$$b_v = \sum_{m=1}^2 \int_{t_j}^{t_{j+1}} \tilde{\mathbf{K}}_{vm}^{(j+1)} dt'' \frac{\nabla \phi_m^{(j+1)}}{n}$$

then

$$\mathbf{X} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \mathbf{E}[\vec{a}^{(i+1)} (\vec{b}^{(j+1)})^\dagger]$$

where

$$(\vec{a}^{(i+1)} (\vec{b}^{(j+1)})^\dagger)_{uv} = \sum_{l=1}^2 \sum_{m=1}^2 \int_{t_i}^{t_{i+1}} \int_{t_j}^{t_{j+1}} \tilde{\mathbf{K}}_{ul}^{(i+1)} \tilde{\mathbf{K}}_{vm}^{(j+1)} dt' dt'' \frac{\nabla \phi_l^{(i+1)}}{n} \frac{\nabla \phi_m^{(j+1)}}{n}$$

If it is assumed that the components $\mathbf{E}[\tilde{\mathbf{K}}_{ul}^{(i+1)} \tilde{\mathbf{K}}_{vm}^{(j+1)}]$ are constant over blocks, then

$$(\mathbf{X})_{uv} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \mathbf{E} \left[\sum_{l=1}^2 \sum_{m=1}^2 \tilde{\mathbf{K}}_{ul}^{(i+1)} \tilde{\mathbf{K}}_{vm}^{(j+1)} \frac{\nabla \phi_l^{(i+1)}}{n} \Delta t_{i+1} \frac{\nabla \phi_m^{(j+1)}}{n} \Delta t_{j+1} \right]$$

And, using the relationship that

$$\begin{aligned} \mathbf{E}[\vec{r}^{(i)}] &= \Delta t_i \mathbf{E}[\vec{V}^{(i)}] = \frac{\Delta t_i}{n} \mathbf{E}[\mathbf{K}^{(i)}] \nabla \phi^{(i)} \\ \Rightarrow \quad \frac{\nabla \phi^{(i)}}{n} \Delta t_i &= \mathbf{E}[\mathbf{K}^{(i)}]^{-1} \mathbf{E}[\vec{r}^{(i)}] \end{aligned}$$

The displacement covariance matrix becomes

$$(\mathbf{X})_{uv} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \mathbf{E} \left[\sum_{l=1}^2 \tilde{\mathbf{K}}_{ul}^{(i+1)} \left(\mathbf{E}[\mathbf{K}^{(i+1)}]^{-1} \mathbf{E}[\vec{r}^{(i+1)}] \right)_l \sum_{m=1}^2 \tilde{\mathbf{K}}_{vm}^{(j+1)} \left(\mathbf{E}[\mathbf{K}^{(j+1)}]^{-1} \mathbf{E}[\vec{r}^{(j+1)}] \right)_m \right]$$

or

$$\mathbf{X} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \mathbf{E} \left[\left(\tilde{\mathbf{K}}^{(i+1)} \mathbf{E}[\mathbf{K}^{(i+1)}]^{-1} \mathbf{E}[\vec{r}^{(i+1)}] \right) \left(\tilde{\mathbf{K}}^{(j+1)} \mathbf{E}[\mathbf{K}^{(j+1)}]^{-1} \mathbf{E}[\vec{r}^{(j+1)}] \right)^\dagger \right]$$

To get an adequate representation of exactly where in the domain different magnitudes of dispersion are to be expected, ensembles of particles must be tracked. Such a representation is given in Figure 4. In this example, 5 particles are tracked from two adjacent numerical grid blocks located in the center of the domain.

These particle paths are used to identify the numerical grid blocks that are most likely to be reached by a tracer plume that emanates from the grid blocks containing the origin of the plume. Once the grid blocks most likely to be reached have been identified, the previous formulation can be applied on a grid block by grid block basis to estimate how dispersion will develop over time. From Equation[2.24], page 71,

$$\begin{aligned} \mathbf{D}_{ij} &= \int_0^t \rho_{ij} \left(\mathbf{E}[\vec{V}]t', \mathbf{E}[\vec{V}]t \right) dt' = \int_0^t \mathbf{E} \left[\vec{V}'_i \left(\mathbf{E}[\vec{V}]t' \right) \vec{V}'_j \left(\mathbf{E}[\vec{V}]t \right) \right] dt' \\ &\approx \frac{1}{\|\mathbf{E}[\vec{V}]\|} \int_{\vec{\sigma}_T} \rho_{ij} ds \end{aligned}$$

As time t' increases from 0 to t , the particle moves from its original position at time $t = 0$ to its position at time t , in block (m) . As it moves, it traverses different grid blocks on its way to block (m) . Using the arguments given previously,

$$\begin{aligned}
\mathbf{D}_{uv} &= \sum_{i=0}^{m-1} \mathbf{E} \left[\int_{t_i}^{t_{i+1}} \left\{ \left(\mathbf{K}^{(i+1)} - \mathbf{E} [\mathbf{K}^{(i+1)}] \right) \frac{\nabla \phi^{(i+1)}}{n} \right\}_u \right. \\
&\quad \left. \times \left\{ \left(\mathbf{K}^{(m)} - \mathbf{E} [\mathbf{K}^{(m)}] \right) \frac{\nabla \phi^{(m)}}{n} \right\}_v dt' \right] \\
&= \sum_{i=0}^{m-1} \mathbf{E} \left[\left(\tilde{\mathbf{K}}^{(i+1)} \frac{\nabla \phi^{(i+1)}}{n} \Delta t_{i+1} \right)_u \left(\tilde{\mathbf{K}}^{(m)} \frac{\nabla \phi^{(m)}}{n} \right)_v \right] \\
&= \sum_{i=0}^{m-1} \mathbf{E} \left[\left(\tilde{\mathbf{K}}^{(i+1)} \mathbf{E} [\mathbf{K}^{(i+1)}]^{-1} \mathbf{E} [\vec{r}^{(i+1)}] \right)_u \left(\tilde{\mathbf{K}}^{(m)} \mathbf{E} [\mathbf{K}^{(m)}]^{-1} \mathbf{E} [\vec{V}^{(m)}] \right)_v \right] \\
&= \sum_{i=0}^{m-1} \mathbf{E} \left[\left(\hat{\mathbf{K}}^{(i+1)} \mathbf{E} [\vec{r}^{(i+1)}] \right)_u \left(\hat{\mathbf{K}}^{(m)} \mathbf{E} [\vec{V}^{(m)}] \right)_v \right]
\end{aligned}$$

Or, in terms of the full tensor

$$\mathbf{D} = \sum_{i=0}^{m-1} \mathbf{E} \left[\left(\hat{\mathbf{K}}^{(i+1)} \mathbf{E} [\vec{r}^{(i+1)}] \right) \left(\hat{\mathbf{K}}^{(m)} \mathbf{E} [\vec{V}^{(m)}] \right)^\dagger \right]$$

Since it may be difficult to identify exactly the numerical grid blocks where a local source is originating, or if more than one point source is involved, an entire column of grid blocks can be used to determine the particle paths. Figure 5 shows the result of tracking a particle from each of the grid blocks in a column in the center of the domain. Of course, in a simulation, more than one particle from each grid block would be tracked.

2.4.3 Molecular Diffusion

It is reasonable to have some level below which these dispersivity coefficients cannot fall. This level would, in effect, represent the level of molecular diffusion. In order to specify such a level, an example from Batchelor[9] is used. For a solute of NaCl in water, the coefficient of diffusion is found to be 1.1×10^{-5} cm²/sec at 15° C and for any concentration. For molecules such as potassium permanganate, KMnO₄, which are much larger than water molecules, the coefficient of diffusion is found to vary with the level of concentration. Since this is not a problem with sodium chloride, the level of

Figure 4 - Ensemble Of Particle Paths

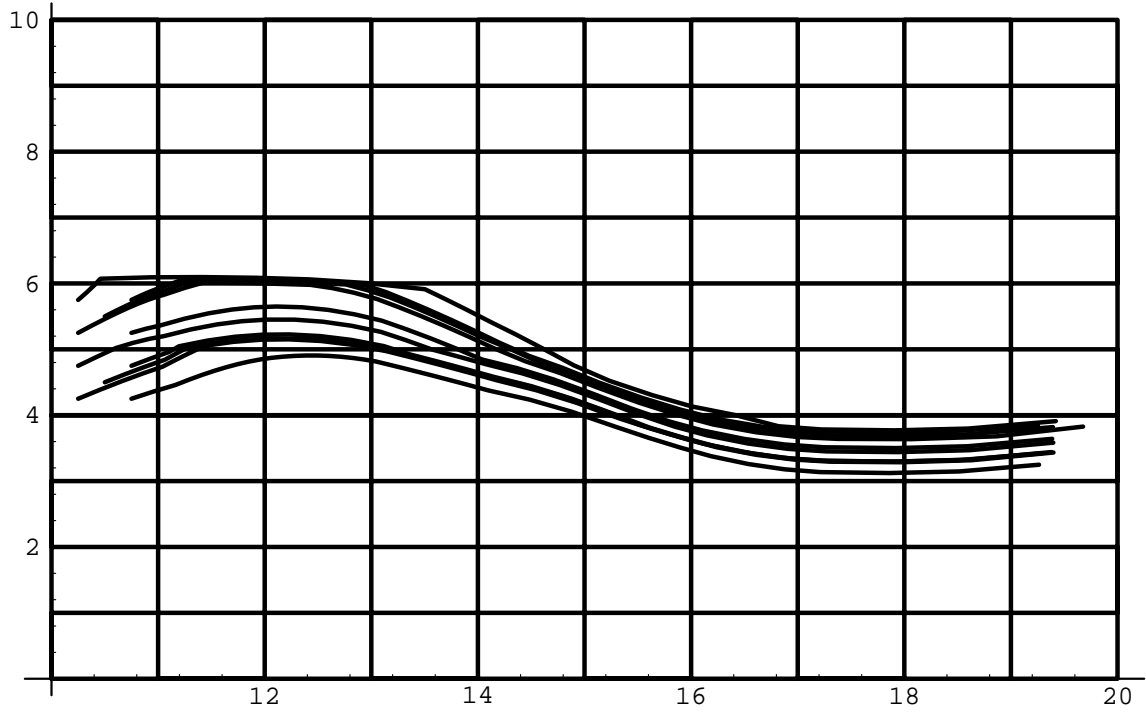
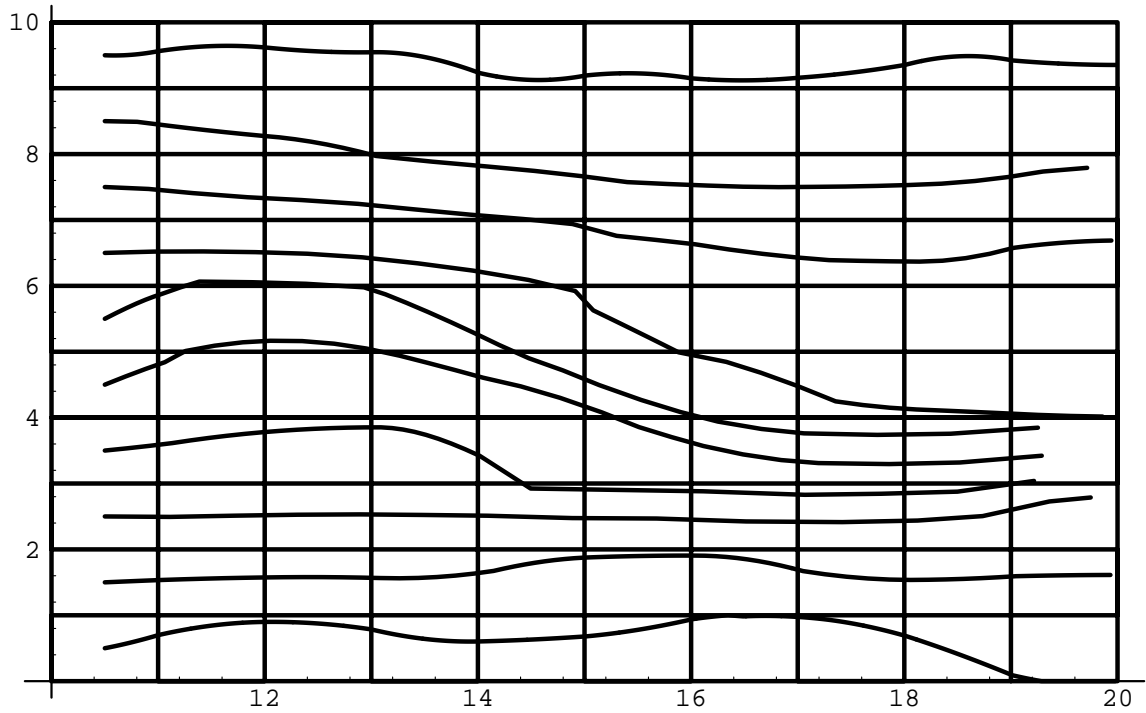


Figure 5 - Tracks From A Column Of Grid Blocks



10^{-5} cm²/sec, or the equivalent level of 6×10^{-4} cm²/min, will be used in the examples.

2.5 Random Variable Generation

2.5.1 Independent Random Variables

In order to generate the sample path of the stochastic process, it is assumed that the stochastic process is Gaussian and stationary, *i.e.*, at each time t , the random variable $X(t, \omega)$ has the same mean and variance and that these two moments are sufficient to describe the random variables's distribution. A sample path can then be generated by sampling from a Gaussian distribution with the specified mean and variance. The following analysis shows that by generating two independent uniform random variables on the interval $[0, 1]$ and using the Box-Muller transformation, Equation[2.25], two independent Gaussian random variables can be produced.

Let U_1 and U_2 be two independent uniform random variables with the same density function on the interval $[0,1]$. Define the random variables

$$\begin{aligned} X_1 &= (-2 \ln(U_1))^{\frac{1}{2}} \cos(2\pi U_2) \\ X_2 &= (-2 \ln(U_1))^{\frac{1}{2}} \sin(2\pi U_2) \end{aligned} \tag{2.25}$$

The inverse relations are given by

$$\begin{aligned} U_1 &= \exp \left[\frac{-(X_1^2 + X_2^2)}{2} \right] \\ U_2 &= \frac{1}{2\pi} \arctan \left(\frac{X_2}{X_1} \right) \end{aligned} \tag{2.26}$$

Then, by taking derivatives, it follows that

$$\begin{aligned} \frac{\partial X_1}{\partial U_1} &= \frac{-\cos(2\pi U_2)}{(-2 \ln(U_1))^{\frac{1}{2}} U_1} \\ \frac{\partial X_1}{\partial U_2} &= (-2 \ln(U_1))^{\frac{1}{2}} (-2\pi \sin(2\pi U_2)) \end{aligned}$$

And, that

$$\begin{aligned}\frac{\partial X_2}{\partial U_1} &= \frac{-\sin(2\pi U_2)}{(-2 \ln(U_1))^{\frac{1}{2}} U_1} \\ \frac{\partial X_2}{\partial U_2} &= (-2 \ln(U_1))^{\frac{1}{2}} (2\pi \cos(2\pi U_2))\end{aligned}$$

The Jacobian of the transformation is then

$$\begin{aligned}J &= \begin{vmatrix} \frac{-\cos(2\pi U_2)}{(-2 \ln(U_1))^{\frac{1}{2}} U_1} & (-2 \ln(U_1))^{\frac{1}{2}} (-2\pi \sin(2\pi U_2)) \\ \frac{-\sin(2\pi U_2)}{(-2 \ln(U_1))^{\frac{1}{2}} U_1} & (-2 \ln(U_1))^{\frac{1}{2}} (2\pi \cos(2\pi U_2)) \end{vmatrix} \\ &= \frac{-2\pi \cos^2(2\pi U_2)}{U_1} - \frac{2\pi \sin^2(2\pi U_2)}{U_1} \\ &= \frac{-2\pi}{U_1}\end{aligned}$$

From Equation[2.26],

$$U_1 = \exp \left[-\frac{(X_1^2 + X_2^2)}{2} \right]$$

Hence,

$$|J| = \frac{2\pi}{\exp \left[-\frac{(X_1^2 + X_2^2)}{2} \right]}$$

If $g(U_1, U_2)$ is the joint density function of U_1 and U_2 , then the joint density function of X_1 and X_2 is given by

$$f(X_1, X_2) = \frac{g(U_1, U_2)}{|J|}$$

From, Equation[2.26], it follows that

$$f(X_1, X_2) = \frac{g \left(\exp \left[-\frac{(X_1^2 + X_2^2)}{2} \right], \frac{1}{2\pi} \arctan \left(\frac{X_2}{X_1} \right) \right)}{\frac{2\pi}{\exp \left[-\frac{(X_1^2 + X_2^2)}{2} \right]}}$$

And, we have that

$$\forall X_1, X_2, \quad \exp \left[-\frac{(X_1^2 + X_2^2)}{2} \right] \in (0, 1)$$

and,

$$\forall X_1, X_2, \quad \frac{1}{2\pi} \arctan \left(\frac{X_2}{X_1} \right) \in (0, 1)$$

Since g is the joint density of two independent uniform random variables on $(0, 1)$, it follows that

$$\forall X_1, X_2 \quad g \left(\exp \left[-\frac{(X_1^2 + X_2^2)}{2} \right], \frac{1}{2\pi} \arctan \left(\frac{X_2}{X_1} \right) \right) = 1$$

Hence,

$$\begin{aligned} f(X_1, X_2) &= \frac{1}{2\pi} \exp \left[-\frac{(X_1^2 + X_2^2)}{2} \right] \\ &= \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{X_1^2}{2} \right] \cdot \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{X_2^2}{2} \right] \end{aligned}$$

which shows that X_1 and X_2 are independent and each has the standard normal distribution. Hence, the two $U(0, 1)$ independent random variables are used to produce two independent $N(0, 1)$ random variables.

Once an $N(0, 1)$ distribution has been produced, an $N(\mu, \sigma^2)$ distribution can be produced by the transformation

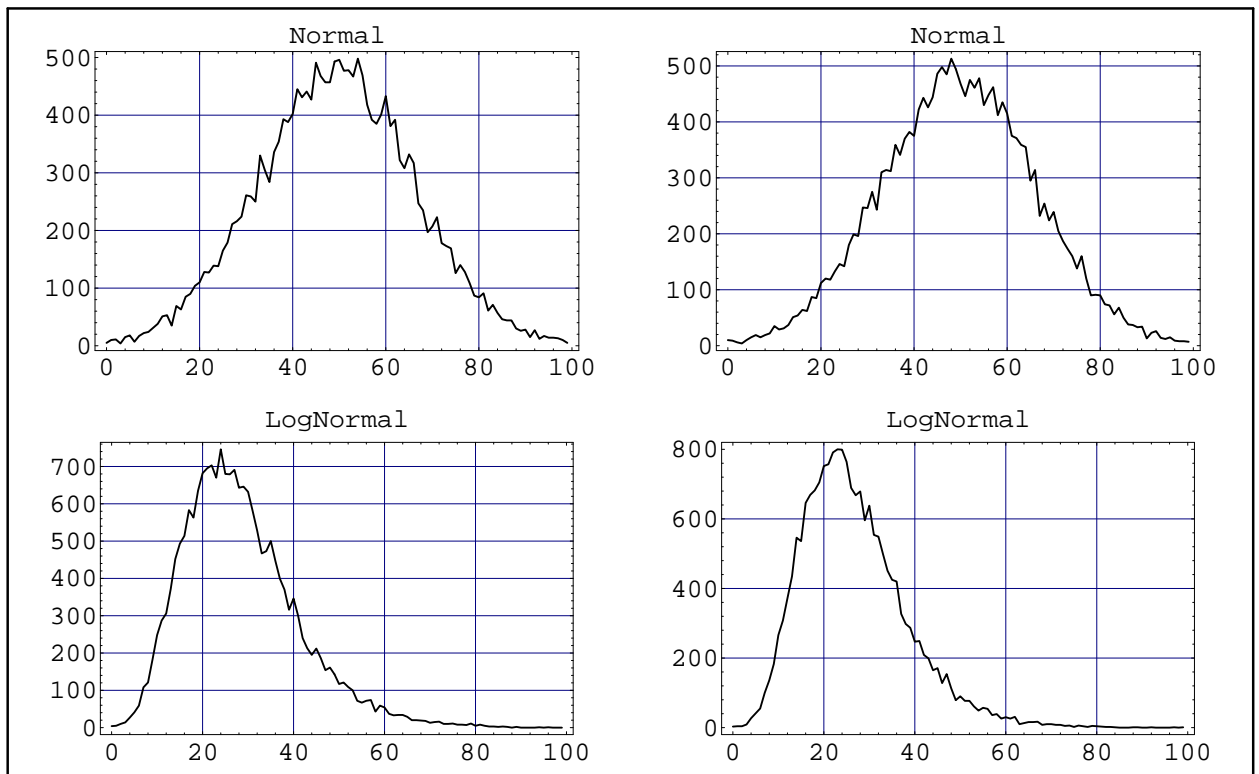
$$Y = \mu + \sigma X \quad X \sim N(0, 1)$$

Figure 6 illustrates this method by showing the results of generating 20000 samples from a Gaussian distribution with mean 3.0 and standard deviation 0.3. Plots of the two normal samples are shown along with their associated lognormal plots. The sample lognormal points are found from the formula

$$\text{lgn}[i] = \exp(\text{n}[i])$$

where $\text{n}[i]$, $i = 1, \dots, m$ are the sample points from the normal distribution.

Figure 6 - Sample Distributions



2.5.2 Correlated Random Variables

Once two independent normal random variables have been produced, it is possible via a linear transformation to produce two correlated random variables. Correlated random variables will be required in order to model an *anisotropic* porous medium.

Given two independent standard normal random variables, $X_1 \sim N(0, 1)$ and $X_2 \sim N(0, 1)$, two new random variables can be defined by letting

$$X_3 = a + bX_1 + cX_2 \quad a, b, c \in \mathfrak{R}^1$$

$$X_4 = d + eX_1 + fX_2 \quad d, e, f \in \mathfrak{R}^1$$

then $\mathbf{E}[X_3] = a$, $\mathbf{E}[X_4] = d$ and

$$\mathbf{E}[X_3X_3] = (a + b)^2 + c^2$$

$$\mathbf{E}[X_4X_4] = (d + e)^2 + f^2$$

$$\mathbf{E}[X_3X_4] = ad + be + cf$$

Example:

Suppose we want to define two random variables X_3 and X_4 such that

$$\mathbf{E}[X_3] = 0 \quad \mathbf{E}[X_4] = 0$$

$$\mathbf{E}[X_3X_3] = 1 \quad \mathbf{E}[X_4X_4] = \frac{1}{3} \quad \mathbf{E}[X_3X_4] = \frac{1}{2}$$

then $a = d = 0$ and

$$b^2 + c^2 = 1$$

$$e^2 + f^2 = \frac{1}{3}$$

$$be + cf = \frac{1}{2}$$

Letting $c = 1 \Rightarrow b = 0 \Rightarrow f = \frac{1}{2} \Rightarrow e = \frac{1}{2\sqrt{3}}$, so that

$$\mathbf{E}[X_3X_3] = (a + b)^2 + c^2 = 1$$

$$\mathbf{E}[X_4X_4] = (d + e)^2 + f^2 = \frac{1}{3}$$

$$\mathbf{E}[X_3X_4] = ad + be + cf = \frac{1}{2}$$

as required. Chapter 7 of Law and Kelton[61] gives many approaches to generating random variables, both correlated and uncorrelated. If a multivariate normal distribution is to be used, then a particularly simple algorithm exists for generating a multivariate normal vector

$$\vec{X} = (X_1, X_2, \dots, X_n)^\dagger$$

The reason for this is that in the joint density function, Equation[2.19], page 64, the covariance matrix, \mathbf{V} , is symmetric and positive definite. Hence, it can be factored as

$$\mathbf{V} = \mathbf{C}\mathbf{C}^\dagger$$

where \mathbf{C} is lower triangular. The algorithm consists of the following two steps:

(1) Generate Z_1, Z_2, \dots, Z_n as independent identically distributed $N(0, 1)$ random variables following the procedure in Section 2.5.1.

(2) For $i = 1, 2, \dots, n$, set $X_i = \mu_i + \sum_{j=1}^i \mathbf{C}_{ij}Z_j$.

It then follows that

$$\vec{X} = \vec{\mu} + \mathbf{C}\vec{Z}$$

In the study of porous media systems, the measurement of the physical properties of the system at each point of the domain is a practical impossibility. In the case of these systems, it is common to assume that the physical properties of the system, *i.e.*, hydraulic conductivity, etc., are realizations of a underlying random field. Since the development of our dispersion estimates is based on assumptions about the covariance function describing the hydraulic conductivities in the spatial domain, a method of generating random fields that takes into consideration the degree of variance, correlation lengths, cross-correlations, anisotropies, etc. of the hydraulic conductivity is necessary. Such

a method is provided by the *Spectral Turning Bands* method. In this method, simulations are performed along several lines using a unidimensional covariance function, $C_1(\cdot)$, that corresponds to the 2 or 3-dimensional covariance function given for the spatial domain. Given two spatial points \vec{x}_1, \vec{x}_2 in the domain, this correspondence is given by

$$C(\vec{x}_1, \vec{x}_2) = C(\vec{h}) = \int_{\Sigma} C_1(\vec{h} \cdot \vec{u}) f(\vec{u}) d\vec{u}$$

where Σ represents the unit circle or unit sphere, $f(\vec{u})$ is the probability density function of \vec{u} , and $\vec{h} = \vec{x}_1 - \vec{x}_2$. The value assigned to a point in the domain is given by an average of the values generated for the projection of the point onto the various lines used in the simulation.

Finally, the algorithm described in Robin, *et al*[80] is capable of co-generating pairs of three dimensional cross correlated random variables.

2.6 Summary

In this section, the all important dispersion tensor component of the transport equation was investigated under some simplifying assumptions that allow a better understanding of the concept. Section 2.1 looked at the dispersion concept from an Eulerian point of view which lead to its characterization as half the time rate of change of the second spatial moment around the centroid. Section 2.3 introduced the Lagrangian approach. Here the motion of a fluid particle was described by a stochastic differential equation, Equation[2.17], page 60, and dispersion was characterized as describing the uncertainty surrounding the path of a fluid particle as it proceeds through the porous medium.

Under the assumption that the trajectory of the fluid particle has a Gaussian probability density function, it was shown that the dispersion tensor is equal to half the time rate of change of the covariances of the displacements in the fluid particle's position, which was then related to the velocity covariances of the particle. Special consideration was given to a method due to Dagan that allows the specification of the dispersion on a grid block by grid block basis.

In Section 2.4, a one dimensional result that describes the dispersivity as a distance dependent entity was extended to higher dimensions. This method shows that the symmetry of the dispersion tensor is with respect to the average velocity vector. If the dispersion tensor is changed on a grid block by grid block basis, then the dispersion tensor must be recomputed to take into consideration changes in the expected velocity on the block. Section 2.4.2 applies the results of Section 2.4.1, which are applicable to locally homogeneous

grid blocks, to the calculation of global dispersivity coefficients. This is done by using a particle tracking algorithm to identify the grid blocks that will be visited by fluid particles along their paths and applying the results of Section 2.4.1 on a grid block by grid block basis.

Finally, Section 2.5 reviewed some methods being used to generate random variables. These methods will be used extensively in our numerical calculations.

3. Methods Of Solution

3.1 General

In this section, methods of solving the groundwater flow/pollution problem are discussed. Whenever two miscible fluids come together to form one phase, there is the potential for either the density of the phase or the viscosity of the phase to change. This change can be brought about by changes in the concentration and/or changes in the pressure. In the pollution problem, a change in the density can be caused by the pollutant mixing with the water in an aquifer. In the Enhanced Oil Recovery (EOR) problem, a change in the viscosity can be brought on by a surfactant mixing with the oil in the reservoir. The solution of an EOR problem may assume a viscosity/concentration relationship such as the quarter power rule, Russell and Wheeler[84]

$$\mu(c) = [(1 - c) + M^{\frac{1}{4}}c]^{-4}$$

where c is the concentration of the surfactant and M is the ratio of the viscosity of oil to the viscosity of surfactant, ie, the mobility ratio.

In the case that the density is dependent on the concentration of the pollutant, a coupled system of PDE's is obtained that must be solved simultaneously. However, there are probably many cases where the density is not dependent on the concentration or that this dependence is weak and it can be assumed that there is no dependence. In this case, the flow and concentration equations may be solved separately. The groundwater flow equation yields estimates of the piezometric head and velocity estimates are derived from Darcy's law. This is the case in tracer experiments. Then, by using either relationships between the velocity covariances and dispersion or between permeability covariances and dispersion, dispersivities can be developed for use in the concentration equation. By starting with the simpler uncoupled case, the total system can be tested for reasonableness without having to deal immediately with the more difficult coupled case. But, since the ultimate goal is to eventually handle the coupled case, the next section investigates some methods of handling the simultaneous solution.

3.1.1 Simultaneous Solution

The groundwater flow equation used in tracer experiments has the form

$$S \frac{\partial \phi^*}{\partial t} = \nabla \cdot \mathbf{T} \nabla \phi^*$$

where S is the storativity of the aquifer, \mathbf{T} is the transmissivity and ϕ^* is the piezometric head. If this equation is not used to determine velocities, then an alternative expression must be used to obtain the velocities. These velocities can be obtained from Darcy's law in the following way: In general, Darcy's law can be written as

$$\vec{q} = \mathbf{K} \vec{J}$$

where \vec{q} is the specific discharge, the volume of fluid flowing per unit time through a *unit* cross-sectional area normal to the flow, \mathbf{K} is the hydraulic conductivity tensor and

$$\vec{J} = -\nabla \phi^*$$

is the hydraulic gradient. Because the cross-sectional area used in the definition is a unit, the vector \vec{q} has the dimension $\left(\frac{L}{t}\right)$, and so is considered a velocity. The Dupuit-Forchheimer equation accounts for the fact that the flow is only through the void part of the solid matrix by dividing the specific discharge by the porosity, n .

$$\vec{V} = \frac{\vec{q}}{n}$$

so that

$$\vec{V} = -\frac{1}{n} \mathbf{K} \nabla \phi^*$$

Hubbert defined the potential ϕ^* of a *compressible* fluid as

$$\phi^* = z + \int_{p_0}^p \frac{dp}{g\rho(p)}$$

where z is the elevation, p is the pressure and g is the gravitational constant. In general, the density ρ depends on temperature, concentration of dissolved matter and pressure. So, the potential ϕ^* is for a compressible fluid under isothermal conditions. The *coefficient of compressibility of a fluid* is expressed as, Bear[12],

$$\beta = \frac{1}{\rho} \frac{\partial \rho}{\partial p}$$

and its solution is given by

$$\rho = \rho_0 \exp[\beta(p - p_0)]$$

so that from the Hubbert potential

$$\begin{aligned} \phi^* &= z + \int_{p_0}^p \frac{dp}{g\rho_0 \exp[\beta(p - p_0)]} \\ &= z + \frac{1}{g\rho_0\beta} - \frac{1}{g\rho_0\beta} e^{-\beta(p-p_0)} \\ &= z + \frac{1}{g\rho_0\beta} - \frac{1}{g\beta\rho} \end{aligned}$$

Using this relationship and differentiating,

$$\nabla\phi^* = \nabla z + \frac{1}{g\beta\rho^2} \nabla\rho \quad \text{and} \quad \nabla\rho = \beta\rho\nabla p$$

which leads to

$$\nabla\phi^* = \nabla z + \frac{1}{g\rho} \nabla p$$

so that

$$\vec{J} = -\nabla\phi^* = \frac{-1}{g\rho} (\nabla p + \rho g \nabla z)$$

Taking Z to be depth, $Z = -z$, it follows from Darcy's law that,

$$\vec{q} = \mathbf{K}\vec{J} = -\mathbf{K}\nabla\phi^* = \frac{-\mathbf{K}}{g\rho} (\nabla p - g\rho\nabla Z)$$

The tensor \mathbf{K} can be expressed as

$$\mathbf{K} = \frac{\mathbf{k}g\rho}{\mu}$$

which means that

$$\vec{V} = \frac{\vec{q}}{n} = \frac{-\mathbf{K}}{ng\rho} (\nabla p - g\rho\nabla Z) = \frac{-\mathbf{k}}{n\mu} (\nabla p - g\rho\nabla Z) \quad (3.1)$$

where \mathbf{k} is the permeability tensor and μ is the viscosity.

In Chapter 2, the basic form of the concentration equation was illustrated using assumptions about the stochastic nature of the variables. Here the concentration equation is developed from a more traditional point of view using the conservation of mass law while at the same time retaining a stochastic flavor. Let $c(\vec{x}, t)$ represent the concentration of the pollutant, ie, $c(\vec{x}, t)$ is the mass of the pollutant per unit volume of the single phase. The conservation of mass law then states that

$$\frac{\partial}{\partial t} \int_{\Omega} nc(\vec{x}, t)d\vec{x} = - \int_{\partial\Omega} c(\vec{x}, t)\vec{u} \cdot \vec{\nu}d\vec{x} + \int_{\Omega} qd\vec{x} \quad (3.2)$$

where

n = Porosity, The Fraction Of The Volume Ω Available For Flow

\vec{u} = Darcy Velocity or Volumetric Flow Rate Across A

q = Mass Flow Rate Per Unit Volume Injected Into Ω

Using the Divergence Theorem,

$$\frac{\partial}{\partial t} \int_{\Omega} nc(\vec{x}, t)d\vec{x} = - \int_{\Omega} \nabla \cdot c(\vec{x}, t)\vec{u}d\vec{x} + \int_{\Omega} qd\vec{x}$$

The total flux of the pollutant is given by the product

$$c(\vec{x}, t)\vec{V}(\vec{x}, t)$$

On a representative elementary volume, REV, the concentration and velocity can be represented as

$$c(\vec{x}, t) = \mathbf{E}[c(\vec{x}, t)] + c'(\vec{x}, t) \quad \mathbf{E}[c'(\vec{x}, t)] = 0$$

$$\vec{V}(\vec{x}, t) = \mathbf{E}[\vec{V}(\vec{x}, t)] + \vec{V}'(\vec{x}, t) \quad \mathbf{E}[\vec{V}'(\vec{x}, t)] = 0$$

then by multiplying and taking expected values,

$$\mathbf{E}[c\vec{V}] = \underbrace{\mathbf{E}[c]\mathbf{E}[\vec{V}]}_{convective\ flux} + \underbrace{\mathbf{E}[c'\vec{V}']}_{dispersive\ flux}$$

Assuming, at least locally, *e.g.*, on a grid block, that the dispersive flux is Fickian, then

$$\mathbf{E}[c' \vec{V}'] = -\mathbf{D} \nabla \mathbf{E}[c]$$

The velocity \vec{V} represents the velocity of the fluid in the void part of the matrix, and so we can write, on average, that

$$\mathbf{E}[c\vec{u}] = n\mathbf{E}[c\vec{V}] = n \left(\mathbf{E}[c]\mathbf{E}[\vec{V}] - \mathbf{D} \nabla \mathbf{E}[c] \right)$$

Hence, by taking expected values in Equation[3.2], it follows that

$$\frac{\partial}{\partial t} \int_{\Omega} n\mathbf{E}[c(\vec{x}, t)] d\vec{x} = - \int_{\Omega} \nabla \cdot n \left(\mathbf{E}[c]\mathbf{E}[\vec{V}] - \mathbf{D} \nabla \mathbf{E}[c] \right) d\vec{x} + \int_{\Omega} \mathbf{E}[q] d\vec{x}$$

Since this equation is in terms of volume averaged variables, the integrals can be removed, see Gray[49], so that

$$\frac{\partial}{\partial t} n\mathbf{E}[c(\vec{x}, t)] = -\nabla \cdot n \left(\mathbf{E}[c]\mathbf{E}[\vec{V}] - \mathbf{D} \nabla \mathbf{E}[c] \right) + \mathbf{E}[q]$$

So, for the case of the density being dependent on the concentration, $\rho = \rho(c)$ the system of equations can be written as (dropping expected value symbols)

$$\begin{aligned} \frac{\partial}{\partial t}(nc) &= -\nabla \cdot n \left(c\vec{V} - \mathbf{D} \nabla c \right) + q \\ \vec{V} &= -\frac{-\mathbf{k}}{n\mu} (\nabla p - g\rho \nabla Z) \\ \frac{\partial(n\rho)}{\partial t} &= -\nabla \cdot n\rho\vec{V} \end{aligned} \tag{3.3}$$

Clearly, this system assumes that the porosity is given by $n(\vec{x}, t)$ and that the four unknowns are $c(\vec{x}, t)$, $\vec{V}(\vec{x}, t)$, $\rho(\vec{x}, t)$ and $p(\vec{x}, t)$.

3.1.2 The Mixed Model

Using the mixed method, the velocity and pressure can be solved for at the same time. The Mixed Model introduces the space

$$H(\text{div}; \Omega) = \{\vec{q} \in (L^2(\Omega))^2 : \text{div}(\vec{q}) \in L^2(\Omega)\}$$

with norm

$$\|\vec{q}\|_{H(\text{div};\Omega)} = (\|\vec{q}\|_{0,\Omega}^2 + \|\text{div}(\vec{q})\|_{0,\Omega}^2)^{1/2}$$

The *Mixed Variational Formulation* is then defined as:

Find a pair of functions

$$(\vec{u}, p) \in H(\text{div};\Omega) \times L^2(\Omega)$$

such that

$$\forall \vec{q} \in H(\text{div};\Omega), \quad \int_{\Omega} \vec{u} \cdot \vec{q} dx + \int_{\Omega} p \text{div}(\vec{q}) dx = 0$$

and

$$\forall v \in L^2(\Omega), \quad \int_{\Omega} v(\text{div}(\vec{u}) + f) dx = 0, \quad f \in L^2(\Omega)$$

The *Discrete Mixed Formulation* requires that there exist two finite dimensional spaces \vec{Q}_h and V_h such that

$$\vec{Q}_h \subset H(\text{div};\Omega) \quad V_h \subset L^2(\Omega)$$

and a pair $(\vec{u}_h, p_h) \in \vec{Q}_h \times V_h$ is sought such that

$$\forall \vec{q}_h \in \vec{Q}_h, \quad \int_{\Omega} \vec{u}_h \cdot \vec{q}_h dx + \int_{\Omega} p_h \text{div}(\vec{q}_h) dx = 0$$

and

$$\forall v_h \in V_h, \quad \int_{\Omega} v_h(\text{div}(\vec{u}_h) + f) dx = 0, \quad f \in L^2(\Omega)$$

A key element in the existence theorem of the mixed finite element analysis is the *Babuška-Brezzi Condition*.

Theorem (Existence):

Assume that $\vec{q}_h \in \vec{Q}_h$ and that

$$\forall v_h \in V_h, \quad \int_{\Omega} v_h \text{div}(\vec{q}_h) dx = 0 \quad \Rightarrow \quad \text{div}(\vec{q}_h) = 0$$

and that \exists a constant $\alpha > 0$ such that

$$\forall v_h \in V_h, \quad \sup_{\vec{q}_h \in \vec{Q}_h} \frac{\int_{\Omega} v_h \operatorname{div}(\vec{q}_h) dx}{\|\vec{q}_h\|_{H(\operatorname{div}; \Omega)}} \geq \alpha \|v_h\|_{0, \Omega}$$

Then the problem (P_h) has a unique solution $(\vec{u}_h, p_h) \in \vec{Q}_h \times V_h$ and \exists a constant $\tau > 0$, depending on α , such that

$$\|\vec{u} - \vec{u}_h\|_{H(\operatorname{div}; \Omega)} + \|p - p_h\|_{0, \Omega} \leq \tau \left[\inf_{\vec{q}_h \in \vec{Q}_h} \|\vec{u} - \vec{q}_h\|_{H(\operatorname{div}; \Omega)} + \inf_{v_h \in V_h} \|p - v_h\|_{0, \Omega} \right] \quad \square$$

Starting from the second line in Equation[3.3], page 99,

$$\vec{V} = \frac{-\mathbf{k}}{n\mu} (\nabla p - g\rho \nabla Z)$$

and taking the divergence,

$$\nabla \cdot \vec{V} = \nabla \cdot \left(\frac{-\mathbf{k}}{n\mu} \nabla p \right) + \nabla \cdot (g\rho \nabla Z)$$

which is written as

$$\nabla \cdot \left(-\frac{\mathbf{k}}{n\mu} \nabla p \right) = f \quad (3.4)$$

Letting

$$\vec{u}_h = \frac{-\mathbf{k}}{n\mu} \nabla p_h$$

and multiplying this equation by any $\vec{q}_h \in \vec{Q}_h$ and integrating over Ω ,

$$\int_{\Omega} n\mu \mathbf{k}^{-1} \vec{u}_h \cdot \vec{q}_h d\vec{x} + \int_{\Omega} \nabla p_h \cdot \vec{q}_h d\vec{x} = 0$$

Integrating the second integral by parts using the Green's formula

$$\int_{\Omega} p \nabla \cdot \vec{q} d\vec{x} + \int_{\Omega} \nabla p \cdot \vec{q} d\vec{x} = \oint_{\partial\Omega} p \vec{q} \cdot \vec{\nu} d\gamma$$

and assuming no flow boundary conditions, $\vec{q}_h \cdot \vec{\nu} = 0$ on $\partial\Omega$, it follows that

$$\int_{\Omega} n\mu\mathbf{k}^{-1}\vec{u}_h \cdot \vec{q}_h d\vec{x} - \int_{\Omega} p_h \nabla \cdot \vec{q}_h d\vec{x} = 0 \quad (3.5)$$

And, from Equation[3.4] with $v_h \in V_h$

$$\int_{\Omega} v_h (\nabla \cdot \vec{u}_h - f) d\vec{x} = 0 \quad (3.6)$$

Hence, Equations[3.5] and [3.6] represent the Discrete Mixed Formulation for the pressure equation.

3.1.3 The Spaces \vec{Q}_h And V_h

In order to define bases for the Raviart-Thomas finite element spaces \vec{Q}_h and V_h for the 3 dimensional case, first let Ω be the 3-simplex $[0, L_1] \times [0, L_2] \times [0, L_3]$ and define the meshes:

$$\Delta X : 0 = x_0 < x_1 < \cdots < x_k = L_1$$

$$\Delta Y : 0 = y_0 < y_1 < \cdots < y_j = L_2$$

$$\Delta Z : 0 = z_0 < z_1 < \cdots < z_j = L_3$$

And, define the piecewise polynomial space

$$\mathcal{M}_q^r(\Delta) = \{v \in C^q[0, L] : v \text{ is a polynomial of degree } \leq r \text{ on each subinterval of } \Delta\}$$

where $q = -1$ refers to discontinuous functions.

