

WRC RESEARCH REPORT NO. 77

ANALYSIS OF LIQUID-WASTE INJECTION WELLS  
IN ILLINOIS BY MATHEMATICAL MODELS

By

Manoutchehr Heidari

and

Keros Cartwright

ILLINOIS STATE GEOLOGICAL SURVEY

and

Paul E. Saylor

DEPARTMENT OF COMPUTER SCIENCE

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN  
Urbana, Illinois

F I N A L R E P O R T  
Project A-058-ILL

The work upon which this publication is based was supported by funds provided by the U. S. Department of the Interior as authorized under the Water Resources Research Act of 1964, P.L. 880379 Agreement No. 14-31-0001-3813

UNIVERSITY OF ILLINOIS  
WATER RESOURCES CENTER  
2535 Hydrosystems Laboratory  
Urbana, Illinois 61801

February 1974

## ABSTRACT

## ANALYSIS OF LIQUID-WASTE INJECTION WELLS IN ILLINOIS BY MATHEMATICAL MODELS

This report contains the results of a preliminary theoretical study of the fate of liquid industrial wastes injected into deep geologic formations. The Jones and Laughlin Corporation well was used as a model and the geology of the area was idealized into a 15-layered homogeneous and anisotropic mathematical model. The finite element method was tested and proved to be an effective mathematical tool in the solution of the equation of flow. The flow and pressure build-up show that the rocks are capable of receiving greater volumes of waste than are now being injected without endangering the integrity of the aquifer or the confining layer.

The mass-transport equation for large and complex ground-water reservoir systems was investigated, and it was concluded that the dispersion and diffusion parts of the equation are relatively insignificant, and under extreme conditions the dispersed zone will not be more than a few feet wide. Therefore, it was concluded that a more practical approach to the problem would be the solution of a system with a moving interface boundary in which mass transport results mainly from convection.

To overcome difficulties encountered with computer time and memory in the solution of the mass-transport equation for large complex systems, an iterative method is proposed for the solution of the equations, which substantially reduces these difficulties.

Heidari, M., Saylor, P., and Cartwright, K.

ANALYSIS OF LIQUID-WASTE INJECTION WELLS IN ILLINOIS BY MATHEMATICAL MODELS

University of Illinois Water Resources Center Report No. 77

KEYWORDS - mass-transport equation/ flow equation/ mathematical modeling/  
finite element method/ multi-layered systems/ solution of large set of linear  
equations/ adaptive-Chebyshev-factorization method

## CONTENTS

	page
Abstract -----	ii
List of Figures -----	v
List of Tables -----	ix
Acknowledgments -----	x
I. INTRODUCTION -----	1
A. Objectives -----	1
B. Background -----	1
C. Previous Works -----	2
II. SOLUTIONS OF MASS TRANSPORT AND FLOW EQUATIONS -----	10
A. Importance of the Dispersion and Diffusion in the Mass Transport Equation -----	11
B. Solution of Flow Equation for Multi-Layered Systems -----	18
1. Finite Element Method -----	18
2. Boundary Conditions -----	30
3. Implementation of the Finite Element Method -----	30
a. Properties of Matrices D and AA -----	30
b. Mesh Construction -----	32
c. Change of Time Increment -----	34
d. Solution of the Linear Equations -----	34
4. Test of the Computer Program -----	34
a. Homogeneous and Isotropic Systems -----	34
b. Layered Systems -----	38
i) Two-Layered System with Constant Potential Boundary at the Well -----	39

	page
ii) Three-Layered System with Constant Injection	
Rate Boundary Condition -----	39
c. Application of the Finite Element Method to a	
Field Problem -----	48
i) Geology -----	48
ii) Solution by Finite Element Method -----	60
III. AN ADAPTIVE-CHEBYSHEV-FACTORIZATION METHOD FOR THE SOLUTION OF LINEAR	
EQUATIONS ARISING FROM THE NUMERICAL SOLUTION OF FLOW EQUATION -----	73
A. Introduction -----	73
B. Factorization Methods -----	76
C. Adaptive Algorithms -----	79
D. Approximate Factorizations -----	80
E. The Finite-Difference Case -----	82
F. Approximate Factorization of Finite Element Matrices -----	88
G. The LU Factors of $A + B$ -----	94
H. Experimental Results -----	104
IV. REFERENCES -----	110

## LIST OF FIGURES

	page
Figure 1. Schematic column setup and typical measurements for an instantaneous point source injection -----	5
Figure 2. Schematic representation of the zone of mixed fluids or gradual boundary -----	6
Figure 3. Concentration distribution obtained with the above data (see Figure 3) and Hoopes and Harleman's approximate solution without molecular diffusion -----	13
Figure 4. Concentration distribution obtained with the above data (see Figure 4) and Hoopes and Harleman's approximate solution with molecular diffusion -----	14
Figure 5. Concentration distribution obtained with the above data (see Figure 5) and Hoopes and Harleman's approximate solution -----	17
Figure 6-a. Division of two-dimensional region into triangular elements -----	23
Figure 6-b. An axisymmetric element with constant triangular cross-section -----	23
Figure 7. Schematic representation of flow through node n -----	28
Figure 8. Schematic representation of the nodes which contribute to the elements $a_{nm}$ and $d_{nm}$ , where $m = i, i + 1, n - 1, n, n + 1, k - 1,$ and $k$ -----	31
Figure 9. Schematic representation of irregular elements -----	33
Figure 10-a. Results obtained by finite element method as compared with analytical solutions for impermeable and constant potential	

	page
outer boundary system, with well of zero radius and constant rate of injection -----	36
Figure 10b. Results obtained by finite element method as compared with analytical solutions for infinite radial system, with well of radius 0.25 ft. and constant rate of injection -----	37
Figure 11. Results of the finite element solution for a two-layered system adjacent to a pond with constant head and radius $r_w = 80.0$ ft. ( $r_w/H = 10.0$ , $r_e/r_w = 5.0$ , $H_2/H_1 =$ $1.0$ , and $K_2/K_1 = 2.0$ ) as compared with the analytical solutions -----	40
Figure 12. Schematic representation of a three-layered system -----	41
Figure 13. Results obtained by finite element method as compared with analytical solutions for a three-layered leaky system with $\beta_{12} = \beta_{32} = 0.1$ and constant injection rate -----	45
Figure 14. Results obtained by finite element method as compared with analytical solutions for a three-layered leaky system with $\beta_{12} = \beta_{32} = 1.0$ with constant injection rate -----	46
Figure 15. Results obtained by finite element method as compared with analytical solutions for a three-layered leaky system with $\beta_{12} = \beta_{32} = 1.0$ and constant injection rate -----	47
Figure 16. Results obtained by finite element method for a three- layered system with $\beta_{12} = \beta_{32} = 0.01$ , constant injection rate, and decay period timed continuously after build-up period -----	49

	page
Figure 17. Approximate contours of equal $\Delta h$ for the three-layered leaky system at the end of build-up period -----	50
Figure 18. Approximate contours of equal $\Delta h$ for the three-layered leaky system at the end of decay period -----	51
Figure 19. Possible disposal reservoirs of liquid wastes, and water quality of deep sandstones -----	53
Figure 20. Thickness of the Mt. Simon sandstone -----	54
Figure 21. Structure on top of Mt. Simon Formation -----	55
Figure 22. Geologic cross-section through the Hennepin Region, showing the section modeled in this study -----	57
Figure 23. Geologic section of the Jones and Laughlin Corporation's waste disposal well, Hennepin, Illinois -----	58
Figure 24. Details of the modeled zone in the Jones and Laughlin Corporation's well at Hennepin, Illinois -----	59
Figure 25. Schematic representation of the idealized layered system used to analyze the Jones and Laughlin Corporation's injection well at Hennepin, Illinois -----	61
Figure 26. Results of finite element method for the grid spacing used in the Jones and Laughlin Corporation's injection well with homogeneous and isotropic properties, as compared with analytical solution -----	64
Figure 27. Change of pressure vs. time calculated at the interface between Ironton-Galesville and Eau Claire for different distances from the center of model -----	65
Figure 28. Change of pressure vs. time calculated at the interface between Eau Claire and Mt. Simon for different distances from the center of model -----	66

Figure 29.	Change of pressure vs. time calculated at the middle of Mt. Simon for different distances from the center of model -----	67
Figure 30.	Approximate contours of equal $\Delta p$ (psi) for the model of Jones and Laughlin Corporation's injection well at the end of the build-up period -----	71
Figure 31.	Approximate contours of equal $\Delta p$ (psi) for the model of Jones and Laughlin Corporation's injection well at the end of the decay period -----	72
Figure 32.	Triangularization of domain -----	105
Figure 33.	Non-zero elements of sparse matrix -----	106
Figure 34.	Magnitude of error vs. number of iterations for 570, 1600, 8732, and 10660 unknowns -----	107



LIST OF TABLES

	page
Table 1. Analysis of data of Figure 3 -----	16
Table 2. Analysis of data of Figure 4 -----	16
Table 3. Physical and geometrical properties of the model of Jones and Laughlin Corporation's injection well -----	62

## Acknowledgments

The authors would like to express their gratitude to Dr. Robert Bergstrom, Mr. Peter Sarapuka, Dr. Edward Holley, and Dr. Paul Heigold for their review of the original manuscript and their comments. Dr. Holley participated in this project as a consultant. His stimulating suggestions helped the authors in the study of the topics studied in the course of this investigation.