Since in the error analysis of Raviart-Thomas[77] there is no requirement that the elements of the subspace V_h be continuous across inter-element boundaries, the finite dimensional subspace for pressure used in Russell and Wheeler[84] for the 2-dimensional case is extended to the 3-dimensional case

$$V_h = \mathcal{M}_{-1}^1(\Delta X) \otimes \mathcal{M}_{-1}^1(\Delta Y) \otimes \mathcal{M}_{-1}^1(\Delta Z)$$

It is further assumed in Raviart-Thomas[77] that given any $v_h \in V_h$ there exists a $\vec{q}_h \in \vec{Q}_h$ such that

$$\operatorname{div} \vec{q}_h = v_h$$

which leads to the definition of the subspace \vec{Q}_h as

$$(\vec{Q}_h)_x = \mathcal{M}_0^2(\Delta X) \otimes \mathcal{M}_{-1}^1(\Delta Y) \otimes \mathcal{M}_{-1}^1(\Delta Z)$$

$$(\vec{Q}_h)_y = \mathcal{M}_{-1}^1(\Delta X) \otimes \mathcal{M}_0^2(\Delta Y) \otimes \mathcal{M}_{-1}^1(\Delta Z)$$

$$(\vec{Q}_h)_z = \mathcal{M}_{-1}^1(\Delta X) \otimes \mathcal{M}_{-1}^1(\Delta Y) \otimes \mathcal{M}_0^2(\Delta Z)$$

So that,

$$\vec{Q}_h = (\vec{Q}_h)_x \times (\vec{Q}_h)_y \times (\vec{Q}_h)_z$$

Theorem 5 of Raviart-Thomas[77] then gives the error estimate

$$\|\vec{q} - \vec{q}_h\|_{H(\operatorname{div};\Omega)} + \|p - p_h\|_{0,\Omega} \leq \mathcal{K} h^{r+1} (|p|_{r+1,\Omega} + |p|_{r+2,\Omega} + |\Delta p|_{r+1,\Omega})$$

where \mathcal{K} is a constant independent of h , and from which it follows that for $r = 0$

$$\|\vec{q} - \vec{q}_h\|_{H(\operatorname{div};\Omega)} \leq O(h)$$

and

$$\|p - p_h\|_{0,\Omega} \leq O(h)$$

so that the error estimates for the pressure and velocity are of the same order.

For the general case where density and viscosity are dependent on the concentration, the following set of *coupled* equations holds:

$$\frac{\partial}{\partial t}(nc) = -\nabla \cdot n(c\vec{V} - \mathbf{D}\nabla c) + q$$

$$\vec{V} = -\frac{\mathbf{k}}{n\mu}(\nabla p - g\rho\nabla Z)$$

$$\frac{\partial(n\rho)}{\partial t} = -\nabla \cdot (n\rho\vec{V})$$

In the tracer case, where the density is constant, the equations can be written in the following *uncoupled* form:

$$S \frac{\partial \phi^*}{\partial t} = \nabla \cdot \mathbf{T} \nabla \phi^*$$

$$\frac{\partial}{\partial t}(nc) = -\nabla \cdot n (c\vec{V} - \mathbf{D}\nabla c) + q$$

where S is the storativity of the aquifer, \mathbf{T} is the transmissivity, and ϕ^* is the piezometric head.

3.2 2D Finite Element Solution

3.2.1 General

The goal of this current investigation is to study the system of equations specified by Equation[3.3], page 99, with the coefficients of the dispersivity tensor calculated in terms of the velocity covariances or permeability covariances. The dispersivity tensor would then be a piecewise constant tensor specified on a grid block basis. Clearly, the system specified by Equation[3.3] is a *coupled* system and must be solved simultaneously. This is because the density of the single phase is considered to be a function of the concentration of the pollutant. It is conceivable that in many cases the density may be considered to be constant, in which case the flow equation and the concentration equation may be solved separately. It is this case that will be studied first.

3.2.2 Finite Elements

A finite element in \mathfrak{R}^n is defined, Ciarlet[24], as a triple (T, Ψ, Σ) such that

- (1) T is a closed subset of \mathfrak{R}^n with nonempty interior and Lipschitz-continuous boundary.
- (2) Ψ is a finite dimensional space of real valued functions defined over the set T with $\dim \Psi = N$.
- (3) Σ is a set of linear functionals, σ_i , $1 \leq i \leq N$ defined on the space Ψ such that $\exists \psi_j \in \Psi$, $1 \leq j \leq N$, with the property that

$$\sigma_i(\psi_j) = \delta_{ij}$$

Also, the following holds

$$\psi = \sum_{i=1}^N \sigma_i(\psi) \psi_i \quad \forall \psi \in \Psi$$

The sets $\{\sigma_i\}_{i=1}^N$ and $\{\psi_i\}_{i=1}^N$ are dual bases. Suppose that there are two finite elements $(D, \hat{\Psi}, \hat{\Sigma})$ and (Ω_e, Ψ, Σ) that are related by the invertible affine mapping

$$\begin{aligned} F : D &\rightarrow \Omega_e && \ni \\ F(\vec{x}) &= \mathbf{A}\vec{x} + \vec{b} \end{aligned}$$

then the finite elements are *affinely equivalent* if

- (1) $F(D) = \Omega_e$
- (2) $\psi = \hat{\psi} \circ F^{-1}$
- (3) $\vec{a}_i = F(\vec{\hat{a}}_i)$ where the $\vec{\hat{a}}_i$ are the element nodes.

The following relationship between mappings is defined

$$\begin{array}{ccc} \Omega_e & \xrightarrow{G} & D \\ & \searrow & \downarrow \hat{\psi} \\ & & \mathfrak{R}^1 \end{array}$$

$G^* \hat{\psi} = \hat{\psi} \circ G$

Figure 7A

The mapping G^* in Figure 7A is called a *pull-back* since it pulls a function defined on D back to a function defined on Ω_e . It follows from (2) that if $G = F^{-1}$, then $\psi = \hat{\psi} \circ G = G^* \hat{\psi}$ where G^* is a *pull-back*.

Similarly, the mapping $G^* : \hat{\Psi} \rightarrow \Psi$ allows the definition of a *push-forward* mapping $G_* : \Sigma \rightarrow \hat{\Sigma}$ according to the following diagram:

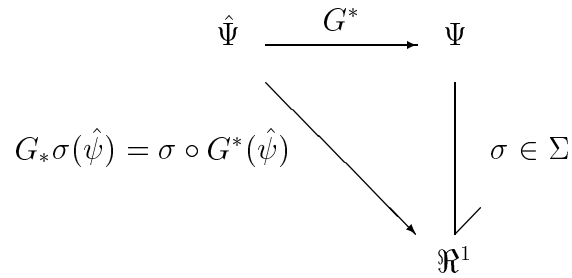


Figure 7B

The function G_* pushes a functional defined on Ψ forward to a functional defined on $\hat{\Psi}$.

3.2.3 Rectangular Elements

In implementing the *uncoupled* version of the 2D finite element model, the elements are assumed to be rectangular with their edges aligned with the x and y axes as shown in Figure 8:

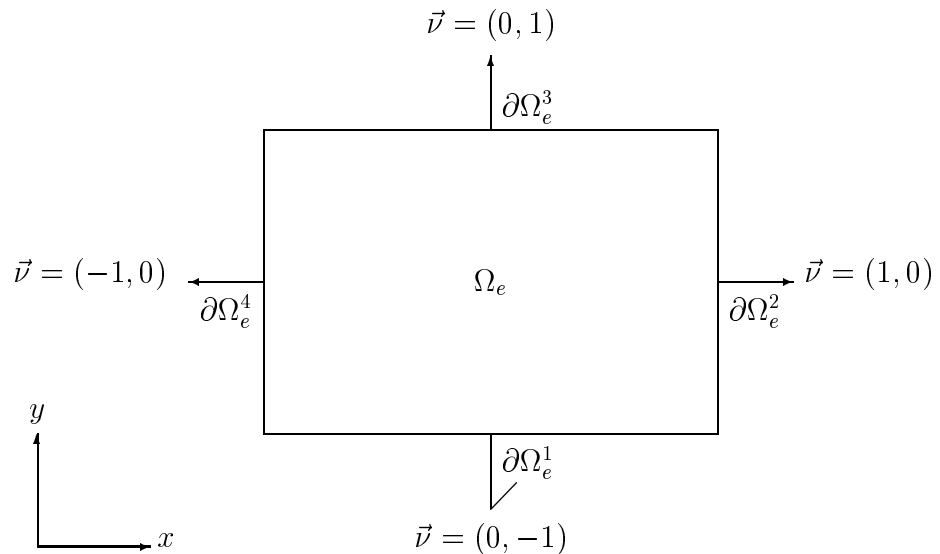


Figure 8 - Rectangular Element

Here $\vec{\nu} = \nu_x \vec{i} + \nu_y \vec{j} = \cos(\alpha) \vec{i} + \sin(\alpha) \vec{j}$ where α is the angle between the outward normal to the boundary and the x -axis.

For integration purposes, using Gaussian formulas, it is convenient to use as a reference element the following rectangle:

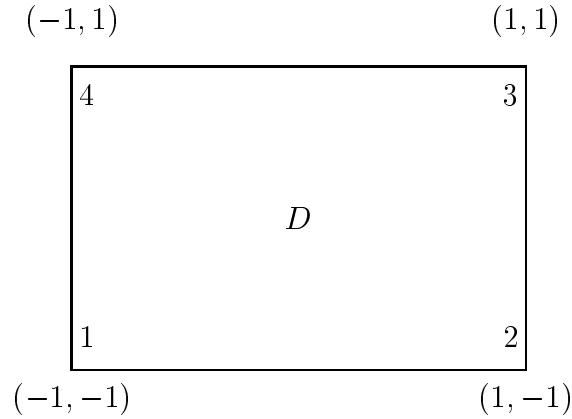


Figure 9 - Reference Element

The numbers in the corners of the element in Figure 9 denote the numbering of the local nodes. On the element D , the first order (linear) Lagrange shape functions are given by

$$\begin{aligned}
 \hat{\psi}_1 &= \frac{1}{4}(1 - \xi)(1 - \eta) \\
 \hat{\psi}_2 &= \frac{1}{4}(1 + \xi)(1 - \eta) \\
 \hat{\psi}_3 &= \frac{1}{4}(1 + \xi)(1 + \eta) \\
 \hat{\psi}_4 &= \frac{1}{4}(1 - \xi)(1 + \eta)
 \end{aligned}$$

These reference functions can be transformed to any rectangular element whose side is of length a in the x -direction and of length b in the y -direction. Letting (x_1, y_1) represent the *global* coordinate of node 1, the transformation is given by

$$\begin{bmatrix} x \\ y \end{bmatrix} = F \left(\begin{bmatrix} \xi \\ \eta \end{bmatrix} \right) = \begin{bmatrix} \frac{a}{2} & 0 \\ 0 & \frac{b}{2} \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix} + \begin{bmatrix} \frac{a+2x_1}{2} \\ \frac{b+2y_1}{2} \end{bmatrix}$$

For example, the 4 corners of the reference rectangle map in the following way:

$$\begin{aligned} \text{Node 1 :} & \quad (-1, -1) \rightarrow (x_1, y_1) \\ \text{Node 2 :} & \quad (1, -1) \rightarrow (a + x_1, y_1) \\ \text{Node 3 :} & \quad (1, 1) \rightarrow (x_1 + a, y_1 + b) \\ \text{Node 4 :} & \quad (-1, 1) \rightarrow (x_1, y_1 + b) \end{aligned}$$

The inverse mapping then becomes

$$\begin{bmatrix} \xi \\ \eta \end{bmatrix} = F^{-1} \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) = \begin{bmatrix} \frac{2}{a} & 0 \\ 0 & \frac{2}{b} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \frac{-2x_1 - a}{a} \\ \frac{-2y_1 - b}{b} \end{bmatrix}$$

Using this to transform the linear Lagrangian shape functions to the transformed element Ω_e , results in

$$\begin{aligned} \hat{\psi}_1 & \rightarrow \left(1 - \frac{x-x_1}{a}\right) \left(1 - \frac{y-y_1}{b}\right) = \psi_1 \\ \hat{\psi}_2 & \rightarrow \left(\frac{x-x_1}{a}\right) \left(1 - \frac{y-y_1}{b}\right) = \psi_2 \\ \hat{\psi}_3 & \rightarrow \left(\frac{x-x_1}{a}\right) \left(\frac{y-y_1}{b}\right) = \psi_3 \\ \hat{\psi}_4 & \rightarrow \left(1 - \frac{x-x_1}{a}\right) \left(\frac{y-y_1}{b}\right) = \psi_4 \end{aligned}$$

3.2.4 Numerical Integration

Let D be the reference element whose corners are given by

$$\{(-1, -1), (1, -1), (-1, 1), (1, 1)\}$$

and let

$$F : D \rightarrow \Omega_e$$

such that

$$\begin{bmatrix} x \\ y \end{bmatrix} = F \left(\begin{bmatrix} \xi \\ \eta \end{bmatrix} \right) = \begin{bmatrix} \frac{a}{2} & 0 \\ 0 & \frac{b}{2} \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix} + \begin{bmatrix} \frac{a+2x_1}{2} \\ \frac{b+2y_1}{2} \end{bmatrix}$$

then

$$\int_{\Omega_e} f(x, y) d\Omega_e = \int_D f(x(\xi, \eta), y(\xi, \eta)) \left| \frac{\partial(x, y)}{\partial(\xi, \eta)} \right| d\xi d\eta$$

The quadrature formulas for evaluating integrals over the reference rectangle, D , are derived from the quadrature formulas for the 1D case. Letting

$$\Psi(\xi, \eta) = f(x(\xi, \eta), y(\xi, \eta)) \left| \frac{\partial(x, y)}{\partial(\xi, \eta)} \right|$$

then

$$\begin{aligned} \int_D \Psi(\xi, \eta) d\xi d\eta &= \int_{-1}^1 \left(\int_{-1}^1 \Psi(\xi, \eta) d\eta \right) d\xi \\ &\approx \int_{-1}^1 \left(\sum_{j=1}^n \Psi(\xi, \eta_j) W_j \right) d\xi \\ &\approx \sum_{i=1}^m \sum_{j=1}^n \Psi(\xi_i, \eta_j) W_i W_j \end{aligned}$$

where m and n are the number of quadrature points in the ξ and η directions, (ξ_i, η_j) are the Gauss points and the W_i and W_j are the Gauss weights.

The number of Gauss points to use is based on the result that if there are $n + 1$ Gauss points, then the formula

$$\sum_{i=1}^n a_i f(x_i)$$

is exact for a polynomial of degree $2n + 1$. If the polynomial has degree p , then

$$2n + 1 = p \quad \Rightarrow \quad n + 1 = \frac{p + 1}{2}$$

so that the number of Gauss points (n_{gp}) is given by

$$n_{gp} = \begin{cases} \frac{p+1}{2} & \text{if } p+1 \text{ is even} \\ \frac{p}{2} + 1 & \text{if } p+1 \text{ is odd} \end{cases}$$

For the linear Lagrange functions, the highest degree in both the ξ and η directions is 2. So, in the numerical integrations $m = n = 2$.

3.2.5 Groundwater Flow Equation

The 2D groundwater flow equation is given by

$$S \frac{\partial \phi^*}{\partial t} = \nabla \cdot \mathbf{T} \nabla \phi^*$$

where S is the storativity of the aquifer, \mathbf{T} is the transmissivity, and ϕ^* is the piezometric head. The semi-discrete variational formulation takes the form with respect to a four-node rectangular element Ω_e

$$0 = \int_{\Omega_e} v S \frac{\partial \phi^*}{\partial t} d\Omega_e - \int_{\Omega_e} v \nabla \cdot \mathbf{T} \nabla \phi^* d\Omega_e$$

The second integral on the right side can be integrated via Green's formula

$$\int_{\Omega} \vec{q} \cdot \nabla p d\Omega + \int_{\Omega} p \nabla \cdot \vec{q} d\Omega = \oint_{\partial\Omega} p \vec{q} \cdot \vec{\nu} d\gamma$$

to give

$$0 = \int_{\Omega_e} v S \frac{\partial \phi^*}{\partial t} d\Omega_e + \int_{\Omega_e} \mathbf{T} \nabla \phi^* \cdot \nabla v d\Omega_e - \oint_{\partial\Omega_e} v \mathbf{T} \nabla \phi^* \cdot \vec{\nu} d\gamma$$

Assuming a uniform time step, Δt , and a backward Euler estimate of the time derivative,

$$\frac{\partial \phi^*}{\partial t} \approx \frac{\phi^{*n} - \phi^{*n-1}}{\Delta t}$$

the fully discrete variational formulation is

$$\begin{aligned} 0 &\approx \int_{\Omega_e} v S \left(\frac{\phi^{*n} - \phi^{*n-1}}{\Delta t} \right) d\Omega_e + \int_{\Omega_e} \mathbf{T}^n \nabla \phi^{*n} \cdot \nabla v d\Omega_e \\ &\quad - \oint_{\partial\Omega_e} v \mathbf{T}^n \nabla \phi^{*n} \cdot \vec{\nu} d\gamma \end{aligned}$$

where the superscript n represents the n^{th} timestep. Then making the substitutions

$$\begin{aligned} \phi^{*n} &= \sum_{j=1}^r \phi_j^*(t_n) \psi_j(\vec{x}) & \vec{x} &= (x, y) \\ v &= \psi_i(\vec{x}) \end{aligned}$$

and letting

$$\mathbf{T}^n = \begin{bmatrix} T_{11}^n & T_{12}^n \\ T_{21}^n & T_{22}^n \end{bmatrix}$$

The system of equations for the e^{th} element becomes

$$\begin{aligned} & \sum_{j=1}^r \frac{S\phi_j^*(t_n)}{\Delta t} \int_{\Omega_e} \psi_i \psi_j d\Omega_e + \sum_{j=1}^r T_{11}^n \phi_j^*(t_n) \int_{\Omega_e} \frac{\partial \psi_j}{\partial x} \frac{\partial \psi_i}{\partial x} d\Omega_e \\ & + \sum_{j=1}^r T_{12}^n \phi_j^*(t_n) \int_{\Omega_e} \frac{\partial \psi_j}{\partial y} \frac{\partial \psi_i}{\partial x} d\Omega_e \\ & + \sum_{j=1}^r T_{21}^n \phi_j^*(t_n) \int_{\Omega_e} \frac{\partial \psi_j}{\partial x} \frac{\partial \psi_i}{\partial y} d\Omega_e \\ & + \sum_{j=1}^r T_{22}^n \phi_j^*(t_n) \int_{\Omega_e} \frac{\partial \psi_j}{\partial y} \frac{\partial \psi_i}{\partial y} d\Omega_e \\ & - \oint_{\partial\Omega_e} v \mathbf{T}^n \nabla \phi^{*n} \cdot \vec{\nu} d\gamma \\ & \approx \sum_{j=1}^r \frac{S\phi_j^*(t_{n-1})}{\Delta t} \int_{\Omega_e} \psi_i \psi_j d\Omega_e \quad i = 1, 2, \dots, r \end{aligned}$$

The surface integral in this equation takes the form

$$\begin{aligned} \oint_{\partial\Omega_e} v \mathbf{T}^n \nabla \phi^{*n} \cdot \vec{\nu} d\gamma &= T_{11}^n \oint_{\partial\Omega_e} v \frac{\partial \phi^{*n}}{\partial x} \nu_1 d\gamma \\ &+ T_{12}^n \oint_{\partial\Omega_e} v \frac{\partial \phi^{*n}}{\partial y} \nu_1 d\gamma \\ &+ T_{21}^n \oint_{\partial\Omega_e} v \frac{\partial \phi^{*n}}{\partial x} \nu_2 d\gamma \\ &+ T_{22}^n \oint_{\partial\Omega_e} v \frac{\partial \phi^{*n}}{\partial y} \nu_2 d\gamma \end{aligned} \quad (3.7)$$

On $\partial\Omega_e^1$, $\vec{\nu} = (\nu_1, \nu_2) = (0, -1)$ this becomes

$$- \left(T_{21}^n \sum_{j=1}^r \phi_j^*(t_n) \oint_{\partial\Omega_e^1} \psi_i \frac{\partial \psi_j}{\partial x} dx + T_{22}^n \sum_{j=1}^r \phi_j^*(t_n) \oint_{\partial\Omega_e^1} \psi_i \frac{\partial \psi_j}{\partial y} dx \right)$$

And, with the no flow boundary condition on $\partial\Omega_e^1$, this becomes

$$-T_{21}^n \sum_{j=1}^r \phi_j^*(t_n) \oint_{\partial\Omega_e^1} \psi_i \frac{\partial\psi_j}{\partial x} dx$$

On $\partial\Omega_e^2$, $\vec{\nu} = (\nu_1, \nu_2) = (1, 0)$ Equation[3.7] becomes

$$T_{11}^n \sum_{j=1}^r \phi_j^*(t_n) \oint_{\partial\Omega_e^2} \psi_i \frac{\partial\psi_j}{\partial x} dy + T_{12}^n \sum_{j=1}^r \phi_j^*(t_n) \oint_{\partial\Omega_e^2} \psi_i \frac{\partial\psi_j}{\partial y} dy$$

And, since ϕ^* is constant along this boundary, this becomes

$$T_{11}^n \sum_{j=1}^r \phi_j^*(t_n) \oint_{\partial\Omega_e^2} \psi_i \frac{\partial\psi_j}{\partial x} dy$$

On $\partial\Omega_e^3$, $\vec{\nu} = (\nu_1, \nu_2) = (0, 1)$ Equation[3.7] becomes

$$T_{21}^n \sum_{j=1}^r \phi_j^*(t_n) \oint_{\partial\Omega_e^3} \psi_i \frac{\partial\psi_j}{\partial x} dx + T_{22}^n \sum_{j=1}^r \phi_j^*(t_n) \oint_{\partial\Omega_e^3} \psi_i \frac{\partial\psi_j}{\partial y} dx$$

And, with the no flow boundary condition, this becomes

$$T_{21}^n \sum_{j=1}^r \phi_j^*(t_n) \oint_{\partial\Omega_e^3} \psi_i \frac{\partial\psi_j}{\partial x} dx$$

On $\partial\Omega_e^4$, $\vec{\nu} = (\nu_1, \nu_2) = (-1, 0)$ Equation[3.7] becomes

$$-\left(T_{11}^n \sum_{j=1}^r \phi_j^*(t_n) \oint_{\partial\Omega_e^4} \psi_i \frac{\partial\psi_j}{\partial x} dy + T_{12}^n \sum_{j=1}^r \phi_j^*(t_n) \oint_{\partial\Omega_e^4} \psi_i \frac{\partial\psi_j}{\partial y} dy \right)$$

And, since ϕ^* is constant along this boundary, this is

$$-T_{11}^n \sum_{j=1}^r \phi_j^*(t_n) \oint_{\partial\Omega_e^4} \psi_i \frac{\partial\psi_j}{\partial x} dy$$

where $\vec{\nu} = \nu_x \vec{i} + \nu_y \vec{j} = \cos(\alpha) \vec{i} + \sin(\alpha) \vec{j}$ and α is the angle between the outward normal to the boundary and the x -axis.

Furthermore, it is not necessary to compute the boundary integrals on a boundary between two elements in the interior of the global domain. Consider the following two elements that have a common boundary that lies in the interior of the global domain.

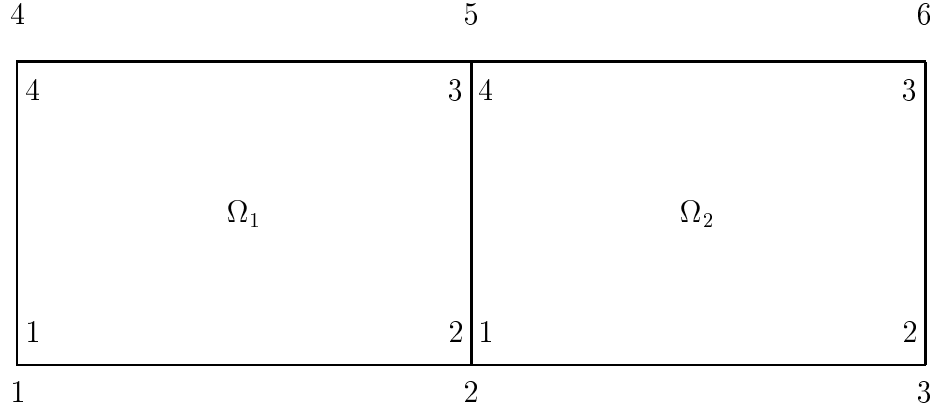


Figure 10 - Boundary Between Elements

Along this common boundary in element Ω_1 , we have the integral

$$\oint_{\partial\Omega_1^2} v \mathbf{T}^n \nabla \phi^{*n} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} d\gamma$$

And, along the common boundary in element Ω_2 , we have

$$\oint_{\partial\Omega_2^4} v \mathbf{T}^n \nabla \phi^{*n} \cdot \begin{pmatrix} -1 \\ 0 \end{pmatrix} d\gamma$$

These integrals are to be interpreted as unoriented flux integrals, so that $\partial\Omega_1^2 = \partial\Omega_2^4$, and the same tensor \mathbf{T}^n is used in both. This means that in the case that boundary coincides with the interface of two physical blocks of differing permeabilities, sufficient mixing has occurred along the interface to insure that there is not a discontinuity of permeabilities along the interface.

In the element Ω_1 of Figure 10, the shape functions have the following form on the common boundary:

$$\psi_1 = 0$$

$$\psi_2 = 1 - \frac{(y - y_1)}{b}$$

$$\psi_3 = \frac{y - y_1}{b}$$

$$\psi_4 = 0$$

And, in the element Ω_2 of Figure 10, the shape functions have the following form on the common boundary:

$$\psi_1 = 1 - \frac{(y - y_1)}{b}$$

$$\psi_2 = 0$$

$$\psi_3 = 0$$

$$\psi_4 = \frac{y - y_1}{b}$$

So that,

$$\psi_1 \text{ on } \Omega_2 = \psi_2 \text{ on } \Omega_1$$

and,

$$\psi_4 \text{ on } \Omega_2 = \psi_3 \text{ on } \Omega_1$$

Making the substitutions

$$\phi^{*n} = \sum_{j=1}^r \phi_j^*(t_n) \psi_j(\vec{x}) \quad \vec{x} = (x, y)$$

$$v = \psi_i(\vec{x})$$

and letting

$$\mathbf{T}^n = \begin{bmatrix} T_{11}^n & T_{12}^n \\ T_{21}^n & T_{22}^n \end{bmatrix}$$

it follows that

$$\begin{aligned} \oint_{\partial\Omega_1^2} v \mathbf{T}^n \nabla \phi^{*n} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} d\gamma &= \oint_{\partial\Omega_1^2} \psi_i T_{12}^n \left[\phi_2^*(t_n) \left(\frac{-1}{b} \right) \right] d\gamma \\ &+ \oint_{\partial\Omega_1^2} \psi_i T_{12}^n \left[\phi_3^*(t_n) \left(\frac{1}{b} \right) \right] d\gamma \end{aligned} \quad (3.8)$$

where $\phi_2^*(t_n)$ represents global point 2, $\phi_3^*(t_n)$ represents global point 5 and for non-zero results, ψ_i can be either ψ_2 or ψ_3 .

And,

$$\begin{aligned} \oint_{\partial\Omega_2^4} v \mathbf{T}^n \nabla \phi^{*n} \cdot \begin{pmatrix} -1 \\ 0 \end{pmatrix} d\gamma &= - \oint_{\partial\Omega_2^4} \psi_i T_{12}^n \left[\phi_1^*(t_n) \left(\frac{-1}{b} \right) \right] d\gamma \\ &- \oint_{\partial\Omega_2^4} \psi_i T_{12}^n \left[\phi_4^*(t_n) \left(\frac{1}{b} \right) \right] d\gamma \end{aligned} \quad (3.9)$$

where $\phi_1^*(t_n)$ represents global point 2, $\phi_4^*(t_n)$ represents global point 5 and for non-zero results, ψ_i can be either ψ_2 or ψ_3 . It is clear that these fluxes are equal but of opposite sign. Consequently, the contributions to either global point 2 or global point 5 cancel each other. Hence, it is only necessary to compute the boundary integrals when the local boundary coincides with the global boundary.

Letting

$$M_{ij} = \int_{\Omega_e} \psi_i \psi_j d\Omega_e$$

$$S_{ij}^{11} = \int_{\Omega_e} \frac{\partial \psi_i}{\partial x} \frac{\partial \psi_j}{\partial x} d\Omega_e \quad S_{ij}^{12} = \int_{\Omega_e} \frac{\partial \psi_i}{\partial x} \frac{\partial \psi_j}{\partial y} d\Omega_e$$

$$S_{ij}^{21} = \int_{\Omega_e} \frac{\partial \psi_i}{\partial y} \frac{\partial \psi_j}{\partial x} d\Omega_e \quad S_{ij}^{22} = \int_{\Omega_e} \frac{\partial \psi_i}{\partial y} \frac{\partial \psi_j}{\partial y} d\Omega_e$$

$$G_{ij}^{11} = \oint_{\partial\Omega_e} \psi_i \frac{\partial \psi_j}{\partial x} \nu_x d\gamma \quad G_{ij}^{21} = \oint_{\partial\Omega_e} \psi_i \frac{\partial \psi_j}{\partial x} \nu_y d\gamma$$

$$G_{ij}^{12} = \oint_{\partial\Omega_e} \psi_i \frac{\partial \psi_j}{\partial y} \nu_x d\gamma \quad G_{ij}^{22} = \oint_{\partial\Omega_e} \psi_i \frac{\partial \psi_j}{\partial y} \nu_y d\gamma$$

then letting

$$K_{ij} = T_{11}^n S_{ij}^{11} + T_{12}^n S_{ij}^{12} + T_{21}^n S_{ij}^{21} + T_{22}^n S_{ij}^{22}$$

and

$$G_{ij} = T_{11}^n G_{ij}^{11} + T_{12}^n G_{ij}^{12} + T_{21}^n G_{ij}^{21} + T_{22}^n G_{ij}^{22}$$

then in matrix form it follows that

$$\left(\frac{S}{\Delta t} [M_{ij}] + [K_{ij}] - [G_{ij}] \right) \{ \phi_j^*(t_n) \} \approx \frac{1}{\Delta t} [M_{ij}] \{ \phi_j^*(t_{n-1}) \} \quad (3.10)$$

3.2.6 Matrix Assembly

This section describes the assembly of the system of equations on the local rectangular elements into a global system for the entire domain. On a rectangular element Ω_e , the local nodes will be designated u_j , $j = 1, 2, 3, 4$. On the global domain, the mesh nodes will be designated U_j , $j = 1, \dots, n_{np}$, where n_{np} is the number of global mesh nodes.

The element nodal data is stored in the *element nodes array*, nod , which relates local node numbers to global node numbers by

$$nod[e][a] = A$$

where e is the *element number*, a is the *local node number* and A is the *global node number*. This array is set up from the input data.

Figure 11 illustrates the subdivision of a rectangular domain into 4 rectangular sub-domains, $\Omega_1, \Omega_2, \Omega_3$ and Ω_4 . In each sub-domain the corners of the rectangle are labeled 1, 2, 3, 4 in a counter-clockwise way. These numbers represent the local node numbers. And, the mesh nodes of the global domain are labeled 1 - 9 in a left to right and bottom to top way.

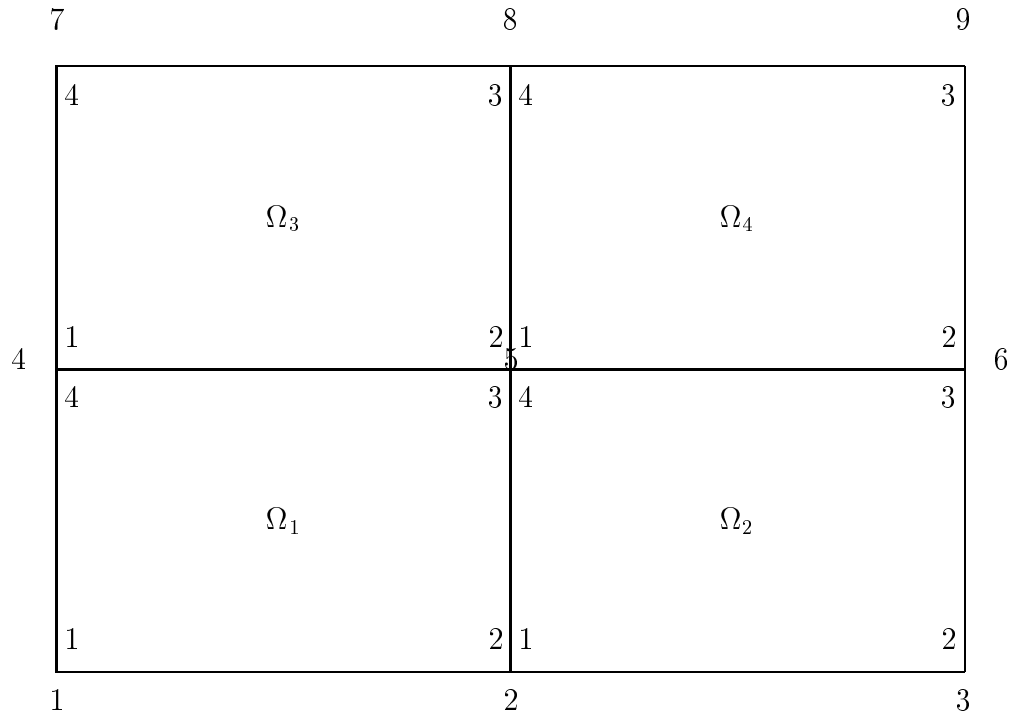


Figure 11 - Subdivision Of A Rectangular Domain

The element nodes array is then given by

$$\begin{aligned}
 \text{nod}[1][1] &= 1 & \text{nod}[1][2] &= 2 & \text{nod}[1][3] &= 5 & \text{nod}[1][4] &= 4 \\
 \text{nod}[2][1] &= 2 & \text{nod}[2][2] &= 3 & \text{nod}[2][3] &= 6 & \text{nod}[2][4] &= 5 \\
 \text{nod}[3][1] &= 4 & \text{nod}[3][2] &= 5 & \text{nod}[3][3] &= 8 & \text{nod}[3][4] &= 7 \\
 \text{nod}[4][1] &= 5 & \text{nod}[4][2] &= 6 & \text{nod}[4][3] &= 9 & \text{nod}[4][4] &= 8
 \end{aligned} \tag{3.11}$$

In this way each local node of each local rectangular element is associated with a global node in the mesh for the global domain.

Equation[3.10], page 116, describes the system of equations on a rectangular element Ω_e . If we let

$$A_{ij}^e = \frac{S}{\Delta t} [M_{ij}] + [K_{ij}] - [G_{ij}]$$

$$F_{ij}^e = \frac{1}{\Delta t} [M_{ij}] \{ \phi_j^*(t_{n-1}) \}$$

then the system on the element Ω_e can be written as

$$\begin{bmatrix} A_{11}^e & A_{12}^e & A_{13}^e & A_{14}^e \\ A_{21}^e & A_{22}^e & A_{23}^e & A_{24}^e \\ A_{31}^e & A_{32}^e & A_{33}^e & A_{34}^e \\ A_{41}^e & A_{42}^e & A_{43}^e & A_{44}^e \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} F_1^e \\ F_2^e \\ F_3^e \\ F_4^e \end{bmatrix}$$

Using the relationship [3.11] between the local nodes and the global nodes, the systems on Ω_1 and Ω_2 can be written as

$$\begin{bmatrix} A_{11}^1 & A_{12}^1 & A_{13}^1 & A_{14}^1 \\ A_{21}^1 & A_{22}^1 & A_{23}^1 & A_{24}^1 \\ A_{31}^1 & A_{32}^1 & A_{33}^1 & A_{34}^1 \\ A_{41}^1 & A_{42}^1 & A_{43}^1 & A_{44}^1 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \end{bmatrix} = \begin{bmatrix} F_1^1 \\ F_2^1 \\ F_3^1 \\ F_4^1 \end{bmatrix}$$

and

$$\begin{bmatrix} A_{11}^2 & A_{12}^2 & A_{13}^2 & A_{14}^2 \\ A_{21}^2 & A_{22}^2 & A_{23}^2 & A_{24}^2 \\ A_{31}^2 & A_{32}^2 & A_{33}^2 & A_{34}^2 \\ A_{41}^2 & A_{42}^2 & A_{43}^2 & A_{44}^2 \end{bmatrix} \begin{bmatrix} U_2 \\ U_3 \\ U_6 \\ U_5 \end{bmatrix} = \begin{bmatrix} F_1^2 \\ F_2^2 \\ F_3^2 \\ F_4^2 \end{bmatrix}$$

The formula for incorporating the elements of the local systems into the global system is derived from the equation

$$\text{nod}[e][a] = A$$

For example, if we have the associations

$$\text{nod}[e][a_1] = A_1$$

$$\text{nod}[e][a_2] = A_2$$

and if A_{ij} represents the ij^{th} element in the global matrix, then A_{ij} is built up with the mappings of the form

$$A_{A_1 A_1} \leftarrow A_{A_1 A_1} + A_{a_1 a_1}^e$$

$$A_{A_1 A_2} \leftarrow A_{A_1 A_2} + A_{a_1 a_2}^e$$

etc.

By either reading the associations off the previous diagram or by using the element nodes array, Equation[3.11], page 117, along with the associations just described, the coefficient matrix of the global system has the form:

$$\begin{bmatrix} A_{11}^1 & A_{12}^1 & 0 & A_{14}^1 & & A_{13}^1 & 0 & 0 & 0 & 0 \\ A_{21}^1 & A_{22}^1 + A_{11}^2 & A_{12}^2 & A_{24}^1 & & A_{23}^1 + A_{14}^2 & A_{13}^2 & 0 & 0 & 0 \\ 0 & A_{21}^2 & A_{22}^2 & 0 & & A_{24}^2 & A_{23}^2 & 0 & 0 & 0 \\ A_{41}^1 & A_{42}^1 & 0 & A_{44}^1 + A_{11}^3 & & A_{43}^1 + A_{12}^3 & 0 & A_{14}^4 & A_{13}^3 & 0 \\ A_{31}^1 & A_{32}^1 + A_{41}^2 & A_{42}^2 & A_{34}^1 + A_{21}^3 & A_{33}^1 + A_{44}^2 + A_{22}^3 + A_{11}^4 & A_{43}^2 + A_{12}^4 & A_{24}^3 & A_{23}^3 + A_{14}^4 & A_{13}^4 & 0 \\ 0 & A_{31}^2 & A_{32}^2 & 0 & & A_{34}^2 + A_{21}^4 & A_{33}^2 + A_{22}^4 & 0 & A_{24}^4 & A_{23}^4 \\ 0 & 0 & 0 & A_{41}^3 & & A_{42}^3 & 0 & A_{44}^3 & A_{43}^3 & 0 \\ 0 & 0 & 0 & A_{31}^3 & & A_{32}^3 + A_{41}^4 & A_{42}^4 & A_{34}^3 & A_{33}^3 + A_{44}^4 & A_{43}^4 \\ 0 & 0 & 0 & 0 & & A_{31}^4 & A_{32}^4 & 0 & A_{34}^4 & A_{33}^4 \end{bmatrix}$$

3.2.7 Velocity Field

Once the piezometric head has been computed using the flow equation, the associated velocity field can be computed from the equation

$$\vec{V} = -\frac{\mathbf{K}}{n} \nabla \phi^*$$

If \mathbf{K} is the hydraulic conductivity and B is the average depth of the tank, then we can write

$$\mathbf{K}' = \frac{1}{n} \mathbf{K} = \frac{1}{nB} \mathbf{T}$$

Hence,

$$\vec{V} = - \begin{bmatrix} K'_{11} & K'_{12} \\ K'_{21} & K'_{22} \end{bmatrix} \begin{bmatrix} \frac{\partial \phi^*}{\partial x} \\ \frac{\partial \phi^*}{\partial y} \end{bmatrix} = \begin{bmatrix} -K'_{11} \frac{\partial \phi^*}{\partial x} - K'_{12} \frac{\partial \phi^*}{\partial y} \\ -K'_{21} \frac{\partial \phi^*}{\partial x} - K'_{22} \frac{\partial \phi^*}{\partial y} \end{bmatrix} = \begin{bmatrix} v_x \\ v_y \end{bmatrix}$$

Here, ϕ^* on a given element, Ω_e is given by the equation

$$\phi^* = \sum_{i=1}^r u_i \psi_i$$

where

$$\begin{aligned}
\psi_1 &= \left(1 - \frac{(x - x_1)}{a}\right) \left(1 - \frac{(y - y_1)}{b}\right) \\
\psi_2 &= \left(\frac{x - x_1}{a}\right) \left(1 - \frac{(y - y_1)}{b}\right) \\
\psi_3 &= \left(\frac{x - x_1}{a}\right) \left(\frac{y - y_1}{b}\right) \\
\psi_4 &= \left(1 - \frac{(x - x_1)}{a}\right) \left(\frac{y - y_1}{b}\right)
\end{aligned}$$

In these equations, a is the horizontal length of the element, b is the vertical height of the element and the point (x_1, y_1) is the coordinate of the lower left-hand corner of the element.

3.2.8 The Transport Equation

The 2D transport equation for an incompressible porous medium is given by

$$\frac{\partial u}{\partial t} - \nabla \cdot (\mathbf{D}\nabla u) + \vec{V} \cdot \nabla u = f \quad (\vec{x}, t) \in \Omega \times T$$

where \mathbf{D} is the dispersion tensor which may depend on time or distance from the source and may have a component that is a stochastic process. The velocity vector, \vec{V} , is assumed to be in $(L^\infty(\Omega \times T))^2$, but may also have a random component.

The semi-discrete variational formulation takes the form with respect to a four-node rectangular element Ω_e

$$0 = \int_{\Omega_e} v \frac{\partial u}{\partial t} d\Omega_e - \int_{\Omega_e} v \nabla \cdot (\mathbf{D}\nabla u) d\Omega_e + \int_{\Omega_e} v \vec{V} \cdot \nabla u d\Omega_e - \int_{\Omega_e} v f d\Omega_e$$

The second integral on the right side can be integrated via Green's formula

$$\int_{\Omega} \vec{q} \cdot \nabla p d\Omega + \int_{\Omega} p \nabla \cdot \vec{q} d\Omega = \oint_{\partial\Omega} p \vec{q} \cdot \vec{\nu} d\gamma$$

to give

$$\begin{aligned}
0 &= \int_{\Omega_e} v \frac{\partial u}{\partial t} d\Omega_e + \int_{\Omega_e} \mathbf{D} \nabla u \cdot \nabla v d\Omega_e \\
&+ \int_{\Omega_e} v \vec{V} \cdot \nabla u d\Omega_e - \int_{\Omega_e} v f d\Omega_e \\
&- \oint_{\partial\Omega_e} v \mathbf{D} \nabla u \cdot \vec{\nu} d\gamma
\end{aligned}$$

Assuming a uniform time step, Δt , and a backward Euler estimate of the time derivative,

$$\frac{\partial u}{\partial t} \approx \frac{u^n - u^{n-1}}{\Delta t}$$

the fully discrete variational formulation is

$$\begin{aligned}
0 &\approx \int_{\Omega_e} v \left(\frac{u^n - u^{n-1}}{\Delta t} \right) d\Omega_e + \int_{\Omega_e} \mathbf{D}^n \nabla u^n \cdot \nabla v d\Omega_e \\
&+ \int_{\Omega_e} v \vec{V} \cdot \nabla u^n d\Omega_e - \int_{\Omega_e} v f^n d\Omega_e \\
&- \oint_{\partial\Omega_e} v \mathbf{D} \nabla u^n \cdot \vec{\nu} d\gamma
\end{aligned}$$

where the superscript n represents the n^{th} timestep. Then making the substitutions

$$\begin{aligned}
u^n &= \sum_{j=1}^r u_j(t_n) \psi_j(\vec{x}) & \vec{x} &= (x, y) \\
v &= \psi_i(\vec{x})
\end{aligned}$$

and letting

$$\mathbf{D}^n = \begin{bmatrix} \mathbf{D}_{11}^n & \mathbf{D}_{12}^n \\ \mathbf{D}_{21}^n & \mathbf{D}_{22}^n \end{bmatrix}$$

The system of equations for the e^{th} element becomes

$$\begin{aligned}
& \sum_{j=1}^r \frac{u_j(t_n)}{\Delta t} \int_{\Omega_e} \psi_i \psi_j d\Omega_e + \sum_{j=1}^r \mathbf{D}_{11}^n u_j(t_n) \int_{\Omega_e} \frac{\partial \psi_j}{\partial x} \frac{\partial \psi_i}{\partial x} d\Omega_e \\
& + \sum_{j=1}^r \mathbf{D}_{12}^n u_j(t_n) \int_{\Omega_e} \frac{\partial \psi_j}{\partial y} \frac{\partial \psi_i}{\partial x} d\Omega_e \\
& + \sum_{j=1}^r \mathbf{D}_{21}^n u_j(t_n) \int_{\Omega_e} \frac{\partial \psi_j}{\partial x} \frac{\partial \psi_i}{\partial y} d\Omega_e \\
& + \sum_{j=1}^r \mathbf{D}_{22}^n u_j(t_n) \int_{\Omega_e} \frac{\partial \psi_j}{\partial y} \frac{\partial \psi_i}{\partial y} d\Omega_e \\
& + \sum_{j=1}^r V_x u_j(t_n) \int_{\Omega_e} \frac{\partial \psi_j}{\partial x} \psi_i d\Omega_e \\
& + \sum_{j=1}^r V_y u_j(t_n) \int_{\Omega_e} \frac{\partial \psi_j}{\partial y} \psi_i d\Omega_e \\
& - \oint_{\partial\Omega_e} v \mathbf{D}^n \nabla u^n \cdot \vec{\nu} d\gamma \\
& \approx \sum_{j=1}^r \frac{u_j(t_{n-1})}{\Delta t} \int_{\Omega_e} \psi_i \psi_j d\Omega_e \\
& + \int_{\Omega_e} \psi_i f^n d\Omega_e \quad i = 1, 2, \dots, r
\end{aligned}$$

3.2.9 Imposition Of The Boundary Condition

In our problem, no flow boundary conditions are assumed to exist along the sides of the tank and zero concentrations are assumed to exist at the ends of the tank. In addition, a pulsed-input is allowed to take place at one of the injection ports in the tank for a specified period of time. The following example illustrates how the system of equations is modified to handle constant concentrations at grid points. Suppose that the concentration is to remain constant at a grid point on the left hand boundary, *i.e.*,

$$u_1 = \alpha$$

Suppose that the system to be solved is

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ & & \vdots & \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_n \end{bmatrix}$$

Then by letting $A_{11} = 1$, $A_{1i} = 0$, $i = 2, \dots, n$ and $F_1 = \alpha$, u_1 is forced to be equal to α and the following system emerges

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ A_{21} & A_{22} & \cdots & A_{2n} \\ & & \vdots & \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} \alpha \\ F_2 \\ \vdots \\ F_n \end{bmatrix}$$

But, since $u_1 = \alpha$, the known terms can be moved to the right hand side to yield

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & A_{22} & \cdots & A_{2n} \\ & & \vdots & \\ 0 & A_{n2} & \cdots & A_{nn} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} \alpha \\ F_2 - A_{21}\alpha \\ \vdots \\ F_n - A_{n1}\alpha \end{bmatrix}$$

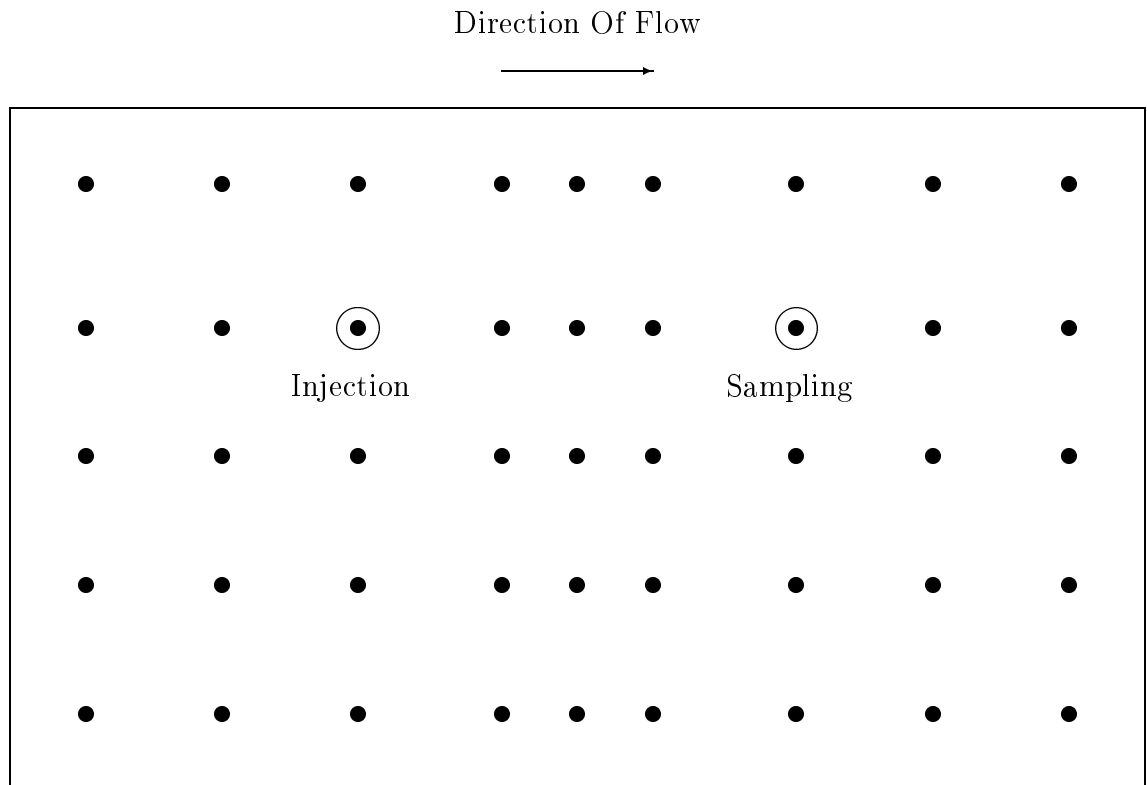


Figure 12 - Boulder's Experimental Tank

3.2.10 The Boulder Experiments

Next, consider the test environment used in the Boulder experiments. It is similar to a confined aquifer into which a tracer (pollutant) is injected. It is assumed that the tracer immediately mixes locally with the water in the aquifer to form one phase. The Boulder experiment's horizontal tank has 45 port locations where tracer injections can be made or samples can be taken. Constant head conditions are assumed at the ends of the tank and no flow conditions are assumed on the sides of the tank. Figure 12 illustrates the layout of the tank. The tank can be packed with various types of sand. In the homogeneous experiments, the tank is packed with a single type of sand. In the heterogeneous experiments, the tank is packed in a block arrangement with each block having the dimensions of 12.2 cm \times 12.2 cm, and there are 200 blocks in a 20 \times 10 rectangular array. Five types of sand are used with the following hydraulic conductivities:

Sand #	Hydraulic- Conductivity (m/day)
1	6.05
2	20.74
3	125.28
4	371.52
5	1036.5

Figure 13 shows conceptually the locations of the five different types of sands by relative level of hydraulic conductivity. The actual hydraulic conductivity levels were not used to produce this plot because if the actual levels were used, the resulting plot would not capture all of the features clearly because of the wide spread between the lowest and highest values. Figure 14 is a contour plot of Figure 13, and probably shows in better detail the arrangement of the 5 sand types. In this figure, the white areas represent the sands with the highest hydraulic conductivity and the darker shades of gray representing progressively lower hydraulic conductivities with the black areas the lowest hydraulic conductivity. With this arrangement of the sand blocks there is a significant amount of heterogeneity in the tank.

Using the finite element model explained in Chapter 3, a simulation of the Boulder horizontal tank is conducted that is essentially described by Figure 15. This methodology is a Monte Carlo method. The first thing that is done is

Figure 13 - Hydraulic Conductivity Levels

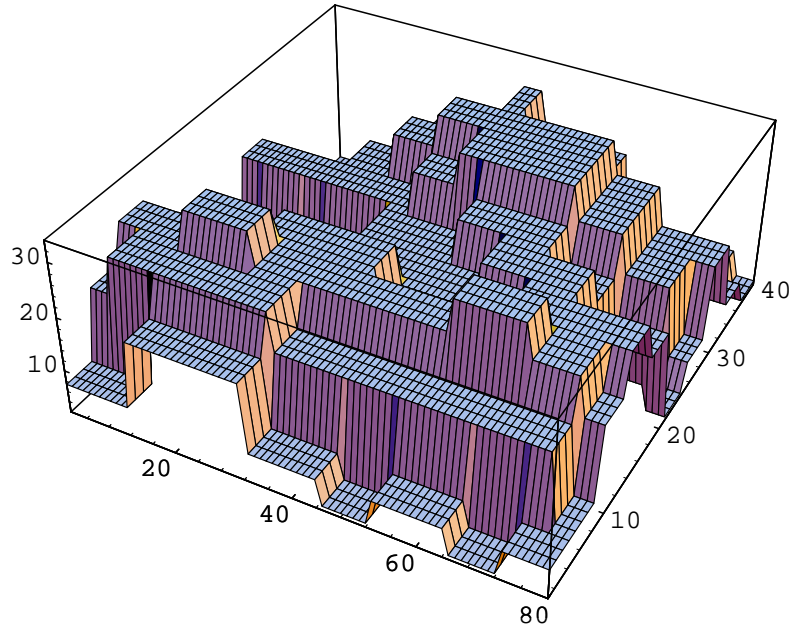
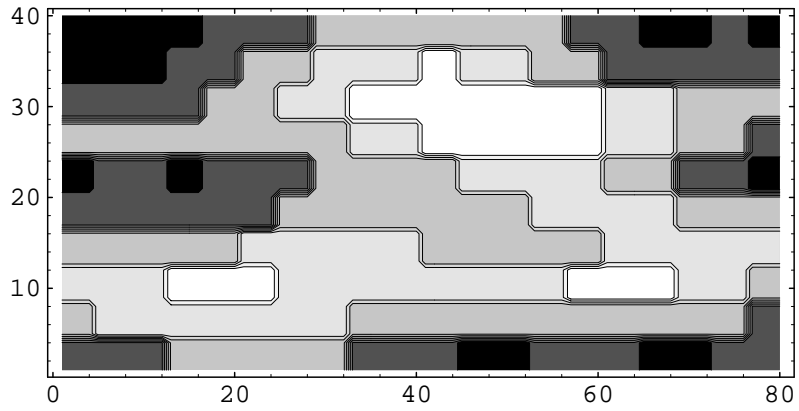


Figure 14 - Hydraulic Conductivity Contours



for each of the 5 sands a sample hydraulic conductivity distribution is calculated according to the method explained in Section 2.5.1. The samples generated are then distributed over the domain of the tank by sand type. Following this, the flow equation is solved for the steady state piezometric heads according to the finite element scheme described in Section 3.2.5.

Flow/Concentration Solution Method

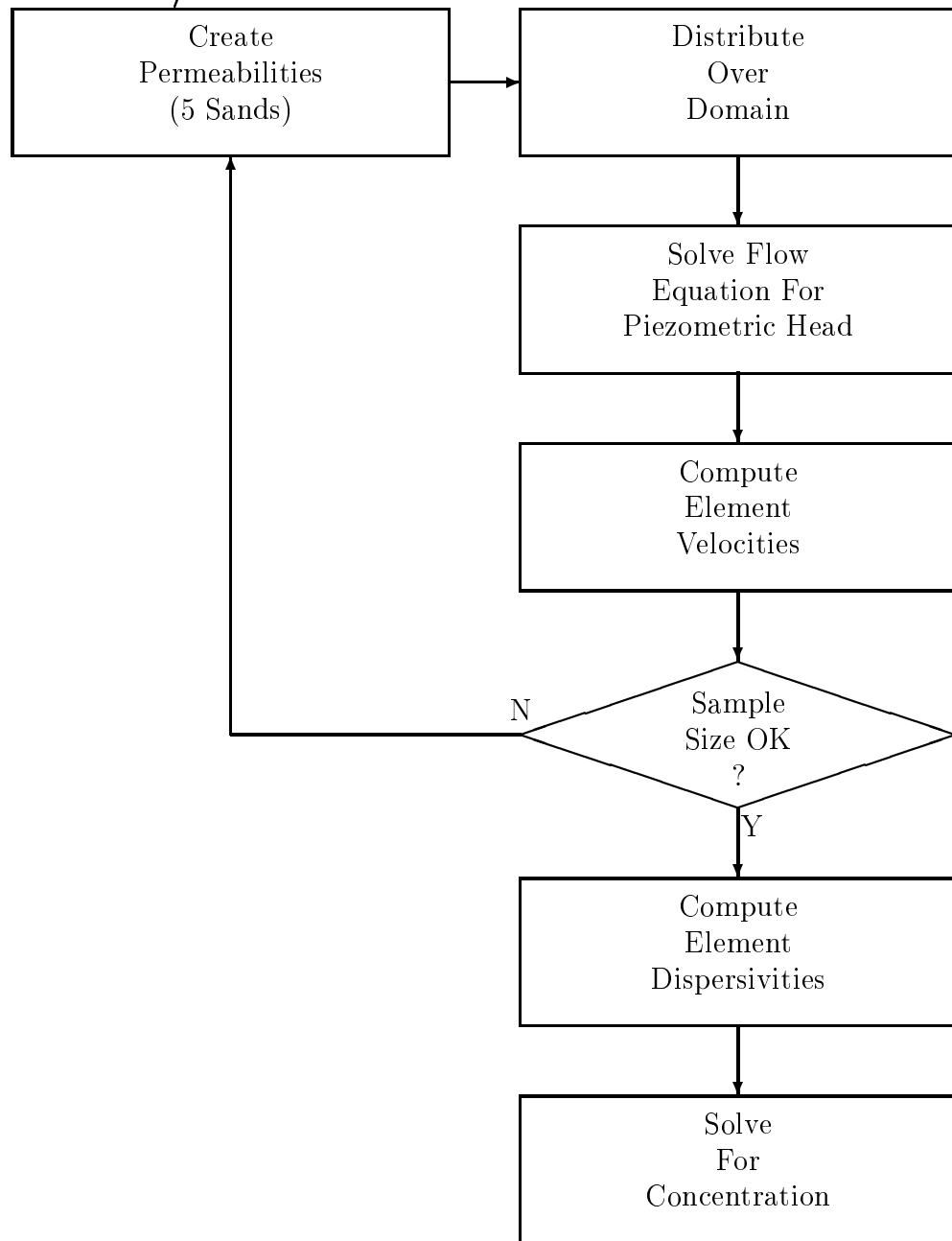


Figure 15 - Solution Methodology

Figure 16 shows the computed pressure contours for one of these runs, which agrees with the pressure contours as measured in one of the tank experiments. The contours show very similar patterns. Figure 17 overlays the computed velocity field, Section 3.2.7, on the computed pressure contours showing that the velocity vectors are orthogonal to the pressure contours. Flow lines from the tank experiments demonstrate velocity patterns similar to the computed velocities shown in Figure 17.

Following the calculation of the velocities, the program cycles back to generate another sample distribution of hydraulic conductivities. This process is continued until a predetermined number of sample runs have been made. The program then computes the dispersion coefficients for the numerical grid blocks according to the methodology specified in Section 2.4. These dispersion coefficients are used along with the average velocities to compute tracer plumes created by injecting a tracer into the tank at a selected point. The finite element scheme for this is outlined in Section 3.2.8. Figure 18 shows the development of a computed plume. The *dog-leg* appearance is to be expected as can be inferred from looking at the flow lines of the velocity field.

3.3 Summary

Chapter 3 begins by defining the two basic forms of the flow/transport problem. In the general case, either the density or the viscosity of the single phase created by two mixing fluids is dependent on the concentration and/or the pressure. In this case, the flow/transport problem takes the form of a coupled system of PDE's that must be solved simultaneously. The pressure/velocity pair can be solved for together using mixed finite elements, followed by a solution of the transport equation by some method. The code that is currently available to solve the coupled system in this fashion is the SEGMIX code. It uses the mixed method to solve for the pressure/velocity pair and the modified method of characteristics (MMOC) to solve the transport equation. However, in order to use this code to study the horizontal tank experiments, the treatment of the boundary conditions would have to be modified.

In the tracer case, the problem is less complicated, and the uncoupled system can be solved using the standard Galerkin finite elements. Section 3.2 contains descriptions of the implementation of the various components of the finite element method. Section 3.2.2 to 3.2.4 explain the numerical grid and the numerical integration scheme used to calculate the integrals arising

Figure 16 - Computed Pressure Contours

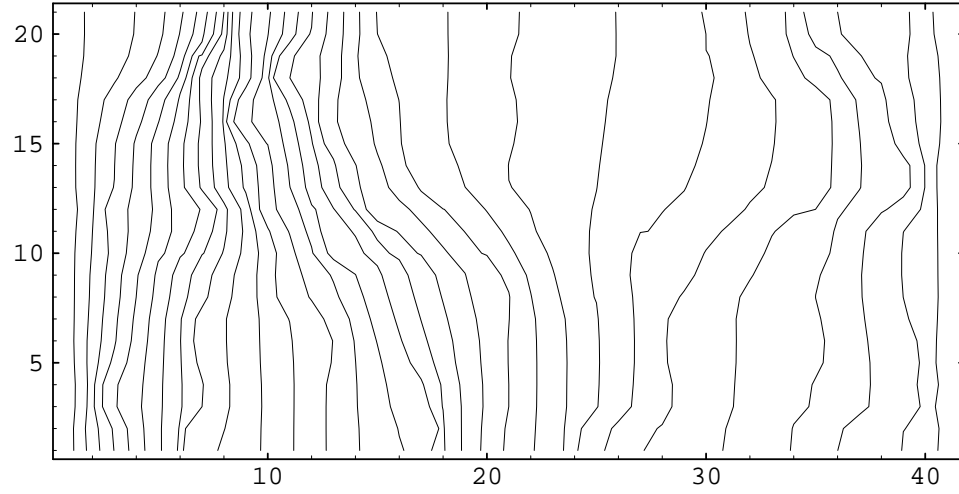


Figure 17 - Computed Velocity Field On Pressure Contours

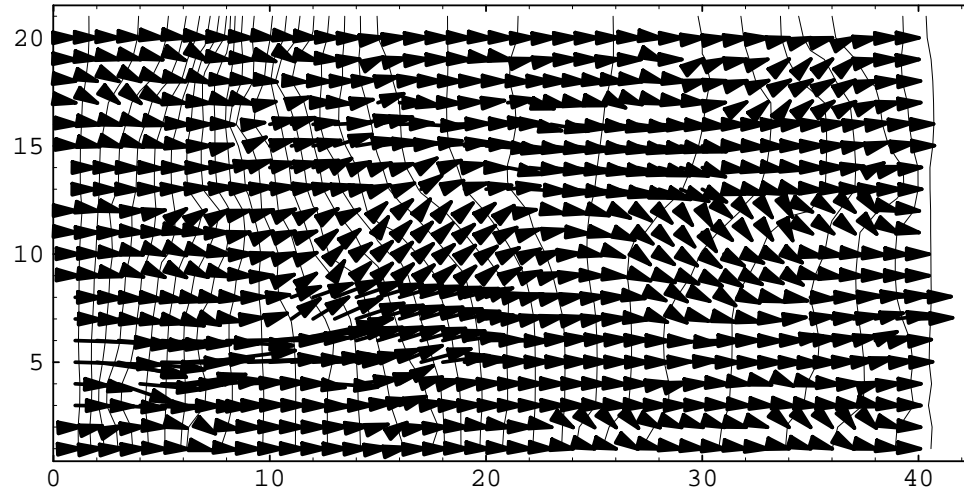
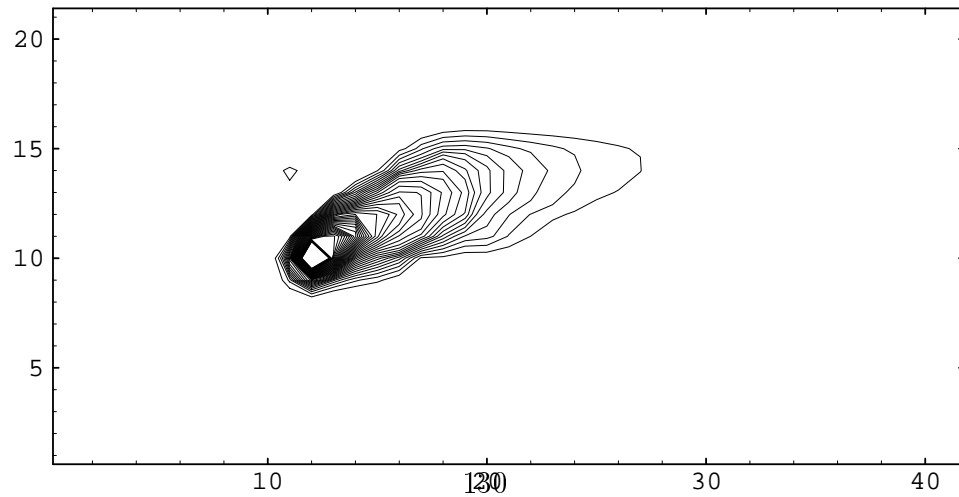


Figure 18 - Computed Tracer Plume



from the variational formulation. Section 3.2.5 contains a derivation of the local system of equations for the flow equation, and Section 3.2.6 describes the assembly of the local systems into a global system. The derivation of the velocity field from the computed piezometric heads is contained in Section 3.2.7. Section 3.2.8 shows the derivation of the system of equations that follow from the variational formulation of the *transport* equation. Since there are constant boundary conditions at the ends of the tank, and a pulsed-input is allowed to take place at an injection point, it is necessary to allow constant concentration conditions to exist at some grid points. The modification of the global system of equations to allow certain grid points to maintain a constant level of concentration is explained in Section 3.2.9. Because of the pulsed nature of the injection process, additional discretization is required around the injection point in order to obtain acceptable Peclet numbers.

Section 3.2.10 describes in more detail the horizontal test tank used in the Boulder experiments. Two types of experiments are conducted in the tank. The *homogeneous* experiments are those in which the tank is packed with a single type of sand as rated by its hydraulic conductivity. In the *heterogeneous* experiments, the tank is packed in a block arrangement with 5 different types of sand. The hydraulic conductivities of the sands range from 6.05 m/day for Sand #1 to 1036.5 m/day for Sand #5. With this wide span of hydraulic conductivities, a significant amount of heterogeneity is represented in the tank. The block arrangement of the sands in the tank is represented graphically in Figures 13 and 14 in Section 3.2.10. Figure 15 provides a flowchart of the basic program components used and how they interact. Comparisons of computer simulation results shown in Figures 16 and 17 to actual tank measurements show very good agreement. Figure 18 illustrates a computed tracer plume.

4. Moment Equations

Permeability, density and viscosity are related through hydraulic conductivity. So, both soil properties and fluid properties are represented in hydraulic conductivity. Furthermore, hydraulic conductivity determines the velocity field of the water in the aquifer, and if a solute is introduced into the aquifer, the path a solute particle takes through the aquifer is determined by two components. First, the path has a component that is due to molecular diffusion and, secondly, a component that is due to the mechanical mixing that results from the convective transport. This means that the developing plume is dispersing about a path that is changing due to the influence of the convective transport determined by the large-scale heterogeneities of the aquifer's domain. In Chapter 2, one method of determining the dispersion used in the transport equation was discussed, and in Chapter 3 that method was implemented to estimate the expected value of the concentration of a tracer injected into the tank. The entire approach was based on stochastic descriptions of the hydraulic conductivities of the 5 different types of sands with which the tank is packed. As explained in Chapter 3, the domain is subdivided into rectangular grid blocks, each of which is assigned its own average velocity and dispersion coefficient computed according to the methodology explained in Chapter 2. This approach provides the domain of the aquifer with a velocity field that mimicks large-scale changes in the permeabilities of the different sections.

However, only the expected value of the concentration is predicted. What is needed in addition to this is a system that will provide information on higher order moments. In general, the more moments that are known, the better the probability density can be described. It would be desirable to at least know something about the second moments.

The purpose of this section is to analyze three methods of providing information on higher moments. In the first method, the second moments are derived from the transport equation by a method of distributed parameters. The second and third methods both involve the theory of stochastic differential equations. For the second method, an approach using the Itô calculus, specifically Itô's lemma, is used. Finally, the third method seeks to find the solution in terms of the evolution operator. In certain circumstances, this method is a semigroup method.

4.1 Moments Derived From Distributed Parameters

The first method is based on the work of Graham And McLaughlin[48] who derive *unconditional* ensemble moments, ie, moments that do not depend on concentration observations. Starting from the transport equation,

$$\frac{\partial c}{\partial t} + \nabla \cdot (c\vec{V}) - \nabla \cdot [\mathbf{D}\nabla c] = 0 \quad (4.1)$$

and letting

$$c = \mathbf{E}[c] + c' \quad \vec{V} = \mathbf{E}[\vec{V}] + \vec{V}' \quad \mathbf{E}[c'] = 0 \quad \mathbf{E}[\vec{V}'] = 0 \quad (4.2)$$

we get by substituting Equation[4.2] into Equation[4.1]

$$\begin{aligned} \frac{\partial \mathbf{E}[c]}{\partial t} + \frac{\partial c'}{\partial t} + \nabla \cdot [\mathbf{E}[c]\mathbf{E}[\vec{V}] + \mathbf{E}[c]\vec{V}' + \mathbf{E}[\vec{V}]c' + c'\vec{V}'] \\ - \nabla \cdot [\mathbf{D}\nabla \mathbf{E}[c] + \mathbf{D}\nabla c'] = 0 \end{aligned} \quad (4.3)$$

and by taking expectations and using $\mathbf{E}[c'] = \mathbf{E}[\vec{V}'] = 0$, we get the *mean concentration equation*

$$\frac{\partial \mathbf{E}[c]}{\partial t} + \nabla \cdot (\mathbf{E}[c]\mathbf{E}[\vec{V}]) - \nabla \cdot (\mathbf{D}\nabla \mathbf{E}[c]) + \nabla \cdot \mathbf{E}[c'\vec{V}'] = 0 \quad (4.4)$$

Subtracting Equation[4.4] from Equation[4.3] we get an equation that involves the perturbations of the concentration and the velocity field

$$\begin{aligned} \frac{\partial c'}{\partial t} + \nabla \cdot \mathbf{E}[c]\vec{V}' + \nabla \cdot \mathbf{E}[\vec{V}]c' + \nabla \cdot (c'\vec{V}') \\ - \nabla \cdot \mathbf{D}\nabla c' - \nabla \cdot \mathbf{E}[c'\vec{V}'] = 0 \end{aligned} \quad (4.5)$$

Multiply Equation[4.5] by the perturbed velocity vector at a point \vec{x}' different from \vec{x} . Since the velocity perturbation depends only on the spatial variable and \vec{x}' is different from \vec{x} and the derivatives are taken at \vec{x} , it follows that

$$\begin{aligned} \frac{\partial (c'\vec{V}'(\vec{x}'))}{\partial t} + (\nabla_{\vec{x}} \cdot \mathbf{E}[c]\vec{V}'(\vec{x})) \vec{V}'(\vec{x}') + \nabla_{\vec{x}} \cdot \mathbf{E}[\vec{V}]c'\vec{V}'(\vec{x}') \\ + \nabla_{\vec{x}} \cdot (c'\vec{V}'(\vec{x})) \vec{V}'(\vec{x}') - \nabla_{\vec{x}} \cdot \mathbf{D}\nabla_{\vec{x}} c'\vec{V}'(\vec{x}') \\ - \nabla_{\vec{x}} \cdot \mathbf{E}[c'\vec{V}'(\vec{x})]\vec{V}'(\vec{x}') = 0 \end{aligned}$$

The i^{th} component of this equation is written

$$\begin{aligned}
\frac{\partial(c'V_i'(\vec{x}'))}{\partial t} &+ \left(\frac{\partial}{\partial x_j}(\mathbf{E}[c]V_j'(\vec{x})) \right) V_i'(\vec{x}') \\
&+ \frac{\partial}{\partial x_j}(\mathbf{E}[V_j]c') V_i'(\vec{x}') + \frac{\partial}{\partial x_j}(c'V_j'(\vec{x})) V_i'(\vec{x}') \\
&- \frac{\partial}{\partial x_j} \left(D_{jk} \frac{\partial}{\partial x_k}(c') \right) V_i'(\vec{x}') - \frac{\partial}{\partial x_j}(\mathbf{E}[c'V_j'(\vec{x})]) V_i'(\vec{x}') = 0
\end{aligned}$$

Since the components V_i' are evaluated at \vec{x}' , they are constants with respect to the $\frac{\partial}{\partial x_j}$ operator, and so we can write

$$\begin{aligned}
\frac{\partial(c'V_i'(\vec{x}'))}{\partial t} &+ \frac{\partial}{\partial x_j}(\mathbf{E}[c]V_j'(\vec{x})V_i'(\vec{x}')) \\
&+ \frac{\partial}{\partial x_j}(\mathbf{E}[V_j]c'V_i'(\vec{x}')) + \frac{\partial}{\partial x_j}(c'V_j'(\vec{x})V_i'(\vec{x}')) \\
&- \frac{\partial}{\partial x_j} \left(D_{jk} \frac{\partial}{\partial x_k}(c'V_i'(\vec{x}')) \right) - \frac{\partial}{\partial x_j}(\mathbf{E}[c'V_j'(\vec{x})]V_i'(\vec{x}')) = 0
\end{aligned} \tag{4.6}$$

Define

$$\begin{aligned}
C_{V_i c}(\vec{x}', \vec{x}, t) &= \mathbf{E}[V_i'(\vec{x}')c(\vec{x}, t)] \\
C_{V_i V_j}(\vec{x}', \vec{x}) &= \mathbf{E}[V_i'(\vec{x}')V_j'(\vec{x})] \\
C_{cV_i V_j}(\vec{x}', \vec{x}, t) &= \mathbf{E}[c'(\vec{x}, t)V_i'(\vec{x}')V_j'(\vec{x})] \\
C_{cc}(\vec{x}', \vec{x}, t) &= \mathbf{E}[c'(\vec{x}', t)c'(\vec{x}, t)]
\end{aligned}$$

Taking expectations of Equation[4.6] yields the *velocity-concentration covariance* equation.

$$\begin{aligned}
\frac{\partial}{\partial t}C_{cV_i}(\vec{x}', \vec{x}, t) &+ \frac{\partial}{\partial x_j}\mathbf{E}[c]C_{V_i V_j}(\vec{x}', \vec{x}) \\
&+ \frac{\partial}{\partial x_j}\mathbf{E}[V_j]C_{cV_i}(\vec{x}', \vec{x}, t) \\
&+ \frac{\partial}{\partial x_j}\mathbf{E}[c'(\vec{x}, t)V_j'(\vec{x})V_i'(\vec{x}')] \\
&- \frac{\partial}{\partial x_j} \left(D_{jk} \frac{\partial}{\partial x_k}C_{cV_i}(\vec{x}', \vec{x}, t) \right) = 0
\end{aligned} \tag{4.7}$$

Note: The last term in Equation[4.6] vanishes because $\mathbf{E}[V'_i(\vec{x}')] = 0$.

Taking Equation[4.5] and multiplying by $c'(\vec{x}', t)$ where $\vec{x}' \neq \vec{x}$ and using the assumption that $\mathbf{E}[c'(\vec{x}', t)] = 0$ and the covariance definitions we get

$$\begin{aligned} \mathbf{E} \left[\frac{\partial}{\partial t} (c'(\vec{x}, t)) c'(\vec{x}', t) \right] &+ \frac{\partial}{\partial x_j} \mathbf{E}[c(\vec{x}, t)] C_{V_j c}(\vec{x}', \vec{x}, t) \\ &+ \frac{\partial}{\partial x_j} \mathbf{E}[V_j(\vec{x})] C_{cc}(\vec{x}', \vec{x}, t) \\ &+ \frac{\partial}{\partial x_j} \mathbf{E}[c'(\vec{x}, t) c'(\vec{x}', t) V'_j(\vec{x})] \\ &- \frac{\partial}{\partial x_j} \left[D_{jk} \frac{\partial}{\partial x_k} C_{cc}(\vec{x}', \vec{x}, t) \right] = 0 \end{aligned} \quad (4.8)$$

Interchanging the roles of \vec{x}' and \vec{x} we get

$$\begin{aligned} \mathbf{E} \left[\frac{\partial}{\partial t} (c'(\vec{x}', t)) c'(\vec{x}, t) \right] &+ \frac{\partial}{\partial x'_j} \mathbf{E}[c(\vec{x}', t)] C_{V_j c}(\vec{x}, \vec{x}', t) \\ &+ \frac{\partial}{\partial x'_j} \mathbf{E}[V_j(\vec{x}')] C_{cc}(\vec{x}, \vec{x}', t) \\ &+ \frac{\partial}{\partial x'_j} \mathbf{E}[c'(\vec{x}', t) c'(\vec{x}, t) V'_j(\vec{x}')] \\ &- \frac{\partial}{\partial x'_j} \left[D_{jk} \frac{\partial}{\partial x'_k} C_{cc}(\vec{x}, \vec{x}', t) \right] = 0 \end{aligned} \quad (4.9)$$

Adding Equations [4.8] and [4.9] and using the product rule of differentiation

$$\begin{aligned} \frac{\partial}{\partial t} C_{cc}(\vec{x}', \vec{x}, t) &= \frac{\partial}{\partial t} \mathbf{E}[c'(\vec{x}', t) c'(\vec{x}, t)] \\ &= \mathbf{E} \left[\frac{\partial}{\partial t} (c'(\vec{x}', t)) c'(\vec{x}, t) \right] + \mathbf{E} \left[\frac{\partial}{\partial t} (c'(\vec{x}, t)) c'(\vec{x}', t) \right] \end{aligned}$$

the following equation for the *concentration covariance* is obtained

$$\frac{\partial}{\partial t} C_{cc}(\vec{x}', \vec{x}, t) + \frac{\partial}{\partial x_j} \mathbf{E}[V_j(\vec{x})] C_{cc}(\vec{x}', \vec{x}, t)$$

$$\begin{aligned}
& + \frac{\partial}{\partial x'_j} \mathbf{E}[V_j(\vec{x}')] C_{cc}(\vec{x}, \vec{x}', t) - \frac{\partial}{\partial x_j} \left[D_{jk} \frac{\partial}{\partial x_k} C_{cc}(\vec{x}, \vec{x}, t) \right] \\
& - \frac{\partial}{\partial x'_j} \left[D_{jk} \frac{\partial}{\partial x'_k} C_{cc}(\vec{x}, \vec{x}', t) \right] + \frac{\partial}{\partial x_j} \mathbf{E}[c(\vec{x}, t)] C_{V_j c}(\vec{x}', \vec{x}, t) \\
& + \frac{\partial}{\partial x'_j} \mathbf{E}[c(\vec{x}', t)] C_{V_j c}(\vec{x}, \vec{x}', t) + \frac{\partial}{\partial x_j} \mathbf{E}[c'(\vec{x}, t) c'(\vec{x}', t) V'_j(\vec{x})] \\
& + \frac{\partial}{\partial x'_j} \mathbf{E}[c'(\vec{x}', t) c'(\vec{x}, t) V'_j(\vec{x}')] = 0 \tag{4.10}
\end{aligned}$$

The *mean concentration equation* can be written using the covariance notation as

$$\begin{aligned}
\frac{\partial \mathbf{E}[c]}{\partial t} + \sum_{i=1}^n \frac{\partial}{\partial x_i} (\mathbf{E}[c] \mathbf{E}[V_i]) & - \sum_{i=1}^n \frac{\partial}{\partial x_i} \left[\sum_{j=1}^n D_{ij} \frac{\partial}{\partial x_j} \mathbf{E}[c] \right] \\
& + \sum_{i=1}^n \frac{\partial}{\partial x_i} C_{cV_i} = 0
\end{aligned}$$

This equation has the form of a transport equation with a forcing term that involves the concentration-velocity covariance.

The *velocity-concentration equation*[4.7] has the form of a transport equation that involves a forcing term that consists of one term that contains the product of the mean concentration and the velocity covariances and one term that involves the expected value of the product of the perturbation of the concentration and velocities. The mean concentration equation and the velocity-concentration covariance are coupled through the $\mathbf{E}[c]$ variable and the C_{cV_i} variable.

The *concentration-covariance equation*[4.10] also has the form of a transport equation with a forcing term consisting of the last four terms in Equation[4.10]. The coupling to the other two equations is through the terms $\mathbf{E}[c]$ and C_{cV_i} .

In order to solve this system, the mean velocities and the velocity covariances are required as inputs. The terms, then, that have to be dealt with to form a closed system are

$$\begin{aligned}
& \frac{\partial}{\partial x_j} \mathbf{E}[c'(\vec{x}, t) V_j'(\vec{x}) V_i'(\vec{x}')] \\
& \frac{\partial}{\partial x_j} \mathbf{E}[c'(\vec{x}, t) c'(\vec{x}', t) V_j'(\vec{x})] \\
& \frac{\partial}{\partial x_j} \mathbf{E}[c'(\vec{x}', t) c'(\vec{x}, t) V_j'(\vec{x}')]
\end{aligned} \tag{4.11}$$

These terms are considered to be small, and therefore neglected. To say that these terms can be considered to be small and therefore can be neglected does not seem convincing. By saying that the perturbations are small would certainly imply that the expectations of the products of the perturbations are small, but these terms involve the spatial derivatives of the perturbations and there is no reason to believe that they are small.

These terms can be eliminated if the assumption is made that they come from a multivariate Gaussian distribution. The multivariate Gaussian probability density function for n dependent random variables is given by

$$f(z_1, z_2, \dots, z_n) = \frac{|\mathbf{V}|^{-\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \exp \left\{ -\frac{1}{2} (\vec{z} - \vec{\mu})^\dagger (\mathbf{V}^{-1}) (\vec{z} - \vec{\mu}) \right\}$$

where

$$\vec{z} = (z_1, z_2, \dots, z_n)^\dagger$$

$$\vec{\mu} = (\mathbf{E}[z_1], \mathbf{E}[z_2], \dots, \mathbf{E}[z_n])^\dagger$$

$$\mathbf{V} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ & & \vdots & \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{pmatrix}$$

So that \mathbf{V} is the variance-covariance matrix.

The multivariate characteristic function is given by, Springer[93], page 75,

$$\phi(\zeta_1, \zeta_2, \dots, \zeta_n) = \exp \left(-\frac{1}{2} \vec{\zeta}^\dagger \mathbf{V} \vec{\zeta} \right) \exp \left(i \vec{\zeta}^\dagger \vec{\mu} \right)$$

Applying this theory to our problem, consider the trivariate case where

$$\vec{z} = (c'(\vec{x}, t), c'(x^T, t), V_j'(\vec{x}))^\dagger$$

Since the variations are assumed to have zero means, it follows that

$$\vec{\mu} = 0$$

And, the trivariate characteristic function has the form

$$\phi(\zeta_1, \zeta_2, \zeta_3) = \exp\left(-\frac{1}{2}\vec{\zeta}^\dagger \mathbf{V} \vec{\zeta}\right)$$

The expression $\vec{\zeta}^\dagger \mathbf{V} \vec{\zeta}$ is a quadratic form, and when expanded is equal to

$$\vec{\zeta}^\dagger \mathbf{V} \vec{\zeta} = (\zeta_1, \zeta_2, \zeta_3) \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix} \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix} = \sum_{i=1}^3 \sum_{j=1}^3 \sigma_{ij} \zeta_i \zeta_j$$

So, the trivariate characteristic function is given by

$$\phi(\zeta_1, \zeta_2, \zeta_3) = \exp\left(-\frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 \sigma_{ij} \zeta_i \zeta_j\right)$$

The reason for introducing the multivariate characteristic function is that moments can be generated from it by taking derivatives. In particular,

$$\mathbf{E}[z_1 z_2 z_3] = \left[\frac{1}{i^3} \frac{\partial^3 \phi(\zeta_1, \zeta_2, \zeta_3)}{\partial \zeta_1 \partial \zeta_2 \partial \zeta_3} \right]_{\zeta_1 = \zeta_2 = \zeta_3 = 0}$$

On taking the partial derivatives and using the condition that $\zeta_1 = \zeta_2 = \zeta_3 = 0$, it follows that

$$\mathbf{E}[z_1 z_2 z_3] = 0$$

So, with the assumption of a joint Gaussian distribution the terms in Equation[4.11] can be removed.

4.2 An Itô Calculus Approach

4.2.1 System Definition

A solution is sought for the following system:

$$\begin{aligned} \frac{\partial u}{\partial t} + A(x, t, \omega)u &= g(x, t, \omega) & (x, t, \omega) \in G \times [0, t] \times \Omega \\ Q(x, t, \omega) &= J(\omega) & (x, t, \omega) \in \partial G \times [0, T] \times \Omega \\ u(x, 0, \omega) &= u_0(x, \omega) & (x, \omega) \in G \times \Omega \end{aligned} \quad (4.12)$$

where $g \in L_2(\Omega, \mathcal{B}, P)$ the space of second order random functions.

$G \subset \mathfrak{R}^n$ is an open domain with a Lipschitz continuous boundary, ∂G , and $t \in (0, \infty)$.

The operator A is defined as

$$Au = \sum_{|k|, |l| \leq m} (-1)^{|k|} D^k (p_{kl}(x, t, \omega) D^l u)$$

The operator D represents weak differentiation and the solution $u \in L^2(0, T; V)$, where

$$L^2(0, T; V) = \left\{ f : [0, T] \rightarrow V : \int_0^t \|f\|_V^2 dt < \infty \right\}$$

The Hilbert space V represents an m^{th} order Sobolev space of $L^2(\Omega)$ -valued random functions on the set G . The space V will be more completely specified in the sub-section entitled **Existence Theory**.