The authors would also like to acknowledge with thanks the research funds provided by the Water Resources Center at the University of Illinois under Agreement #14-31-0001-3813.

We are grateful to Mrs. Edna Yeargin and Mrs. Margaret Gidel for their painstaking care in typing the manuscript.

## I. INTRODUCTION

### A. Objectives

This is the completion report on Phase I of a project which had originally been planned for two years. The basic objective of the two-year project was to investigate numerical and field methods to predict the fate of liquid waste injected into underground rock formations in Illinois. From these studies, it was hoped to develop engineering and hydrogeological criteria that would help the Illinois State Geological Survey assist the State Environmental Protection Agency in the evaluation of permit applications by industries throughout the State to inject liquid waste underground.

From the beginning of the project, two criteria were considered vital to this study: practicality and economy. One cannot definitely predict the problems which may be encountered in a model based on actual field data by studying only a small and hypothetical model. Therefore, it had been planned to base the bulk of the conclusions in this study on test runs with field data. In this regard, we faced a budgetary limitation; i.e., the analysis of field problems by numerical methods requires substantial amounts of computer time and computer memory. Thus, considerable effort was expended on the development of the new techniques that would reduce the cost of analysis.

Because the project was not funded for the second year, many of these objectives were not achieved. What follows is a report on the highlights and implications of the different subjects studied in one year and their evaluation.

### B. Background

The process of injection of liquid waste into a subsurface formation is usually accomplished by the displacement of the native fluids by some liquid waste.

In time, the liquid waste moves in accordance with the gradient developed in the receiving reservoir. The process of liquid waste injection has been used in the oil industry for many years (20,000 brine injection wells in Texas alone by 1966) where brines separated from pumped oil are re-injected into the oil producing reservoirs. An advantage that the oil industries have over other industries in the design of disposal wells is that substantial physical and chemical data are generally available for oil reservoirs. By the use of these data, the design and operation of a brine injection well may be studied, and many important factors such as pressure head, rate of injection, duration of continuous injection, etc., may be decided in advance. Other industries which seek to inject liquid wastes commonly have less data available.

The proper storage of waste liquids in the underground pore spaces requires hydrologic, geologic, mathematical, and engineering studies. These studies must be undertaken in conjunction with one another. The review of the literature reveals that in many instances the development of a highly sophisticated mathematical model was based on the small hypothetical models. There are geologic reports which efficiently describe only the geologic parameters needed for the storage of liquid waste in the subsurface formations. Attempts to integrate the geological, hydrological and engineering studies done in waste disposal management have been made in the recent years, and with the growing awareness of environmental problems one would expect that these efforts will be expanded.

### C. Previous Works

The literature dealing with the movement of fluids in porous media is very extensive. Many of these publications are relevant to subsurface waste injection directly or indirectly. The U. S. Environmental Protection Agency (1972)

has compiled an annotated bibliography of 106 publications dealing directly with subsurface waste injection. The purpose of this section is not to review all these papers but to focus on specific studies which contributed to this investigation.

Van Everdingen and Freeze (1971) have compiled a list of 188 references dealing with subsurface liquid waste storage as well as underground gas storage and hydrodynamic dispersion. They have categorized the problems presented by the subsurface injection of "natural" liquids such as saline water, waste brine, etc., into: (1) hydrodynamics, (2) stress mechanics, and (3) fluid compressibility. For foreign liquid wastes the problems are not confined to these three areas; rather it is realized that very little is known about the behavior of these liquids when they come in contact with native fluids under reservoir conditions generated by long periods of injection.

The subject of earth tremors generated by injection of liquids into disposal wells has been brought to the attention of the investigators in recent years. The Rocky Mountain Arsenal disposal well is perhaps the first case for which enough geophysical data is available to correlate injection with earth tremors. Van Poolen and Hoover (1970) have reviewed some of the mechanisms proposed to explain these tremors. Among these mechanisms the interstitial-pressure mechanisms postulated by Hubbert and Rubey (1959), hydraulic fracturing, reduction of coefficient of friction, thermal effects, and chemical effects are discussed briefly.

The movement of contaminants through porous media was investigated as far back as 1905 by Slichter. In this study a slug of salt solution was injected into a well and it was noticed that its arrival at a nearby well was gradual and nonuniform. Perhaps Slichter's observations formed the basis for the development

of the theoretical models to describe the movement of foreign liquids through porous media. In general, three factors contribute to the movement of any substance through porous media; dispersion, diffusion, and convection. Hoopes and Harleman (1965) have set up a schematic representation of dispersion and diffusion, as shown in Fig. 1-a. In a saturated vertical column with steady flow, a slug of dye is injected at a point source. It is observed that as the slug travels down the column, its size increases and mixing with the native fluid takes place. This mixing produces a concentration profile as shown in Figs. 1-b and 1-c at time  $t$ , as a function of distance,  $z_1$ , down the column from the point source, and distance from the column axis. These concentration profiles are associated with dispersion and diffusion. Dispersion is caused by the mechanical mixing of the individual foreign fluid particles with the native fluids. The mixing is a result of variable velocities through the pore spaces of the medium. Diffusion, on the other hand, is the result of thermal or molecular motion of individual fluid molecules. The contribution of the dispersion process to a concentration profile such as Figs. 1-b and 1-c in some parts of the flow field where velocities are high, is many times greater than that of the diffusion process. In a porous medium being injected with a tracer, these two processes cause the boundary between the tracer and the native fluids not to be a sharp and well-defined boundary. Rather, the ratio of the concentration of the tracer in the zone of mixed fluids to the concentration of the tracer at the fluid decreases from 100% to 0 as one moves across this zone of mixed fluids or gradual boundary as shown in Fig. 2.