4.2.2 Types Of Problems

The following is a list of the different types of problems that can potentially be handled using the stochastic evolution equation formulation:

- The random initial value problem; u_0 is random

- The random boundary value problem; J is random
- The random forcing problem; g is random
- The random operator problem; A or Q is random
- The random geometry problem; G is random
- Combinations of the above

In this report, the groundwater flow problem will be treated as a random forcing term problem, ie, g is allowed to be random. And, the groundwater transport problem will be treated as a random operator problem, ie, the operator A will be allowed to have a random component. The interest here is in techniques for solving the stochastic evolution equations and in determining their first and second moment equations.

First, the problem of existence of solutions has to be addressed. It is necessary to be able to state conditions under which solutions will exist, and be able to specify the spaces that will contain the solutions.

4.2.3 Existence Theory

The existence theory in this section is compiled from Becus[13], Sawaragi, Soeda, Omatu[85], Serrano, Unny, Lennox[90] and Oden and Reddy[72].

Let (Ω, \mathcal{B}, P) be a complete probability space and define

$$L_2(\Omega) = L_2(\Omega, \mathcal{B}, P)$$

to be the space of second order random functions on Ω . A probability space is complete if the measure P is complete, *i.e.*, if any subset of a set, $B \in \mathcal{B}$, with $P(B) = 0$, also belongs to \mathcal{B} . The space $L_2(\Omega)$ is a Hilbert space with inner product

$$(f, g)_\Omega = \int_\Omega fg dP = \mathbf{E}[fg]$$

Next the following set M is defined

$$M = \{f : G \rightarrow L_2(\Omega)\}$$

to be the set of second order random functions on $G \subset \mathfrak{R}^n$.

Using the set M , the following spaces are defined:

$$H = L_2[G; L_2(\Omega)] = \{f \in M : \|f\|_\Omega \in L^2(G)\}$$

where $L^2(G)$ are the square-integrable functions on G .

H is a Hilbert space with inner product

$$\begin{aligned} (f, g)_H &= \int_G (f, g)_\Omega dG = \int_G E[fg] dG \\ &= \int_G \int_\Omega fgdPdG \end{aligned}$$

And, for $m \geq 0$,

$$H^m = H^m(G; L_2(\Omega)) = \{f \in M : D^\alpha f \in H, |\alpha| \leq m\}$$

H^m is a Hilbert space with inner product

$$(f, g)_{H^m} = \sum_{|\alpha| \leq m} (D^\alpha f, D^\alpha g)_H$$

Hence, H^m is the m^{th} order Sobolev space of $L_2(\Omega)$ -valued functions on G . Let V be a real separable Hilbert space such that

$$\bar{V} = H$$

and the injection

$$i : V \rightarrow H$$

is continuous. It then follows that the following diagram can be established

$$\begin{array}{ccc} V & \xrightarrow{i} & H \\ Z_V \downarrow & & \downarrow Z_H \\ V' & \xleftarrow{i'} & H' \end{array}$$

where the mappings Z_V and Z_H are the Riesz maps between the Hilbert spaces V and its dual V' and between H and its dual H' , respectively.

And, by identifying H with its dual, H' , it can be shown that

$$V \subset H = H' \subset V'$$

and that H' is densely embedded in V' . Using the Hahn-Banach theorem, the duality pairing on $V' \times V$ can be identified with the unique extension of the duality pairing on $H' \times H$, $\langle q, u \rangle_H$. And, by the Riesz Representation theorem, $\forall q \in H', \exists v_q \in H$ such that

$$\langle q, u \rangle_H = (v_q, u)_H \quad \forall u \in H$$

where $(\cdot, \cdot)_H$ is the inner product on H . So, the duality pairing on $V' \times V$ can be identified with the unique extension of the inner product on H .

Given this, the norm on V' can be represented as

$$\begin{aligned} \|\phi\|_{V'} &= \sup_{\substack{u \in V \\ u \neq 0}} \frac{|\langle \phi, u \rangle_{V'}|}{\|u\|_V} \\ &= \sup_{\substack{u \in V \\ u \neq 0}} \frac{|(v_\phi, u)_H|}{\|u\|_V} \end{aligned}$$

For $0 < T < \infty$, define

$$L^2(0, T; V) = \left\{ f : [0, T] \rightarrow V : \int_0^T \|f\|_V^2 dt < \infty \right\}$$

And, if $f \in L^2(0, T; V)$, then $D_t f$ is the derivative of f in the sense of V -valued distributions, ie,

$$D_t f \in V'$$

Define

$$W(0, T) = \left\{ f \in L^2(0, T; V) : D_t f \in L^2(0, T; V') \right\}$$

$W(0, T)$ is a Hilbert space with norm

$$\|f\|_W^2 = \int_0^T (\|f\|_V^2 + \|D_t f\|_{V'}^2) dt$$

Becus[13] recasts the stochastic evolution equation in its variational form, and letting

$$(A(x, t, \omega)u, v)_H = a(u, v)$$

satisfy

$$|a(u, v)| \leq M\|u\|_V\|v\|_V \quad \exists M > 0$$

and the ellipticity condition that $\exists \lambda$ such that $\forall v \in V$ and for some $\alpha > 0$

$$a(v, v) + \lambda\|v\|_H^2 \geq \alpha\|v\|_V^2$$

for almost all $t \in [0, T]$, proves the following existence theorem.

Theorem: There exists a unique stochastic process $u \in W(0, T)$ as a solution of the system [4.12]. Also, this solution is continuously dependent on the data, ie, the mapping

$$\{g, u_0\} \rightarrow u$$

is continuous from $L^2(0, T; V') \times H$ to $L^2(0, T; V)$. □

4.2.4 Stochastic Integration

At this point I will return to the concept of stochastic integration that was discussed briefly in Chapter 2. That discussion characterized the stochastic integral in terms of a Wiener or Brownian motion process. Doob[36] generalizes this somewhat to define the stochastic integral in terms of a *martingale*. This term is not very descriptive. In fact, the primary definition in Webster's dictionary is that of a part of a harness for a horse. However, it is also used to describe a system of betting strategies. Of course, probability theory makes this form of the definition more precise. Following Doob[36], Burrill[17] and Jazwinski[54], the major ideas are outlined below. As a matter of convenience the *Radon-Nikodym Theorem* is stated as found in Burrill[17].

Radon-Nikodym Theorem Let the measure μ and the absolutely continuous additive function ϕ be σ -finite. Then there is a finite valued measurable function g such that

$$\phi(E) = \int_E g d\mu$$

for each measurable set E . \square

Given a probability space (Ω, \mathcal{E}, P) and an integrable random variable X on Ω , define the function

$$\phi(E) = \int_E X dP$$

which is an additive function and absolutely continuous relative to P , ($\phi(E) = 0$ if $P(E) = 0$). The set E belongs to a σ -algebra \mathcal{F} contained in \mathcal{E} . So, by the Radon-Nikodym Theorem there is an \mathcal{F} -measurable function denoted by $\mathbf{E}[X|\mathcal{F}]$ such that

$$\phi(E) = \int_E \mathbf{E}[X|\mathcal{F}] dP = \int_E X dP$$

for each $E \in \mathcal{F}$ and called the *conditional expectation of X given \mathcal{F}* . In fact, $\mathbf{E}[X|\mathcal{F}]$ represents an equivalence class of integrable random variables such that any member of the equivalence class is measurable with respect to \mathcal{F} and has the same integral as X over any $E \in \mathcal{F}$.

Next, let

$$T \subset I \cup \{-\infty, +\infty\}$$

where I is the set of integers, and let $\{\mathcal{F}_t : t \in T\}$ be a collection of σ -algebras such that

$$\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{E} \quad \text{for } s < t$$

and, finally, let $\{X(t) : t \in T\}$ be a collection on integrable random variables such that $X(t)$ is measurable relative to \mathcal{F}_t for each t . In probability theory, the sets in \mathcal{F}_t are called *events*, and the measurability of $X(t)$ with respect to the σ -algebra \mathcal{F}_t can be interpreted to mean that the values of $X(t)$ are detectable by the events in \mathcal{F}_t .

Definition: The collection $\{X(t) : t \in T\}$ is a *martingale* relative to $\{\mathcal{F}_t : t \in T\}$ if

$$X(s) = \mathbf{E}[X(t)|\mathcal{F}_s] \quad s < t$$

and a *semi-martingale* relative to $\{\mathcal{F}_t : t \in T\}$ if

$$X(s) \leq \mathbf{E}[X(t)|\mathcal{F}_s] \quad s < t$$

Furthermore, the following theorem holds

Theorem: The collection $\{X(t) : t \in T\}$ of integrable random variables is a *martingale* iff for all $s, t \in T$ with $s < t$ and all $E \in \mathcal{F}_s$

$$\int_E X(s)dP = \int_E \mathbf{E}[X(t)|\mathcal{F}_s]dP = \int_E X(t)dP$$

and a *semi-martingale* iff for all $s, t \in T$ with $s < t$ and all $E \in \mathcal{F}_s$

$$\int_E X(s)dP \leq \int_E \mathbf{E}[X(t)|\mathcal{F}_s]dP \leq \int_E X(t)dP \quad \square$$

The relationship between martingales and the Wiener process is given by the following theorem from Doob[36]

Theorem: Let $\{X(t), \mathcal{F}_t, a \leq t \leq b\}$ be a real martingale, and suppose that almost all sample paths of the process are continuous. Suppose that

$$\mathbf{E}[X(t)^2] < \infty \quad a \leq t \leq b$$

and that for each pair s, t with $s < t$

$$E[(X(t) - X(s))^2|\mathcal{F}_s] = t - s$$

with probability 1. Then it follows that the $X(t)$ process has independent increments and is a Wiener process. \square

Doob[36] defines the stochastic integral

$$\int_E \Phi(t, \omega)d\beta(t)$$

by assuming that the process $\beta(t)$ is a martingale (evidently it can be extended to include $\beta(t)$ as a semi-martingale). The Itô integral discussed in Chapter 2 follows as a special case from the preceding theorem.

Jazwinski[54] defines the Itô integral as the mean square limit of step function processes in the following manner:

Definition A *step function*, $g(t, \omega)$, is defined as

$$g(t, \omega) = \begin{cases} 0 & t < a_1 \\ g_j(\omega) & a_j \leq t < a_{j+1} \quad j \leq n-1 \\ 0 & a_n \leq t \end{cases}$$

where $a_1 < \dots < a_n$, and $g_j(\omega)$ is measurable with respect to \mathcal{F}_{a_j} and $E[|g_j(\omega)|^2] < \infty$ and $g_j(\omega)$ is independent of

$$\{\beta(a_k) - \beta(a_l) : a_j \leq a_l \leq a_k \leq a_n\}$$

This is a condition of *nonanticipativeness*. One way of interpreting this is that the function g_j is independent of the Wiener process in future time t . In other words, the values of g_j are observable only by events prior to a_j .

Let $\{g^n(t, \omega)\}$ be a sequence of step function processes converging to the process $g(t, \omega)$ in the sense that

$$\int_T E[|g(t, \omega) - g^n(t, \omega)|^2] dt \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

then the Itô integral of the process $g(t, \omega)$ with respect to the Wiener process $\beta(t, \omega)$ is defined to be

$$\int_T g(t, \omega) d\beta(t) = (m^2) \lim_{n \rightarrow \infty} \int_T g^n(t, \omega) d\beta(t)$$

Stochastic integrals are defined in the sense of mean squared convergence which implies convergence in measure P , because if $\epsilon > 0$ and

$$\Omega_n = \left\{ \omega : \left| \int_T g^n(t, \omega) d\beta(t) - \int_T g(t, \omega) d\beta(t) \right| \geq \epsilon \right\}$$

then

$$\int_{\Omega_n} \left| \int_T g^n(t, \omega) d\beta(t) - \int_T g(t, \omega) d\beta(t) \right|^2 dP \geq \int_{\Omega_n} \epsilon^2 dP = \epsilon^2 P(\Omega_n)$$

Hence, from the mean convergence it follows that

$$\begin{aligned}
P(\Omega_n) &\leq \frac{1}{\epsilon^2} \int_{\Omega_n} \left| \int_T g^n(t, \omega) d\beta(t) - \int_T g(t, \omega) d\beta(t) \right|^2 dP \\
&\leq \frac{1}{\epsilon^2} \int_{\Omega} \left| \int_T g^n(t, \omega) d\beta(t) - \int_T g(t, \omega) d\beta(t) \right|^2 dP \\
&\rightarrow 0 \quad \text{as } n \rightarrow \infty
\end{aligned}$$

In order to extend these results to a Hilbert space, H , it is necessary to define a Wiener process in a Hilbert space, Falb[37], Curtain and Falb[26], [27], and Sawaragi, Soeda, Omatu[85]. If $W(t)$ is an H -valued Wiener process, then there are complex random processes $\{\beta_i\}_{i=0}^{\infty}$ such that

$$W(t) = \sum_{i=0}^{\infty} \beta_i(t) e_i$$

almost everywhere in (t, ω) . Here, $\{e_i\}_{i=0}^{\infty}$ is an orthonormal basis of H . And, $\Re(\beta_i(t))$ and $\Im(\beta_i(t))$ are real Wiener processes.

Itô stochastic integration is extended to the Hilbert space setting as follows: First, in Section 1.3, a complex-valued second order random variable was defined in terms of the modulus function, $|\cdot|$. In the case of a Hilbert space valued random variable, the H -valued random variable, $X(\omega)$, is second order if

$$\mathbf{E} \left[\|X(\omega)\|_H^2 \right] < \infty$$

where the modulus function is now replaced by the H -norm, $\|\cdot\|_H$. Secondly, the mean squared convergence is done in terms of the $\|\cdot\|_H$ norm instead of the $|\cdot|$ function.

Let H be a Hilbert space and $W(t)$ an H -valued Wiener process. Also, let $g(t, \omega)$ be a step function from T into $\mathcal{L}(H, H)$

$$g(t, \omega) = \begin{cases} 0 & t < a_1 \\ g_j(\omega) & a_j \leq t < a_{j+1} \quad j \leq n-1 \\ 0 & a_n \leq t \end{cases}$$

where $a_1 < \dots < a_n$, and if $g^n(t, \omega)$ is a sequence of step functions converging to $g(t, \omega)$, then

$$\begin{aligned}\int_T g(t, \omega) dW(t) &= (m^2) \lim_{n \rightarrow \infty} \int_T g^n(t, \omega) dW(t) \\ &= (m^2) \lim_{n \rightarrow \infty} \sum_{j=1}^n g_j^n(\omega) [W(t_j) - W(t_{j-1})]\end{aligned}$$

Or,

$$\lim_{n \rightarrow \infty} \mathbf{E} \left[\left\| \int_T g^n(t, \omega) dW(t) - \int_T g(t, \omega) dW(t) \right\|_H^2 \right] = 0$$

Because the Itô integral is applicable to a wider class of functions, it is used in this analysis even though new rules of Itô calculus must be devised.

4.2.5 Itô's Lemma In Hilbert Space

The most important new rule required for the solution of the stochastic evolution equations is Itô's lemma in Hilbert space. It is a kind of change of variable formula. The one-dimensional version of the Itô formula was described in Section 1.1, where the relationship

$$dW(t)^2 = dt$$

was used to develop it. Using the relationship

$$\mathbf{E} [d\vec{W}(t) d\vec{W}(t)^\dagger] = Q(t)dt$$

where

$$\vec{W}(t) = [W^1(t), \dots, W^m(t)]^\dagger$$

is an m -dimensional Wiener process, and the mappings

$$\vec{a} : [t_0, T] \times \mathfrak{R}^d \rightarrow \mathfrak{R}^d$$

$$\mathbf{b} : [t_0, T] \times \mathfrak{R}^d \rightarrow \mathfrak{R}^{d \times m}$$

the d -dimensional vector stochastic differential equation is

$$d\vec{X}(t) = \vec{a}(t, \vec{X}(t)) + \mathbf{b}(t, \vec{X}(t)) d\vec{W}(t)$$

Then, Jazwinski[54], Kloeden, et al[59], for a sufficiently smooth function

$$g : [t_0, T] \times \mathfrak{R}^d \rightarrow \mathfrak{R}^k$$

of the solution $\vec{X}(t)$, $t_0 \leq t \leq T$, there is a k -dimensional process

$$\vec{Y}(t) = g(t, \vec{X}(t)) \quad t_0 \leq t \leq T$$

such that for the p^{th} component of the vector process $\vec{Y}(t)$,

$$\begin{aligned} dY^p(t) &= \left(\frac{\partial g^p}{\partial t} + \sum_{i=1}^d a^i \frac{\partial g^p}{\partial x_i} + \frac{1}{2} \sum_{j,k=1}^d \sum_{i,l=1}^m b_{ki} Q_{il} b_{jl} \frac{\partial^2 g^p}{\partial x_j \partial x_k} \right) dt \\ &+ \sum_{l=1}^m \sum_{i=1}^d b_{il} \frac{\partial g^p}{\partial x_i} dW^l(t) \quad p = 1, \dots, k \end{aligned}$$

where all terms are evaluated at the points $(t, \vec{X}(t))$. This is the finite dimensional vector version of Itô's formula.

As for the infinite dimensional version, let

$$\begin{aligned} \mu_1(H, K) &= \{S(t, \omega) : S(t, \omega) \in \mathcal{L}(H, K) \\ &\text{and } \int_0^T E[\|S(t, \omega)\|_{\mathcal{L}(H, K)}^2] < \infty\} \end{aligned}$$

Itô's Lemma Let H , K , and G be Hilbert spaces and let $W(t)$ be an H -valued Wiener process. Suppose that $g(t, c)$ is a continuous map of $[0, T] \times K$ into G and that $u(t)$ is a K -valued stochastic process with stochastic differential

$$du(t) = A(t, \omega)dt + C(t, \omega)dW(t)$$

such that

- $g_t(t, c)$ is continuous on $[0, T] \times K$
- $g(t, \cdot)$ is twice continuously differentiable on K for each fixed $t \in [0, T]$.
- $g_c(t, c)$ and $g_{cc}(t, c)$ are continuous in (t, c) on $[0, T] \times K$.
- $A(t, \omega)$ is a K -valued stochastic process which is measurable relative to \mathcal{F}_t , $t \in [0, T]$, and integrable on $[0, T]$, with probability 1.
- $C(t, \omega) \in \mu_1(H, K)$ and $\int_0^T E[\|C(t, \omega)\|^4] dt < \infty$
- $W(t)$ is real, and g_t and g_c denote the partial and Frechet derivatives.

Then

$$z(t) = g(t, u(t))$$

has the G -valued stochastic differential

$$\begin{aligned} dz(t) &= \left\{ g_t(t, u(t)) + g_c(t, u(t))[A(t, \omega)] + \frac{1}{2} \tilde{tr}(g_{cc}(t, u(t))[C(t)\xi_W] \right\} dt \\ &+ g_c(t, u(t))[C(t)]dW(t) \end{aligned}$$

Here \tilde{tr} represents a trace operator which is defined as

$$\tilde{tr}(g_{cc}(t, u(t))[C(t)\xi_W] \equiv \sum_{i=1}^{\infty} g_{cc}(t, u(t))[C(t)\sqrt{\lambda_i}e_i, C(t)\sqrt{\lambda_i}e_i] \quad (4.13)$$

where $\xi_W = \sum_{i=1}^{\infty} \sqrt{\lambda_i}e_i$ and the $\{e_i\}$ is an orthonormal basis of H of eigenvectors of Q , the covariance operator associated with the Wiener process, $W(t)$, with corresponding eigenvalues $\{\lambda_i\}$. The existence of these eigenvalues and eigenvectors follows from the definition of an H -valued Wiener process, Falb[37], Curtain and Falb[26, 27] and Sawaragi, Soeda and Omatu[85], where the covariance operator, Q , is assumed to be compact.

A Corollary that will be more useful is:

Corollary: Suppose that in addition to the hypothesis of the theorem that we let $G = \mathfrak{R}$. Then $dz(t)$ can be written as

$$\begin{aligned} dz(t) &= \{g_t(t, u(t)) + (A(t, \omega), \nabla_c g(t, u(t))) \\ &+ \frac{1}{2} tr(C(t, \omega)Q(t)C^*(t, \omega)\Theta_{cc}g(t, u(t)))\} dt \\ &+ (C^*(t, \omega)\nabla_c g(t, u(t)), dW(t)) \end{aligned}$$

where $\nabla_c g$ and $\Theta_{cc}g$ are the gradient and Hessian of g with respect to c .

Versions of these results on Itô's lemma are found in Curtain and Falb[26], Bensoussan[14] and Sawaragi, Soeda, Omatu[85].

4.2.6 Small o Notation

Let X be a B-space with dual X' , then $\langle y', x \rangle$ is the duality pairing of y' and x . Let $x_1 \in X$, $y'_1 \in X'$ and define the mapping

$$x_1 \circ y'_1 : X \rightarrow X$$

such that

$$(x_1 \circ y'_1)x = x_1 \langle y'_1, x \rangle \quad \forall x \in X$$

Theorem: Let X be a B-space and let ψ be the mapping of $X \oplus X'$ into $\mathcal{L}(X, X)$ defined by

$$\psi(x_1, y'_1) = x_1 \circ y'_1$$

Then ψ has the following properties:

- ψ is continuous
- ψ is linear in both x_1 and y'_1
- $(x_1 \circ y'_1)' = y'_1 \circ x_1$ if X is reflexive

Note: If $X = \mathfrak{R}^n$, then $\vec{x}_1 \circ \vec{y}'_1$ can be identified with the matrix $\vec{x}_1 \vec{y}'_1$.

The small \circ notation can be used to define the concept of covariance on a Hilbert space H . The inner product $(h, X(\omega))_H$ is a linear random functional on H' . So, if $X(\omega) \in H$, then

$$(\cdot, X(\omega))_H : H' \rightarrow \mathfrak{R}^1$$

and $(h, X(\omega))_H$ is a real random variable. Hence, $\mathbf{E}[(h_1, X_1)_H (h_2, X_2)_H]$ represents the covariance of X_1 and X_2 . Let $h_1, h_2 \in H'$ and let $X_1, X_2 \in H$, then by identifying $H = H'$ it follows that

$$\begin{aligned} (h_1, (X_1 \circ X_2')h_2)_H &= (h_1, (X_1 \circ X_2)h_2)_H \\ &= (h_1, X_1(X_2, h_2)_H)_H \\ &= (h_1, X_1)_H (h_2, X_2)_H \end{aligned}$$

Since $(X_1 \circ X_2')h_2 \in H$, by taking expectations it follows that

$$\mathbf{E}[(h_1, X_1)_H (h_2, X_2)_H] = (h_1, \mathbf{E}[(X_1 \circ X_2')h_2])_H$$

Assuming for the moment that $\mathbf{E}[X_1] = \mathbf{E}[X_2] = 0$ and that the mapping

$$\Xi : H' \times H' \longrightarrow \mathfrak{R}^1 \quad \ni$$

$$\Xi(h_1, h_2) = \mathbf{E}[(h_1, x_1)_H (h_2, x_2)_H] = (h_1, \text{Cov}[x_1, x_2] h_2)_H$$

Then, by the Riesz Representation theory, there is a unique Riesz map $\Lambda \in \mathcal{L}(H', H)$ such that

$$\Xi(h_1, h_2) = (h_1, \Lambda h_2)_H$$

where Λ is called the covariance operator, and since Λ is unique, $\Lambda = \mathbf{E}[X_1 \circ X_2']$. In the case that the expected values of X_1 and X_2 are not zero, the covariance operator of X_1 and X_2' is defined as, Falb[37],

$$\text{Cov}[X_1, X_2] = \text{Cov}[X_1, X_2'] = \mathbf{E}[X_1 \circ X_2'] - \mathbf{E}[X_1] \circ \mathbf{E}[X_2']$$

4.2.7 Hilbert Space Structures

Kadison[55] gives the following definitions for a *Direct Sum Of Hilbert Spaces* and for a *Direct Sum Of Operators*:

Let $\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_n$ be Hilbert spaces and \mathcal{K} be the set of all n -tuples $\{x_1, x_2, \dots, x_n\}$ with $x_i \in \mathcal{H}_i$. Then there is a Hilbert space structure on \mathcal{K} with the following definitions:

Algebraic Operations:

$$a\{x_1, \dots, x_n\} + b\{y_1, \dots, y_n\} = \{ax_1 + by_1, \dots, ax_n + by_n\}$$

Inner Product:

$$\langle \{x_1, \dots, x_n\}, \{y_1, \dots, y_n\} \rangle = (x_1, y_1) + \dots + (x_n, y_n)$$

Norm:

$$\|\{x_1, \dots, x_n\}\| = [\|x_1\|^2 + \dots + \|x_n\|^2]^{\frac{1}{2}}$$

The resulting Hilbert space \mathcal{K} is called a *Hilbert direct sum of $\mathcal{H}_1, \dots, \mathcal{H}_n$* and is denoted

$$\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n \equiv \sum_{i=1}^n \oplus \mathcal{H}_i$$

Let \mathcal{H}_i and \mathcal{K}_i be Hilbert spaces and $T_i \in \mathcal{B}(\mathcal{H}_i, \mathcal{K}_i), i = 1, \dots, n$, then the equation

$$T\{x_1, \dots, x_n\} = \{T_1x_1, \dots, T_nx_n\} \quad x_1 \in \mathcal{H}_1, \dots, x_n \in \mathcal{H}_n$$

defines a linear operator T such that

$$T : \sum_{i=1}^n \oplus \mathcal{H}_i \rightarrow \sum_{i=1}^n \oplus \mathcal{K}_i$$

where

$$T \equiv \sum_{i=1}^n \oplus T_i$$

The following notation will be used:

$$\{x_1, \dots, x_n\} \equiv \sum_{i=1}^n \oplus x_i$$

The direct sum of operators has the following properties:

$$\begin{aligned} \left(\sum_{i=1}^n \oplus T_i \right)^* &= \sum_{i=1}^n T_i^* \\ \sum_{i=1}^n \oplus (aS_i + bT_i) &= a \left(\sum_{i=1}^n \oplus S_i \right) + b \left(\sum_{i=1}^n \oplus T_i \right) \\ \left(\sum_{i=1}^n \oplus R_i \right) \left(\sum_{i=1}^n \oplus S_i \right) &= \sum_{i=1}^n \oplus R_i S_i \end{aligned}$$

Frechet Derivatives

In order to derive the first moment equation, let

$$g(t, \nu) = (h, \nu)_H \quad h \in H'$$

So that

$$g : [0, T] \times H \rightarrow \mathfrak{R}$$

And, for a fixed t ,

$$g : H \rightarrow \mathfrak{R}$$

If the Frechet derivative exists, then the Gateaux derivative exists and the two are equal. Denote the Frechet derivative by the symbol

$$\frac{\partial g}{\partial \nu} \in \mathcal{L}(H, \mathfrak{R})$$

By the Riesz Representation theorem, the Frechet derivative can be represented by the inner product on H . So, for a fixed h , the Frechet derivative is defined as

$$\frac{\partial g}{\partial \nu} \eta = \left(\frac{\partial g}{\partial \nu}, \eta \right)_H = \lim_{\alpha \rightarrow 0} \frac{(h, \nu + \alpha \eta)_H - (h, \nu)_H}{\alpha}$$

From this definition it follows that

$$\left(\frac{\partial g}{\partial \nu}, \eta \right)_H = (h, \eta)_H \quad \Rightarrow \quad \frac{\partial g}{\partial \nu} = h \in \mathcal{L}(H, \mathfrak{R})$$

Since h and η are fixed, $(h, \eta)_H$ is a constant. So, the second Frechet derivative of g is zero.

To derive the second moment equation, let

$$g(t, \nu) = (h_1, \nu)_H (h_2, \nu)_H$$

Then,

$$\begin{aligned} \left(\frac{\partial g}{\partial \nu}, \eta \right)_H &= \lim_{\epsilon \rightarrow 0} \frac{(h_1, \nu + \epsilon \eta)(h_2, \nu + \epsilon \eta) - (h_1, \nu)(h_2, \nu)}{\epsilon} \\ &= \lim_{\epsilon \rightarrow 0} \frac{(h_1, \epsilon \eta)(h_2, \nu) + (h_1, \nu)(h_2, \epsilon \eta) + (h_1, \epsilon \eta)(h_2, \epsilon \eta)}{\epsilon} \\ &= (h_1, \eta)(h_2, \nu) + (h_1, \nu)(h_2, \eta) \end{aligned}$$

Or, in operator notation if we let

$$T_1 = (h_1, \cdot)h_2 \quad \text{and} \quad T_2 = (h_2, \cdot)h_1$$

Then,

$$T = T_1 \oplus T_2 : H \oplus H \rightarrow H \oplus H$$

and, using the definitions for the inner products we have,

$$\langle (h_1, \nu)h_2 \oplus (h_2, \nu)h_1, \eta \oplus \eta \rangle = (h_1, \eta)(h_2, \nu) + (h_1, \nu)(h_2, \eta) = \left(\frac{\partial g}{\partial \nu}, \eta \right)_H$$

So, we can make the identification

$$\frac{\partial g}{\partial \nu} = (h_1, \nu)h_2 \oplus (h_2, \nu)h_1$$

For the second derivative, we can write

$$\frac{\partial^2 g}{\partial \nu^2} \in \mathcal{L}(H, \mathcal{L}(H, \mathfrak{R})) \cong \mathcal{L}(H \oplus H, \mathfrak{R})$$

where \cong represents an isometry. Again, by the Riesz Representation theorem, the second derivative is given by

$$\begin{aligned} \frac{\partial^2 g}{\partial \nu^2}(\zeta, \eta) &= \left(\frac{\partial^2 g}{\partial \nu^2} \zeta \oplus \zeta, \eta \oplus \eta \right)_{H \oplus H} \\ &= \lim_{\epsilon \rightarrow 0} \frac{[(h_1, \nu + \epsilon \zeta)(h_2, \eta) + (h_2, \nu + \epsilon \zeta)(h_1, \eta)] - [(h_1, \nu)(h_2, \eta) + (h_2, \nu)(h_1, \eta)]}{\epsilon} \\ &= \lim_{\epsilon \rightarrow 0} \frac{[(h_1, \nu + \epsilon \zeta)(h_2, \eta) - (h_1, \nu)(h_2, \eta)] + [(h_2, \nu + \epsilon \zeta)(h_1, \eta) - (h_2, \nu)(h_1, \eta)]}{\epsilon} \\ &= (h_1, \zeta)(h_2, \eta) + (h_2, \zeta)(h_1, \eta) \quad \zeta, \eta \in H \end{aligned}$$

And, if we identify H with its dual H' we can write this operator in terms of the small o notation, so that

$$(h_1, \zeta)h_2 = (h_2 \circ h_1)\zeta$$

With this notation we can write

$$\begin{aligned}\frac{\partial^2 g}{\partial \nu^2}(\zeta, \eta) &= (h_1, \zeta)(h_2, \eta) + (h_2, \zeta)(h_1, \eta) \\ &= \langle (h_2 \circ h_1) \oplus (h_1 \circ h_2)\zeta \oplus \zeta, \eta \oplus \eta \rangle\end{aligned}$$

Hence,

$$\frac{\partial^2 g}{\partial \nu^2} = h_1 \circ h_2 \oplus h_2 \circ h_1$$

4.2.8 Moment Equation Derivation

In this section, equations for which the forcing term has a Gaussian white noise component are discussed, References for this section are Åström[7], Bensoussan[14], Chow[22] and Serrano, Unny, Lennox[90].

In the finite dimensional case, Åström[7] shows that the linear stochastic differential equation

$$dx = -Axdt + f + dv$$

where dv is a white noise process, has moment equations

$$\frac{dM_1}{dt} = -AM_1 + f$$

and

$$\frac{dM_2}{dt} = -(AM_2 + M_2A^\dagger) + fM_1 + M_1f^\dagger + R_1 \quad (4.14)$$

where $\mathbf{E}[v(t)v(t)] = R_1t$. A result similar to this will be derived for the infinite dimensional case.

Consider the equation

$$\frac{du}{dt} = -Au + f + \zeta$$

where u is a function of (t, x, ω) and belongs to a Hilbert space H ; A is a spatially elliptic operator; f is deterministic; and ζ is a Gaussian white noise process.

In integral form, this equation is

$$u(t) = u(0) + \int_0^t (-Au + f)ds + \int_0^t \zeta(s)ds$$

Then, $W(t) = \int_0^t \zeta(s)ds$ is an H -valued Wiener process and the original equation can be written as

$$du(t) = (-Au + f)dt + dW(t)$$

For more generality, introduce the stochastic operator $?(t)$ such that $?(t) \in \mathcal{L}(H, H)$ and

$$\int_0^t \|?(s)\|_{\mathcal{L}(H,H)}^2 ds < \infty$$

So that we have

$$du(t) = (-Au + f)dt + ?(t)dW(t)$$

Using Itô's lemma,

$$\begin{aligned} dg(t, u(t)) &= \left\{ \frac{\partial g}{\partial t}(t, u(t)) + \left(-Au + f, \frac{\partial g}{\partial u}(t, u(t)) \right) \right. \\ &\quad \left. + \frac{1}{2} tr ? Q(t) ?^* \frac{\partial^2 g}{\partial u^2} \right\} dt + \left(\frac{\partial g}{\partial u}, ? dW \right) \end{aligned} \quad (4.15)$$

And, $Q(t) \in \mathcal{L}^\infty(0, T; \mathcal{L}(H, H))$ and is called the *covariance operator*.

Allowing $g(0, u(0)) = 0$, Equation[4.15] can be interpreted in the stochastic differential equation sense as

$$\begin{aligned} g(t, u(t)) &= \int_0^t \left\{ \frac{\partial g}{\partial s}(s, u(s)) + \left(-Au + f, \frac{\partial g}{\partial u}(s, u(s)) \right) \right. \\ &\quad \left. + \frac{1}{2} tr ? Q(s) ?^* \frac{\partial^2 g}{\partial u^2} \right\} ds + \int_0^t \left(\frac{\partial g}{\partial u}(s, u(s)), ? dW(s) \right) \end{aligned}$$

where $dW(s)$ is to be interpreted as a Gaussian white noise.

Now, if $g = (h, u(t))_H$, $h \in H'$, we have from the Frechet derivative

$$\frac{\partial g}{\partial u} = h \quad \frac{\partial^2 g}{\partial u^2} = 0$$

Furthermore, since g depends on t only through u ,

$$\frac{\partial g}{\partial t} = 0$$

Taking expectations and using the result that $\mathbf{E}[(h, u)_H] = (h, \mathbf{E}[u])_H$, it follows that

$$(h, \mathbf{E}[u])_H = \int_0^t \mathbf{E}[(-Au + f, h)_H] ds + \int_0^t \mathbf{E}[(h, ? dW(s))_H]$$

To evaluate the last integral,

$$\begin{aligned} \int_0^t \mathbf{E}[(h, ? dW(s))_H] &= (m^2) \lim_{n \rightarrow \infty} \sum_{j=1}^n \mathbf{E} \left[(h, g_j^n(\omega) [W(t_{j+1}) - W(t_j)])_H \right] \\ &= 0 \end{aligned}$$

since $g_j^n(\omega)$ and $W(t_{j+1}) - W(t_j)$ are independent which follows from the nonanticipativeness of the operator $g_j^n(\omega)$ with respect to $W(t_{j+1}) - W(t_j)$. So, if $M_1 = \mathbf{E}[u]$, then assuming that A is deterministic

$$\left(h, \frac{dM_1}{dt} \right) = -(h, AM_1) + (h, f) \quad (4.16)$$

Or, in this weak sense, the *first moment equation* is

$$\frac{dM_1}{dt} = -AM_1 + f$$

To obtain the moment equation for the second moment, let

$$g = (h_1, u)_H (h_2, u)_H$$

From the Frechet derivatives we have

$$\frac{\partial g}{\partial u} = (h_1, u)h_2 \oplus (h_2, u)h_1$$

And,

$$\frac{\partial^2 g}{\partial u^2} = h_2 \circ h_1 \oplus h_1 \circ h_2$$

Hence, it follows from Itô's lemma

$$\begin{aligned} \frac{d}{dt}(h_1, u)(h_2, u) &= - \langle (h_1, u)h_2 \oplus (h_2, u)h_1, Au \oplus Au \rangle \\ &+ \langle (h_1, u)h_2 \oplus (h_2, u)h_1, f \oplus f \rangle \\ &+ \frac{1}{2} \text{tr}[?^*(h_2 \circ h_1) \oplus (h_1 \circ h_2)?Q(t)] \\ &+ \langle (h_1, u)h_2 \oplus (h_2, u)h_1, ?dW \oplus ?dW \rangle \end{aligned}$$

Expanding this equation we get

$$\begin{aligned} \frac{d}{dt}(h_1, u)(h_2, u) &= -[(h_1, u)(h_2, Au) + (h_2, u)(h_1, Au)] \\ &+ (h_1, u)(h_2, f) + (h_2, u)(h_1, f) \quad (4.17) \\ &+ \frac{1}{2} \text{tr}[?^*(h_2 \circ h_1) \oplus (h_1 \circ h_2)?Q(t)] \\ &+ (h_1, u)(h_2, ?dW) + (h_2, u)(h_1, ?dW) \end{aligned}$$

Taking expectations and using the following definition of the correlation operator,

$$\mathbf{E}[(h_1, X(\omega))(h_2, Y(\omega))] = (h_1, R_{XY}h_2)$$

where $R_{XY} = M_2$, if $X = Y$, it follows that for $M_2 = M_2^*$ and A deterministic,

$$\begin{aligned} \mathbf{E}[(h_1, u)(h_2, Au)] &= \mathbf{E}[(h_1, u)(A^*h_2, u)] = (h_1, M_2A^*h_2) \\ \mathbf{E}[(h_2, u)(h_1, Au)] &= \mathbf{E}[(h_2, u)(A^*h_1, u)] = (h_2, M_2A^*h_1) = (h_1, AM_2h_2) \\ \mathbf{E}[(h_1, u)(h_2, f)] &= (h_1, R_{uf}h_2) = (h_1, M_1fh_2) \end{aligned}$$

And,

$$\begin{aligned} \frac{d}{dt}(h_1, M_2h_2) &= -[(h_1, M_2A^*h_2) + (h_1, AM_2h_2)] \\ &+ (h_1, M_1fh_2) + (h_1, M_1fh_2) \quad (4.18) \\ &+ \frac{1}{2} \mathbf{E}[\text{tr}?^*(h_2 \circ h_1) \oplus (h_1 \circ h_2)?Q(t)] \end{aligned}$$

Interchanging the roles of h_1 and h_2 , we get

$$\begin{aligned}
\frac{d}{dt}(h_2, M_2 h_1) &= -[(h_2, M_2 A^* h_1) + (h_2, A M_2 h_1)] \\
&+ (h_2, M_1 f h_1) + (h_2, M_1 f h_1) \\
&+ \frac{1}{2} \mathbf{E} [tr?^*(h_1 \circ h_2) \oplus (h_2 \circ h_1)? Q(t)]
\end{aligned} \tag{4.19}$$

Adding these Equations [4.18] and [4.19] together and using the definition of the inner product on a direct sum of Hilbert spaces,

$$\begin{aligned}
\left\langle h_1 \oplus h_2, \frac{dM_2}{dt} \oplus \frac{dM_2}{dt} h_2 \oplus h_1 \right\rangle &= \langle h_1 \oplus h_2, -(A M_2 \oplus M_2 A^*) h_2 \oplus h_1 \rangle \\
&+ \langle h_1 \oplus h_2, (M_1 f \oplus (M_1 f)^*) h_2 \oplus h_1 \rangle \\
&+ \frac{1}{2} \mathbf{E} [tr?^*(h_2 \circ h_1) \oplus (h_1 \circ h_2)? Q(t)] \\
&+ \frac{1}{2} \mathbf{E} [tr?^*(h_1 \circ h_2) \oplus (h_2 \circ h_1)? Q(t)]
\end{aligned}$$

where, as before, the last term of Equation[4.17] vanishes on taking expectations.