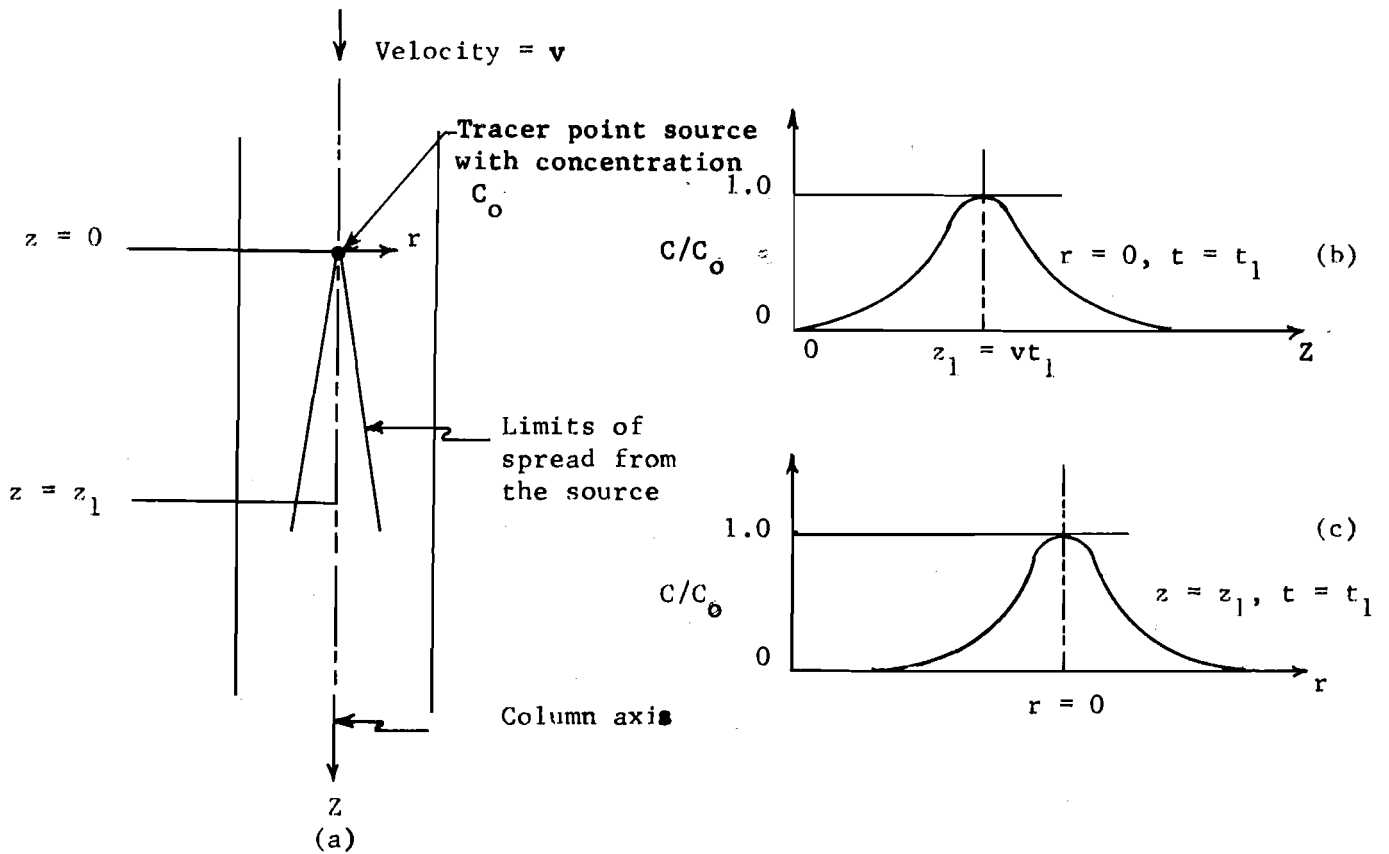


Figure 1: Schematic column setup (a) and typical measurements (b) and (c) for an instantaneous point source injection. (After Hoopes and Harleman, 1965)

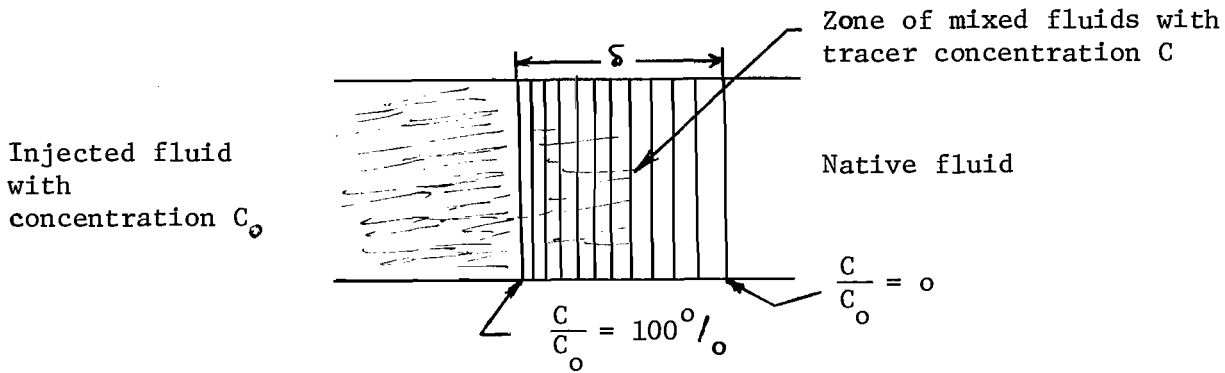


Figure 2: Schematic representation of the zone of mixed fluids or gradual boundary.

The extent of the zone of mixed fluids,  $\delta$ , depends on the local velocity of the fluids and physical properties of the porous media. If no dispersion and diffusion take place in the medium, the foreign fluid moves under the influence of convection only, and the boundary between the foreign fluid and the native fluids will be a sharp one, i.e.,  $\delta = 0$ .

Many studies have been made on the solution of the mass-transport equation containing dispersion and diffusion. Hoopes and Harleman (1965) derived the transport equation and, assuming a steady flow, analytical solutions were found for longitudinal as well as lateral dispersion. In cylindrical coordinates their convective-dispersion equation for steady flow from a line source perpendicular to the  $r - \beta$  plane is:

$$\frac{\partial c}{\partial t} + \frac{A}{r} \frac{\partial c}{\partial r} = \alpha_1 \frac{A}{r} \frac{\partial^2 c}{\partial r^2} + \alpha_2 \frac{A}{r} \frac{1}{r^2} \frac{\partial^2 c}{\partial \beta^2} \quad (1)$$

where:  $c$  = concentration of the tracer; i.e., foreign liquid

$t$  = time

$A = Q/2 \pi b \psi$



Q = total flow

b = thickness of the geologic formation into which injection is taking place

$\psi$  = porosity

$\beta$  = angular coordinate with limits of 0 to  $2\pi$

r = radial distance

$\alpha_1' = D_1' / q = \text{constant, a function of media structure}$

$D_1' = \text{coefficient of longitudinal dispersion in nonuniform flow}$

q = seepage velocity

$\alpha_2' = D_2' / q = \text{constant, a function of media structure}$

$D_2' = \text{coefficient of lateral dispersion in nonuniform flow}$

Then, assuming a symmetrical concentration distribution with respect to angular coordinate, i.e., no lateral dispersion takes place, equation (1) becomes:

$$\frac{\partial c}{\partial t} + \frac{A}{r} \frac{\partial c}{\partial r} = \alpha_1' \frac{A}{r} \frac{\partial^2 c}{\partial r^2} \quad (2)$$

This would be the equation for the concentration distribution resulting from flow from a recharge well which fully penetrates a homogeneous and isotropic aquifer. The solution of (2) may be somewhat simplified by assuming that the effect of the longitudinal dispersion term, i.e., the right-hand side, in (2) is small in comparison with the convective term in (2). Then (2) becomes:

$$\frac{\partial c}{\partial t} \approx - \frac{A}{r} \frac{\partial c}{\partial r} \quad \text{or} \quad \frac{\partial c}{\partial r} \approx - \frac{r}{A} \frac{\partial c}{\partial t} \quad (3)$$

Using (3) to calculate  $\frac{\partial^2 c}{\partial r^2}$  on the right-hand side of (2), one obtains:

$$\frac{\partial c}{\partial t} + \frac{A}{r} \frac{\partial c}{\partial r} \approx \alpha_1 \frac{r}{A} \frac{\partial^2 c}{\partial t^2} \quad (4)$$

The development of these equations is more fully discussed by Hoopes and Harleman (1965). Replacing the approximately equal sign with the equal sign, the solution of (4) for the case where a tracer with concentration  $c_0$  is injected continuously at the injection source,  $r = 0$ , is:

$$\frac{c}{c_0} = \frac{1}{2} \operatorname{erf}_c \left[ \frac{(r^2/2) - At}{\sqrt{4/3 \alpha_1 r^3}} \right] \quad (5)$$

where:  $c_0$  = concentration of the tracer at the origin

$c$  = concentration of the tracer at a distance  $r$  from the injection source

$r$  = distance from the injection source

$\operatorname{erf}_c$  = complementary error function of the quantity in bracket

Equation (5) gives an approximate solution for the concentration of the tracer as affected by convection and longitudinal dispersion at any distance  $r$  from the origin. In order to include the effect of molecular diffusion which becomes of some significance at very large distances from the injection source, the term  $\frac{D_m}{A^2} r^2 \frac{\partial^2 c}{\partial t^2}$  may be added to the right-hand side of (4). Then the solution of this equation becomes:

$$\frac{c}{c_0} = \frac{1}{2} \operatorname{erf}_c \left[ \frac{(r^2/2) - At}{\sqrt{4/3 \alpha_1 r^3 + \frac{D_m}{A} r^4}} \right] \quad (6)$$

where:  $D_m$  = coefficient of molecular diffusion, and rest of the terms have already been defined.