Even though this equation has a weak sense formulation, it has a form similar to the simpler case Equation[4.14], page 155, above. Using Equation[4.13], page 149, and the fact that $Qe_i = \lambda_i e_i$, the trace term can be put into a more usable form by expanding and using Parseval's relation and the definition of the small o notation

$$\begin{aligned}
\frac{1}{2} \tilde{tr} \left(\frac{\partial^2 g}{\partial u^2} \right) [?\xi_W] &= \frac{1}{2} \sum_{i=1}^{\infty} \frac{\partial^2 g}{\partial u^2} \left[? \sqrt{\lambda_i} e_i, ? \sqrt{\lambda_i} e_i \right] \\
&= \frac{1}{2} \sum_{i=1}^{\infty} (h_1 \circ h_2) \oplus (h_2 \circ h_1) \left[? \sqrt{\lambda_i} e_i, ? \sqrt{\lambda_i} e_i \right] \\
&= \frac{1}{2} \sum_{i=1}^{\infty} \left[(h_1, ? \sqrt{\lambda_i} e_i) + (h_2, ? \sqrt{\lambda_i} e_i)(h_2, ? \sqrt{\lambda_i} e_i)(h_1, ? \sqrt{\lambda_i} e_i) \right] \\
&= \frac{1}{2} \sum_{i=1}^{\infty} [(?^* h_1, \lambda_i e_i)(?^* h_2, e_i) + (?^* h_2, \lambda_i e_i)(?^* h_1, e_i)]
\end{aligned}$$

$$\begin{aligned}
& \stackrel{\text{Parseval}}{=} \frac{1}{2} [(Q?^* h_1, ?^* h_2) + (Q?^* h_2, ?^* h_1)] \\
& = (h_1, ? Q?^* h_2) \tag{4.20}
\end{aligned}$$

This theory will be demonstrated in the next two examples. The first example will use the theory to derive the mean concentration equation, Equation[4.4], page 132, the velocity-concentration equation, Equation[4.7], page 133, and the concentration-covariance equation, Equation[4.10], page 135. Consider the transport equation

$$\frac{\partial c}{\partial t} + \nabla \cdot (c\vec{V}) - \nabla \cdot (\mathbf{D}\nabla c) = 0$$

Suppose that the tensor \mathbf{D} has been specified in a deterministic manner such as specified in Chapter 2, and the velocity and concentration are expressed as

$$\vec{V}(\vec{x}, \omega) = \mathbf{E}[\vec{V}(\vec{x})] + \vec{V}'(\vec{x}) \quad ; \quad \mathbf{E}[\vec{V}'(\vec{x})] = 0$$

$$c(\vec{x}, t) = \mathbf{E}[c(\vec{x}, t)] + c'(\vec{x}, t) \quad ; \quad \mathbf{E}[c'(\vec{x}, t)] = 0$$

In terms of Equation[4.12], page 138,

$$A(\vec{x}, t, \omega) = \nabla \cdot ((\cdot)(\mathbf{E}[\vec{V}(\vec{x}, \omega)] + \vec{V}'(\vec{x}, \omega))) - \nabla \cdot (\mathbf{D}\nabla(\cdot))$$

$$g(\vec{x}, t, \omega) = 0$$

From now on, ω will not be specifically stated. Then, from Equation[4.16] page 157,

$$\begin{aligned}
\left(h, \frac{\partial c}{\partial t} \right) &= -(h, Ac) \\
&= -\left(h, \nabla \cdot [(\mathbf{E}[c] + c')(\mathbf{E}[\vec{V}] + \vec{V}')] - \nabla \cdot (\mathbf{D}\nabla(\mathbf{E}[c] + c')) \right) \\
&= -\left(h, \nabla \cdot [(\mathbf{E}[c]\mathbf{E}[\vec{V}] + \mathbf{E}[c]\vec{V}' + c'\mathbf{E}[\vec{V}] + c'\vec{V}')] - \nabla \cdot (\mathbf{D}\nabla(\mathbf{E}[c] + c')) \right)
\end{aligned}$$

Taking expectations, and using $\mathbf{E}[c'] = \mathbf{E}[\vec{V}'] = 0$,

$$\left(h, \frac{\partial \mathbf{E}[c]}{\partial t} \right) = -\left(h, \nabla \cdot [(\mathbf{E}[c]\mathbf{E}[\vec{V}] + \mathbf{E}[c'\vec{V}'])] - \nabla \cdot (\mathbf{D}\nabla(\mathbf{E}[c])) \right)$$

So that in the weak sense,

$$\frac{\partial \mathbf{E}[c]}{\partial t} + \nabla \cdot (\mathbf{E}[c] \mathbf{E}[\vec{V}]) - \nabla \cdot (\mathbf{D} \nabla \mathbf{E}[c]) + \nabla \cdot \mathbf{E}[c' \vec{V}'] = 0$$

which agrees with Equation[4.4], page 132. For the second moment equations, let \vec{x} and \vec{x}' be two different coordinate systems, then from Equation[4.17], page 158, with $h_1, h_2 \in H'$,

$$\begin{aligned} \frac{\partial}{\partial t} (h_1, c(\vec{x}, t))(h_2, c(\vec{x}', t)) &= - \left[(h_1, c(\vec{x}, t))(h_2, A_{\vec{x}'} c(\vec{x}', t)) \right. \\ &\quad \left. + (h_2, c(\vec{x}', t))(h_1, A_{\vec{x}} c(\vec{x}, t)) \right] \\ &= (h_1, c(\vec{x}, t)) \left(h_2, \frac{\partial c(\vec{x}', t)}{\partial t} \right) + (h_2, c(\vec{x}', t)) \left(h_1, \frac{\partial c(\vec{x}, t)}{\partial t} \right) \end{aligned} \quad (4.21)$$

as would be expected since the Wiener process is excluded from playing a role in this example. Expanding the left hand side of Equation[4.21], taking expectations, using $\mathbf{E}[c'(\vec{x}, t)] = \mathbf{E}[c'(\vec{x}', t)] = 0$ and letting $C_{cc}(\vec{x}, \vec{x}', t) = \mathbf{E}[c'(\vec{x}, t)c'(\vec{x}', t)]$

$$\begin{aligned} \mathbf{E} \left[\frac{\partial}{\partial t} (h_1, c(\vec{x}, t))(h_2, c(\vec{x}', t)) \right] &= \mathbf{E} \left[\frac{\partial}{\partial t} (h_1, \mathbf{E}[c(\vec{x}, t)] + c'(\vec{x}, t)) \right. \\ &\quad \left. \times (h_2, \mathbf{E}[c(\vec{x}', t)] + c'(\vec{x}', t)) \right] \\ &= \frac{\partial}{\partial t} \left[\mathbf{E}(h_1, \mathbf{E}[c(\vec{x}, t)])(h_2, \mathbf{E}[c(\vec{x}', t)]) + (h_1, C_{cc}(\vec{x}, \vec{x}', t)h_2) \right] \\ &= \frac{\partial}{\partial t} \left[(h_1, \mathbf{E}[c(\vec{x}, t)]\mathbf{E}[c(\vec{x}', t)]h_2) + (h_1, C_{cc}(\vec{x}, \vec{x}', t)h_2) \right] \\ &= \left(h_1, \frac{\partial \mathbf{E}[c(\vec{x}, t)]}{\partial t} \mathbf{E}[c(\vec{x}', t)]h_2 \right) + \left(h_1, \frac{\partial \mathbf{E}[c(\vec{x}', t)]}{\partial t} \mathbf{E}[c(\vec{x}, t)]h_2 \right) \\ &\quad + \frac{\partial}{\partial t} (h_1, C_{cc}(\vec{x}, \vec{x}', t)h_2) \end{aligned}$$

Setting this equal to the expected value of the right hand side of Equation[4.21]

$$\frac{\partial}{\partial t} (h_1, C_{cc}(\vec{x}, \vec{x}', t)h_2) + \left(h_1, \mathbf{E} \left[\frac{\partial \mathbf{E}[c(\vec{x}, t)]}{\partial t} c(\vec{x}', t) + \frac{\partial \mathbf{E}[c(\vec{x}', t)]}{\partial t} c(\vec{x}, t) \right] h_2 \right)$$

$$= \left(h_1, \mathbf{E} \left[\frac{\partial c(\vec{x}', t)}{\partial t} c(\vec{x}, t) \right] h_2 \right) + \left(h_1, \mathbf{E} \left[\frac{\partial c(\vec{x}, t)}{\partial t} c(\vec{x}', t) \right] h_2 \right)$$

so that on rearranging terms,

$$\begin{aligned} \frac{\partial}{\partial t} (h_1, C_{cc}(\vec{x}, \vec{x}', t) h_2) &= - \left(h_1, \mathbf{E} \left[c(\vec{x}', t) A_{\vec{x}} c(\vec{x}, t) + \frac{\partial \mathbf{E}[c(\vec{x}, t)]}{\partial t} c(\vec{x}', t) \right] h_2 \right) \\ &\quad (4.22) \\ &\quad - \left(h_1, \mathbf{E} \left[c(\vec{x}, t) A_{\vec{x}'} c(\vec{x}', t) + \frac{\partial \mathbf{E}[c(\vec{x}', t)]}{\partial t} c(\vec{x}, t) \right] h_2 \right) \end{aligned}$$

From the first moment equation,

$$\frac{\partial \mathbf{E}[c(\vec{x}, t)]}{\partial t} = -\nabla_{\vec{x}} \cdot (\mathbf{E}[c(\vec{x}, t)] \mathbf{E}[\vec{V}(\vec{x})]) + \nabla_{\vec{x}} \cdot (\mathbf{D} \nabla_{\vec{x}} \mathbf{E}[c(\vec{x}, t)]) - \nabla_{\vec{x}} \cdot \mathbf{E}[c'(\vec{x}, t) \vec{V}'(\vec{x})]$$

and,

$$\begin{aligned} \frac{\partial \mathbf{E}[c(\vec{x}', t)]}{\partial t} &= -\nabla_{\vec{x}'} \cdot (\mathbf{E}[c(\vec{x}', t)] \mathbf{E}[\vec{V}'(\vec{x}')]) + \nabla_{\vec{x}'} \cdot (\mathbf{D} \nabla_{\vec{x}'} \mathbf{E}[c(\vec{x}', t)]) \\ &\quad - \nabla_{\vec{x}'} \cdot \mathbf{E}[c'(\vec{x}', t) \vec{V}'(\vec{x}')] \end{aligned}$$

Since

$$A_{\vec{x}}(\cdot) = \nabla_{\vec{x}} \cdot ((\cdot) \mathbf{E}[\vec{V}(\vec{x})] + \vec{V}'(\vec{x})) - \nabla_{\vec{x}} \cdot (\mathbf{D} \nabla_{\vec{x}}(\cdot))$$

it follows from the first term on the right hand side of Equation[4.22] and by letting $c(\vec{x}, t) = \mathbf{E}[c(\vec{x}, t)] + c'(\vec{x}, t)$ that

$$\begin{aligned} &- \left(h_1, \mathbf{E} \left[c(\vec{x}', t) A_{\vec{x}} c(\vec{x}, t) + \frac{\partial \mathbf{E}[c(\vec{x}, t)]}{\partial t} c(\vec{x}', t) \right] h_2 \right) \\ &= - \left(h_1, \mathbf{E} \left[c(\vec{x}', t) \left[\nabla_{\vec{x}} \cdot (\mathbf{E}[c(\vec{x}, t)] \vec{V}'(\vec{x})) + c'(\vec{x}, t) \mathbf{E}[\vec{V}(\vec{x})] \right. \right. \right. \\ &\quad \left. \left. \left. + c'(\vec{x}, t) \vec{V}'(\vec{x}) - \nabla_{\vec{x}} \cdot (\mathbf{D} \nabla_{\vec{x}} c'(\vec{x}, t)) - \nabla_{\vec{x}} \cdot \mathbf{E}[c'(\vec{x}, t) \vec{V}'(\vec{x})] \right] \right] h_2 \right) \end{aligned}$$

letting $c(\vec{x}', t) = \mathbf{E}[c(\vec{x}', t)] + c'(\vec{x}', t)$ and expanding the previous result,

$$\begin{aligned}
&= - \left(h_1, \mathbf{E} \left[\nabla_{\vec{x}} \cdot \left(\mathbf{E}[c(\vec{x}', t)] \mathbf{E}[c(\vec{x}, t)] \vec{V}'(\vec{x}) + c'(\vec{x}', t) \mathbf{E}[c(\vec{x}, t)] V'(\vec{x}) \right. \right. \right. \\
&+ \left. \left. \mathbf{E}[c(\vec{x}', t)] c'(\vec{x}, t) \mathbf{E}[\vec{V}(\vec{x})] + c'(\vec{x}', t) c'(\vec{x}, t) \mathbf{E}[\vec{V}(\vec{x})] \right. \right. \\
&+ \left. \left. \mathbf{E}[c(\vec{x}', t)] c'(\vec{x}, t) \vec{V}'(\vec{x}) + c'(\vec{x}', t) c'(\vec{x}, t) \vec{V}'(\vec{x}) \right) \right. \\
&- \left. \nabla_{\vec{x}} \cdot \left(\mathbf{D} \nabla_{\vec{x}} c'(\vec{x}, t) \mathbf{E}[c(\vec{x}', t)] \right) - \nabla_{\vec{x}} \cdot \left(\mathbf{D} \nabla_{\vec{x}} c'(\vec{x}, t) c'(\vec{x}', t) \right) \right) \\
&- \left. \nabla_{\vec{x}} \cdot \mathbf{E}[c'(\vec{x}, t) \vec{V}'(\vec{x})] \mathbf{E}[c(\vec{x}', t)] - \nabla_{\vec{x}} \cdot \mathbf{E}[c'(\vec{x}, t) \vec{V}'(\vec{x})] c'(\vec{x}', t) \right] h_2)
\end{aligned}$$

taking expectations, cancelling terms and using $\mathbf{E}[c'] = \mathbf{E}[\vec{V}'] = 0$

$$\begin{aligned}
&= - \left(h_1, \left[\nabla_{\vec{x}} \cdot \mathbf{E}[c(\vec{x}, t)] C_{c\vec{V}}(\vec{x}, \vec{x}', t) + \nabla_{\vec{x}} \cdot \mathbf{E}[\vec{V}(\vec{x})] C_{cc}(\vec{x}, \vec{x}', t) \right. \right. \\
&\left. \left. + \nabla_{\vec{x}} \cdot \mathbf{E}[c'(\vec{x}', t) c'(\vec{x}, t) \vec{V}'(\vec{x})] - \nabla_{\vec{x}} \cdot \left(\mathbf{D} \nabla_{\vec{x}} C_{cc}(\vec{x}, \vec{x}', t) \right) \right] h_2 \right)
\end{aligned} \tag{4.23}$$

Also, from the second term on the right hand side of Equation[4.22] an equation exactly like this one can be derived in the same way only with the vectors \vec{x} and \vec{x}' interchanged.

$$\begin{aligned}
&- \left(h_1, \left[\nabla_{\vec{x}'} \cdot \mathbf{E}[c(\vec{x}', t)] C_{c\vec{V}}(\vec{x}', \vec{x}, t) + \nabla_{\vec{x}'} \cdot \mathbf{E}[\vec{V}(\vec{x}')] C_{cc}(\vec{x}', \vec{x}, t) \right. \right. \\
&\left. \left. + \nabla_{\vec{x}'} \cdot \mathbf{E}[c'(\vec{x}, t) c'(\vec{x}', t) \vec{V}'(\vec{x}')] - \nabla_{\vec{x}'} \cdot \left(\mathbf{D} \nabla_{\vec{x}'} C_{cc}(\vec{x}', \vec{x}, t) \right) \right] h_2 \right)
\end{aligned} \tag{4.24}$$

Equations[4.22], [4.23] and [4.24] give the result for the concentration covariance equation as

$$\begin{aligned}
\frac{\partial}{\partial t} (h_1, C_{cc}(\vec{x}, \vec{x}', t) h_2) &= - \left(h_1, \left[\nabla_{\vec{x}} \cdot \mathbf{E}[c(\vec{x}, t)] C_{c\vec{V}}(\vec{x}, \vec{x}', t) + \nabla_{\vec{x}} \cdot \mathbf{E}[\vec{V}(\vec{x})] C_{cc}(\vec{x}, \vec{x}', t) \right. \right. \\
&+ \left. \nabla_{\vec{x}} \cdot \mathbf{E}[c'(\vec{x}', t) c'(\vec{x}, t) \vec{V}'(\vec{x})] - \nabla_{\vec{x}} \cdot \left(\mathbf{D} \nabla_{\vec{x}} C_{cc}(\vec{x}, \vec{x}', t) \right) \right] h_2) \\
&- \left(h_1, \left[\nabla_{\vec{x}'} \cdot \mathbf{E}[c(\vec{x}', t)] C_{c\vec{V}}(\vec{x}', \vec{x}, t) + \nabla_{\vec{x}'} \cdot \mathbf{E}[\vec{V}(\vec{x}')] C_{cc}(\vec{x}', \vec{x}, t) \right. \right. \\
&+ \left. \left[\nabla_{\vec{x}'} \cdot \mathbf{E}[c'(\vec{x}, t) c'(\vec{x}', t) \vec{V}'(\vec{x}')] - \nabla_{\vec{x}'} \cdot \left(\mathbf{D} \nabla_{\vec{x}'} C_{cc}(\vec{x}', \vec{x}, t) \right) \right] h_2 \right)
\end{aligned}$$

Comparing this with Equation[4.10], page 135, it is seen that this equation is the vector form of Equation[4.10].

Finally, returning to Equation[4.16], page 157, and multiplying both sides by $(h, \vec{V}_i(\vec{x}^t))$, where $\vec{V}_i(\vec{x}^t)$ is the i^{th} component of the vector $\vec{V}(\vec{x}^t)$, we have

$$\frac{\partial}{\partial t}(h_1, c(\vec{x}, t))(h_2, \vec{V}_i(\vec{x}^t)) = - \left[(h_1, A_{\vec{x}}c(\vec{x}, t))(h_2, \vec{V}_i(\vec{x}^t)) \right]$$

Next, write

$$\begin{aligned} \frac{\partial}{\partial t}(h_1, c'(\vec{x}, t))(h_2, \vec{V}'_i(\vec{x}^t)) &= \frac{\partial}{\partial t}(h_1, c(\vec{x}, t) - \mathbf{E}[c(\vec{x}, t)])(h_2, \vec{V}_i(\vec{x}^t) - \mathbf{E}[\vec{V}_i(\vec{x}^t)]) \\ &= \frac{\partial}{\partial t} \left[(h_1, c(\vec{x}, t))(h_2, \vec{V}_i(\vec{x}^t)) - (h_1, c(\vec{x}, t))(h_2, \mathbf{E}[\vec{V}_i(\vec{x}^t)]) \right. \\ &\quad \left. - (h_1, \mathbf{E}[c(\vec{x}, t)])(h_2, \vec{V}_i(\vec{x}^t)) + (h_1, \mathbf{E}[c(\vec{x}, t)])(h_2, \mathbf{E}[\vec{V}_i(\vec{x}^t)]) \right] \end{aligned}$$

Taking expectations and differentiating,

$$\begin{aligned} \frac{\partial}{\partial t} (h_1, C_{c\vec{V}_i}(\vec{x}, \vec{x}^t, t)h_2) &= \frac{\partial}{\partial t} \mathbf{E}[(h_1, c(\vec{x}, t))(h_2, \vec{V}_i(\vec{x}^t))] \\ &\quad - \mathbf{E} \left[\left(h_1, \frac{\partial}{\partial t} c(\vec{x}, t) \right) (h_2, \mathbf{E}[\vec{V}_i(\vec{x}^t)]) \right] \\ &= \left(h_1, \left[\mathbf{E} \left[\frac{\partial}{\partial t} c(\vec{x}, t) \vec{V}_i(\vec{x}^t) \right] - \frac{\partial}{\partial t} \mathbf{E}[c(\vec{x}, t)] \mathbf{E}[\vec{V}_i(\vec{x}^t)] \right] h_2 \right) \end{aligned}$$

Substituting expressions for $\frac{\partial}{\partial t} c(\vec{x}, t)$ and $\frac{\partial}{\partial t} \mathbf{E}[c(\vec{x}, t)]$ yields

$$\begin{aligned} \frac{\partial}{\partial t} (h_1, C_{c\vec{V}_i}(\vec{x}, \vec{x}^t, t)h_2) &= \left(h_1, \left[\mathbf{E} \left[- \left\{ \nabla_{\vec{x}} \cdot c(\vec{x}, t) (\mathbf{E}[\vec{V}(\vec{x})] + \vec{V}'(\vec{x})) \right. \right. \right. \right. \\ &\quad \left. \left. \left. - \nabla_{\vec{x}} \cdot (\mathbf{D}\nabla_{\vec{x}}c(\vec{x}, t)) \right\} \vec{V}_i(\vec{x}^t) \right] \right. \\ &\quad \left. + \left\{ \nabla_{\vec{x}} \cdot (\mathbf{E}[c(\vec{x}, t)] \mathbf{E}[\vec{V}(\vec{x})]) - \nabla_{\vec{x}} \cdot (\mathbf{D}\nabla_{\vec{x}} \mathbf{E}[c(\vec{x}, t)]) \right. \right. \\ &\quad \left. \left. + \nabla_{\vec{x}} \cdot \mathbf{E}[c'(\vec{x}, t) \vec{V}'(\vec{x}^t)] \right\} \mathbf{E}[\vec{V}_i(\vec{x}^t)] \right] h_2 \right) \end{aligned}$$

Let $c(\vec{x}, t) = \mathbf{E}[c(\vec{x}, t)] + c'(\vec{x}, t)$ and expand

$$\begin{aligned}
\frac{\partial}{\partial t} (h_1, C_{c\vec{v}_i}(\vec{x}, \vec{x}', t) h_2) &= (h_1, [\mathbf{E} [-\{\nabla_{\vec{x}} \cdot \mathbf{E}[c(\vec{x}, t)]\mathbf{E}[\vec{V}(\vec{x})] + \nabla_{\vec{x}} \cdot c'(\vec{x}, t)\mathbf{E}[\vec{V}(\vec{x})] \\
&\quad + \nabla_{\vec{x}} \cdot \mathbf{E}[c(\vec{x}, t)]\vec{V}'(\vec{x}) + \nabla_{\vec{x}} \cdot c'(\vec{x}, t)\vec{V}'(\vec{x}) \\
&\quad - \nabla_{\vec{x}} \cdot (\mathbf{D}\nabla_{\vec{x}}\mathbf{E}[c(\vec{x}, t)]) - \nabla_{\vec{x}} \cdot (\mathbf{D}\nabla_{\vec{x}}c'(\vec{x}, t))\} \vec{V}_i(\vec{x}')] \\
&\quad + \{\nabla_{\vec{x}} \cdot (\mathbf{E}[c(\vec{x}, t)]\mathbf{E}[\vec{V}(\vec{x})]) - \nabla_{\vec{x}} \cdot (\mathbf{D}\nabla_{\vec{x}}\mathbf{E}[c(\vec{x}, t)]) \\
&\quad + \nabla_{\vec{x}} \cdot \mathbf{E}[c'(\vec{x}, t)\vec{V}'(\vec{x}')]\} \mathbf{E}[\vec{V}_i(\vec{x}')]] h_2)
\end{aligned}$$

And, finally, by letting $\vec{V}_i(\vec{x}') = \mathbf{E}[\vec{V}_i(\vec{x}')] + \vec{V}'_i(\vec{x}')$ and expanding again

$$\begin{aligned}
\frac{\partial}{\partial t} (h_1, C_{c\vec{v}_i}(\vec{x}, \vec{x}', t) h_2) &= (h_1, [\mathbf{E} [-\{\nabla_{\vec{x}} \cdot \mathbf{E}[c(\vec{x}, t)]\mathbf{E}[\vec{V}(\vec{x})]\mathbf{E}[\vec{V}_i(\vec{x}')] \\
&\quad + \nabla_{\vec{x}} \cdot \mathbf{E}[c(\vec{x}, t)]\mathbf{E}[\vec{V}(\vec{x})]\vec{V}'_i(\vec{x}') \\
&\quad + \nabla_{\vec{x}} \cdot c'(\vec{x}, t)\mathbf{E}[\vec{V}(\vec{x})]\mathbf{E}[\vec{V}_i(\vec{x}')] + \nabla_{\vec{x}} \cdot c'(\vec{x}, t)\mathbf{E}[\vec{V}(\vec{x})]\vec{V}'_i(\vec{x}') \\
&\quad + \nabla_{\vec{x}} \cdot \mathbf{E}[c(\vec{x}, t)]\vec{V}'(\vec{x})\mathbf{E}[\vec{V}_i(\vec{x}')] + \nabla_{\vec{x}} \cdot \mathbf{E}[c(\vec{x}, t)]\vec{V}'(\vec{x})\vec{V}'_i(\vec{x}') \\
&\quad + \nabla_{\vec{x}} \cdot c'(\vec{x}, t)\vec{V}'(\vec{x})\mathbf{E}[\vec{V}_i(\vec{x}')] + \nabla_{\vec{x}} \cdot c'(\vec{x}, t)\vec{V}'(\vec{x})\vec{V}'_i(\vec{x}') \\
&\quad - \nabla_{\vec{x}} \cdot (\mathbf{D}\nabla_{\vec{x}}\mathbf{E}[c(\vec{x}, t)]\mathbf{E}[\vec{V}_i(\vec{x}')] - \nabla_{\vec{x}} \cdot (\mathbf{D}\nabla_{\vec{x}}\mathbf{E}[c(\vec{x}, t)]\vec{V}'_i(\vec{x}')) \\
&\quad - \nabla_{\vec{x}} \cdot (\mathbf{D}\nabla_{\vec{x}}c'(\vec{x}, t)\mathbf{E}[\vec{V}_i(\vec{x}')] - \nabla_{\vec{x}} \cdot (\mathbf{D}\nabla_{\vec{x}}c'(\vec{x}, t)\vec{V}'_i(\vec{x}')) \\
&\quad + \{\nabla_{\vec{x}} \cdot (\mathbf{E}[c(\vec{x}, t)]\mathbf{E}[\vec{V}(\vec{x})]) - \nabla_{\vec{x}} \cdot (\mathbf{D}\nabla_{\vec{x}}\mathbf{E}[c(\vec{x}, t)]) \\
&\quad + \nabla_{\vec{x}} \cdot \mathbf{E}[c'(\vec{x}, t)\vec{V}'(\vec{x}')]\} \mathbf{E}[\vec{V}_i(\vec{x}')]] h_2)
\end{aligned}$$

So, by taking expectations, cancelling terms and using the conditions $\mathbf{E}[\vec{V}'] = \mathbf{E}[c'] = 0$, the final equation for the velocity-concentration equation is

$$\begin{aligned}
\frac{\partial}{\partial t} \left(h_1, C_{c\vec{v}_i}(\vec{x}, \vec{x}', t) h_2 \right) &= \left(h_1, \nabla_{\vec{x}} \cdot \mathbf{E}[\vec{V}(\vec{x})] C_{c\vec{v}_i}(\vec{x}, \vec{x}', t) \right. \\
&+ \nabla_{\vec{x}} \cdot \mathbf{E}[c(\vec{x}, t)] C_{\vec{v}_i \vec{v}}(\vec{x}, \vec{x}', t) \\
&+ \nabla_{\vec{x}} \cdot \mathbf{E}[c'(\vec{x}, t) \vec{V}'(\vec{x}') \vec{V}_i(\vec{x}')] \\
&\left. - \nabla_{\vec{x}} \cdot (\mathbf{D} \nabla_{\vec{x}} C_{c\vec{v}_i}(\vec{x}, \vec{x}', t)) h_2 \right)
\end{aligned}$$

which agrees with the previously obtained velocity-concentration Equation[4.7], page 133.

The second example involves measurement uncertainty. This example will show how the trace term in Equation[4.17] can be used. Consider

$$\frac{\partial c}{\partial t} + \nabla \cdot (c\vec{V}) - \nabla \cdot (\mathbf{D}\nabla c) = 0 \quad (4.25)$$

Suppose that there is measurement uncertainty in the laboratory experiment, and that this uncertainty is random. Then, in the laboratory, the experimenters will record the results \bar{c} , and the variables u and \bar{c} will be related by

$$\bar{c}(\vec{x}, t, \omega) = c(\vec{x}, t) + \epsilon(t, \omega) \quad (4.26)$$

Using Equation[4.26] in Equation[4.25],

$$\frac{\partial(\bar{c}(\vec{x}, t, \omega) - \epsilon(t, \omega))}{\partial t} + \nabla \cdot (\bar{c}(\vec{x}, t, \omega)\vec{V}) - \nabla \cdot (\mathbf{D}\nabla\bar{c}(\vec{x}, t, \omega)) = 0 \quad (4.27)$$

The derivatives in Equation[4.27] have to be interpreted in the mean-square sense, ie, if $x(t, \omega)$ is a random function, then

$$\begin{aligned}
\dot{x}(t, \omega) &= (m^2) \lim_{h \rightarrow 0} \frac{x(t+h, \omega) - x(t, \omega)}{h} \\
&\Rightarrow \lim_{h \rightarrow 0} E \left[\left| \frac{x(t+h, \omega) - x(t, \omega)}{h} - \dot{x}(t, \omega) \right|^2 \right] \\
&= \lim_{h \rightarrow 0} \int_{\Omega} \left| \frac{x(t+h, \omega) - x(t, \omega)}{h} - \dot{x}(t, \omega) \right|^2 dP
\end{aligned}$$

And, from the inequality

$$|f + g|^2 \leq 2^2(|f|^2 + |g|^2)$$

it follows that

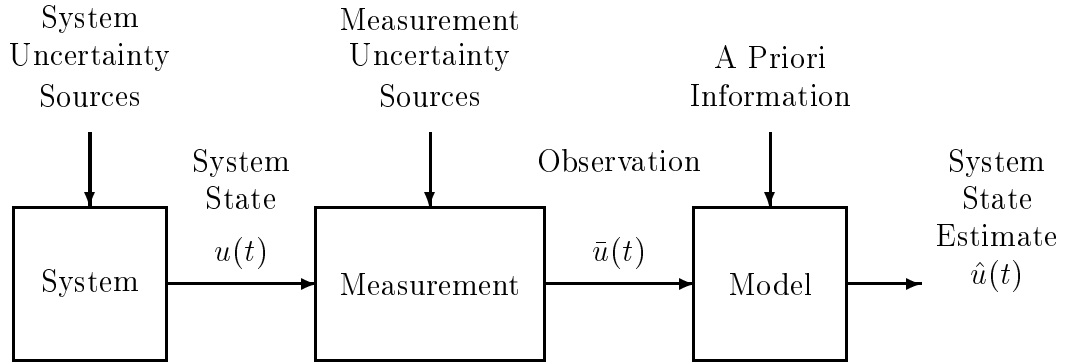
$$\frac{d}{dt}(f + g) = \frac{d}{dt}f + \frac{d}{dt}g$$

Using this, Equation[4.27] can be written as

$$\frac{\partial \bar{c}(\vec{x}, t, \omega)}{\partial t} + \nabla \cdot (\bar{c}(\vec{x}, t, \omega) \vec{V}) - \nabla \cdot (\mathbf{D} \nabla \bar{c}(\vec{x}, t, \omega)) = \frac{d\epsilon(t, \omega)}{dt}$$

where $\frac{d\epsilon}{dt}$ is a stochastic process. Hence, the introduction of measurement uncertainty is equivalent to applying a stochastic forcing term to the equation.

The situation can be characterized by the following diagram similar to one found in Gelb[43]:

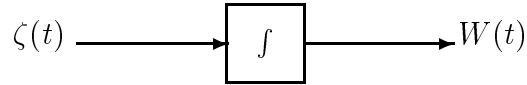


In this diagram, the System Uncertainty Sources are represented by any uncertainty that may exist in the specification of \mathbf{D} and \vec{V} , the Measurement Uncertainty Sources are represented by ϵ .

The random forcing term is assumed to be a Gaussian white noise process. This is equivalent to assuming that the process, $\epsilon(t, \omega)$, is a Wiener process which can be defined as the limit of a *random walk*, or as the integral of a Gaussian white noise process with zero mean.

$$W(t) = \int_0^t \zeta(s) ds$$

The following is a block diagram representation of this equation:



$$E[\zeta(t)\zeta(\tau)] = q(t)\delta(t - \tau)$$

A second feature of the Boulder experiments that must be modeled is the *pulsed input* feature of the experiment. This means that in the tracer experiment, the tracer, benzene, is injected at a rate of 5 ml/min for a period of 4 hours and then the injection pump is turned off. However, samples are taken for a period of 8 hours. This means that at a specified measuring point, the sampling device will see the concentrations of benzene first increase, then level off, and finally decrease to zero.

In the finite element model, this allowed for by imposing a non-zero boundary condition at the origin for a specified number of time steps, and then imposing a zero boundary condition at the origin for the remainder of the time steps of the simulation. The following is a segment of code that performs this task:

```

/*****
/*      Impose The Left Hand Boundary Condition      */
/*****
    if (nt <= bctimesteps)
        impose_bndy_cond();
    else {
        lbdy = 0.0;
        impose_bndy_cond();
    }
  
```

Here, the variable *lbdy* is originally input to the program with a non-zero value. Once the specified number of timesteps for injection of the tracer, *bctimesteps*, has passed, *lbdy* is set to zero and the boundary condition function imposes a zero boundary condition on each succeeding time step.

The two graphs on the next page illustrate the output from the finite

element program with a pulsed input. The surface shown in Figure 19 is a space-time representation of the concentration. Figure 20 shows a time-slice of this surface. This curve has the same shape as the actual measurements when they are plotted.

Figure 21 entitled *Comparison Time Profile* shows the Pulsed Input Time Profile with and without the effects of a random forcing term (measurement uncertainty). The dotted line represents the time profile without measurement uncertainty, and the solid line shows the time profile with measurement uncertainty taken into consideration.

Returning to the equation

$$\frac{\partial \bar{c}(\vec{x}, t, \omega)}{\partial t} + \nabla \cdot (\bar{c}(\vec{x}, t, \omega) \vec{V}) - \nabla \cdot (\mathbf{D} \nabla \bar{c}(\vec{x}, t, \omega)) = \frac{d\epsilon(t, \omega)}{dt}$$

For the sake of simplicity, assume that the parameters \mathbf{D} and \vec{V} are deterministic. This means that there will be no need of a velocity-concentration covariance equation as in the previous example. The equation for the expected value of the concentration takes the form

$$\frac{\partial \mathbf{E}[\bar{c}(\vec{x}, t)]}{\partial t} + \nabla \cdot (\mathbf{E}[\bar{c}(\vec{x}, t)] \vec{V}(\vec{x})) - \nabla \cdot (\mathbf{D} \nabla \mathbf{E}[\bar{c}(\vec{x}, t)]) = 0$$

For the equation of the concentration covariance, let \vec{x} and \vec{x}' be two different coordinate systems, then from Equation[4.17], page 158,

$$\begin{aligned} \frac{\partial}{\partial t} (h_1, \bar{c}(\vec{x}, t))(h_2, \bar{c}(\vec{x}', t)) &= - \left[(h_1, \bar{c}(\vec{x}, t))(h_2, A_{\vec{x}'} \bar{c}(\vec{x}', t)) \right. \\ &+ \left. (h_1, \bar{c}(\vec{x}', t))(h_2, A_{\vec{x}} \bar{c}(\vec{x}, t)) \right] \\ &+ \frac{1}{2} tr [(h_2 \circ h_1) \oplus (h_1 \circ h_2) Q(t)] \end{aligned}$$

where $\circ = I$. And, by proceeding as in the previous example,

$$\begin{aligned} \frac{\partial}{\partial t} (h_1, C_{\bar{c}\bar{c}}(\vec{x}, \vec{x}', t) h_2) &= - \left(h_1, \left[\nabla_{\vec{x}} \cdot \mathbf{E}[\vec{V}(\vec{x})] C_{\bar{c}\bar{c}}(\vec{x}, \vec{x}', t) \right. \right. \\ &- \left. \left. \nabla_{\vec{x}} \cdot (\mathbf{D} \nabla_{\vec{x}} C_{\bar{c}\bar{c}}(\vec{x}, \vec{x}', t)) + \nabla_{\vec{x}'} \cdot \mathbf{E}[\vec{V}(\vec{x}')] C_{\bar{c}\bar{c}}(\vec{x}', \vec{x}, t) \right] \right) \end{aligned}$$

Figure 19 - 1D Pulsed Input Over Time

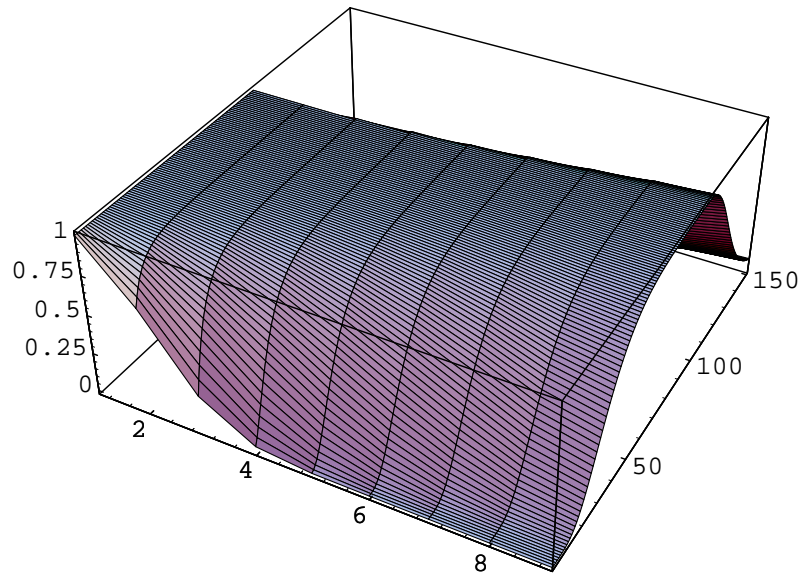


Figure 20 - Pulsed Input Time Slice

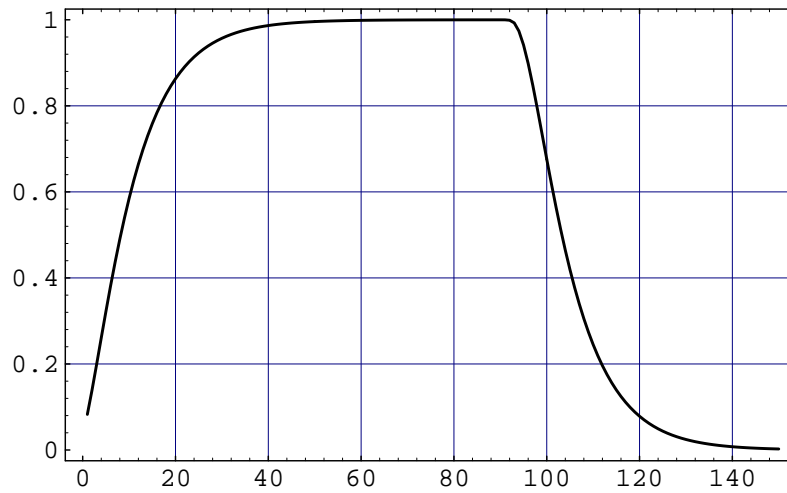
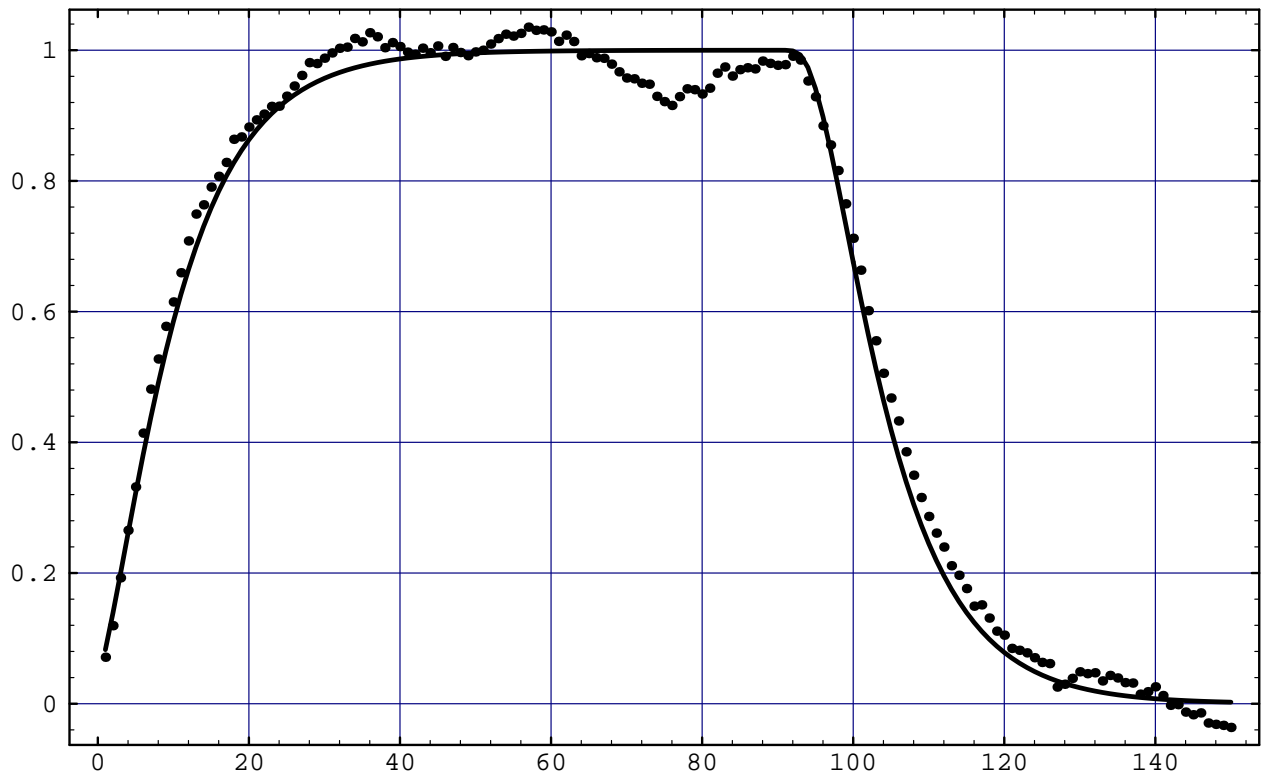


Figure 21 - Comparison Time Profile



$$\begin{aligned}
& - \nabla_{\vec{x}'} \cdot (\mathbf{D}\nabla_{\vec{x}'} C_{\bar{c}\bar{c}}(\vec{x}', \vec{x}, t)) h_2) \\
& + \frac{1}{2} tr [(h_2 \circ h_1) \oplus (h_1 \circ h_2) Q(t)]
\end{aligned}$$

Finally, substituting for the tr term from Equation[4.13], page 149, it follows that

$$\begin{aligned}
\frac{\partial}{\partial t}(h_1, C_{\bar{c}\bar{c}}(\vec{x}, \vec{x}', t)h_2) & = - (h_1, [\nabla_{\vec{x}} \cdot \mathbf{E}[\vec{V}(\vec{x})]C_{\bar{c}\bar{c}}(\vec{x}, \vec{x}', t) \\
& - \nabla_{\vec{x}} \cdot (\mathbf{D}\nabla_{\vec{x}} C_{\bar{c}\bar{c}}(\vec{x}, \vec{x}', t) + \nabla_{\vec{x}'} \cdot \mathbf{E}[\vec{V}(\vec{x}')C_{\bar{c}\bar{c}}(\vec{x}', \vec{x}, t)] \\
& - \nabla_{\vec{x}'} \cdot (\mathbf{D}\nabla_{\vec{x}'} C_{\bar{c}\bar{c}}(\vec{x}', \vec{x}, t)) - Q] h_2)
\end{aligned}$$

4.3 Summary

Chapter 4 actually starts the second part of the thesis. The previous sections have investigated the components of the equations and the forms of the equations. However, only the expected or mean value of the concentration is predicted. Because of the uncertainties involved in specifying the physical characteristics of the porous medium, the concentration of a solute at a given point in time is a random variable, and over a period of time it is a stochastic process. Consequently, in order to more accurately characterize the distribution of the solute concentration, higher order statistical moments such as the variance need to be estimated also. In theory, the more moments that can be predicted, the better this characterization will be. But, in practice, it is usually a difficult problem just to obtain information on the variance or covariance of variables in the system. A much referenced paper in this area is the Graham and McGlaughlin[48] paper which specifies a set of three equations that are to be solved for the mean concentration, the velocity-concentration covariance and the concentration covariance. These equations were presented in Section 4.1 for the purpose of comparison with mean and covariance equations derived from other methods.

Randomness can enter the boundary value problem in many different ways. Equation[4.12], page 138, is a statement of the stochastic boundary value problem, and the discussion following that equation specifies the various ways in which randomness can enter the picture. Existence theory for the

stochastic boundary value problem was covered in Section 4.2.3 and found to be not unlike the nonstochastic case.

Stochastic integration is again addressed in Section 4.2.4, this time from the more general perspective of a *martingale*. The Itô integral then follows from this more general definition as a special case. The use of the Itô integral requires that the rules of calculus have to be modified. The most important new rule is that of Itô's lemma. It is a change of variable formula. The reason the change of variable formula has to be modified is due to changes in differential relationships that were covered in Section 1.1. The Itô formula is a stochastic calculus chain-rule. It can also be extended to martingale type processes, Karatzas[57]. Curtain and Falb[26] have extended Itô's lemma to infinite dimensional Hilbert spaces. It is this form that is used to derive weak forms of the moment equations in Section 4.2.8. For the purpose of illustrating this theory, the key equation is Equation[4.17], page 158, which is applied to two examples. The first example uses this theory to derive mean and covariance equations that in the weak form are identical to those used by Graham and McLaughlin[48]. The second example is cast in terms of accounting for the effects of measurement error that is assumed to enter the experiment as a random perturbation that takes the form of a Wiener process.