In the subsequent sections, these two approximate solutions of equation (1), i.e., equations (5) and (6), will be used to evaluate the effects of dispersion and diffusion on mass transport.

The solution of mass transport equation by numerical methods has encountered some critical problems. Reddell and Sunada (1970) used the finite difference technique to solve the flow equation for pressure in an unsteady, nonuniform flow field with density and viscosity variations between the injected and native fluids, and used this pressure distribution in conjunction with the method of characteristics proposed by Garder, et al. (1964) to solve the mass-transport equation. The method of characteristics is employed to avoid the so-called numerical dispersion which is usually encountered when the finite difference technique is used to solve the mass-transport equation.

Nalluswami (1971) used the finite element method to solve the mass-transport equation in a two-dimensional coordinate system. The second order linear partial differential equation for the mass-transport as used in that study has mixed partial derivatives which result from treating the dispersion coefficients as second order symmetric tensors. Thus, a functional had to be obtained to handle the mixed partial derivatives. The results of the minimization of this functional seem to compare favorably with the analytic solutions, and various boundary conditions may be incorporated into the model.

## II. SOLUTION OF THE FLOW AND MASS-TRANSPORT EQUATIONS

The prediction of the fate of liquid wastes injected into deep subsurface formations may be accomplished by the following steps. First, the subsurface formation must be idealized by a model with nonhomogeneous and anisotropic properties and well-defined boundaries. Then, the flow equation should be solved for this model, with density and viscosity variations between the injected waste and native fluids, to obtain potential distribution throughout the system. This potential distribution may be used in the mass-transport equation to predict the distance that a particle of liquid waste would travel in the direction of decreasing potential in a given increment of time. Based on this approach, first a system of equations consisting of the flow equations and the mass-transport equation was formulated in which the solution of one equation was substituted in other equations to obtain the solution of the second which was then substituted in the first equation to get another solution, and etc. This was rather similar to Reddell and Sunada's (1970) approach. However, in order to avoid the problems associated with finite difference technique and the method of characteristics, the finite element method was proposed as the main tool for the solution of these equations. Javandel and Witherspoon (1968a-b and 1969) and Neuman and Witherspoon (1971) have demonstrated the applicability of the finite element method to the solution of the flow equation in two- and three-layered systems with high permeability contrasts between layers. Nalluswami (1971) has solved the mass-transport equation by finite element method. Due to the lack of the physical data, subsurface geologic formations used for liquid injection may best be idealized by a stratified system with each layer being homogeneous but anisotropic. Such an idealization, together with the references cited above

should provide the answers appropriate to the objectives of this study. However, before the start of this process, the following points had to be resolved.

A. Importance of the Dispersion and Diffusion in the Mass Transport Equation

In order to evaluate the importance of the terms related to the dispersion and diffusion, equation (5) developed by Hoopes and Harleman (1965) was applied to a medium with the following characteristics:

$$\alpha_1 = \text{dispersion coefficient} = .15 \text{ cm} = 4.92 \times 10^{-3} \text{ ft.}$$

$$C'_0 = \text{concentration of liquid injected} = 1.0 \text{ (arbitrarily may be chosen as unit mass per unit volume)}$$

$$Q = \text{rate of injection of liquid waste} = 200 \text{ GPM}$$

$$b = \text{thickness of injection zone} = 1800 \text{ ft.}$$

$$\psi = \text{porosity of the medium} = 0.10$$

These constants approximate the conditions at the Jones and Laughlin Corporation's injection well. With these constants equation (5) was solved for 128.0 days of continuous injection. The results for days 1, 2, 4, 8, 16, 32 and 128, are plotted in Fig. 3. This figure shows that the width of the zone of mixed fluids,  $\delta$ , increases with time.

In order to evaluate the effect of the molecular diffusion, equation (6) was solved for the same period of time and the same constants as in Fig. 3 but with:

$$D_m = \text{molecular diffusion coefficient} = 1.0 \times 10^{-5} \frac{\text{cm}^2}{\text{sec}} \cong 9.3 \times 10^{-4} \frac{\text{ft}^2}{\text{day}}$$

The results are plotted in Fig. 4. Visual examination of these curves leads us to the same conclusion as in Fig. 3. However, in order to (1) make any statement about the comparison of the curves in Figures 3 and 4, (2) evaluate the effect of

time on the width of the zone of mixed fluids,  $\delta$ , and (3) assess the importance of the molecular diffusion, a parameter such as the percentile coefficient of the kurtosis given by Spiegel (1961) may be used:

$$K = \frac{Q}{P_{90} - P_{10}} \quad (7)$$

where:  $K$  = percentile coefficient of kurtosis

$$Q = \frac{1}{2} (Q_3 - Q_1)$$

$Q_1$  and  $Q_3$  = the first and third quartiles (values of  $r$  in Figures 3 and 4 which would give 0.25 and 0.75 of the areas of the respective frequency curves) where frequency curves are equivalent to the derivatives of the distribution curves in Figures 3 and 4

$P_{90}$  and  $P_{10}$  = 90% and 10% percentiles (values of  $r$  in Figures 3 and 4 which would give 0.90 and 0.10 of the areas of the respective frequency curves)

Using equation (7),  $K$  was calculated for every curve in Figures 3 and 4. These calculations are tabulated in Tables 1 and 2 together with the size of the zone of mixed fluids,  $\delta$ , and the ratio of one half of  $\delta$  to  $Q_2$ , where  $Q_2$  is the value of  $r$  which would divide the area under frequency curve into half. In Tables 1 and 2,  $\delta$  is measured from a point where  $C/C_0 = 99.0\%$  to a point where  $C/C_0 = 1.0\%$ . These tables show that the values of  $K$  decrease as time increases, i.e., the flatter the curve the lower the  $K$ . The value of  $K$  for standard normal curve ( $\mu = 0, \sigma = 1$ ) as given by Spiegel is 0.263. Thus, in Tables 1 and 2 curves of up to about 32 days are leptokurtic (a distribution having a relatively high peak), and the curves after this time are platykurtic (a distribution which is flat-topped). Furthermore, one observes that the effect of the molecular diffusion is shown in the smaller values of  $K$  for the same period of time. However, this effect is