5. Stochastic Evolution Equations

5.1 General Theoretical Foundations

As was observed in Chapter 2, the dispersion tensor can be considered a time stochastic process. That result along with the stochastic nature of the velocity field allows the transport equation to be written in a form that allows the separation of the deterministic components from the stochastic components. That is, the stochastic PDE that represents the transport equation can be separated into a sum of a deterministic operator and a stochastic operator as the following 1-D example shows:

$$\begin{aligned} \frac{\partial u}{\partial t} - \mathbf{E}[D(t)] \frac{\partial^2 u}{\partial x^2} + \mathbf{E}[V] \frac{\partial u}{\partial x} &= D'(t, \omega) \frac{\partial^2 u}{\partial x^2} - V'(\omega) \frac{\partial u}{\partial x} \\ u(x, 0) = u_0 & \quad u(\infty, t) = 0 \end{aligned}$$

The left hand side of this equation is the standard form of the transport equation, while the random components associated with dispersion and velocity have been moved to the right hand side in the form of a random operator. It is well known that there is a correspondence between the Cauchy problem and the abstract boundary value problem similar to the left hand side of the equation stated above. The Cauchy problem can be stated as:

Let H be a Hilbert space, $\mathcal{D}(A)$ a subspace of H and let the operator A an unbounded, linear operator from $\mathcal{D}(A)$ to H . Then let

$$\begin{aligned} u'(t) + Au(t) &= f(t) \\ u(0) &= u_0 \end{aligned} \tag{5.1}$$

and the Cauchy problem is to find a function $u(t)$ such that for $t > 0$, $u(t) \in \mathcal{D}(A)$ and satisfies Equation[5.1], so that $u(t)$ is an H -valued function.

Hence, there must exist a correspondence between the H -valued function $u(t)$ and the solution, $u(\vec{x}, t)$ to the boundary value problem. And, a corresponding relationship between the derivative $u'(t)$ and the partial derivative

$\frac{\partial}{\partial t}u(\vec{x}, t)$. Also, the operator A must be related to the boundary value problem.

Following Showalter[91], let $I = [a, b]$ be a closed, bounded interval in \mathfrak{R} and let Ω be a bounded, measurable subset of \mathfrak{R}^n . Let $u(t) \in C[I, L^2(\Omega)]$. Let $a = t_0 < t_1 < \dots < t_n = b$ be a *uniform* partition of I such that

$$u_n(\vec{x}, t) = \begin{cases} u_0(\vec{x}, t_k) & t_k \leq t < t_{k+1}, \quad k = 0, 1, \dots, n-1 \\ u_0(\vec{x}, b) & t = t_n \end{cases}$$

Here $u(t_i) = u_0(\vec{x}, t_i) \in \mathcal{D}(A)$ is a representation of $u(t)$. Then

$$u_n(\vec{x}, t) : \Omega \times I \rightarrow \mathfrak{R}$$

Since each $u_0(\vec{x}, t_k) \in L^2(\Omega)$, for $t \in I$, the t -section of the function $u_n(\vec{x}, t)$ is given by

$$u_n^t(\vec{x}) = u_n(\vec{x}, t) \quad \vec{x} \in \Omega$$

And, since the t -section $u_n^t(\vec{x}) \in L^2(\Omega)$, it is measurable. And if $\vec{x} \in \Omega$ is fixed we get a step function. So, the t -section defines the following mapping:

$$u_n^{t_k}(\vec{x}) : [t_k, t_{k+1}) \times \Omega \rightarrow \mathfrak{R}$$

If $\alpha \in \mathfrak{R}$, then the set

$$P = \{(\vec{x}, t) : u_n^{t_k}(\vec{x}) > \alpha\}$$

is either empty or of the form

$$E \times [t_k, t_{k+1})$$

where E is a measurable subset of Ω . In either case, P is a measurable subset of $\Omega \times I$. And, since $u_n(\vec{x}, t)$ is a finite sum of these functions, it is a measurable function on the product measure space $\Omega \times I$.

Since the partition of I is uniform and since $u(t)$ is uniformly continuous on I , then for $\epsilon > 0 \exists \delta_\epsilon > 0$ such that if the partition n is large enough, then

$$|t_k - t_{k+1}| < \delta_\epsilon \quad k = 0, 1, \dots, n-1$$

Hence, if $t \in I$, then $\exists k$ such that $t \in [t_k, t_{k+1})$ and so

$$|t - t_k| \leq |t_k - t_{k+1}| < \delta_\epsilon$$

$$\Rightarrow \|u_n(\vec{x}, t) - u(t)\|_{0,\Omega} = \|u_0(\vec{x}, t_k) - u(t)\|_{0,\Omega} < \epsilon$$

This means that

$$\lim_{n \rightarrow \infty} \|u_n(\vec{x}, t) - u(t)\|_{0,\Omega} = 0$$

uniformly on I . Hence, using Fubini's theorem,

$$\lim_{n \rightarrow \infty} \|u_n(\vec{x}, t) - u(t)\|_{L^2(\Omega \times I)}^2 = \lim_{n \rightarrow \infty} \int_I \|u_n(\vec{x}, t) - u(t)\|_{0,\Omega}^2 dt = 0$$

Clarkson's inequality states that if $f, g \in L^2(\Omega)$, then

$$\left\| \frac{f+g}{2} \right\|_{0,\Omega}^2 + \left\| \frac{f-g}{2} \right\|_{0,\Omega}^2 \leq \frac{1}{2} \|f\|_{0,\Omega}^2 + \frac{1}{2} \|g\|_{0,\Omega}^2$$

Using this, it follows that

$$\|u_m(\vec{x}, t) - u_n(\vec{x}, t)\|_{0,\Omega}^2 \leq 2 \left(\|u_m(\vec{x}, t) - u(t)\|_{0,\Omega}^2 + \|u_n(\vec{x}, t) - u(t)\|_{0,\Omega}^2 \right)$$

And, since this limit is uniform on I , it follows that

$$\lim_{n,m \rightarrow \infty} \int_I \|u_m(\vec{x}, t) - u_n(\vec{x}, t)\|_{0,\Omega}^2 dt = \lim_{n,m \rightarrow \infty} \|u_m(\vec{x}, t) - u_n(\vec{x}, t)\|_{L^2(\Omega \times I)}^2 = 0$$

So, the sequence $\{u_n(\vec{x}, t)\}$ is a Cauchy sequence in $L^2(\Omega \times I)$, and since $L^2(\Omega \times I)$ is complete, $\exists u(\vec{x}, t) \in L^2(\Omega \times I)$ such that

$$\lim_{n \rightarrow \infty} \|u_n(\vec{x}, t) - u(\vec{x}, t)\|_{L^2(\Omega \times I)}^2 = 0$$

Hence, for $\epsilon > 0$,

$$\|u(\vec{x}, t) - u(t)\|_{L^2(\Omega \times I)} \leq \|u(\vec{x}, t) - u_n(\vec{x}, t)\|_{L^2(\Omega \times I)} + \|u_n(\vec{x}, t) - u(t)\|_{L^2(\Omega \times I)}$$

for n large enough. This implies then that

$$\|u(\vec{x}, t) - u(t)\|_{L^2(\Omega \times I)} = 0$$

Hence, $u(\vec{x}, t) = u(t)$ a.e. on I . And by changing $u(t)$ to $u_0(\vec{x}, t)$ on a set of measure zero, the correspondence between $u(t)$ and $u(\vec{x}, t)$ is established.

To show the relationship between $u'(t)$ and $\frac{\partial}{\partial t}u(\vec{x}, t)$, let $\phi \in C_0^\infty[I, L^2(\Omega)]$ and let $\Phi(\vec{x}, t) \in C_0^\infty[\Omega \times I]$ be a representation of $\phi(t)$. Also, let $u \in C^1[I, L^2(\Omega)]$ and let $v(\vec{x}, t)$ be a representation of $u'(t)$ for almost all $t \in I$. Then, integrating by parts,

$$\begin{aligned} \int_I \int_\Omega u(t)\phi'(t)d\Omega dt &= \int_I \left[\int_{\partial\Omega} u(t)\phi(t)d\Omega - \int_\Omega u'(t)\phi(t)d\Omega \right] dt \\ &= - \int_I \int_\Omega u'(t)\phi(t)d\Omega dt \end{aligned}$$

Then, using the representations, it follows that

$$- \int_I \int_\Omega u(\vec{x}, t)\frac{\partial}{\partial t}\Phi(\vec{x}, t)d\Omega dt = \int_I \int_\Omega v(\vec{x}, t)\Phi(\vec{x}, t)d\Omega dt$$

And, this means that in the weak or distributional sense that

$$u'(t) = v(\vec{x}, t) = \frac{\partial}{\partial t}u(\vec{x}, t)$$

In order to show the relationship of the operator A to the boundary value problem, suppose that V, H , and B are Hilbert spaces, that γ is a linear surjection of V onto B with kernel V_0 , that $\bar{V}_0 = H$, that i is a continuous injection of V_0 into H , that H is a pivot space, and that H is identified with its dual, $H = H'$, then the following diagram can be established

$$\begin{array}{ccccc} V_0 & \xrightarrow{i} & H & & \\ Z_{V_0} & \downarrow & & \downarrow & Z_H \\ V_0' & \xleftarrow{i'} & H' & & \end{array}$$

This means that the following embeddings exist:

$$V_0 \hookrightarrow H = H' \hookrightarrow V_0'$$

Similarly, suppose that $\bar{V} = H$, that i is a continuous injection of V into H , that H is a pivot space, and that H is identified with its dual, $H = H'$, then the following diagram can be established

$$\begin{array}{ccccc}
& V & \xrightarrow{i} & H & \\
Z_V & \updownarrow & & \updownarrow & Z_H \\
& V' & \xleftarrow{i'} & H' &
\end{array}$$

This means that the following embeddings exist:

$$V \hookrightarrow H = H' \hookrightarrow V'$$

Hahn-Banach Theorem: Let X be a normed linear space, M a linear subspace of X and f_1 a continuous linear functional on M . Then there exists a continuous linear functional f defined on X such that

- i $f_1(u) = f(u) \quad \forall u \in M$
- ii $\|f_1\| = \|f\|$ □

If $v \in V$ and $v' \in V'$, then

$$v'(v) = \langle v', v \rangle_V$$

and since V is a linear subspace of H , there is an $h' \in H'$ such that h' is an extension of v' . Hence, the duality pairing on $V' \times V$ can be identified with the duality pairing on H . By the Riesz Representation Theorem, $\exists v_{h'} \in H$ such that

$$\langle h', h \rangle_H = (v_{h'}, h)_H$$

Hence, the duality pairing on $V' \times V$ can be identified with the inner product on H .

The trace operator γ maps the space V onto the Hilbert space B . For example, we might have

$$H = L^2(\Omega); \quad V = H^1(\Omega); \quad V_0 = H_0^1(\Omega); \quad B = H^{\frac{1}{2}}(\partial\Omega)$$

Suppose there is a continuous bilinear form $a_1 : V \times V \rightarrow \mathfrak{R}$. If $(v_1, v_2) \in V \times V$ and v_1 is considered fixed and $v_2 \in V_0$, then $a_1(v_1, v_2)$ is a continuous linear operator on V_0 . So, for each $v_1 \in V$ we can write

$$a_1(v_1, v_2) = A_{v_1}(v_2) \quad v_2 \in V_0$$

The linear functional A_{v_1} depends linearly and continuously on v_1 . This dependence is given formally by

$$Av_1 = A_{v_1}$$

Hence,

$$a_1(v_1, v_2) = \langle Av_1, v_2 \rangle_V \quad v_2 \in V_0$$

And, the operator A is a continuous linear operator from V to V_0' , ie,

$$A \in \mathcal{L}[V, V_0']$$

Let

$$\mathcal{V}_A = \{v_1 \in V : Av_1 \in H\} \quad (H = H' \hookrightarrow V_0')$$

Note: The reason that \mathcal{V}_A is required can be seen by the following example: Let $V = H^1(\Omega)$, then by the trace theorem the operator γ_0 can be extended by continuity to a mapping of $H^1(\Omega)$ onto $H^{\frac{1}{2}}(\partial(\Omega))$, but it says nothing about the mapping γ_1 . In fact, γ_1 cannot be extended to all of $H^1(\Omega)$. A smaller space is required, and \mathcal{V}_A is that space.

Since

$$Av_1 \in H = H' \hookrightarrow V_0' \quad \text{for } v_1 \in \mathcal{V}_A$$

we can write

$$a_1(v_1, v_2) = \langle Av_1, v_2 \rangle_V = (Av_1, v_2)_H \quad \forall v_2 \in V_0$$

Now, let $v_2 \in V$ and define the operator

$$(G_A v_1, v_2)_H = a_1(v_1, v_2) - (Av_1, v_2)_H \quad v_1 \in \mathcal{V}_A, \quad v_2 \in V$$

If $v_2 \in V_0$, it follows that

$$(G_A v_1, v_2)_H = 0$$

and this means that

$$G_A : \mathcal{V}_A \rightarrow V_0^- \subset V_0'$$

The trace operator γ is a continuous linear operator that maps the space V onto the boundary value space B . At this point the following theorem can be applied:

Theorem: Let X, Y be B-spaces, and let $T \in \mathcal{L}(X, Y)$, and $T' \in \mathcal{L}(Y', X')$ its transpose, then the following hold:

- i $\mathcal{N}(T') = \mathcal{R}(T)^-$
- ii $\mathcal{R}(T)$ is dense in Y iff T' is 1-1
- iii $\mathcal{R}(T)$ is closed in Y iff $\mathcal{R}(T')$ is norm-closed in X'
- iv If $\mathcal{R}(T)$ is closed, then $\mathcal{R}(T') = \mathcal{N}(T)^-$ \square

By letting $X = V, Y = B, T = \gamma, T' = \gamma'$, and $\gamma : V$ onto B it follows from part (iii) of the Theorem that since B is closed the transpose γ' maps B' onto $\overline{\mathcal{R}(\gamma')}$. Furthermore, by part (ii) of the Theorem, since $\mathcal{R}(\gamma) = B$, γ' is 1-1. Hence, γ' is an isomorphism of B' onto $\overline{\mathcal{R}(\gamma')}$. Since the $\mathcal{N}(\gamma) = V_0$ and since $\mathcal{R}(\gamma)$ is closed, part (iv) of the theorem gives

$$\mathcal{R}(\gamma') = V_0^-$$

This means that

$$(\gamma')^{-1} : V_0^- \xrightarrow{\text{onto}} B'$$

Define the function

$$\delta = (\gamma')^{-1} G_A : \mathcal{V}_A \rightarrow B'$$

Then,

$$G_A v_1 = (\gamma') (\gamma')^{-1} G_A v_1 = \gamma' \delta v_1 \quad \forall v_1 \in \mathcal{V}_A$$

Hence, if $G_A v_1 \in H \hookrightarrow V_0'$, then

$$\begin{aligned} (G_A v_1, v_2)_H &= (\gamma' \delta v_1, v_2)_H = (\delta v_1, \gamma v_2)_H \\ &= \langle \delta v_1, \gamma v_2 \rangle_B \end{aligned}$$

since $\delta v_1 \in B'$ and $\gamma v_2 \in B$.

So, for $v_1 \in \mathcal{V}_A$ it follows that from the definition of $(G_A v_1, v_2)_H$ that

$$a_1(v_1, v_2) = (A v_1, v_1)_H + \langle \delta v_1, \gamma v_2 \rangle_B \quad (5.2)$$

This equation has the general form of a Green's formula. Also, by specifying the bilinear form $a_1(v_1, v_2)$ the operator A can be derived.

For example, given the equation

$$\frac{\partial u}{\partial t} + \mathcal{A}(\vec{x})u = 0$$

where

$$\mathcal{A}(\vec{x}) = - \sum_{i,j=1}^2 \frac{\partial}{\partial x_i} a_{ij}(\vec{x}) \frac{\partial}{\partial x_j} + \sum_{j=1}^2 b_j(\vec{x}) \frac{\partial}{\partial x_j} + c(\vec{x})$$

then if $H = L^2(\Omega)$, $a_1(v_1, v_2)$ can be taken to be

$$a_1(v_1, v_2) = \int_{\Omega} \left\{ \sum_{i,j=1}^2 a_{ij}(\vec{x}) \frac{\partial}{\partial x_j} v_1 \frac{\partial}{\partial x_i} v_2 + \sum_{j=1}^2 b_j(\vec{x}) \frac{\partial}{\partial x_j} v_1 v_2 + c(\vec{x}) v_1 v_2 \right\} d\Omega$$

Then $a_1(v_1, v_2)$ can be written as, with $\mathbf{A} = \begin{bmatrix} a_{11}(\vec{x}) & a_{12}(\vec{x}) \\ a_{21}(\vec{x}) & a_{22}(\vec{x}) \end{bmatrix}$,

$$\begin{aligned} a_1(v_1, v_2) &= \int_{\Omega} \mathbf{A}(\vec{x}) \nabla v_1 \cdot \nabla v_2 d\Omega + \int_{\Omega} v_2 \vec{b}(\vec{x}) \nabla v_1 d\Omega \\ &+ \int_{\Omega} c(\vec{x}) v_1 v_2 d\Omega \end{aligned}$$

And, if the first term on the right-hand side is integrated by parts it follows that

$$\begin{aligned}
a_1(v_1, v_2) &= \int_{\Omega} \left(-\nabla \cdot (\mathbf{A}(\vec{x})\nabla v_1) + \vec{b}(\vec{x}) \cdot \nabla v_1 + c(\vec{x})v_1 \right) v_2 d\Omega \\
&+ \int_{\partial\Omega} v_2 \mathbf{A}(\vec{x})\nabla v_1 \cdot \vec{n} d\sigma
\end{aligned}$$

By comparing this equation to Equation[5.2], and by making the following associations,

$$\begin{aligned}
\delta v_1 &= \mathbf{A}(\vec{x})\nabla v_1 \cdot \vec{n}|_{\partial\Omega} & \gamma v_2 &= v_2|_{\partial\Omega} \\
Av_1 &= \left(-\nabla \cdot (\mathbf{A}(\vec{x})\nabla v_1) + \vec{b}(\vec{x}) \cdot \nabla v_1 + c(\vec{x}) \right) v_1
\end{aligned}$$

So, that

$$A = \mathcal{A}$$

It is known that if the operator A generates a strongly continuous semigroup, \mathcal{T}_t , then the *mild* solution of the Cauchy problem is given by

$$u(t) = \mathcal{T}_t u_0 + \int_0^t \mathcal{T}_{t-s} f(s) ds \quad (5.3)$$

Pazy[74] defines a *strong solution* as a function u which is differentiable a.e. on $[0, T]$ such that $u' \in L^1(0, T : H)$, $u(0) = u_0$ and $u'(t) = Au(t) + f(t)$ a.e. on $[0, T]$. Furthermore, Pazy[74] Corollary 2.10 and Corollary 2.11, since H is reflexive, if f is Lipschitz continuous on $[0, T]$, then the Cauchy problem has a unique strong solution given by Equation[5.3].

In the above discussion, the operator \mathcal{A} does not depend on the time variable t . In our case, the situation is more complicated because the operator \mathcal{A} not only depends on t , e.g. is temporally inhomogeneous, but also has a random component ω .

As far as the temporally inhomogeneous case is concerned, the fundamental solution, as characterized in Tanabe[94], is an evolution operator $U(t, s)$ which has the following properties:

- (1) $U(t, s)$ is a strongly continuous function, defined on $0 \leq s \leq t \leq T$, and is bounded
- (2) $U(t, r)U(r, s) = U(t, s)$ for $0 \leq s \leq r \leq t \leq T$

$$(3) \quad U(s, s) = I \text{ for each } s \in [0, T]$$

$$(4) \quad \frac{\partial}{\partial t} U(t, s) = A(t)U(t, s)$$

$$(5) \quad \frac{\partial}{\partial s} U(t, s) = -U(t, s)A(s)$$

And, the solution to the initial value problem

$$\frac{du(t)}{dt} = A(t)u(t) + f(t) \quad 0 \leq t \leq T$$

$$u(0) = u_0$$

can be written as

$$u(t) = U(t, 0)u_0 + \int_0^t U(t, s)f(s)ds$$

In Curtain and Falb[27], the authors extend this result to evolution equations of the form

$$du + A(t)u(t)dt = \Phi dW(t)$$

$$u(0) = u_0$$

where for a separable Hilbert space H and a Hilbert space K , $A(t)$ is a closed, possibly unbounded linear operator on K , $\Phi(\cdot, \cdot) \in \mu_2(H, K)$ where

$$\mu_2(H, K) = \left\{ S(t, \omega) : S(t, \omega) \in \mathcal{L}(H, K) \text{ and } \int_0^T \|S(t, \omega)\|_{\mathcal{L}(H, K)}^2 dt < \infty \text{ wp } 1 \right\}$$

$W(t)$ is an H -valued Wiener process and u_0 is a K -valued random variable. Theorem 3.6 of Curtain and Falb[27] gives the solution as

$$u(t) = U(t, 0)u_0 + \int_0^t U(t, s)\Phi(s)dW(s)$$

where $U(t, s)$ is an evolution operator generated by $-A(t)$. Letting $\Phi(s) = I$, the identity operator,

$$u(t) = U(t, 0)u_0 + \int_0^t U(t, s)dW(s)$$

From the definition of the stochastic integral,

$$\mathbf{E} \left[\int_0^t U(t, s)dW(s) \right] = 0$$

Consequently, $\mathbf{E}[u(t)] = U(t, 0)\mathbf{E}[u_0]$, because $U(t, 0)$ is deterministic and u_0 is a random variable. From Section 4.2.6,

$$\text{Cov}[u(t), u(t)] = \mathbf{E}[u(t) \circ u'(t)] - \mathbf{E}[u(t)] \circ \mathbf{E}[u'(t)]$$

and, in the weak sense,

$$\begin{aligned} (h_1, \text{Cov}[u(t), u(t)]h_2) &= (h_1, \mathbf{E}[u(t) \circ u(t)]h_2) \\ &\quad - (h_1, \mathbf{E}[u(t)] \circ \mathbf{E}[u(t)]h_2) \end{aligned} \tag{5.4}$$

Expanding the first term in the right hand side of Equation[5.4],

$$\begin{aligned} (h_1, \mathbf{E}[u(t) \circ u(t)]h_2) &= (h_1, \mathbf{E}[u(t)]) (h_2, \mathbf{E}[u(t)]) \\ &= (h_1, U(t, 0)\mathbf{E}[u_0]) (h_2, U(t, 0)\mathbf{E}[u_0]) \\ &= \mathbf{E} [(h_1, U(t, 0)\mathbf{E}[u_0]) (U^*(t, 0)h_2, u_0)] \\ &= (h_1, U(t, 0)\mathbf{E}[u_0]\mathbf{E}[u_0]u^*(t, 0)h_2) \end{aligned}$$

Assuming nonanticipativeness of u_0 with $W(s)$ for $s > 0$, the second term on the right hand side of Equation[5.4] is

$$\begin{aligned} (h_1, \mathbf{E}[u(t) \circ u(t)]h_2) &= \mathbf{E} \left[\left(h_1, U(t, 0)u_0 + \int_0^t U(t, s)dW(s) \right) \right. \\ &\quad \left. \times \left(h_2, U(t, 0)u_0 + \int_0^t U(t, s)dW(s) \right) \right] \\ &= \mathbf{E} [(h_1, U(t, 0)u_0)(h_2, U(t, 0)u_0) \\ &\quad + \left(h_1, \int_0^t U(t, s)dW(s) \right) \left(h_2, \int_0^t U(t, s)dW(s) \right)] \\ &= (h_1, U(t, 0)\mathbf{E}[u_0u_0]U^*(t, 0)h_2) \\ &\quad + \left(h_1, \mathbf{E} \left[\int_0^t U(t, s)dW(s) \circ \int_0^t U(t, s)dW(s) \right] h_2 \right) \end{aligned}$$

But, from Sawaragi[85], Lemma 2.3,

$$\mathbf{E} \left[\int_0^t U(t, s) dW(s) \circ \int_0^t U(t, s) dW(s) \right] = \int_0^t U(t, s) Q(t) U^*(t, s) ds$$

where $Q(t)$ is the covariance operator associated with the Wiener process $W(t)$. Hence, in the weak sense of Equation[5.4]

$$(h_1, \text{Cov}[u(t), u(t)]h_2) = \left(h_1, \left[U(t, 0) \text{Cov}[u_0, u_0] U^*(t, 0) + \int_0^t U(t, s) Q(t) U^*(t, s) ds \right] h_2 \right)$$

which is in agreement with results from the finite dimensional case, Åström[8]. It should be noted that in these cases, the operator $A(t)$, although allowed to depend on t , is not allowed to have random components. Hence, this theory would apply to the situation where the boundary value problem is allowed to have a random forcing term.

5.2 Application To Transport And Scale-Up

As presented in Section 1.7, the spectral methods used to treat scale-up resulted in analytical expressions of a scaled-up dispersivity tensor. And, as commented on in Section 1.8, these scaled-up dispersivity tensors were all subject to the constraint that the medium involved was assumed to be only mildly heterogeneous. Another way of saying this is that the variance of the log-hydraulic conductivities, Y , is subject to the condition $\sigma_Y^2 < 1$. This assumption allowed linearization techniques to be used in the development of the scaled-up dispersivity tensors.

The approach that makes the most sense in describing dispersion is the Lagrangian framework used by Dagan[29, 30, 31, 32, 33, 33]. In this approach, transport is developed in terms of indivisible solute particles which are convected by the fluid. The second spatial moment about the centroid of the plume, \mathbf{S}_{ij} , characterizes how the solute plume is dispersed about the centroid. The actual dispersivity coefficients are defined as half the time rate of change of the second spatial moment about the centroid,

$$\mathbf{S}_{ij}(t) = \frac{1}{M} \int_{\Omega} n(x_i - R_i)(x_j - R_j) C(\vec{x}, t) d\vec{x} \quad i, j = 1, 2, 3$$

$$\mathbf{D}_{ij}(t) = \frac{1}{2} \frac{d\mathbf{S}_{ij}(t)}{dt}$$

The validity of this definition was verified for a special case in Chapter 2. There are two sources of uncertainty in the expression for \mathbf{S}_{ij} , the exact position of the centroid of the plume, $\vec{R}(t)$, and the exact level of solute concentration, $C(\vec{x}, t)$. Since the concentration has a random component to it, the second spatial moment about the centroid is a stochastic process in time. Because of this, Dagan, Section 2.3.5, defines effective dispersivity coefficients as the expected values of the \mathbf{D}_{ij} 's.

$$\mathbf{D}_{ij}(t) = \frac{1}{2} \frac{dE[\mathbf{S}_{ij}(t)]}{dt}$$

Figure 22 conceptualizes this idea of dispersivity. As the plume spreads out, a new average dispersivity tensor applies at each time step. But in the types of models that are considered in this section, Monte Carlo methods will be used to include both the dispersivity and the velocity as random components of a random differential operator.

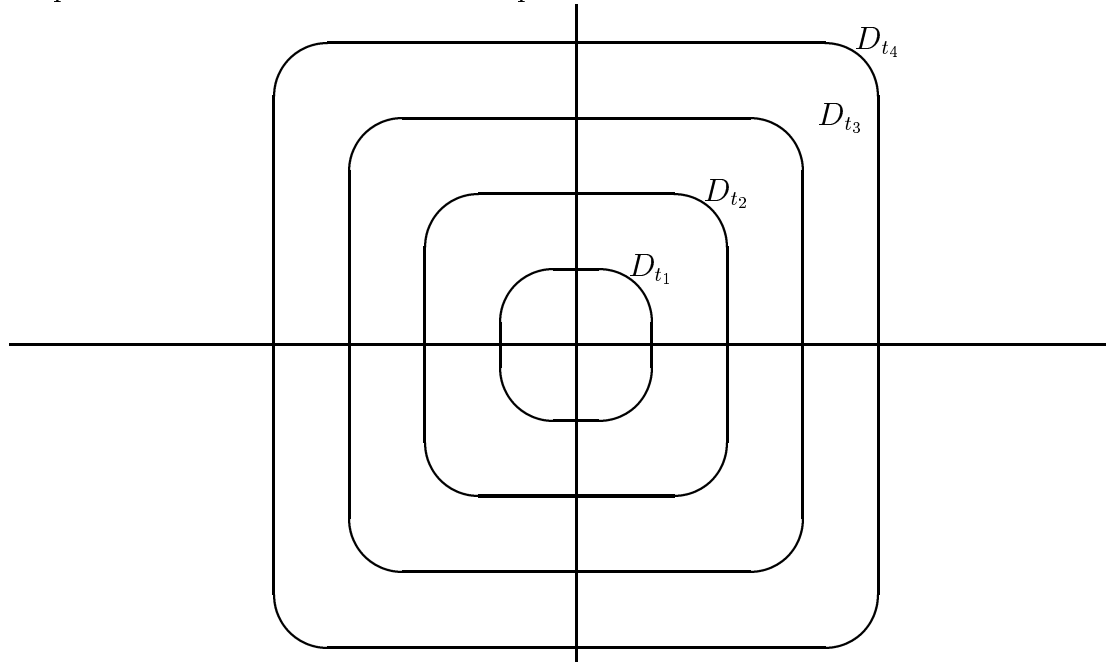


Figure 22 - Conceptual Dispersivity

$$\mathbf{S}_{ij} = \frac{1}{M} \int_{\Omega} n(x_i - R_i)(x_j - R_j)C(\vec{x}, t)d\vec{x} \quad i, j = 1, 2, 3$$

$$\mathbf{D}_{ij} = \frac{1}{2} \frac{d\mathbf{E}[\mathbf{S}_{ij}]}{dt}$$

In the following sections, the time dependent dispersivity coefficient will be incorporated into the stochastic partial differential equation model. In keeping with the basic notion of dispersion being a stochastic quantity, both the dispersion coefficients and the velocity will be allowed to have random components. The approach used is to treat the stochastic partial differential equation as a stochastic evolution equation. The solution of which requires an iterative process. References for this section are Adams[1], Adomian[3], Butzer and Berens[18], Serrano[88, 89], Tanabe[94], Tang and Pinder[95] and Yosida[99].

5.3 Stochastic Parameters

In Section 4.2.1, the stochastic partial differential equation was introduced, Equation[4.12], page 138. And, in Section 4.2.2, one of the many problems that can be treated using the stochastic partial differential equation was identified as the stochastic operator equation. The operator A is stochastic if one or more of its components is a stochastic process.

Consider the equation

$$\frac{\partial u}{\partial t}(x, t, \omega) + A(x, t, \omega)u = g(x, t, \omega)$$

$$u|_{\partial G} = 0 \quad u(x, 0) = u_0(x)$$

Let

$$\begin{aligned} A(x, t, \omega)u &= (\mathbf{E}[\vec{V}] + \vec{V}'(t, \omega))\nabla u - \nabla \cdot ((\mathbf{E}[\mathbf{D}(t)] + \mathbf{D}'(t, \omega))\nabla u) \\ &= \mathbf{E}[\vec{V}]\nabla u - \nabla \cdot (\mathbf{E}[\mathbf{D}(t)]\nabla u) + \vec{V}'(t, \omega)\nabla u - \nabla \cdot (\mathbf{D}'(t, \omega)\nabla u) \end{aligned}$$

So, we can write

$$\frac{\partial u}{\partial t} + \mathbf{E}[\vec{V}]\nabla u - \nabla \cdot (\mathbf{E}[\mathbf{D}(t)]\nabla u) = g - Ru$$

$$u|_{\partial G} = 0 \quad u(x, 0) = u_0$$

where

$$R \equiv \vec{V}'(t, \omega) \nabla(\cdot) - \nabla \cdot (\mathbf{D}'(t, \omega) \nabla(\cdot))$$

Notice that in order to simplify matters, the velocity is assumed to be strictly homogeneous (strictly stationary). Since the solution u appears on both sides of this equation, it cannot be represented explicitly. The solution can, however, be formally represented using an iterative process. The next section applies this series approach to represent the solution.

5.4 Formal Solution

In this section we formally develop a series representation for the inverse of the partial differential operator. Consider the equation

$$\mathcal{L}_{t,x} u = g$$

where

$$\mathcal{L}_{t,x} = L_{t,x} + R_{t,x}$$

and $L_{t,x}$ is a deterministic partial differential operator and $R_{t,x}$ is a *zero mean* stochastic partial differential operator.

If $\mathcal{L}_{t,x}^{-1}$ and $L_{t,x}^{-1}$ exist, then

$$\begin{aligned} L_{t,x} u &= g - R_{t,x} u \\ \Rightarrow u &= L_{t,x}^{-1} g - L_{t,x}^{-1} R_{t,x} u \end{aligned} \tag{5.5}$$

Hence,

$$u = \mathcal{L}_{t,x}^{-1} g = L_{t,x}^{-1} g - L_{t,x}^{-1} R_{t,x} \mathcal{L}_{t,x}^{-1} g$$

So, the operator equation is

$$\mathcal{L}_{t,x}^{-1} = L_{t,x}^{-1} - L_{t,x}^{-1} R_{t,x} \mathcal{L}_{t,x}^{-1}$$

Parametrizing with λ we get

$$\mathcal{L}_{t,x}^{-1} = L_{t,x}^{-1} - \lambda L_{t,x}^{-1} R_{t,x} \mathcal{L}_{t,x}^{-1}$$

Substituting

$$\mathcal{L}_{t,x}^{-1} = \sum_{i=0}^{\infty} \lambda^i H_i$$

we have that

$$\sum_{i=0}^{\infty} \lambda^i H_i = L_{t,x}^{-1} - \lambda L_{t,x}^{-1} R_{t,x} \sum_{i=0}^{\infty} \lambda^i H_i$$

Equating powers of λ , it follows that

$$\begin{aligned} H_0 &= L_{t,x}^{-1} \\ H_1 &= -L_{t,x}^{-1} R_{t,x} H_0 = -L_{t,x}^{-1} R_{t,x} L_{t,x}^{-1} \\ H_2 &= -L_{t,x}^{-1} R_{t,x} H_1 = (L_{t,x}^{-1} R_{t,x})^2 L_{t,x}^{-1} \\ &\vdots \end{aligned}$$

So, H_i can be expressed in terms of H_{i-1} , and since

$$\mathcal{L}_{t,x}^{-1} = \sum_{i=0}^{\infty} \lambda^i H_i$$

it follows with $\lambda = 1$ that

$$\mathcal{L}_{t,x}^{-1} = \sum_{i=0}^{\infty} (-1)^i (L_{t,x}^{-1} R_{t,x})^i L_{t,x}^{-1}$$

Hence,

$$u = \sum_{i=0}^{\infty} (-1)^i (L_{t,x}^{-1} R_{t,x})^i L_{t,x}^{-1} g$$

Since $L_{t,x}$ is the deterministic part of the equation, we can write its inverse in terms of the evolution operator as

$$L_{t,x}^{-1} = U(t, 0)u_0 + \int_0^t U(t, s)(\cdot)ds$$

then from Equation[5.5], the solution can be written as

$$u = U(t, 0)u_0 + \int_0^t U(t, s)g(s)ds - \int_0^t U(t, s)R_{s,x}u(s)ds$$

and parametrizing with λ as

$$u = U(t, 0)u_0 + \int_0^t U(t, s)g(s)ds - \lambda \int_0^t U(t, s)R_{s,x}u(s)ds$$

The two integrals in this expression are stochastic since the integrands are stochastic processes. By selecting sample paths of the random components, the integrals become ordinary integrals.

Letting

$$u = \sum_{i=0}^{\infty} \lambda^i H_i g = \mathcal{L}_{t,x}^{-1} g$$

then, in operator notation it follows that

$$\begin{aligned} \sum_{i=0}^{\infty} \lambda^i H_i(\cdot) &= U(t, 0)u_0 + \int_0^t U(t, s)(\cdot)ds - \lambda \int_0^t U(t, s)R_{s,x} \sum_{i=0}^{\infty} \lambda^i H_i(\cdot)ds \\ &= U(t, 0)u_0 + \int_0^t U(t, s)(\cdot)ds - \int_0^t U(t, s)R_{s,x} \sum_{i=0}^{\infty} \lambda^{i+1} H_i(\cdot)ds \end{aligned}$$

Equating powers of λ ,

$$\begin{aligned} i = 0 &\Rightarrow H_0(\cdot) = U(t, 0)u_0 + \int_0^t U(t, s)(\cdot)ds \\ i = 1 &\Rightarrow H_1(\cdot) = - \int_0^t U(t, s)R_{s,x}H_0(\cdot)ds \\ i = 2 &\Rightarrow H_2(\cdot) = - \int_0^t U(t, s)R_{s,x}H_1(\cdot)ds \\ &\vdots \\ i = n &\Rightarrow H_n(\cdot) = - \int_0^t U(t, s)R_{s,x}H_{n-1}(\cdot)ds \end{aligned}$$

Expanding $H_1(\cdot)$,

$$\begin{aligned}
H_1(\cdot) &= - \int_0^t U(t, s) R_{s,x} \left[U(s, 0) u_0 + \int_0^s U(s, \xi)(\cdot) d\xi \right] ds \\
&= - \int_0^t U(t, s) R_{s,x} U(s, 0) u_0 ds - \int_0^t \int_0^s U(t, s) R_{s,x} U(s, \xi)(\cdot) d\xi ds
\end{aligned}$$

Expanding $H_2(\cdot)$,

$$\begin{aligned}
H_2(\cdot) &= - \int_0^t U(t, s) R_{s,x} H_1(\cdot) ds \\
&= - \int_0^t U(t, s) R_{s,x} \left[- \int_0^s U(s, \xi) R_{\xi,x} U(\xi, 0) u_0 d\xi \right. \\
&\quad \left. - \int_0^s \int_0^\xi U(s, \xi) R_{\xi,x} U(\xi, \zeta)(\cdot) d\zeta d\xi \right] ds \\
&= \int_0^t \int_0^s U(t, s) R_{s,x} U(s, \xi) R_{\xi,x} U(\xi, 0) u_0 d\xi ds \\
&\quad + \int_0^t \int_0^s \int_0^\xi U(t, s) R_{s,x} U(s, \xi) R_{\xi,x} U(\xi, \zeta)(\cdot) d\zeta d\xi ds
\end{aligned}$$

Hence, in this way all of the terms of the series can be expanded. The procedure is demonstrated in the following example:

Example - Transport Equation

Consider the 1-D problem in which the parameters $\mathbf{E}[D]$ and $\mathbf{E}[V]$ are constants.

$$\begin{aligned}
\frac{\partial u}{\partial t} - (\mathbf{E}[D] + D'(t, \omega)) \frac{\partial^2 u}{\partial x^2} + (\mathbf{E}[V] + V'(t, \omega)) \frac{\partial u}{\partial x} &= 0 \\
-\infty \leq x \leq \infty \quad t > 0 \\
u(x, 0) = u_0, \quad u(-\infty, t) = 0, \quad u(\infty, t) = 0
\end{aligned}$$

This equation can be rewritten as

$$\frac{\partial u}{\partial t} - \mathbf{E}[D] \frac{\partial^2 u}{\partial x^2} + \mathbf{E}[V] \frac{\partial u}{\partial x} = D'(t, \omega) \frac{\partial^2 u}{\partial x^2} - V'(t, \omega) \frac{\partial u}{\partial x}$$

Before proceeding, the next two definitions are required.

Definition: Let X be a B-space. If \mathcal{T}_t is an operator such that

$$\mathcal{T}_t : \mathfrak{R}^+ \rightarrow \mathcal{B}(X)$$

and satisfies

- (1) $\mathcal{T}_{t+s} = \mathcal{T}_t \mathcal{T}_s \quad t \geq 0, \quad s \geq 0$
- (2) $\mathcal{T}_0 = I$
- (3) $\lim_{t \rightarrow 0} \|\mathcal{T}_t x - x\|_X = 0 \quad \forall x \in X$

then \mathcal{T}_t is called a *strongly continuous semigroup*.

Definition: The *infinitesimal generator of a semigroup* \mathcal{T}_t is defined by

$$\mathcal{A}x = \lim_{t \rightarrow 0^+} \frac{1}{t} (\mathcal{T}_t x - x)$$

$\mathcal{D}(\mathcal{A})$ is the set $x \in X$ for which the limit exists.

In order to illustrate the meaning of this definition, the following Theorem and Proposition from Rudin[83] regarding bounded or unbounded self-adjoint operators are helpful:

Theorem: To every self-adjoint operator A in H there corresponds a unique resolution E of the identity, on the Borel sets of the real line, such that

$$(Ax, y) = \int_{-\infty}^{\infty} \lambda dE_{x,y}(\lambda) = \int_{-\infty}^{\infty} \lambda d(E_\lambda x, y) \quad (x \in \mathcal{D}(A), y \in H)$$

Also, E is concentrated on $\sigma(A) \subset (-\infty, \infty)$ in the sense that

$$E(\sigma(A)) = I \quad \square$$

Proposition: Let A be self-adjoint. $(Ax, x) \leq 0$ if and only if $\sigma(A) \subset (-\infty, 0]$.

From the Theorem, it is clear that $A = \int_{-\infty}^{\infty} \lambda dE_\lambda$, and from the symbolic calculus for operators, Rudin[83], Friedman[39], if E is a spectral decomposition of the operator A and the spectrum of A is such that

$$\sigma(A) \subset (-\infty, 0]$$

then e^{tA} can be represented as

$$\mathcal{T}_t = e^{tA} = \int_{\sigma(A)} e^{t\lambda} dE_\lambda$$

Then, from the definition of the integral it follows that if $t = 0$, then $e^0 = I$ and

$$\begin{aligned} \lim_{t \rightarrow 0^+} \frac{\int_{\sigma(A)} e^{t\lambda} dE_\lambda x - x}{t} &= \lim_{t \rightarrow 0^+} \frac{\int_{\sigma(A)} e^{t\lambda} dE_\lambda x - \int_{\sigma(A)} e^0 dE_\lambda x}{t} \\ &= \lim_{t \rightarrow 0^+} \int_{\sigma(A)} \left(\frac{e^{t\lambda} - 1}{t} \right) dE_\lambda x \\ &= \int_{\sigma(A)} \left. \frac{d}{dt} e^{t\lambda} \right|_{t=0} dE_\lambda x \\ &= \int_{\sigma(A)} \lambda dE_\lambda x \\ &= Ax \end{aligned}$$

Therefore, $(\mathcal{T}_t)' \Big|_{t=0} = A$, and in this sense, the operator A is the infinitesimal generator of the semigroup \mathcal{T}_t .