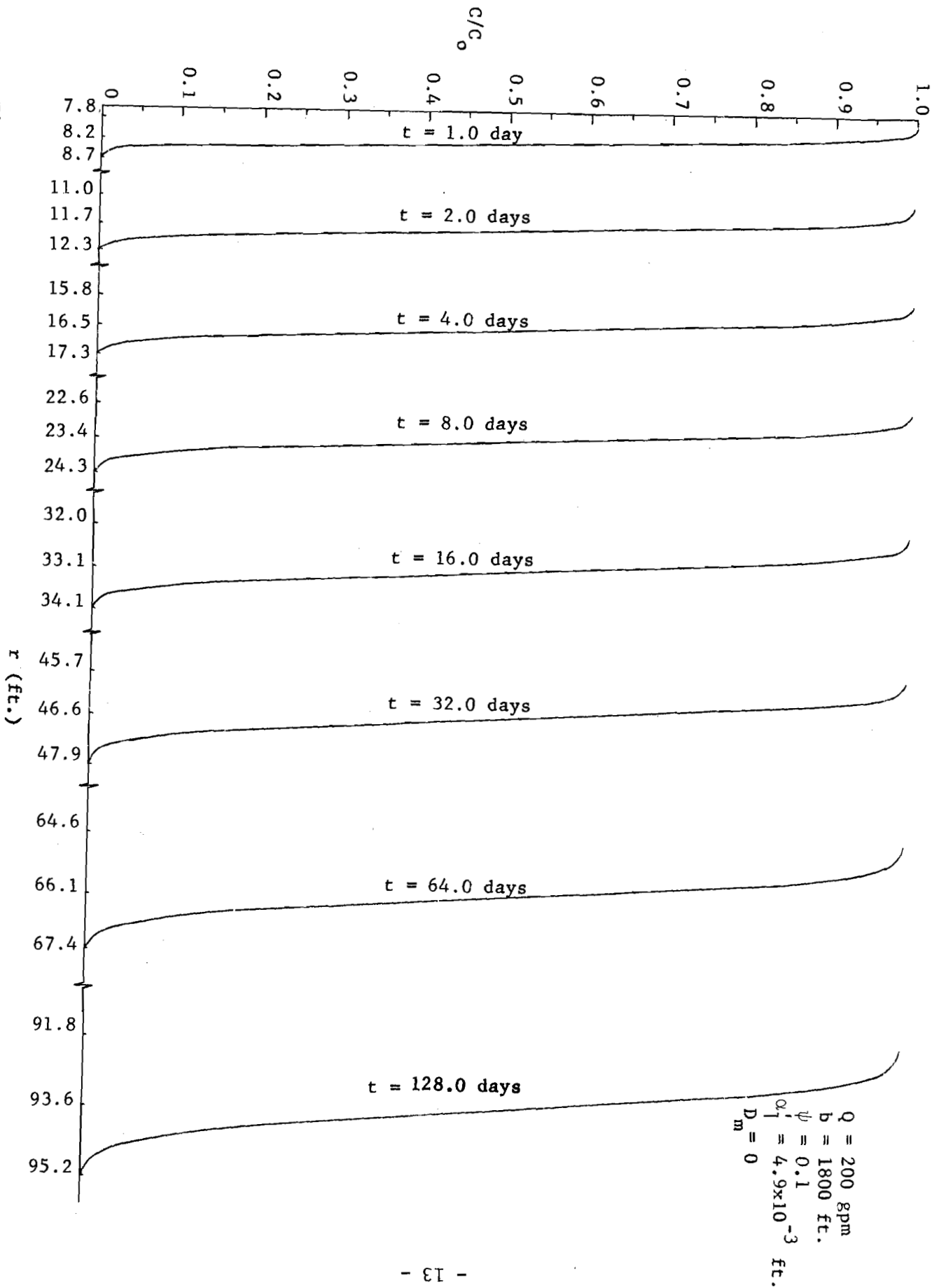
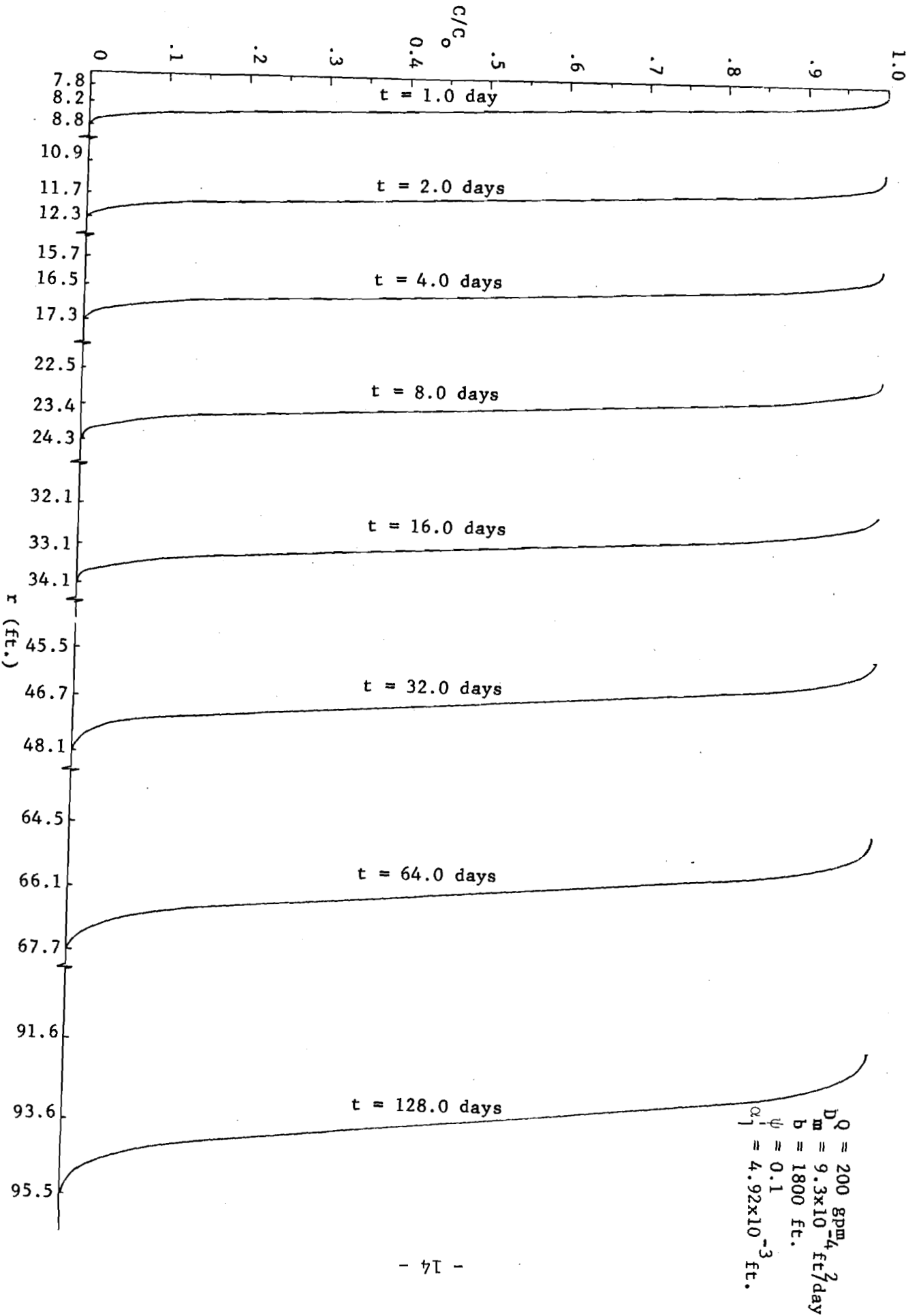


Figure 3 : Concentration distribution obtained with the above data and Hoopes and Hartleman's (1965) approximate solution without molecular diffusion.

Figure 4 : Concentration distributions obtained with the above data and Hoopes and Harleman's (1965) solution with molecular diffusion.





insignificant (around 3% of the value of K after 128 days of continuous injection). The effect of molecular diffusion at the early stages of injection is even less significant. This is due to the high velocities which exist around the injection well, causing the dispersion to be many times more significant than the molecular diffusion.

The width of the zone of mixed fluids,  $\delta$ , in Tables 1 and 2 is another measure of the importance of dispersion and diffusion. For these curves  $\delta$  increases from a minimum of 0.95 feet (after 1 day of injection without molecular diffusion) to a maximum of 3.90 feet (after 128 days of injection with molecular diffusion). Columns (10) and (11) of Tables 1 and 2 show the values of  $Q_2$  (median) or the second quartile which relates to  $C/C_0 = 0.5$ , and the measure of the width of the dispersed zone relative to the distance from the injection source, E. It was noticed that E decreased with time, and may be considered insignificant for practical studies.

Figure 5 shows the results of the computations performed on a set of data which may be considered an extreme case. After 512 days of continuous injection the width of the zone of mixed fluids,  $\delta$ , is expanded to 15 feet with the center of this zone 295.5 feet from the injection zone, i.e.,  $E = 2.54\%$ .

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
t	Q <sub>1</sub>	Q <sub>3</sub>	Q = $\frac{1}{2}(Q_3 - Q_1)$	P <sub>10</sub>	P <sub>90</sub>	P = $P_{90} - P_{10}$	K = Q/P	δ	Q <sub>2</sub>	E = $\frac{1}{2}(\frac{\delta}{Q_2} \times 100)$
(days)	(ft)	(ft)		(ft)	(ft)			(ft)	(ft)	
1.0	8.37	8.17	-0.100	8.40	8.07	-0.33	0.303	0.95	8.23	5.8
2.0	11.83	11.55	-0.140	11.92	11.44	-0.48	0.292	1.35	11.69	5.8
4.0	16.7	16.35	-0.170	16.81	16.22	-0.59	0.288	1.45	16.52	4.4
8.0	23.56	23.18	-0.190	23.71	23.05	-0.66	0.282	1.70	23.38	3.6
16.0	33.30	32.83	-0.230	33.50	32.65	-0.85	0.270	2.14	33.07	3.2
32.0	46.90	46.40	-0.258	47.14	46.18	-0.96	0.269	2.20	46.64	2.3
64.0	66.42	65.80	-0.310	66.72	64.55	-1.17	0.265	2.78	66.12	2.1
128.0	93.99	93.23	-0.380	94.34	92.86	-1.48	0.257	3.40	93.59	1.8

Table 1: Analysis of Data of Figure 3

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
t	Q <sub>1</sub>	Q <sub>3</sub>	Q = $\frac{1}{2}(Q_3 - Q_1)$	P <sub>10</sub>	P <sub>90</sub>	P = $P_{90} - P_{10}$	K = Q/P	δ	Q <sub>2</sub>	E = $\frac{1}{2}(\frac{\delta}{Q_2} \times 100)$
(days)	(ft)	(ft)		(ft)	(ft)			(ft)	(ft)	
1.0	8.42	8.16	-0.13	8.50	8.07	-0.43	0.302	1.00	8.23	6.1
2.0	11.84	11.54	-0.15	11.94	11.42	-0.52	0.289	1.40	11.69	6.0
4.0	16.80	16.45	-0.17	16.93	16.32	-0.61	0.279	1.60	16.52	4.8
8.0	23.58	23.18	-0.20	23.74	23.00	-0.74	0.270	1.80	23.38	3.9
16.0	33.28	32.78	-0.25	33.54	32.60	-0.94	0.266	2.00	33.07	3.0
32.0	47.02	46.49	-0.27	47.22	46.20	-1.02	0.264	2.60	46.74	2.8
64.0	66.44	65.79	-0.32	66.74	65.49	-1.25	0.256	3.20	66.12	2.6
128.0	93.96	93.09	-0.43	94.37	92.65	-1.72	0.250	3.90	93.60	2.1

Table 2: Analysis of Data of Figure 4

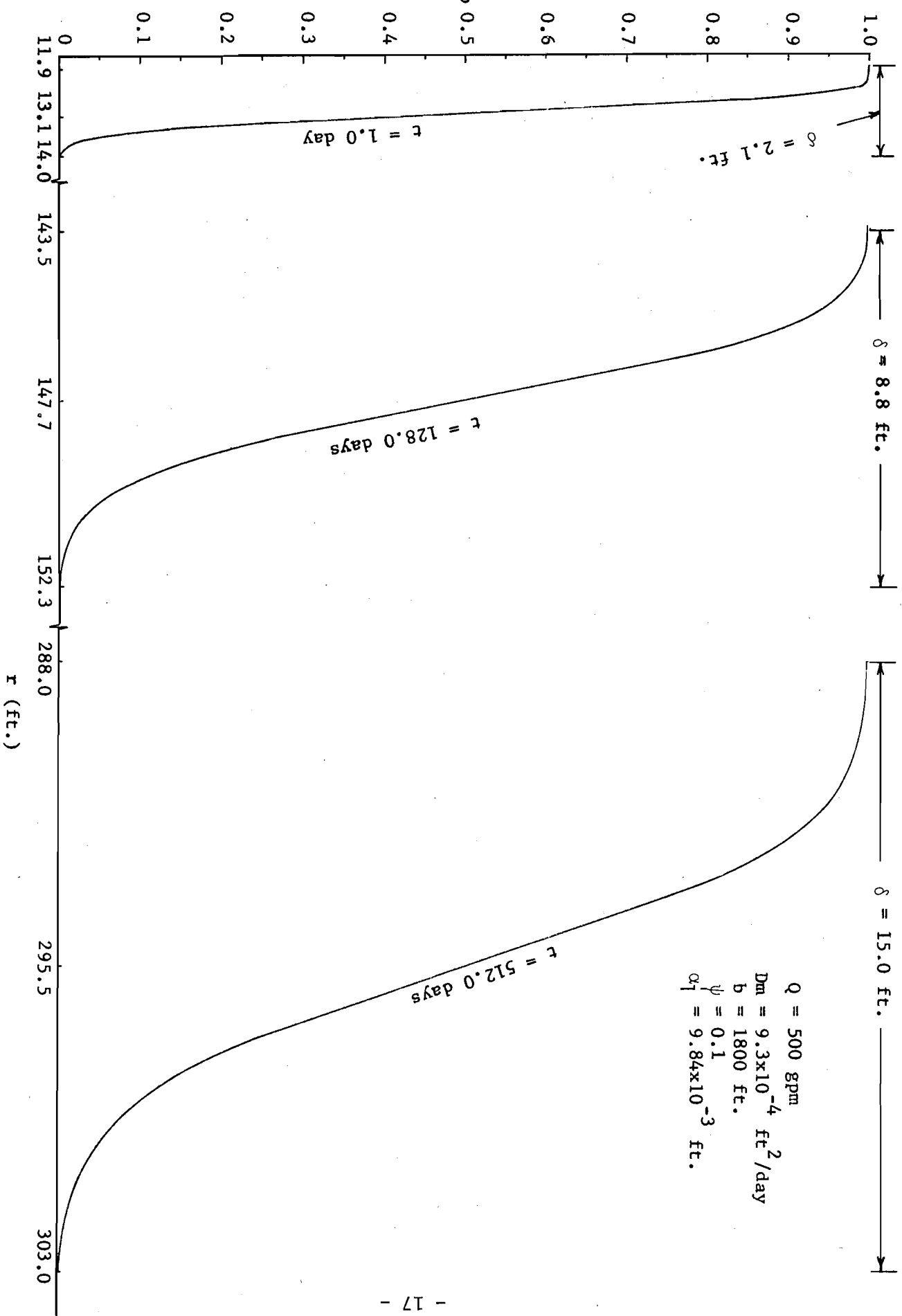


Figure 5: Concentration distribution obtained with the above data and Hoopes and Harleman's (1965) approximate solution.

## B. Solution of Flow Equation for Multi-Layered Systems

As was indicated earlier, an anisotropic and nonhomogeneous system may be idealized by a layered system with each layer being homogeneous but anisotropic. This idealization may lead to a system with high permeability contrast between the layers, and thus make the computation of the permeability derivative terms in the equation of flow by the traditional finite difference technique subject to significant error. The finite element method in this regard seems to have a definite advantage over the finite difference technique. Neuman and Witherspoon (1971), in extending the work of Javandel and Witherspoon (1969) to leaky aquifers, concluded that the solution of the equation of flow for a three-layered leaky system with permeability contrasts of 5000 by the finite element method was in excellent agreement with the analytical solution. Thus, attention was focused on understanding the finite element method.

1. Finite Element Method: The development of the finite element method started in the aircraft industry and structural engineering. It has been adapted to the solution of several types of differential equations. Zienkiewicz and Cheung (1970) were among the first to apply this method to the flow equation. The theory and formulation of the finite element method is well documented (Zienkiewicz and Cheung, 1970, Wilson and Nickell, 1966, Javandel and Witherspoon, 1969). It suffices here to mention that in the finite element method, the differential equation with the initial condition is first replaced by an equivalent functional which when minimized will provide a solution which is a very close approximation to the solution of the initial boundary value problem. In the following paragraphs we shall follow the formulation of Javandel and Witherspoon (1969).

The initial boundary value problem to be solved is the equation of flow of a slightly compressible fluid in a nonhomogeneous and anisotropic porous

medium which is:

$$\frac{\partial}{\partial x_i} \left( K_{ij} \frac{\partial \Phi}{\partial x_j} \right) = S_s \frac{\partial \Phi}{\partial t} \quad (8)$$

with initial condition:

$$\Phi (x_i, t = 0) = \Phi_0 (x_i) \quad (9)$$

and boundary conditions:

$$\Phi (B_1, t) = E_1 (B_1) \quad \text{on boundary } B_1 \quad (10a)$$

and

$$\left. \frac{\partial \Phi}{\partial n} \right|_{B_2} = E_2 (B_2) \quad \text{on boundary } B_2 \quad (10b)$$

where  $\Phi = \text{force potential} = gh = g \left[ z + \frac{P}{\rho g} \right]$

$h$  = hydraulic head

$g$  = acceleration of gravity

$z$  = elevation of the point where the force potential is being measured

$P$  = pressure at the point in question

$\rho$  = density of the fluid at the point in question

$x_k$  &  $x_j$  ( $i, j = 1, 2, 3$ ) = the axes of the space coordinates

$K_{ij}$  = Permeability, a symmetric 3 x 3 matrix which usually is a function of space coordinates, ( $K_{ij} = k_{ij} \frac{\rho g}{\mu}$ )

$k_{ij}$  = specific permeability

$\mu$  = viscosity of the fluid at the point in question

$S_s$  = specific storage, i.e., storage per unit thickness of the formation,  $\psi c \rho g$

$\psi$  = porosity of the formation

$c$  = effective compressibility of the formation

$t$  = time

$n$  = outer normal to boundary  $B_2$

$\phi_0$ ,  $E_1$ , and  $E_2$  = functions or constants depending on the boundary and initial conditions

$B_1$  and  $B_2$  = boundaries on which  $E_1$  and  $E_2$  must be satisfied

By obtaining the Laplace transform of the left hand side of (8) and substituting for  $\phi_0$  from (9) into this transform, after inversion one arrives at:

$$\frac{\partial}{\partial x_i} (K_{ij} * \frac{\partial \phi}{\partial x_j}) = S_s (\phi - \phi_0) \quad (11)$$

where \* stands for convolution as defined by Gurtin (1964):

$$G * H = \int_0^t G(t - \tau) H(\tau) d\tau \quad (12)$$

Equation (11) is equivalent to equations (8) and (9) combined. This step eliminates one equation, namely the one for the initial condition equation (9). Therefore, the problem is reduced to the solution of (11) subject to the boundary conditions (10a) and (10b).