Since $\mathbf{E}[D]$ is a constant, the solution can be expressed in terms of a semigroup as

$$u = \mathcal{T}_t u_0 + \int_0^t \mathcal{T}_{t-s} R_{s,x} u(x, s) ds$$

So, if we let

$$R_{s,x} = D'(s, \omega) \frac{\partial^2}{\partial x^2} - V'(t, \omega) \frac{\partial}{\partial x}$$

the solution can be written as

$$u(x, t) = \mathcal{T}_t u_0 + \int_0^t \mathcal{T}_{t-s} \left(D'(s, \omega) \frac{\partial^2}{\partial x^2} - V'(t, \omega) \frac{\partial}{\partial x} \right) u(x, s) ds$$

The semigroup \mathcal{T}_t is the one generated by the operator

$$\mathcal{A} = -\mathbf{E}[V]\frac{\partial}{\partial x} + \mathbf{E}[D]\frac{\partial^2}{\partial x^2}$$

To find the semigroup, start from the problem

$$\frac{\partial u}{\partial t} + \mathbf{E}[V]\frac{\partial u}{\partial x} - \mathbf{E}[D]\frac{\partial^2 u}{\partial x^2} = 0 \quad (5.6)$$

And,

$$u(-\infty, t) = 0 \quad u(\infty, t) = 0 \quad u(x, 0) = u_0$$

In Section 2.3, it was shown that the solution to this problem was given, with the change of variables $X = x - \mathbf{E}[V]t$ and $T = t$, by

$$\nu(X, T) = \int_{-\infty}^{\infty} K(X - \xi, \mathbf{E}[D]T)u_0(\xi)d\xi$$

where

$$K(X - \xi, \mathbf{E}[D]T) = \frac{1}{(4\pi\mathbf{E}[D]T)^{\frac{1}{2}}}e^{-\frac{(X-\xi)^2}{4\mathbf{E}[D]T}}$$

Substituting for X and T yields

$$\nu(X, T) = \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mathbf{E}[V]t-\xi)^2}{4\mathbf{E}[D]t}}}{2\sqrt{\pi\mathbf{E}[D]t}}u_0(\xi)d\xi$$

And, the candidate for the semigroup operator becomes

$$\mathcal{T}_t(\cdot) \equiv \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mathbf{E}[V]t-\xi)^2}{4\mathbf{E}[D]t}}}{2\sqrt{\pi\mathbf{E}[D]t}}(\cdot)d\xi \quad (5.7)$$

The following verify the semigroup properties: Let $u_0(x) \in L^p(-\infty, \infty)$

Property 1: $\mathcal{T}_0 = I$

$$\mathcal{T}_t u_0(x) = \int_{-\infty}^{\infty} \frac{e^{-\frac{(x - \mathbf{E}[V]t - \xi)^2}{4\mathbf{E}[D]t}}}{2\sqrt{\pi\mathbf{E}[D]t}} u_0(\xi) d\xi$$

Let

$$\xi = x - \mathbf{E}[V]t + 2\sqrt{\mathbf{E}[D]t} s$$

Then,

$$\mathcal{T}_t u_0(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-s^2} u_0(x - \mathbf{E}[V]t + 2\sqrt{\mathbf{E}[D]t} s) ds$$

And, it follows that as $t \rightarrow 0$,

$$\mathcal{T}_t u_0(x) \rightarrow \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-s^2} u_0(x) ds = u_0(x)$$

Hence,

$$\mathcal{T}_0 = I$$

Property 2: $\mathcal{T}_t \mathcal{T}_s = \mathcal{T}_{t+s}$ $t \geq 0, s \geq 0$

$$\begin{aligned} \mathcal{T}_t \mathcal{T}_s u_0(x) &= \int_{-\infty}^{\infty} \frac{e^{-(x - \mathbf{E}[V]t - \rho)^2 / 4\mathbf{E}[D]t}}{2\sqrt{\pi}\sqrt{\mathbf{E}[D]t}} \left[\int_{-\infty}^{\infty} \frac{e^{-(\rho - \mathbf{E}[V]s - \xi)^2 / 4\mathbf{E}[D]s}}{2\sqrt{\pi}\sqrt{\mathbf{E}[D]s}} u_0(\xi) d\xi \right] d\rho \\ &= \int_{-\infty}^{\infty} \frac{1}{4\pi\mathbf{E}[D]\sqrt{ts}} \int_{-\infty}^{\infty} \exp \left[\frac{-(x - \mathbf{E}[V]t - \rho)^2}{4\mathbf{E}[D]t} \right] \\ &\quad \times \exp \left[\frac{-(\rho - \mathbf{E}[V]s - \xi)^2}{4\mathbf{E}[D]s} \right] d\rho u_0(\xi) d\xi \end{aligned}$$

Let $g = \rho - \mathbf{E}[V]s - \xi$, then the inner integral becomes

$$\frac{1}{4\pi\mathbf{E}[D]\sqrt{ts}} \int_{-\infty}^{\infty} \exp \left[\frac{-(x - \mathbf{E}[V](t+s) - \xi - g)^2}{4\mathbf{E}[D]t} \right] \exp \left[\frac{-g^2}{4\mathbf{E}[D]s} \right] d\rho$$

Then using the following property of the Gaussian distribution

$$\frac{1}{\sqrt{4\pi(t_1 + t_2)}} \exp \left(\frac{-x^2}{4(t_1 + t_2)} \right) = \frac{1}{4\pi\sqrt{t_1 t_2}} \int_{-\infty}^{\infty} \exp \left[\frac{-(x - u)^2}{4t_1} \right] \exp \left[\frac{-u^2}{4t_2} \right] du$$

it follows that

$$\begin{aligned} \frac{1}{4\pi\mathbf{E}[D]\sqrt{ts}} & \int_{-\infty}^{\infty} \exp\left[\frac{-(x - \mathbf{E}[V](t+s) - \xi) - g)^2}{4\mathbf{E}[D]t}\right] \exp\left[\frac{-g^2}{4\mathbf{E}[D]s}\right] d\rho \\ & = \frac{1}{2\sqrt{\pi\mathbf{E}[D](t+s)}} \exp\left(\frac{-(x - \mathbf{E}[V](t+s) - \xi)^2}{4\mathbf{E}[D](t+s)}\right) \end{aligned}$$

Hence,

$$\mathcal{T}_t \mathcal{T}_s u_0(x) = \int_{-\infty}^{\infty} \frac{e^{-\frac{(x - \mathbf{E}[V](t+s) - \xi)^2}{4\mathbf{E}[D](t+s)}}}{2\sqrt{\pi\mathbf{E}[D](t+s)}} u_0(\xi) d\xi = \mathcal{T}_{t+s} u_0(x)$$

Property 3: $\lim_{t \rightarrow 0} \|\mathcal{T}_t x - x\|_X = 0 \quad \forall x \in X$

Starting from the definition

$$\mathcal{T}_t u_0(x) = \int_{-\infty}^{\infty} \frac{e^{-\frac{(x - \mathbf{E}[V]t - \xi)^2}{4\mathbf{E}[D]t}}}{2\sqrt{\pi\mathbf{E}[D]t}} u_0(\xi) d\xi$$

If we let $s = x - \mathbf{E}[V]t$ and

$$G(t, s - \xi) = \frac{e^{-\frac{(s - \xi)^2}{4\mathbf{E}[D]t}}}{2\sqrt{\pi\mathbf{E}[D]t}}$$

Then it follows that

$$\mathcal{T}_t u_0(x) = \int_{-\infty}^{\infty} G(t, s - \xi) u_0(\xi) d\xi$$

From Adams[1] we have the following Theorem 4.30 due to Young:

Theorem: Let $1 \leq p < \infty$ and let $u \in L^1(\mathfrak{R}^n)$ and $v \in L^p(\mathfrak{R}^n)$. Then the convolution products

$$u * v(x) = \int_{\mathfrak{R}^n} u(x - y)v(y)dy, \quad v * u(x) = \int_{\mathfrak{R}^n} v(x - y)u(y)dy$$

are well defined and equal for almost all $x \in \mathfrak{R}^n$. Moreover, $u * v \in L^p(\mathfrak{R}^n)$ and

$$\|u * v\|_p \leq \|u\|_1 \|v\|_p \quad \square$$

Applying this to our case, it follows that

$$\|\mathcal{T}_t u_0(x)\|_{L^p(-\infty, \infty)} \leq \|G(t, s - \xi)\|_{L^1(-\infty, \infty)} \|u_0(\xi)\|_{L^p(-\infty, \infty)}$$

But,

$$\|G(t, s - \xi)\|_{L^1(-\infty, \infty)} = \int_{-\infty}^{\infty} \frac{e^{-\frac{(x - \mathbf{E}[V]t - \xi)^2}{4\mathbf{E}[D]t}}}{2\sqrt{\pi\mathbf{E}[D]t}} d\xi = 1$$

Hence, $\|\mathcal{T}_t u_0(x)\|_{L^p(-\infty, \infty)} \leq \|u_0\|_{L^p(-\infty, \infty)} \Rightarrow \|\mathcal{T}_t\| \leq 1$, so that each \mathcal{T}_t is continuous.

Clearly, $u_0(x) = \int_{-\infty}^{\infty} \frac{e^{-\frac{(x - \mathbf{E}[V]t - \xi)^2}{4\mathbf{E}[D]t}}}{2\sqrt{\pi\mathbf{E}[D]t}} u_0(x) d\xi$, so that

$$|\mathcal{T}_t u_0(x) - u_0(x)| = \frac{1}{2\sqrt{\pi\mathbf{E}[D]t}} \left| \int_{-\infty}^{\infty} e^{-\frac{(s - \xi)^2}{4\mathbf{E}[D]t}} (u_0(\xi) - u_0(x)) d\xi \right|$$

Letting $y = \frac{s - \xi}{2\sqrt{\mathbf{E}[D]t}} = \frac{x - \mathbf{E}[V]t - \xi}{2\sqrt{\mathbf{E}[D]t}}$, then

$$\xi = x - \mathbf{E}[V]t - 2\sqrt{\mathbf{E}[D]t}y$$

And,

$$|\mathcal{T}_t u_0(x) - u_0(x)| = \frac{1}{\sqrt{\pi}} \left| \int_{-\infty}^{\infty} e^{-y^2} (u_0(x - \mathbf{E}[V]t - 2\sqrt{\mathbf{E}[D]t}y) - u_0(x)) dy \right|$$

Letting p' be the conjugate exponent of p , ie, $\frac{1}{p} + \frac{1}{p'} = 1$, then

$$|\mathcal{T}_t u_0(x) - u_0(x)| \leq \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{p}} e^{-\frac{y^2}{p'}} |u_0(x - \mathbf{E}[V]t - 2\sqrt{\mathbf{E}[D]t}y) - u_0(x)| dy$$

Applying Hölder's inequality to the right hand side integral

$$\begin{aligned} |\mathcal{T}_t u_0(x) - u_0(x)| &\leq \frac{1}{\sqrt{\pi}} \left(\int_{-\infty}^{\infty} e^{-\frac{y^2}{p}} |u_0(x - \mathbf{E}[V]t - 2\sqrt{\mathbf{E}[D]t}y) - u_0(x)|^p dy \right)^{\frac{1}{p}} \\ &\quad \times \left(\int_{-\infty}^{\infty} e^{-y^2} dy \right)^{\frac{1}{p'}} \\ &= K \left(\int_{-\infty}^{\infty} e^{-\frac{y^2}{p}} |u_0(x - \mathbf{E}[V]t - 2\sqrt{\mathbf{E}[D]t}y) - u_0(x)|^p dy \right)^{\frac{1}{p}} \end{aligned}$$

So that,

$$|\mathcal{T}_t u_0(x) - u_0(x)|^p \leq K^p \int_{-\infty}^{\infty} e^{\frac{-y^2}{p}} |u_0(x - \mathbf{E}[V]t - 2\sqrt{\mathbf{E}[D]ty}) - u_0(x)|^p dy$$

Integrating with respect to x and using Fubini's theorem,

$$\begin{aligned} & \|\mathcal{T}_t u_0(x) - u_0(x)\|_{L^p(-\infty, \infty)}^p \leq \\ & K^p \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{\frac{-y^2}{p}} |u_0(x - \mathbf{E}[V]t - 2\sqrt{\mathbf{E}[D]ty}) - u_0(x)|^p dy dx \\ & = K^p \int_{-\infty}^{\infty} e^{\frac{-y^2}{p}} \int_{-\infty}^{\infty} |u_0(x - \mathbf{E}[V]t - 2\sqrt{\mathbf{E}[D]ty}) - u_0(x)|^p dx dy \end{aligned}$$

Then, using Fatou's lemma,

$$\begin{aligned} & \limsup_{t \downarrow 0} \|\mathcal{T}_t u_0(x) - u_0(x)\|_{L^p(-\infty, \infty)}^p \leq \\ & \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \left(\limsup_{t \downarrow 0} \int_{-\infty}^{\infty} |u_0(x - \mathbf{E}[V]t - 2\sqrt{\mathbf{E}[D]ty}) - u_0(x)|^p dx \right) dy \end{aligned}$$

But,

$$\limsup_{t \downarrow 0} \int_{-\infty}^{\infty} |u_0(x - \mathbf{E}[V]t - 2\sqrt{\mathbf{E}[D]ty}) - u_0(x)|^p dx = 0$$

since this is true for continuous functions with compact support and the integrand can be arbitrarily closely approximated by such functions. So,

$$\lim_{t \downarrow 0} \|\mathcal{T}_t u_0(x) - u_0(x)\|_{L^p(-\infty, \infty)}^p = 0$$

Hence, \mathcal{T}_t is a *strongly continuous semigroup*. The above argument is based on similar arguments given in Tanabe[94] and Yosida[99]. This argument can be extended to \mathfrak{R}^n . Next, a uniqueness argument can be used to show that \mathcal{T}_t is a semigroup generated by the operator \mathcal{A} . To do this we need the following

Theorem: Let \mathcal{T}_t be a strongly continuous semigroup on a B-space X with infinitesimal generator \mathcal{A} . If $x_0 \in \mathcal{D}(\mathcal{A})$, then

- (1) $\mathcal{T}_t x_0 \in \mathcal{D}(\mathcal{A}) \quad \forall t \geq 0$
- (2) $\frac{d}{dt}(\mathcal{T}_t x_0) = \mathcal{A} \mathcal{T}_t x_0 = \mathcal{T}_t \mathcal{A} x_0 \quad t > 0 \quad \square$

If \mathcal{G}_t is a strongly continuous semigroup generated by A , then by the theorem

$$\frac{d}{dt}\mathcal{G}_t x_0 = \mathcal{A}\mathcal{G}_t x_0$$

then by forming the product $\mathcal{G}_{s-t}\mathcal{T}_t x_0$ and differentiating

$$\begin{aligned} \frac{d}{dt}(\mathcal{G}_{s-t}\mathcal{T}_t x_0) &= \mathcal{G}_{s-t}\frac{d}{dt}(\mathcal{T}_t x_0) + \frac{d}{dt}(\mathcal{G}_{s-t})\mathcal{T}_t x_0 \\ &= \mathcal{G}_{s-t}\mathcal{A}\mathcal{T}_t x_0 - \mathcal{A}\mathcal{G}_{s-t}\mathcal{T}_t x_0 \end{aligned}$$

Since \mathcal{G}_t is the semigroup generated by \mathcal{A} , the theorem gives that

$$\mathcal{A}\mathcal{G}_{s-t}\mathcal{T}_t x_0 = \mathcal{G}_{s-t}\mathcal{A}\mathcal{T}_t x_0$$

so that

$$\frac{d}{dt}(\mathcal{G}_{s-t}\mathcal{T}_t x_0) = \mathcal{G}_{s-t}\mathcal{A}\mathcal{T}_t x_0 - \mathcal{G}_{s-t}\mathcal{A}\mathcal{T}_t x_0 = 0$$

Hence, $\mathcal{G}_{s-t}\mathcal{T}_t x_0$ is constant with respect to t . Letting $t = 0$ and $s = t$ it follows that

$$\mathcal{G}_s \mathcal{T}_0 x_0 = \mathcal{G}_s x_0 = \mathcal{G}_0 \mathcal{T}_s x_0 = \mathcal{T}_s x_0$$

Therefore, for all $s > 0$ and $x_0 \in \mathcal{D}(\mathcal{A})$ the following holds:

$$\mathcal{G}_s x_0 = \mathcal{T}_s x_0$$

Since $\overline{\mathcal{D}(\mathcal{A})} = X$, if $x \in X$ and $\{x_n\}_{n=1}^{\infty} \subset \mathcal{D}(\mathcal{A})$ such that $x_n \rightarrow x$, then

$$\begin{aligned} \|\mathcal{G}_s x - \mathcal{T}_s x\|_X &= \|\mathcal{G}_s x - \mathcal{G}_s x_n + \mathcal{T}_s x_n - \mathcal{T}_s x\|_X \\ &\leq \|\mathcal{G}_s x - \mathcal{G}_s x_n\|_X + \|\mathcal{T}_s x_n - \mathcal{T}_s x\|_X \\ &\leq \|\mathcal{G}_s\| \|x - x_n\|_X + \|\mathcal{T}_s\| \|x_n - x\|_X \\ &\rightarrow 0 \end{aligned}$$

as $x_n \rightarrow x$ since $\mathcal{G}_s, \mathcal{T}_s \in \mathcal{B}(X)$. This means that $\mathcal{G}_s x = \mathcal{T}_s x, \forall x \in X$ so that the semigroups are equal. Hence \mathcal{A} is the infinitesimal generator of \mathcal{T}_t .

Using the expression for the solution

$$u = \sum_{i=0}^{\infty} H_i g$$

where for $g = 0$,

$$H_0(0) = \mathcal{T}_t u_0$$

$$H_1(0) = - \int_0^t \mathcal{T}_{t-s} R \mathcal{T}_s u_0 ds$$

$$H_2(0) = \int_0^t \int_0^s \mathcal{T}_{t-s} R \mathcal{T}_{s-r} R \mathcal{T}_r u_0 dr ds$$

And, the solution is given by

$$u(x, t) = \mathcal{T}_t u_0 + \int_0^t \mathcal{T}_{t-s} \left(D'(s, \omega) \frac{\partial^2}{\partial x^2} - V'(t, \omega) \frac{\partial}{\partial x} \right) \sum_{i=0}^{\infty} H_i(0) ds$$

with

$$\mathcal{T}_t(\cdot) \equiv \int_{-\infty}^{\infty} \frac{e^{-\frac{(x - \mathbf{E}[V]t - \xi)^2}{4\mathbf{E}[D]t}}}{2\sqrt{\pi\mathbf{E}[D]t}}(\cdot) d\xi$$

And, for the infinite interval, the solution to the problem [5.6] page 194, is given by

$$\mathcal{T}_t u_0(x) = \int_{-\infty}^{\infty} \frac{e^{-\frac{(x - \mathbf{E}[V]t - \xi)^2}{4\mathbf{E}[D]t}}}{2\sqrt{\pi\mathbf{E}[D]t}} u_0(\xi) d\xi$$

Also, if this operator is restricted to the functions $u(x)$ for which

$$u(-x) = -u(x) \quad \text{for } x > 0$$

then by a change of variables and rearranging terms, it follows that

$$\int_{-\infty}^{\infty} \frac{e^{-\frac{(x - \mathbf{E}[V](t-s) - \xi)^2}{4\mathbf{E}[D](t-s)}}}{2\sqrt{\pi\mathbf{E}[D](t-s)}} u(\xi) d\xi = \int_0^{\infty} \left[\frac{e^{-\frac{(x - \mathbf{E}[V](t-s) - \xi)^2}{4\mathbf{E}[D](t-s)}} - e^{-\frac{(x - \mathbf{E}[V](t-s) + \xi)^2}{4\mathbf{E}[D](t-s)}}}{2\sqrt{\pi\mathbf{E}[D](t-s)}} \right] u(\xi) d\xi$$

So, on this restricted domain, the operator

$$U(t, s) = \mathcal{T}_{t-s}(\cdot) \equiv \int_0^\infty \left[\frac{e^{-\frac{(x-\mathbf{E}[V](t-s)-\xi)^2}{4\mathbf{E}[D](t-s)}} - e^{-\frac{(x-\mathbf{E}[V](t-s)+\xi)^2}{4\mathbf{E}[D](t-s)}}}{2\sqrt{\pi\mathbf{E}[D](t-s)}} \right] (\cdot) d\xi$$

is our original semigroup. Restricting the domain should not effect us, since we are only interested in the non-negative x axis.

The solution for a semi-infinite interval is then given by Guenther and Lee[51] as

$$\mathcal{G}_t u_0 = \int_0^\infty \left[\frac{e^{-\frac{(x-\mathbf{E}[V]t-\xi)^2}{4\mathbf{E}[D]t}} - e^{-\frac{(x-\mathbf{E}[V]t+\xi)^2}{4\mathbf{E}[D]t}}}{2\sqrt{\pi\mathbf{E}[D]t}} \right] u_0(\xi) d\xi$$

which is the solution for the problem

$$\begin{aligned} \frac{\partial \nu}{\partial \tilde{t}} - \mathbf{E}[D] \frac{\partial^2 \nu}{\partial \tilde{x}^2} &= 0 \\ \nu(\tilde{x}, 0) &= u_0 \quad \tilde{x} > 0 \\ \nu(0, \tilde{t}) &= 0 \quad \tilde{t} > 0 \end{aligned}$$

where $\tilde{x} = x - \mathbf{E}[V]t$, $\tilde{t} = t$ and $\nu(\tilde{x}, \tilde{t}) = u(x, t)$. Li[62] uses a Laplace transform to show that the solution is also given by

$$a(x, t) = \frac{u_0}{2} \left\{ \operatorname{erfc} \left(\frac{x - \mathbf{E}[V]t}{2\sqrt{\mathbf{E}[D]t}} \right) + \exp \left(\frac{\mathbf{E}[V]x}{\mathbf{E}[D]} \right) \operatorname{erfc} \left(\frac{x - \mathbf{E}[V]t}{2\sqrt{\mathbf{E}[D]t}} \right) \right\} \quad (5.8)$$

One numerical routine that makes this formulation of the solution easy to work with is a routine for calculating the error function, $\operatorname{erf}(x)$, which is computed by an algorithm due to Hastings[52] which is given by

$$\operatorname{erf}(x) \approx 1 - (a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 t^5) e^{-x^2}$$

where

$$t = \frac{1}{1 + 0.3275911x}, \quad a_1 = 0.254829592, \quad a_2 = -0.284496736$$

$$a_3 = 1.421413741, \quad a_4 = -1.453152027, \quad a_5 = 1.061405429$$

According to Greenberg[50] it is supposed to be uniformly accurate over $0 \leq x < \infty$ to within $\pm 1.5 \times 10^{-7}$.

The function $u(x, t)$, which is given by

$$u(x, t) = \mathcal{G}_t u_0 + \int_0^t \mathcal{G}_{t-s} \left(D'(s, \omega) \frac{\partial^2}{\partial x^2} - V'(t, \omega) \frac{\partial}{\partial x} \right) u(x, s) ds$$

is then the solution of the problem of a long uniform channel which is initially uncontaminated, and at $t = 0$, a contaminant is introduced whose concentration at $x = 0$ is maintained at u_0 .

Taking this equation and parameterizing it with λ ,

$$u(x, t) = \mathcal{G}_t u_0 + \lambda \int_0^t \mathcal{G}_{t-s} \left(D'(s, \omega) \frac{\partial^2}{\partial x^2} - V'(t, \omega) \frac{\partial}{\partial x} \right) u(x, s) ds \quad (5.9)$$

Clearly, as long as $\mathcal{G}_t u_0 \neq 0$, the operator

$$\mathcal{G}_t u_0 + \lambda \int_0^t \mathcal{G}_{t-s} \left(D'(s, \omega) \frac{\partial^2}{\partial x^2} - V'(t, \omega) \frac{\partial}{\partial x} \right) (\cdot) ds$$

is not linear. So. by writing

$$u(x, t) = \sum_{i=0}^{\infty} \lambda^i H_i(g(x, t))$$

where $g(x, t)$ is the forcing term, on substituting, it follows that

$$\sum_{i=0}^{\infty} \lambda^i H_i(g) = \mathcal{G}_t u_0 + \lambda \int_0^t \mathcal{G}_{t-s} \left(D'(s, \omega) \frac{\partial^2}{\partial x^2} - V'(t, \omega) \frac{\partial}{\partial x} \right) \sum_{i=0}^{\infty} \lambda^i H_i(g) ds$$

Equating powers of λ ,

$$\begin{aligned}
i = 0 &\Rightarrow H_0(g) \\
i = 1 &\Rightarrow H_1(g) = \int_0^t \mathcal{G}_{t-s} \left(D'(s, \omega) \frac{\partial^2}{\partial x^2} - V'(t, \omega) \frac{\partial}{\partial x} \right) H_0(g) ds \\
i = 2 &\Rightarrow H_2(g) = \int_0^t \mathcal{G}_{t-s} \left(D'(s, \omega) \frac{\partial^2}{\partial x^2} - V'(t, \omega) \frac{\partial}{\partial x} \right) H_1(g) ds \quad (5.10) \\
&\vdots \\
i = n &\Rightarrow H_n(g) = \int_0^t \mathcal{G}_{t-s} \left(D'(s, \omega) \frac{\partial^2}{\partial x^2} - V'(t, \omega) \frac{\partial}{\partial x} \right) H_{n-1}(g) ds
\end{aligned}$$

$H_0(g)$ has already been determined. Since it is assumed that the solute injection at the left hand boundary is held constant at u_0 , the formula Equation[5.8] will be used as a representation of $H_0(g)$. In order to have a specific realization for $H_i(g)$, it is necessary to fix ω . This will make $V'(t, \omega)$ and $D'(s, \omega)$ sample paths.

Additional questions that have to be answered before the feasibility of this type of approach can be addressed clearly involve the convergence of the series. Some type of convergence analysis has to be made to determine the number of terms that are needed to be retained for a desired accuracy. Also, the additional complexity of extending this approach to 2 and 3 dimensions has to be considered. And, numerical implementations need to be worked out.

5.5 Convergence

This section will consider convergence questions surrounding the infinite series solutions developed above. As discussed earlier, recall that the solution to the problem is given as an infinite series in the form

$$u(x, t) = \sum_{i=0}^{\infty} H_i(g)$$

where the $H_i(g)$ terms are given by Equations[5.10]. In particular, $H_0(g)$ is given as

$$H_0(g) = a(x, t)$$

First off, define the operators

$$\begin{aligned}
\mathcal{L}_{t,x}^{-1} &\equiv \int_0^t \mathcal{T}_{t-s}(\cdot) ds \\
&= \int_0^t \int_0^\infty \left[\frac{e^{-\frac{(x-\mathbf{E}[V](t-s)-\xi)^2}{4\mathbf{E}[D](t-s)}} - e^{-\frac{(x-\mathbf{E}[V](t-s)+\xi)^2}{4\mathbf{E}[D](t-s)}}}{2\sqrt{\pi\mathbf{E}[D](t-s)}} \right] (\cdot) d\xi ds
\end{aligned}$$

And,

$$R_{s,\xi} \equiv D'(s, \omega) \frac{\partial^2}{\partial \xi^2} - V'(t, \omega) \frac{\partial}{\partial \xi}$$

So that

$$\begin{aligned}
\mathcal{L}_{t,x}^{-1} R_{s,\xi}(\cdot) &= \int_0^t \int_0^\infty \left[\frac{e^{-\frac{(x-\mathbf{E}[V](t-s)-\xi)^2}{4\mathbf{E}[D](t-s)}} - e^{-\frac{(x-\mathbf{E}[V](t-s)+\xi)^2}{4\mathbf{E}[D](t-s)}}}{2\sqrt{\pi\mathbf{E}[D](t-s)}} \right] \\
&\quad \times \left(D'(s, \omega) \frac{\partial^2}{\partial \xi^2} - V'(t, \omega) \frac{\partial}{\partial \xi} \right) (\cdot) d\xi ds
\end{aligned}$$

Fix X and T so that

$$(x, t) \in [0, X] \times [0, T]$$

Then, using Equations[5.10] as a reference, the $H_i(g)$'s can be written in the following way:

$$\begin{aligned}
H_0(g(x, t)) &= a(x, t) \\
H_1(g(x, t)) &= \mathcal{L}_{t,x}^{-1} R_{s,\xi} H_0(g(\xi, s)) \\
&= \mathcal{L}_{t,x}^{-1} R_{s,\xi} a(\xi, s)
\end{aligned}$$

Hence,

$$\begin{aligned}
u_1(x, t) &= H_0(g(x, t)) + H_1(g(x, t)) \\
&= a(x, t) + [\mathcal{L}_{t,x}^{-1} R_{s,\xi}]^1 a(\xi, s)
\end{aligned}$$

And, continuing in this way, we can write

$$\begin{aligned} H_2(g(x, t)) &= \mathcal{L}_{t,x}^{-1} R_{s,\xi} H_1(g(\xi, s)) \\ &= [\mathcal{L}_{t,x}^{-1} R_{s,\xi}]^2 a(\xi, s) \end{aligned}$$

So that,

$$\begin{aligned} u_2(x, t) &= H_0(g(x, t)) + H_1(g(x, t)) + H_2(g(x, t)) \\ &= a(x, t) + [\mathcal{L}_{t,x}^{-1} R_{s,\xi}]^1 a(\xi, s) + [\mathcal{L}_{t,x}^{-1} R_{s,\xi}]^2 a(\xi, s) \end{aligned}$$

In general, the n^{th} approximation can be written as

$$H_n(g(x, t)) = [\mathcal{L}_{t,x}^{-1} R_{s,\xi}]^n a(\xi, s)$$

$$u_n = \sum_{i=0}^n H_i(g(x, t)) = a(x, t) + \sum_{i=1}^n [\mathcal{L}_{t,x}^{-1} R_{s,\xi}]^i a(\xi, s)$$

The solution is given by the limit

$$u(x, t) = \lim_{n \rightarrow \infty} u_n(x, t) = a(x, t) + \lim_{n \rightarrow \infty} \sum_{i=1}^n [\mathcal{L}_{t,x}^{-1} R_{s,\xi}]^i a(\xi, s) \quad (5.11)$$

Consider the difference for $n > m$

$$\begin{aligned} \|u_{n+1} - u_n\| &= \left\| \left(a(x, t) + \sum_{i=1}^{n+1} [\mathcal{L}_{t,x}^{-1} R_{s,\xi}]^i a(\xi, s) \right) \right. \\ &\quad \left. - \left(a(x, t) + \sum_{i=1}^n [\mathcal{L}_{t,x}^{-1} R_{s,\xi}]^i a(\xi, s) \right) \right\| \\ &= \| [\mathcal{L}_{t,x}^{-1} R_{s,\xi}]^{n+1} a(\xi, s) \| \\ &\leq \| \mathcal{L}_{t,x}^{-1} R_{s,\xi} \| \| [\mathcal{L}_{t,x}^{-1} R_{s,\xi}]^n a(\xi, s) \| \\ &= \| \mathcal{L}_{t,x}^{-1} R_{s,\xi} \| \| u_n - u_{n-1} \| \end{aligned}$$

So, if we let $\alpha = \|\mathcal{L}_{t,x}^{-1}R_{s,\xi}\|$ then

$$\begin{aligned} \|u_{n+1} - u_n\| &\leq \alpha \|u_n - u_{n-1}\| \\ &\leq \alpha^2 \|u_{n-1} - u_{n-2}\| \\ &\vdots \\ &\leq \alpha^n \|u_1 - u_0\| \end{aligned}$$

More generally,

$$\begin{aligned} \|u_n - u_m\| &= \|(u_n - u_{n-1}) + (u_{n-1} - u_{n-2}) \\ &\quad + \cdots + (u_{m+1} - u_m)\| \\ &\leq \|u_n - u_{n-1}\| + \|u_{n-1} - u_{n-2}\| \\ &\quad + \cdots + \|u_{m+1} - u_m\| \\ &\leq (\alpha^{n-1} + \alpha^{n-2} + \cdots + \alpha^m) \|u_1 - u_0\| \\ &\leq \left(\sum_{i=m}^{\infty} \alpha^i \right) \|u_1 - u_0\| \end{aligned}$$

But,

$$\sum_{i=m}^{\infty} \alpha^i = \frac{1}{1-\alpha} - \frac{1-\alpha^m}{1-\alpha} = \frac{\alpha^m}{1-\alpha}$$

So, it follows that

$$\|u_n - u_m\| \leq \frac{\alpha^m}{1-\alpha} \|u_1 - u_0\|$$

Now, if $\alpha < 1$, ie $\mathcal{L}_{t,x}^{-1}R_{s,\xi}$ is a contraction, then given $\epsilon > 0$, $\exists N_\epsilon$ such that if $n > m > N_\epsilon$, then

$$\alpha^m < \frac{1-\alpha}{\|u_1 - u_0\|} \epsilon$$

and,

$$\|u_n - u_m\| \leq \frac{\alpha^m}{1 - \alpha} \|u_1 - u_0\| < \epsilon$$

The sequence $\{u_n\}$ is a Cauchy sequence and by the completeness of the Hilbert space converges to a unique limit, namely, $u(x, t)$.

Referring to Equation[5.9] on page 202, and as shown on that page, Equation[5.9] is parametrized with λ to obtain

$$u(x, t) = \mathcal{G}_t u_0 + \lambda \int_0^t \mathcal{G}_{t-s} \left(D'(s, \omega) \frac{\partial^2}{\partial x^2} - V'(t, \omega) \frac{\partial}{\partial x} \right) u(x, s) ds \quad (5.12)$$

This equation has the form of a Volterra integral equation of the second kind. Recalling that

$$\begin{aligned} \mathcal{L}_{t,x}^{-1} R_{s,\xi}(\cdot) &= \int_0^t \int_0^\infty \left[\frac{e^{-\frac{(x - \mathbf{E}[V](t-s) - \xi)^2}{4\mathbf{E}[D](t-s)}} - e^{-\frac{(x - \mathbf{E}[V](t-s) + \xi)^2}{4\mathbf{E}[D](t-s)}}}{2\sqrt{\pi\mathbf{E}[D](t-s)}} \right] \\ &\quad \times \left(D'(s, \omega) \frac{\partial^2}{\partial \xi^2} - V'(t, \omega) \frac{\partial}{\partial \xi} \right) (\cdot) d\xi ds \end{aligned}$$

Equation[5.12] can be put in the form

$$u(x, t) = a(x, t) + \lambda \mathcal{L}_{t,x}^{-1} R_{s,\xi} u(\xi, s) ds$$

Therefore, the equation to solve is given by

$$Tu = u$$

with

$$T \equiv a(x, t) + \lambda \mathcal{L}_{t,x}^{-1} R_{s,\xi}$$

If f_0 is some initial estimate of the solution, then

$$Tf_0 = a + \lambda \mathcal{L}_{t,x}^{-1} R_{s,\xi} f_0$$

$$\begin{aligned}
T^2 f_0 &= T[a + \lambda \mathcal{L}_{t,x}^{-1} R_{s,\xi} f_0] \\
&= Ta + \lambda \mathcal{L}_{t,x}^{-1} R_{s,\xi} [a + \lambda \mathcal{L}_{t,x}^{-1} R_{s,\xi} f_0] \\
&= a + \lambda \mathcal{L}_{t,x}^{-1} R_{s,\xi} a + \lambda^2 [\mathcal{L}_{t,x}^{-1} R_{s,\xi}]^2 f_0
\end{aligned}$$

In general,

$$\begin{aligned}
T^n f_0 &= a + \lambda \mathcal{L}_{t,x}^{-1} R_{s,\xi} a + \lambda^2 [\mathcal{L}_{t,x}^{-1} R_{s,\xi}]^2 a \\
&\quad + \cdots + \lambda^n [\mathcal{L}_{t,x}^{-1} R_{s,\xi}]^n f_0
\end{aligned}$$

and if $\mathcal{L}_{t,x}^{-1} R_{s,\xi}$ is a contraction then the solution is given by

$$\begin{aligned}
u &= \lim_{n \rightarrow \infty} T^n f_0 \\
&= a + \sum_{i=1}^{\infty} \lambda^i [\mathcal{L}_{t,x}^{-1} R_{s,\xi}]^i a
\end{aligned}$$

with $\lambda = 1$ then u becomes

$$u(x, t) = a(x, t) + \lim_{m \rightarrow \infty} \sum_{i=1}^m [\mathcal{L}_{t,x}^{-1} R_{s,\xi}]^i a(\xi, s)$$

which is the same as Equation[5.11], page 205. So, the convergence theory is the same as that for the Volterra integral equation of the second kind.

Following methods similar to the one illustrated in Figure 12, Section 3.2.10, Monte Carlo methods can be used to construct concentration profiles from which concentration means and variances can be constructed.

As a test problem, consider the 1-D example of a long channel of uniform sand which has an established flow through it. The channel is initially uncontaminated and at $t = 0$ a contaminant or tracer is introduced whose concentration at $x = 0$ is maintained at u_0 . The physical parameters used in the model are as follows:

Distance From Origin	1, 2, 5, 10 Meters
Time	10 Days
Velocity	0.3 m/day
Dispersion	0.1 m ² /Day
Concentration At x = 0	1.0 mgrm/liter
Δt	0.05
Δx	0.05

The following tables give some results of applying this iterative method to this problem. The numbers in the columns labeled *Concentration* and *Increment* are interpreted as follows: In row i , add the number in the *Increment* column to the number in the *Concentration* column. This will produce the number in the *Concentration* column in row $i + 1$.

Example 1:

Time = 10.0 Days			
Distance From Origin Meters	Iteration	Concentration	Increment
1.0	1	0.968878174061	0.000018964773
	2	0.968897138835	0.000000847497
	3	0.968897986332	0.000000338709
	4	0.968898325040	0.00000002449
	5	0.968898327489	-0.000000004363
	6	0.968898323127	
2.0	1	0.843295693603	-0.000014472127
	2	0.843281221477	-0.000002613105
	3	0.843278608371	0.000000035747
	4	0.843278644118	0.000000044050
	5	0.843278688168	0.000000002064
	6	0.843278690232	
5.0	1	0.101972946853	-0.000002073268
	2	0.101970873586	0.000000148862
	3	0.101971022447	-0.000000034350
	4	0.101970988097	0.000000005056
	5	0.101970993154	0.000000000575
	6	0.101970993728	
10.0	1	0.000000211885	-0.000000141955
	2	0.000000197690	0.000000000034
	3	0.000000197724	0.000000000000
	4	0.000000197724	0.000000000000
	5	0.000000197724	0.000000000000
	6	0.000000197724	

In this example, the dispersion is allowed to have a random component and the velocity is not. The sample path of the stochastic process $D'(t, \omega)$ was generated from a Gaussian distribution with a mean of zero and a standard deviation of 0.03.

Example 2:

Time = 10.0 Days			
Distance From Origin Meters	Iteration	Concentration	Increment
1.0	1	0.971451709277	0.000027765168
	2	0.971479474446	-0.000000288321
	3	0.971479186125	0.000000008773
	4	0.971479194898	-0.000000000252
	5	0.971479194646	0.000000000007
	6	0.971479194653	
2.0	1	0.835760450146	-0.000031500611
	2	0.835728949535	0.000001117351
	3	0.835730066886	-0.000000023517
	4	0.835730043368	0.000000000398
	5	0.835730043767	-0.000000000004
	6	0.835730043762	
5.0	1	0.101035253556	0.000023792755
	2	0.101059046311	0.000000049911
	3	0.101059096222	-0.000000007834
	4	0.101059088388	0.000000000358
	5	0.101059088746	-0.000000000013
	6	0.101059088733	
10.0	1	0.000000249522	-0.000000007750
	2	0.000000241772	-0.000000003842
	3	0.000000237930	-0.000000000018
	4	0.000000237912	0.000000000000
	5	0.000000237912	0.000000000000
	6	0.000000237912	0.000000000000

In this example, the velocity is allowed to have a random component and the dispersion is not. The sample path of the stochastic process $V'(t, \omega)$ was generated from a Gaussian distribution with a mean of zero and a standard deviation of 0.03.

Example 3:

Time = 10.0 Days			
Distance From Origin Meters	Iteration	Concentration	Increment
1.0	1	0.972000869383	0.000001256451
	2	0.972002125834	-0.000002943310
	3	0.971999182524	-0.000000029912
	4	0.971999152612	0.000000008036
	5	0.971999160648	0.000000000603
	6	0.971999161252	
2.0	1	0.836701222644	-0.000060724077
	2	0.836640498567	0.000001026103
	3	0.836641524670	-0.000000493583
	4	0.836641031086	-0.000000062218
	5	0.836640968868	0.000000005893
	6	0.836640974761	
5.0	1	0.099088557472	0.000025447894
	2	0.099114005366	0.000001381694
	3	0.099115387060	-0.000000291936
	4	0.099115095124	0.000000027379
	5	0.099115122503	0.000000005205
	6	0.099115127708	
10.0	1	0.000000089299	-0.000000236624
	2	-0.000000147325	0.000000001784
	3	-0.000000145541	-0.000000000044
	4	-0.000000145585	-0.000000000001
	5	-0.000000145585	0.000000000000
	6	-0.000000145585	

In this example, both the velocity and the dispersion are allowed to have random components. The sample paths of the stochastic processes $V'(t, \omega)$ and $D'(t, \omega)$ were generated from a Gaussian distribution with a mean of zero and a standard deviation of 0.03. At the 10.0 meter mark, the process has the pathological behavior of converging to a negative concentration.

In Section 2.3.3, the transport equation was solved using the Fourier transform. As part of that process, a first order equation, Equation[2.21], page 66, was obtained. Since the dispersion tensor contained constant coefficients, the autonomous first order equation had an easy solution in terms of the Fourier transform,

$$\hat{v}(\omega, T) = \hat{f}(\omega) e^{-\tilde{\omega}^T \tilde{\mathbf{D}} \omega T}$$

However, the situation changes considerably if the dispersion tensor is allowed to depend on time. The first order equation, Equation[2.21] is now a nonautonomous equation because the dispersion term $\tilde{\mathbf{D}}$ depends on time. In order to solve this equation, we need to know the analytical form of $\tilde{\mathbf{D}}$. Referring to Equation[1.13], page 37, it can be seen that in the case of steady state or ergodic flow, the dispersion is essentially constant and so can be handled by the semigroup approach presented earlier. In the case of non-ergodic flow, the problem becomes much more difficult since the dispersion tensor is now time dependent. The equation that needs to be solved for the fundamental solution is now a nonautonomous equation and has the form

$$\frac{du(t)}{dt} = A(t) u(t)$$

where $A(t) \equiv -\mathbf{E}[V] \frac{\partial}{\partial t} + \mathbf{E}[D(t)] \frac{\partial^2}{\partial x^2}$.

This case can be reduced to the autonomous case by creating the system

$$\begin{cases} \frac{du(t)}{dt} = A(\tau) u(t) & u(0) = u_0 \\ \frac{d\tau}{dt} = 1 & \tau(0) = 0 \end{cases}$$

Solving this system yields the solution

$$\mathcal{T}_t u_0(x) = \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mathbf{E}[V]t-\xi)^2}{4\mathbf{E}[D(t)]t}}}{2\sqrt{\pi\mathbf{E}[D(t)]t}} u_0(\xi) d\xi$$

If the operator $U(t, 0)$ is defined as

$$U(t, 0)(\cdot) \equiv \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mathbf{E}[V]t-\xi)^2}{4\mathbf{E}[D(t)]t}}}{2\sqrt{\pi\mathbf{E}[D(t)]t}} (\cdot) d\xi$$

Then this integral operator can be used to study the case of the expected dispersion being time dependent.

5.6 Summary

Chapter 5 investigated the application of the theory of stochastic evolution equations to the problems of flow and transport. Section 5.1 reviewed the connections between the boundary value problem and the abstract evolution equation for both the autonomous and nonautonomous cases. Curtain and Falb[27] extend these results to the case where the forcing term of the abstract evolution equation contains an H -valued Wiener process. From the Curtain and Falb result, weak forms of the mean and covariance equations are found. Section 5.2 considered the time dependent forms of the dispersion tensor and Section 5.3 gave the form of the stochastic PDE in terms of a sum of deterministic and stochastic operators.

Section 5.4 represented the solution of the stochastic PDE as an integral equation involving an evolution operator that derived from treating the deterministic part of the stochastic PDE as an abstract differential equation. An iterative approach to solving this integral equation was presented and the convergence property of the series solution was given in Section 5.5. A test problem was also solved in order to investigate the speed of convergence of the series.

6. Future Research

6.1 General

The Neumann expansion procedure outlined in Sections 5.4 and 5.5 can be extended using the Karhunen-Loève expansion for a stochastic process and the Galerkin method, thus turning it into a *stochastic finite element method*, Ghanem and Spanos[46]. However, the method continues to be subject to the convergence criterion

$$\|\mathcal{L}_{t,x}^{-1}R_{s,\xi}\| < 1$$

In order to avoid this restriction, a method using the Homogeneous Chaoses of Wiener is outlined next. The Karhunen-Loève expansion is a Fourier-type expansion of the form

$$\xi(\vec{x}, \omega) = \sum_{k=1}^{\infty} (\lambda_k)^{\frac{1}{2}} \xi_k(\omega) \phi_k(\vec{x})$$

The $\{\lambda_k\}$ is a sequence of constants, the $\{\phi_k(\vec{x})\}$ is a sequence of orthonormal *deterministic* functions and the $\{\xi_k(\omega)\}$ is a sequence of random variables given by

$$\xi_k(\omega) = (\lambda_k)^{-\frac{1}{2}} \int_D \xi(\vec{x}, \omega) \overline{\phi_k(\vec{x})} d\vec{x}$$

The details of the existence of the constants $\{\lambda_k\}$ and the orthonormal sequence $\{\phi_k(\vec{x})\}$ is covered in Loève[63]. Basically, the covariance function of the process $\xi(\vec{x}, \omega)$ can be written as

$$C(\vec{x}_1, \vec{x}_2) = \sum_{k=0}^{\infty} \lambda_k \phi_k(\vec{x}_1) \phi_k(\vec{x}_2)$$

And, since the ϕ_k 's form an orthonormal sequence, it follows that

$$\int_D C(\vec{x}_1, \vec{x}_2) \overline{\phi_k(\vec{x}_1)} d\vec{x}_1 = \lambda_k \phi_k(\vec{x}_2)$$

so that the constants $\{\lambda_k\}$ and the functions $\{\phi_k(\vec{x})\}$ are solutions of this Fredholm type two integral equation. And, the random variables $\xi_k(\omega)$ are determined by these solutions.

In Section 5.4, the formal solution of the equation

$$\mathcal{L}_{t,x}u = g$$

was developed by first splitting the operator $\mathcal{L}_{t,x}$ into a deterministic part and a *zero mean* stochastic part, *i. e.*,

$$\mathcal{L}_{t,x} \equiv L_{t,x} + R_{t,x}$$

For example, the stochastic transport equation can be written as

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} [\mathbf{E}[V(x)]u] - \mathbf{D}(t) \frac{\partial^2 u}{\partial x^2} + \frac{\partial(uV'(x, \omega))}{\partial x} = 0$$

or,

$$\frac{\partial u}{\partial t} + \mathbf{E}[V(x)] \frac{\partial u}{\partial x} - \mathbf{D}(t) \frac{\partial^2 u}{\partial x^2} + \frac{\partial(uV'(x, \omega))}{\partial x} = \frac{\partial}{\partial x} \mathbf{E}[V(x)]u$$

Since $\frac{\partial}{\partial x} \mathbf{E}[V(x)]$ can be computed and $u(t, x, \omega)$ can be estimated from the previous time step, $t - \Delta t$, by letting

$$f(t, x, \omega) = \frac{\partial}{\partial x} \mathbf{E}[V(x)]u(t, x, \omega)$$

the following holds approximately,

$$\frac{\partial u}{\partial t} + \mathbf{E}[V(x)] \frac{\partial u}{\partial x} - \mathbf{D}(t) \frac{\partial^2 u}{\partial x^2} + \frac{\partial(uV'(x, \omega))}{\partial x} \approx \tilde{f}$$

where $\tilde{f} = \frac{\partial}{\partial x} \mathbf{E}[V(x)]u(t - \Delta t, x, \omega)$. In this example, it is assumed that the dispersion tensor, \mathbf{D} , has been dealt with as explained in Section 2.4. This leaves as the only random coefficient the random velocity term, $V'(x, \omega)$ which *does not* depend on t . By the Karhunen-Loève expansion, $V'(x, \omega)$ can be expanded as

$$V'(x, \omega) \approx \sum_{k=1}^m (\lambda_k)^{\frac{1}{2}} \xi_k(\omega) \phi_k(x)$$

where the sum has been truncated after m terms, the $\{\xi_k(\omega)\}$ are random variables and $\{\phi_k(x)\}$ are deterministic functions of x . Then, it follows that

$$L_{t,x}u + \sum_{k=1}^m (\lambda_k)^{\frac{1}{2}} \xi_k(\omega) R(u\phi_k(x)) \approx \tilde{f}$$

or,

$$L_{t,x}u + \sum_{k=1}^m (\lambda_k)^{\frac{1}{2}} \xi_k(\omega) \phi_k(x) Ru \approx \tilde{f} - \sum_{k=1}^m (\lambda_k)^{\frac{1}{2}} \xi_k(\omega) u R \phi_k(x) = \hat{f}$$

where

$$L_{t,x}(\cdot) \equiv \frac{\partial}{\partial t}(\cdot) + \mathbf{E}[V(x)] \frac{\partial}{\partial x}(\cdot) - \mathbf{D}(t) \frac{\partial^2}{\partial x^2}(\cdot)$$

$$R(\cdot) \equiv \frac{\partial}{\partial x}(\cdot)$$

6.2 Homogeneous Chaos

Clearly, the Karhunen-Loève expansion requires a knowledge of the covariance function. This can be done for the random coefficients in the operator equation, but not for the solution process since its covariance function is not known. What is needed is a way of representing the solution process that does not require knowledge of its covariance function. In order to circumvent this problem, the Homogeneous Chaoses first introduced by Wiener in 1938 can be used. First, the following results are provided for convenience.

Definition: Let D be a subset of \mathfrak{R}^n . Let X denote the complete inner product space of functions defined on D . A function of two variables \vec{x}_1 and \vec{x}_2 in D , $K(\vec{x}_1, \vec{x}_2)$ is called a *Reproducing Kernel Function* for the space X if

- for each fixed $\vec{x}_2 \in D$, $K(\vec{x}_1, \vec{x}_2)$ considered as a function of \vec{x}_1 is in X .
- for each function $f(\vec{x}_1) \in X$ and every point $\vec{x}_2 \in D$, the *reproducing property*

$$f(\vec{x}_1) = (f(\vec{x}_1), K(\vec{x}_1, \vec{x}_2))_X$$

where $(\cdot, \cdot)_X$ is the inner product on X and \vec{x}_2 is held constant.

The following results due to Aronszajn provide existence and uniqueness:

Theorem: A necessary and sufficient condition that X have a reproducing kernel function is that for each fixed $\vec{x}_2 \in D$, the linear functional

$$L(f) = f(\vec{x}_2)$$

is bounded

$$|L(f)| \leq c \|f\| \quad \forall f \in X$$

Theorem: If X possesses a reproducing kernel, it is unique.

Let $\xi(\vec{x}, \omega)$, $\vec{x} \in D$ be a real Gaussian process defined on a probability space (Ω, \mathcal{B}, P) . Define the σ -algebra \mathcal{B}^ξ to be the completion with respect to the measure P of

$$\sigma\{\xi(\vec{x}) : \vec{x} \in D\}$$

where $\sigma\{\xi(\vec{x}) : \vec{x} \in D\}$ denotes the smallest σ -algebra with respect to which the $\xi(\vec{x})$ are measurable.

Assume that

$$\mathbf{E}[\xi(\vec{x})] = 0$$

and

$$C(\vec{x}_1, \vec{x}_2) = \mathbf{E}[\xi(\vec{x}_1)\xi(\vec{x}_2)]$$

Then, $H(C)$ is the Hilbert space determined by the kernel $C(\vec{x}_1, \vec{x}_2)$. $H(C)$ is called the *Reproducing Kernel Hilbert Space* (RKHS) of $C(\vec{x}_1, \vec{x}_2)$ or of the process $\xi(\vec{x})$. For each $\vec{x} \in D$, $C(\cdot, \vec{x}_2) \in H(C)$ and

$$(f, C(\cdot, \vec{x}_2))_{H(C)} = f(\vec{x}_2) \quad \forall f \in H(C)$$

The following notation for tensor products is required for what follows. Let $H_i, i = 1, \dots, p$ be Hilbert spaces, then the tensor product Hilbert space is given by $H_1 \otimes \dots \otimes H_p$ or the shorthand

$$H_1 \otimes \dots \otimes H_p \equiv \otimes^p H \quad \text{if } H_i = H, \quad i = 1, \dots, p$$

Let $\{e_i\}_{i=1}^\infty$ be a complete orthonormal sequence in $H(C)$. Also, let $L_1(\xi)$ be the closed linear subspace of $L_2(\Omega, \mathcal{B}^\xi, P)$ spanned by all finite, real linear combinations

$$\sum_{i=1}^n c_i \xi(\vec{x}_i)$$

then it can be shown, Kallianpur[56], that there is an isometric isomorphism between $L_1(\xi)$ and $H(C)$. Clearly, $e_{i_1} \otimes \dots \otimes e_{i_p}$ as $i_1 \dots i_p$ range independently from 1 to ∞ form a complete orthonormal system for $\otimes^p H(C)$. And, if $h_1 \otimes \dots \otimes h_p \in \otimes^p H$, then the tensor

$$\sigma(h_1 \otimes \cdots \otimes h_p) = \frac{1}{p!} \sum_{\pi} h_{\pi_1} \otimes \cdots \otimes h_{\pi_p}$$

where $\pi = (\pi_1, \dots, \pi_p)$ is a permutation of the integers $(1, \dots, p)$, is a completely symmetric tensor.

The Hilbert space $\sigma(\otimes^p H)$ is then taken to be the closed linear subspace of $\otimes^p H$ generated by elements of the form

$$\sum_{k=1}^n c_k \sigma(h_1^k \otimes \cdots \otimes h_p^k)$$

Let $H = H(C)$ be the RKHS of the Gaussian process $\xi(\vec{x}, \omega)$, $\vec{x} \in D$. Furthermore, let $\{\xi_i\}_{i=1}^{\infty}$ be a sequence of random variables over $(\Omega, \mathcal{B}^{\xi}, P)$ such that ξ_i is the element in $L_1(\xi)$ which corresponds to e_i in $H(C)$ by the isometry above. Then, from Kallianpur[56, Lemma 6.5.1], a complete orthonormal system denoted $\{e_{\lambda_1, \dots, \lambda_r}^{n_1, \dots, n_r}\}$ exists for the space $\sigma(\otimes^p H(C))$ where $\lambda_1, \dots, \lambda_r$ are distinct integers in the sequence i_1, \dots, i_p with λ_i occurring n_i times and $\sum_{i=1}^r n_i = p$.

Next, define the following linear subspaces of $L_2(\Omega, \mathcal{B}^{\xi}, P)$. First, define $\hat{?}_p$ to be the space of all polynomials in $\{\xi_i(\omega)\}_{i=1}^{\infty}$ of degree not exceeding p . Then let $?_p$ to be the set of all polynomials in $\hat{?}_p$ orthogonal to $\hat{?}_{p-1}$, sometimes written $?_p = \hat{?}_p \ominus \hat{?}_{p-1}$. Finally, let $\bar{?}_p$ be the space spanned by $?_p$. The subspace $\bar{?}_p$ of $L_2(\Omega, \mathcal{B}^{\xi}, P)$ is called the p^{th} *Homogeneous Chaos*. The set $?_p$ is called the *Polynomial Chaos of Order p*.

It can then be shown, Kallianpur[56, Lemma 6.6.1], that the following representation of elements of $?_p$ holds:

Lemma: A random variable γ belongs to $?_p$ ($p > 1$) iff it is of the form

$$\gamma(\omega) = \sum a_{m_1, \dots, m_r} h_{m_1}[\xi_{\lambda_1}(\omega)] \cdots h_{m_r}[\xi_{\lambda_r}(\omega)]$$

for some choice of distinct integers $\lambda_1, \dots, \lambda_r$. The summation is over $m_i \geq 0$ and $\lambda_1, \dots, \lambda_r$ are fixed, $h_n(x)$ is the n^{th} normalized Hermite polynomial and the coefficients satisfy

$$a_{m_1, \dots, m_r} = 0 \quad \text{if} \quad \sum_{i=1}^r m_i \neq p$$

The desired homogeneous chaos decomposition of the L_2 -space of a Gaussian process is given in the following theorem from Kallianpur[56, Theorem 6.6.1]:

Theorem:

$$L_2(\Omega, \mathcal{B}^\xi, P) = \sum_{p \geq 0} \oplus \overline{?}_p$$

For any $u \in L_2(\Omega, \mathcal{B}^\xi, P)$, the expansion

$$u = \sum_{p \geq 0} \sum_{n_1 + \dots + n_p = p} \sum_{\lambda_1 < \dots < \lambda_r} \left(F_p, e_{\lambda_1, \dots, \lambda_r}^{n_1, \dots, n_r} \right)_p h_{n_1}[\xi_{\lambda_1}(\omega)] \cdots h_{n_r}[\xi_{\lambda_r}(\omega)] \quad (6.1)$$

holds where $F_p \in \sigma(\otimes^p H(C))$.

Polynomial chaoses of any order p consist of all orthogonal polynomials of order p involving any combination of the random variables $\{\xi_i(\omega)\}_{i=1}^\infty$. Ghanem and Spanos[47, 46] rewrite the previous expression for $u(\omega) \in L_2(\Omega, \mathcal{B}^\xi, P)$ as

$$\begin{aligned} u(\omega) &= a_{i_0} ?_0 + \sum_{i_1=1}^\infty a_{i_1} ?_1(\xi_{i_1}(\omega)) \\ &+ \sum_{i_1=1}^\infty \sum_{i_2=1}^{i_1} a_{i_1 i_2} ?_2(\xi_{i_1}(\omega), \xi_{i_2}(\omega)) \\ &+ \sum_{i_1=1}^\infty \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} a_{i_1 i_2 i_3} ?_3(\xi_{i_1}(\omega), \xi_{i_2}(\omega), \xi_{i_3}(\omega)) + \dots \end{aligned}$$

Each polynomial chaos is a function of the countably infinite set $\{\xi_i(\omega)\}$ and is therefore an infinite dimensional polynomial. For practical purposes, this must be reduced to a finite dimensional subspace. The n -dimensional polynomial chaos of order p is the subset of the polynomial chaos of order p which involves only n of the uncorrelated random variables $\xi_i(\omega)$. The convergence properties will then depend on n and the choice of the subset $\{\xi_{\lambda_i}\}_{i=1}^n$.

In Ghanem and Spanos[47, 46] the polynomial chaoses for orders 0 to 3 are found to be

$$\xi_0 = 1$$

$$\xi_1(\xi_i) = \xi_i$$

$$\xi_2(\xi_{i_1}, \xi_{i_2}) = \xi_{i_1}\xi_{i_2} - \delta_{i_1 i_2}$$

$$\xi_3(\xi_{i_1}, \xi_{i_2}, \xi_{i_3}) = \xi_{i_1}\xi_{i_2}\xi_{i_3} - \xi_{i_1}\delta_{i_2 i_3} - \xi_{i_2}\delta_{i_1 i_3} - \xi_{i_3}\delta_{i_1 i_2}$$

where δ_{ij} is the Kronecker delta.

6.3 Stochastic Finite Elements

Starting from the equation

$$L_{t,x}u + \sum_{k=1}^m (\lambda_k)^{\frac{1}{2}} \xi_k(\omega) \phi_k(x) Ru \approx \hat{f} \quad (6.2)$$

with boundary conditions of the form

$$\Sigma(x)u(t, x, \omega) = 0 \quad x \in \partial D$$

assume that for a fixed t , $u(t, x, \omega)$ is a second order random variable, hence its Karhunen-Loève expansion can be expressed as

$$u(t, x, \omega) = \sum_{j=1}^l e_j(t) \chi_j(t, \omega) b_j(t, x) \quad (6.3)$$

where

$$\chi_j(t, \omega) = \frac{1}{e_j(t)} \int_D u(t, x, \omega) b_j(t, x) dx$$

Since at time t the covariance function is not known, Equation[6.3] is of little use in this form. However, using the polynomial chaoses, the random variable $\chi(t, \omega)$ can be expressed as

$$\begin{aligned} \chi_j(t, \omega) &= a_{i_0}^{(j)}(t) \xi_0 + \sum_{i_1=1}^{\infty} a_{i_1}^{(j)}(t) \xi_1(\xi_{i_1}) \\ &+ \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} a_{i_1 i_2}^{(j)}(t) \xi_2(\xi_{i_1}, \xi_{i_2}) + \dots \end{aligned}$$

where $a_{i_1 \dots i_p}^{(j)}(t)$ are *deterministic* constants and $\Psi_p(\xi_{i_1}, \dots, \xi_{i_p})$ is the p^{th} order polynomial chaos. Truncating after the P^{th} polynomial gives

$$\chi_j(t, \omega) = \sum_{i=0}^P x_i^{(j)}(t) \Psi_j[\{\xi_r\}] \quad (6.4)$$

where

$$x_0^{(j)}(t) = a_{i_0}^{(j)}(t); \quad x_1^{(j)}(t) = a_{i_1}^{(j)}(t); \quad x_2^{(j)}(t) = a_{i_1 i_2}^{(j)}(t); \quad \dots$$

and

$$\Psi_0[\{\xi_r\}] = \Psi_0; \quad \Psi_1[\{\xi_r\}] = \Psi_1(\xi_{i_1}); \quad \Psi_2[\{\xi_r\}] = \Psi_2(\xi_{i_1}, \xi_{i_2}); \quad \dots$$

Substituting Equation[6.4] into Equation[6.3] and letting $c_j(t, x) = e_j(t)b_j(t, x)$ yields

$$\begin{aligned} u(t, x, \omega) &= \sum_{j=1}^l \sum_{i=0}^P x_i^{(j)}(t) \Psi_i[\{\xi_r\}] c_j(t, x) \\ &= \sum_{i=0}^P d_i(t, x) \Psi_i[\{\xi_r\}] \end{aligned} \quad (6.5)$$

where

$$d_i(t, x) = \sum_{j=1}^l x_i^{(j)}(t) c_j(t, x)$$

Then, substituting Equation[6.5] into Equation[6.2] gives

$$\left[L_{t,x} + \sum_{k=1}^m (\lambda_k)^{\frac{1}{2}} \xi_k(\omega) \phi_k(x) R \right] \sum_{i=0}^P d_i(t, x) \Psi_i[\{\xi_r\}] \approx \hat{f} \quad (6.6)$$

Next, expanding $d_i(t, x)$ in the space C^2 as

$$d_i(t, x) \approx \sum_{j=1}^n d_{ji}(t) g_j(x) \quad (6.7)$$

and substituting it into Equation[6.6] results in

$$\sum_{i=0}^P \sum_{j=1}^n \Psi_i[\{\xi_r\}] L_{t,x} d_{ji}(t) g_j(x) + \sum_{i=0}^P \sum_{k=1}^m (\lambda_k)^{\frac{1}{2}} \xi_k(\omega) \Psi_i[\{\xi_r\}] \sum_{j=1}^n d_{ji}(t) R g_j(x) \phi_k(x) \approx \hat{f}$$

The terms can be rearranged to form the following

$$\sum_{i=0}^P \sum_{j=1}^n \left[\Psi_i[\{\xi_r\}] L_{t,x} (d_{ji}(t) g_j(x)) + d_{ji}(t) \sum_{k=1}^m (\lambda_k)^{\frac{1}{2}} \xi_k(\omega) \Psi_i[\{\xi_r\}] R (g_j(x)) \phi_k(x) \right] \approx \hat{f}$$

Multiplying both sides by $g_l(x)$ and integrating over D gives the system of equations

$$\sum_{i=0}^P \sum_{j=1}^n \left[\Psi_i[\{\xi_r\}] \left\{ \frac{\partial d_{ji}(t)}{\partial t} L_{jl}^{(1)} + d_{ji}(t) L_{jl}^{(2)} \right\} + d_{ji}(t) \sum_{k=1}^m (\lambda_k)^{\frac{1}{2}} \xi_k(\omega) \Psi_i[\{\xi_r\}] R_{ikl} \right] \approx f_l$$

where $l = 1, \dots, n$ and

$$L_{jl}^{(1)} = \int_D g_j(x) g_l(x) dx$$

$$L_{jl}^{(2)} = \int_D \left[\mathbf{E}[V(x)] \frac{\partial g_j(x)}{\partial x} - \mathbf{D}(t) \frac{\partial^2 g_j(x)}{\partial x^2} \right] g_l(x) dx$$

$$R_{ikl} = \int_D R(g_i(x)) g_l(x) \phi_k(x) dx$$

$$f_l = \int_D \hat{f} g_l(x) dx$$

In this system of equations, i spans the number of polynomial chaoses and j spans the number of basis functions from C^2 . In order to find the vector of entries $\{d_{ji}(t)\}$ at time t , the polynomial chaoses have to be replaced with numbers. This is done by multiplying by $\Psi_m[\{\xi_r\}]$, taking expected values and using the orthonormal relationships

$$\mathbf{E}[\Psi_j[\{\xi_r\}] \Psi_m[\{\xi_r\}]] = \delta_{jm}$$

and the computable quantities

$$\mathbf{E}[\xi_k(\omega) \Psi_j[\{\xi_r\}] \Psi_m[\{\xi_r\}]]$$

This yields an $nP \times nP$ system. Once this system is solved for the vector $\{d_{ji}(t)\}$, the $d_i(t, x)$ coefficients can be calculated from Equation[6.7], and $u(t, x, \omega)$ can then be subsequently represented by Equation[6.5].

6.4 The Covariance Function

The success of the procedure outlined above depends in large part in describing the covariance function and obtaining its eigenvalues and eigenfunctions. The problems of flow and transport as presented in this thesis assume a knowledge of the porous medium that is very difficult to obtain in actual practice. For that reason, special assumptions regarding the stochastic properties of the hydraulic conductivities are often made. In particular, Equation[1.21], page 43, relates the spectrum of the velocity to the spectrum of the log-hydraulic conductivity. By taking Fourier transforms, a relationship between the covariance of the velocity and the covariance of the log-hydraulic conductivity can then be established. The problem is then in choosing a spectrum that can be representative of the log-hydraulic conductivity. In order to illustrate the procedure, one choice that has been used is the *Whittle* spectrum for two-dimensional spatial processes. This spectrum has the form, Mizell[67],

$$S(\vec{k}) = \frac{\sigma^2 \alpha^2}{\pi (k_1^2 + k_2^2 + \alpha^2)^2}; \quad \vec{k} = (k_1, k_2)^\dagger; \quad \alpha = \frac{\pi}{2\lambda}$$

where λ is the integral scale and σ^2 the variance of the process. Graham and McLaughlin[48] use a slightly modified form of this spectrum called Spectrum A by Mizell[67] which has the form

$$S(\vec{k}) = \frac{2\sigma^2 \alpha^2 (k_1^2 + k_2^2)}{\pi (k_1^2 + k_2^2 + \alpha^2)^3}$$

Substituting this for $S_{Y'Y'}(\vec{k})$ in Equation[1.21], page 43, gives an expression for the velocity spectrum of

$$S_{\vec{q}\vec{q}}(\vec{k}) \approx (\mathbf{I} - \|\vec{k}\|^{-2} \vec{k}\vec{k}^\dagger) \mathbf{E}[\vec{q}] \mathbf{E}[\vec{q}]^\dagger (\mathbf{I} - \|\vec{k}\|^{-2} \vec{k}\vec{k}^\dagger) \frac{2\sigma^2 \alpha^2 (k_1^2 + k_2^2)}{\pi (k_1^2 + k_2^2 + \alpha^2)^3}$$

In order to obtain the velocity covariance matrix, the Fourier transform of this expression must be taken, which yields for this two dimensional example

$$\begin{aligned} C(\vec{x}_1, \vec{x}_2) &= R_{\vec{q}\vec{q}}(\vec{\xi}) \\ &\approx \int_{\mathbb{R}^2} (\mathbf{I} - \|\vec{k}\|^{-2} \vec{k}\vec{k}^\dagger) \mathbf{E}[\vec{q}] \mathbf{E}[\vec{q}]^\dagger (\mathbf{I} - \|\vec{k}\|^{-2} \vec{k}\vec{k}^\dagger) \frac{2\sigma^2 \alpha^2 (k_1^2 + k_2^2)}{\pi (k_1^2 + k_2^2 + \alpha^2)^3} e^{i\vec{\xi}\cdot\vec{k}} d\vec{k} \end{aligned}$$

where as described in Section 1.6, the vector $\vec{\xi}$ is the separation vector extending from \vec{x}_1 to \vec{x}_2 . This particular 2×2 symmetric velocity covariance matrix

is evaluated in Graham and McLaughlin[48]. As noted in Section 1.6, this relationship is the result of the first order approximation Equation[1.19], page 42. However, higher order estimates have been calculated, one of the more recent studies in this area is given in Deng and Cushman[35]. Also, instead of analytical representations of the velocity covariance, numerical estimates can be established by Monte Carlo methods.

In order to obtain estimates of the eigenvalues and eigenfunctions for the problem

$$\lambda\phi(\vec{x}_1) = \int_D C(\vec{x}_1, \vec{x}_2)\phi(\vec{x}_2)d\vec{x}_2$$

a Galerkin approach is used. Let $\{h_i(\vec{x})\}$ be a complete set of functions in the space of square integrable functions on D . Each eigenfunction of the kernel $C(\vec{x}_1, \vec{x}_2)$ can be approximated by the finite sum

$$\phi_k(\vec{x}) \approx \sum_{i=1}^N \gamma_i^{(k)} h_i(\vec{x})$$

On substituting this finite sum into the preceding equation produces an error. The error, ϵ_N , is then assumed to be orthogonal to the approximating subspace, or equivalently,

$$(\epsilon_N, h_j(\vec{x})) = \int_D \epsilon_N h_j(\vec{x})d\vec{x} = 0 \quad j = 1, \dots, N$$

which yields the following system of equations:

$$0 = \sum_{i=1}^N \gamma_i^{(N)} \left[\int_D \left[\int_D C(\vec{x}_1, \vec{x}_2) h_i(\vec{x}_2) d\vec{x}_2 \right] h_j(\vec{x}_1) d\vec{x}_1 - \lambda_k^{(N)} \int_D h_i(\vec{x}_1) h_j(\vec{x}_1) d\vec{x}_1 \right]$$

$$j = 1, 2, \dots, N$$

Letting

$$C_{ij} = \int_D \int_D C(\vec{x}_1, \vec{x}_2) h_i(\vec{x}_2) h_j(\vec{x}_1) d\vec{x}_1 d\vec{x}_2$$

$$B_{ij} = \int_D h_i(\vec{x}_1) h_j(\vec{x}_1) d\vec{x}_1$$

$$D_{ij} = d_i^{(j)}$$

$$\Lambda_{ij} = \delta_{ij} \lambda_i^{(N)}$$

yields the system

$$(C - \Lambda B)D = 0$$

which is a generalized eigenvalue-eigenfunction problem. If the $\{h_i(\vec{x})\}$ is an orthonormal sequence, then an ordinary eigenvalue-eigenfunction problem is the result.

REFERENCES

- [1] Adams, R. A., *Sobolev Spaces*, Academic Press, 1975
- [2] Abraham, R., Marsden, J. E. and Ratiu, T., *Manifolds, Tensor Analysis, and Applications*, Addison-Wesley, 1983
- [3] Adomian, George, *Stochastic Systems*, Academic Press, 1983
- [4] Adomian, George and Malakian, K., Operator-Theoretic Solution Of Stochastic Systems, *Journal Of Mathematical Analysis and Applications*, 76, 183-201, 1980
- [5] Adomian, George and Malakian, K., Inversion Of Stochastic Partial Differential Operators - The Linear Case, *Mathematical Analysis and Applications*, 77, 505-512, 1980
- [6] Ash, Robert B., *Real Analysis and Probability*, Academic Press, 1972
- [7] Åström, K.J. and Wittenmark, B., *Computer Controlled Systems*, Prentice-Hall, 1984
- [8] Åström, K.J., *Introduction To Stochastic Control Theory*, Academic Press, 1970
- [9] Batchelor, G. K., *An Introduction To Fluid Dynamics*, Cambridge University Press, 1967
- [10] Bear, Jacob, *Hydraulics Of Groundwater*, McGraw-Hill, 1979
- [11] Bear, Jacob and Verruijt, Arnold, *Modeling Groundwater Flow And Pollution*, D. Reidel Publishing Company, 1987
- [12] Bear, Jacob, *Dynamics of Fluids in Porous Media*, American Elsevier Publishing, New York, 1972
- [13] Becus, G. A., Variational Formulations Of Some Problems For The Random Heat Equation, *Applied Stochastic Processes*, Academic Press, 1980

- [14] Bensoussan, A., Control Of Stochastic Partial Differential Equations, In *Distributed Parameter Systems*, Edited By W.H. Ray and D.G. Lainiotis, Marcel Dekker, 209-246, 1979
- [15] Brenner, S. C. and Scott, L. R., *The Mathematical Theory Of Finite Element Methods*, Springer-Verlag, 1994
- [16] Brezzi, F., *On The Existence, Uniqueness And Approximation Of Saddle Point Problems Arising From Lagrangian Multipliers*, RAIRO Anal. Numer., 2(1974), pp. 129-151
- [17] Burrill, Claude W., *Measure, Integration and Probability*, McGraw-Hill, 1972
- [18] Butzer, P.L. and Berens, H., *Semi-Groups of Operators and Approximations*, Springer-Verlag, 1967
- [19] Chin, D. A. and Wang, T., *An Investigation of the Validity of First-Order Stochastic Dispersion Theories in Isotropic Porous Media*, Water Resources Research, Vol. 28, No. 6, pp. 1531-1542, 1992
- [20] Chin, D. A., *Estimation of Dispersion Coefficients From Particle Paths*, Journal of Environmental Engineering, Vol 120, No. 1, 1994
- [21] Chow, Pao-Liu, Stochastic Partial Differential Equations In Turbulence Related Problems, In *Probabilistic Analysis And Related Topics*, Edited By A. T. Bharucha-Reid, Academic Press, 1978
- [22] Chow, Pao-Liu, Approximate Solution Of Random Evolution Equations, In *Approximate Solutions Of Random Equations*, Edited By A.T. Bharucha-Reid, Elsevier North Holland, 1979
- [23] Chung, Kai Lai, *A Course In Probability Theory*, Second Edition, Academic Press, 1974
- [24] Ciarlet, P. G. and Lions, J. L., Editors, *Handbook Of Numerical Analysis*, Vol II, Part 1, Finite Element Methods, North-Holland, 1991
- [25] Curtain, R.F. and Pritchard, A.J., *Functional Analysis In Modern Applied Mathematics*, Academic Press, 1977
- [26] Curtain, R.F. and Falb, P.L., Itô's Lemma In Infinite Dimensions, *Journal Of Mathematical Analysis And Applications*, 31, 434-448, 1970

- [27] Curtain, R.F. and Falb, P.L., Stochastic Differential Equations In Hilbert Space, *Journal Of Differential Equations*, 10, 412-430, 1971
- [28] Curtain, R. F., Stochastic Evolution Equations with General White Noise Disturbance, *Journal Of Mathematical Analysis And Applications*, 60, 570-595, 1977
- [29] Dagan, Gedeon, *Flow and Transport in Porous Formations*, Springer-Verlag, 1989
- [30] Dagan, Gedeon, *Dispersion of a passive solute in non-ergodic transport by steady velocity fields in heterogeneous formations*, *J. Fluid Mech.* (1991), vol. 233, pp. 197-210
- [31] Dagan, Gedeon, *Theory Of Solute Transport By Groundwater*, *Ann. Rev. Fluid Mech.*, 1987, 19, pp. 183-215
- [32] Dagan, Gedeon, *Upscaling Of Dispersion Coefficients in Transport through Heterogeneous Formations*, Prepared for the GXG Meeting, Sandia National Laboratory, March 1993
- [33] Dagan, Gedeon, *Solute transport in heterogeneous porous formations*, *J. Fluid Mech.* (1984), vol. 145, pp. 151-177
- [34] Dagan, Gedeon, *Transport in Heterogeneous Porous Formations: Spatial Moments, Ergodicity, and Effective Dispersion*, *Water Resources Research*, Vol. 26, No. 6, pp. 1281-1290, 1990
- [35] Deng, F. W. and Cushman, J. H., In Higher-Order Corrections to the Flow Velocity Covariance Tensor, *Water Resources Research*, 31(7), 1659-1672, July 1995
- [36] Doob, J. L., *Stochastic Processes*, John Wiley & Sons, 1953
- [37] Falb, P. L., Infinite-Dimensional Filtering: The Kalman-Bucy Filter In Hilbert Space, *Information And Control*, 11, 102-137, 1967
- [38] Flanders, H., *Differential Forms With Applications To The Physical Sciences*, Academic Press, 1963
- [39] Friedman, A., *Partial Differential Equations*, Holt, Reinhart, Winston, 1969

- [40] Friedman, A., *Stochastic Differential Equations And Applications, v. 1*, Academic Press, 1975
- [41] Fuhrmann, P. A., *Linear Systems and Operations In Hilbert Space*, McGraw-Hill, 1981
- [42] Gardiner, C. W., *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Second Edition, Springer-Verlag, Third printing, 1994
- [43] Gelb, Arthur, Editor, *Applied Optimal Estimation*, The M.I.T. Press, 1984
- [44] Gelhar, Lynn W., *Stochastic Subsurface Hydrology*, Prentice Hall, 1993
- [45] Gelhar, L.W. and Axness, C.L., Three-Dimensional Stochastic Analysis Of Macrodispersion In Aquifers, *Water Resour. Res.*, 19(1), 161-180, 1983
- [46] Ghanem, R. G. and Spanos, P. D., *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, 1991
- [47] Ghanem, R. G. and Spanos, P. D., Polynomial Chaos in Stochastic Finite Elements, *Journal of Applied Mechanics, ASME*, Vol. 57, No. 1, 197-202, March 1990
- [48] Graham, W. and McLaughlin, D., Stochastic Analysis Of Nonstationary Subsurface Solute Transport, 1. Unconditional Moments, *Water Resour. Res.*, 25(2), 215-232, 1989
- [49] Gray, William G., *A Derivation Of The Equations For Multi-phase Transport*, Chemical Engineering Science, 1975, Vol. 30, pp. 229-233
- [50] Greenberg, M.D., *Foundations Of Applied Mathematics*, Prentice Hall, 1978
- [51] Guenther, Ronald B. and Lee, John W., *Partial Differential Equations Of Mathematical Physics And Integral Equations*, Prentice Hall, 1988
- [52] Hastings Jr., C., *Approximations For Digital Computers*, Princeton University Press, 1955
- [53] Hochstadt, Harry, *Integral Equations*, Wiley-Interscience, 1973

- [54] Jazwinski, A. H., *Stochastic Processes And Filtering Theory*, Academic Press, 1970
- [55] Kadison, R.V. and Ringrose, J.R., *Fundamentals Of The Theory Of Operator Algebras*, Volume 1, Academic Press, 1983
- [56] Kallianpur, G., *Stochastic Filtering Theory*, Springer-Verlag, 1980
- [57] Karatzas, Ioannis and Shreve, Steven E., *Brownian Motion and Stochastic Calculus*, Second Edition, Springer-Verlag, 1991
- [58] Kitanidis, P. K., Prediction By The Method Of Moments of Transport In A Heterogeneous Formation, *Journal Of Hydrology*, 102, pp 453-473, 1988
- [59] Kloeden, P. E., Platen, E. and Schurz, H., *Numerical Solution Of SDE Through Computer Experiments*, Springer-Verlag, 1991
- [60] Ladas, G. E. and Lakshmikantham, V., *Differential Equations in Abstract Spaces*, Academic Press, 1972
- [61] Law, A. M. and Kelton, W. D., *Simulation Modeling and Analysis*, McGraw-Hill, 1982
- [62] Li, Wen-Hsiung, *Differential Equations of Hydraulic Transients, Dispersion, and Groundwater Flow*, Prentice Hall, 1972
- [63] Loève, M., *Probability Theory II*, Fourth Edition, Springer-Verlag, 1977
- [64] Lumley, J.L. and Panofsky, H.A., *The Structure Of Atmospheric Turbulence*, John Wiley, New York, 1964
- [65] Lundgren, T. S. and Pointin, Y. B., Turbulent Diffusion, *Physics of Fluids*, 19(3), 355-358, March 1976
- [66] Marlow, W. H., *The Mathematics For Operations Research*, John Wiley & Sons, 1978
- [67] Mizell, S. A., Gutjahr, A. L. and Gelhar, L. W., Stochastic Analysis of Spatial Variability in Two-Dimensional Steady Groundwater Flow Assuming Stationary and Nonstationary Heads, *Water Resources Research*, 18(4), 1053-1067, August 1982

- [68] Monin, A. S. and Yaglom, A. M., *Statistical Fluid Mechanics: Mechanics of Turbulence*, Vol 2, The MIT Press, 1975
- [69] Neuman, S.P., Winter, C.L., and Newman, C.M., Stochastic Theory Of Field-Scale Fickian Dispersion In Anisotropic Porous Media, *Water Resour. Res.*, 23(3), 453-466, 1987
- [70] Neuman, S.P. and Zhang, You-Kuan, A Quasi-Linear Theory Of Non-Fickian And Fickian Subsurface Dispersion, 1. Theoretical Analysis With Application To Isotropic Media, *Water Resour. Res.*, 26(5), 887-902, 1990
- [71] Neuman, S.P., Universal Scaling Of Hydraulic Conductivities And Dispersivities In Geologic Media, *Water Resour. Res.*, 26(8), 1749-1758, 1990
- [72] Oden, J. T. and Reddy, J. N., *An Introduction To The Mathematical Theory Of Finite Elements*, John Wiley and Sons, 1976
- [73] Ortega, James M., *Matrix Theory*, Plenum Press, 1989
- [74] Pazy, A., *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983
- [75] Papoulis, A., *Probability, Random Variables And Stochastic Processes*, McGraw-Hill, 1984
- [76] Pinder, G. F. and Gray, W. G., *Finite Element Simulation In Surface And Subsurface Hydrology*, Academic Press, 1977
- [77] Raviart, P. A. and Thomas, J. M., *A Mixed Finite Element Method For 2nd Order Elliptic Problems*, in Mathematical Aspects of the Finite Element Method, Rome, 1975, Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1977
- [78] Reddy, J. N., *An Introduction To The Finite Element Method*, McGraw-Hill, 1984
- [79] Riesz, F. and Sz.-Nagy, B., *Functional Analysis*, Ungar, New York, 1955
- [80] Robin, M. J. L., Gutjahr, A. L., Sudicky, E. A., and Wilson, J. L., Cross-Correlated Random Field Generation With The Direct Fourier Transform Method, *Water Resources Research*, 29(7), 2385-2397, 1993

- [81] Royden, H. L., *Real Analysis*, Second Edition, The Macmillan Company, New York, 1968
- [82] Rubin, Y., Stochastic Modeling Of Macrodispersion In Heterogeneous Porous Media, *Water Resour. Res.*, 26(1), 133-141, 1990
- [83] Rudin, W., *Functional Analysis*, McGraw-Hill, 1991
- [84] Russell, T. F. and Wheeler, M. F., *Finite Element and Finite Difference Methods for Continuous Flows in Porous Media*, in The Mathematics Of Reservoir Simulation, SIAM, Philadelphia, 1983
- [85] Sawaragi, Y., Soeda, T., and Omatu, S., Modeling, Estimation and Their Applications For Distributed Parameter Systems, *Lecture Notes In Control And Information Sciences*, Springer-Verlag, 1978
- [86] Serrano, S.E. and Unny, T.E., Boundary Element Solution Of The Two-Dimensional Groundwater Flow Equation With Stochastic Free Surface Boundary Condition, *Numerical Methods For Partial Differential Equations*, 2, 251-258, 1986
- [87] Serrano, S.E. and Unny, T.E., Predicting Groundwater Flow In A Phreatic Aquifer, *Journal Of Hydrology*, 95, 241-268, 1987
- [88] Serrano, S.E., General Solution To Random Advective-Dispersive Equation In Porous Media, Part 1, *Stochastic Hydrology And Hydraulics*, 2, 79-98, 1988
- [89] Serrano, S.E., General Solution To Random Advective-Dispersive Equation In Porous Media, Part 2, *Stochastic Hydrology And Hydraulics*, 2, 79-98, 1988
- [90] Serrano, S.E., Unny, T.E., Lennox, W.C., Analysis Of Groundwater Flow Problems, Parts I, II and III, *Journal Of Hydrology*, 82, 247-306, 1985
- [91] Showalter, R. E., *Hilbert Space Methods For Partial Differential Equations*, Pitman Press, 1977
- [92] Spanos, P. D. and Ghanem, R., Stochastic Finite Element Expansion For Random Media, *Journal of the Engineering Mechanics Division, ASCE*, Vol. 115, No. 5, 1035-1053, May 1989

- [93] Springer, M. D., *The Algebra Of Random Variables*, Wiley-Interscience, 1979
- [94] Tanabe, H., *Equations Of Evolution*, Pitman Press, 1979
- [95] Tang, D.H. and Pinder, G.F., Analysis Of Mass Transport With Uncertain Parameters, *Water Resources Research*, 15(5), 1147-1155, 1979
- [96] Todorovic, P., *An Introduction To Stochastic Processes And Their Applications*, Springer-Verlag, 1992
- [97] Walker, James S., *Fourier Analysis*, Oxford University Press, 1988
- [98] Wentzell, A.D., *A Course In The Theory Of Stochastic Processes*, McGraw-Hill, 1981
- [99] Yosida, K., *Functional Analysis*, Sixth Edition, Springer-Verlag, 1980