The functional representing this problem is:

$$\Omega_t(\phi) = \int_V \left[ S_s \phi * \phi + K_{ij} * \left( \frac{\partial \phi}{\partial x_i} \right) * \left( \frac{\partial \phi}{\partial x_j} \right) - 2 S_s \phi * \phi_0 \right] dV - 2 \int_{B_2} K_{ij} * C_2 * \phi d B \quad (13)$$

where: V represents the volume of the medium, and i,j = 1,2,3.

The conditions that (13) has to satisfy in order to be the functional of (11) subject to (10a) and (10b) are given by Zienkiewicz and Cheung (1970).

Equation (13), which is the variational problem of (11) subject to (10a) and (10b), when minimized will provide a solution which is the solution of (11) too.

It represents a homogeneous but anisotropic medium. In order to apply equation (13) to a heterogeneous and anisotropic system, the medium may first be divided into small anisotropic and homogeneous elements and equation (13) may be discretized, i.e., replaced by the summation of the functionals of each individual element. In this fashion functional (13) for a specific element e is:

$$\Omega_t^e(\phi) = \int_{V^e} \left[ S_s \phi * \phi + K_{ij} * \left( \frac{\partial \phi}{\partial x_i} \right) * \left( \frac{\partial \phi}{\partial x_j} \right) - 2 S_s \phi * \phi_0 \right] dV - 2 \int_{B_2^e} K_{ij} * C_2 * \phi d B \quad (14)$$

Using equation (14) for every element, equation (13) may be replaced by

$$\Omega_t(\phi) = \sum_e \Omega_t^e(\phi) \quad (15)$$

To minimize (15),  $\phi$  for each element must be approximated by a function. If the elements are sufficiently small,  $\phi$  over each element for a two-dimensional system, may be approximated by a linear function such as

$$\phi_e (r,z,t) = a + br + cz \quad (16)$$

where a,b,c = coefficients which are functions of time, and r and z = coordinates of a specific point within element e. The coefficients a,b, and c for each element must be chosen in such a way that  $\phi_e$  satisfy the boundary conditions of element e, and be continuous with respect to the two-dimensional system over element e. The simplest element is one with a triangular cross section. For axisymmetric systems, rings with a triangular cross section (concentric with the z-axis) may be considered. Figures 6-a and 6-b represent a two-dimensional triangularization and a ring with triangular cross section. An element e and potential within it may be totally defined by the three nodes 1,2, and 3 and their respective potentials  $\phi_1, \phi_2, \phi_3$ . This means that one may write equation (16) for the three nodes 1,2, and 3 and solve for a, b, and c in terms of the  $\phi, r, s$  and  $z, s$ . Upon substitution of these values of a, b, and c into equation (16) we obtain:

$$\phi_e (r,z,t) = \frac{1}{2\Delta} [(\alpha_1 + \beta_1 r + \gamma_1 z) \phi_1 + (\alpha_2 + \beta_2 r + \gamma_2 z) \phi_2 + (\alpha_3 + \beta_3 r + \gamma_3 z) \phi_3] \quad (17)$$

where:

$$\begin{array}{lll} \alpha_1 = r_2 z_3 - r_3 z_2 & \beta_1 = z_2 - z_3 & \gamma_1 = r_3 - r_2 \\ \alpha_2 = r_3 z_1 - r_1 z_3 & \beta_2 = z_3 - z_1 & \gamma_2 = r_1 - r_3 \\ \alpha_3 = r_1 z_2 - r_2 z_1 & \beta_3 = z_1 - z_2 & \gamma_3 = r_2 - r_1 \end{array}$$



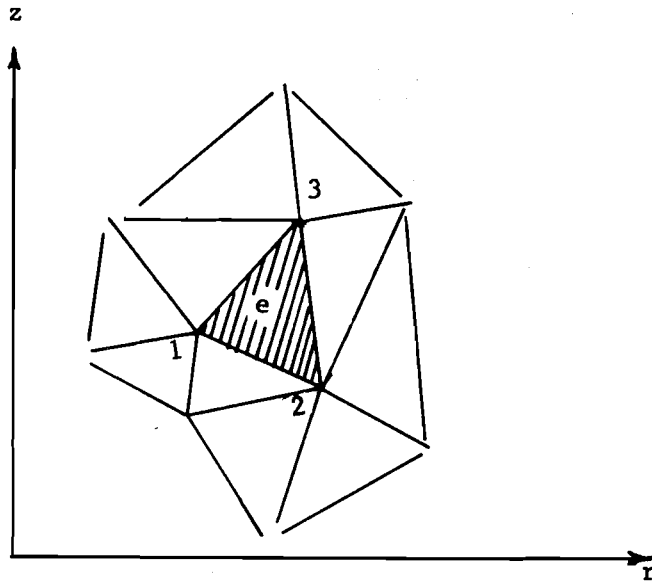


Figure 6-a: Division of two-dimensional region into triangular elements.  
(After Zienkiewicz and Cheung, 1970)

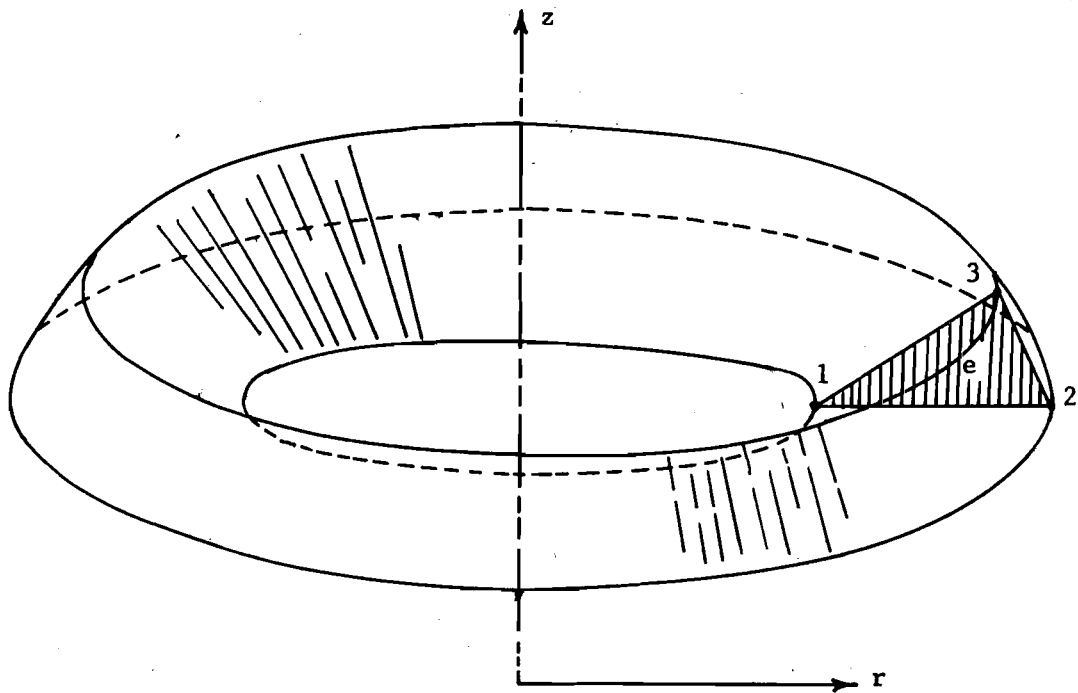


Figure 6-b: An axisymmetric element with constant triangular cross-section.  
(After Neuman and Witherspoon, 1971)

and

$2 \Delta =$  twice the area of triangle 1,2,3 =

$$r_2 (z_3 - z_1) + r_1 (z_2 - z_3) + r_3 (z_1 - z_2)$$

In (17) all the variables except  $\phi_1$ ,  $\phi_2$ , and  $\phi_3$  are functions of space.

$\phi_1$ ,  $\phi_2$ , and  $\phi_3$  are fixed in position and are functions of time only. In matrix form (17) may be written as

$$\phi_e (r,z,t) = [N_1, N_2, N_3] [\phi (t)]^e = [N_n] [\phi_n] \quad (18)$$

where  $N_n = \frac{1}{2\Delta} (\alpha_n + \beta_n r + \gamma_n z)$  is the only non-zero element in a 1 x 3 matrix

and  $[\phi_n] = [\phi(t)]^e$  is a column matrix with elements being the potentials

at nodes 1, 2, and 3.

Substituting (18) into (14)

$$\begin{aligned} \Omega_t^e (\phi) = & \int_{V^e} \{ S_s [N_n] [\phi_n] * [N_m] [\phi_m] + K_{ij} * \left[ \frac{\partial N_n}{\partial x_i} \right] [\phi_n] \\ & * \left[ \frac{\partial N_m}{\partial x_j} \right] [\phi_m] - 2S_s [N_n] * [N_m] [\phi_{om}] \quad dV - \\ & - 2 \int_{B_2^e} K_{ij} * C_2 * [N_n] [\phi_n] \quad dB \end{aligned} \quad (19)$$

where n, m = 1, 2, 3 for element e.

Now, the values of  $\phi_n$  in (19) are to be calculated so that they minimize the functional (15). This cannot be achieved by taking the first variation of  $\Omega_t(\phi)$  with respect to  $\phi_n$ , because  $\phi$  is a function of both space and time, and  $\Omega_t(\phi)$  contains a convolution, i.e., integration with respect to time, and integration with respect to space coordinates. Following the method used in calculus of variation and presented by Neuman and Witherspoon (1971),  $\phi_n(t)$  may be replaced by  $\phi_n(t) + E\eta(t)$ , where  $E$  is an arbitrary constant and  $\eta(t)$  is a time dependent continuously differentiable function which vanishes at  $t = 0$  and  $t = t$ . Then (19) becomes:

$$\begin{aligned} \Omega_n(E) = & \sum_e \left\{ \int_V e \int_0^t \{ S_s [N_n] [\phi_n(\tau) + E\eta(\tau)] [N_m] [\phi_m(t - \tau)] \right. \\ & + K_{in} * \frac{\partial N_n}{\partial x_i} [\phi_n(\tau) + E\eta(\tau)] \left[ \frac{\partial N_m}{\partial x_j} \right] [\phi_m(t - \tau)] \\ & - 2S_s [N_n] [\phi_n(\tau) + E\eta(\tau)] [\phi_{om}] \} d\tau dV \\ & - 2 \int_{B_2^e} \int_0^t K_{ij} * C_2 [N_n] [\phi_n(\tau) + E\eta(\tau)] d\tau dB \} \end{aligned} \quad (20)$$

which may be minimized now with respect to  $\phi_n$  by

$$\left. \frac{d\Omega_n(E)}{dE} \right|_{E=0} = 0$$

or

$$\begin{aligned} \left. \frac{d\Omega_n(E)}{dE} \right|_{E=0} = & 2\eta * \sum_e \left\{ \int_V e \{ S_s [N_n] [N_m] [\phi_m] \right. \\ & + K_{ij} * \left[ \frac{\partial N_n}{\partial x_i} \right] \left[ \frac{\partial N_m}{\partial x_j} \right] [\phi_m] - S_s [N_n] [N_m] [\phi_{om}] dV - \int_{B_2^e} K_{ij} * C_2 [N_n] db \} = 0 \end{aligned}$$

By assuming that  $\eta$  is an arbitrary function, one may write:

$$\sum_e \left\{ \int_{V^e} \{ S_s [N_n] [N_m] [\phi_m] + K_{ij} * \begin{bmatrix} \partial N_n \\ \partial x_i \end{bmatrix} \begin{bmatrix} \partial N_m \\ \partial x_j \end{bmatrix} [\phi_m] - S_s [N_n] [N_m] [\phi_{om}] \} dV - \int_{B_2^e} K_{ij} * C_2 [N_n] dB \right\} = 0 \quad (21)$$

Equation (21) may be written for  $n = 1, 2, \dots, N$ , where  $N$  is the total number of nodes in the system, thus creating  $N$  equations with the unknowns,  $\phi_m$  ( $m = 1, 2, \dots, N$ ). It must be noticed that when writing (21) for node  $n$ , only elements which have node  $n$  in common will contribute to this equation. This will produce a symmetric sparse matrix whose band width is determined by the numbering system of nodes.

To simplify the computations of (21), let

$$D_{nm}^e = S_s \int_{V^e} [N_n] [N_m] dV^e \quad (22)$$

$$AA_{nm}^e = K_{ij} \int_{V^e} \begin{bmatrix} \partial N_n \\ \partial x_i \end{bmatrix} \begin{bmatrix} \partial N_m \\ \partial x_j \end{bmatrix} dV \quad (23)$$

Thus (20) for node  $n$  becomes:

$$\sum_e [D_{nm}^e (\phi_m - \phi_{om}) + AA_{nm}^e * \phi_m - \int_{B_2^e} K_{ij} * C_2 [N_n] dB_2^e] = 0 \quad (24)$$

The values of integral (22) and (23) as formulated by Wilson and Nickell (1966) and proved by Neuman and Witherspoon (1971) for elements with triangular cross section are:

$$D_{nm}^e = \begin{cases} \frac{S_s \Delta}{12} & \text{if } n \neq m \\ \frac{S_s \Delta}{6} & \text{if } n = m \end{cases} \quad \text{for planar triangle}$$

$$D_{nm}^e = \begin{cases} \pi \bar{r} \frac{S_s \Delta}{6} & \text{if } n \neq m \\ \pi \bar{r} \frac{S_s \Delta}{3} & \text{if } n = m \end{cases} \quad \text{for concentric ring with z-axis}$$

where:

$$\bar{r} = \frac{r_1 + r_2 + r_3}{3}$$

$\Delta$  = area of the triangular cross section

and

$$\begin{aligned} AA_{nm}^e &= \int_{\Delta} K_{ij} \left[ \frac{\partial N_n}{\partial x_i} \right] \left[ \frac{\partial N_m}{\partial x_j} \right] d\Delta = \int_{\Delta} \left\{ K_m \left[ \frac{\partial N_n}{\partial r} \right] \left[ \frac{\partial N_m}{\partial r} \right] + K_{rz} \left[ \frac{\partial N_n}{\partial r} \right] \left[ \frac{\partial N_m}{\partial z} \right] \right. \\ &+ K_{zr} \left[ \frac{\partial N_n}{\partial z} \right] \left[ \frac{\partial N_m}{\partial r} \right] + K_{zz} \left[ \frac{\partial N_n}{\partial z} \right] \left[ \frac{\partial N_m}{\partial z} \right] \left. \right\} drdz \\ &= \int_{\Delta} \frac{1}{4\Delta^2} (K_{rr} \beta_n \beta_m + K_{rz} \beta_n \gamma_m + K_{zr} \gamma_n \beta_m + K_{zz} \gamma_n \gamma_m) drdz \\ &= \frac{1}{4\Delta} (K_{rr} \beta_n \beta_m + K_{rz} \beta_n \gamma_m + K_{zr} \gamma_n \beta_m + K_{zz} \gamma_n \gamma_m) \text{ for planar triangle} \end{aligned}$$

and  $AA_{nm}^e = \pi r \frac{1}{2\Delta} (K_{rr}\beta_n\beta_m + K_{rz}\beta_n\gamma_m + K_{zr}\gamma_n\beta_m + K_{zz}\gamma_n\gamma_m)$  for concentric ring.

The last integral in (23) is a measure of the net flow through the entire boundary  $B_2$ . This integral vanishes for internal boundaries between elements, since for these boundaries the outflow from one element will be the inflow to another element, and thus the net effect is zero. Therefore, this integral needs to be evaluated only for the external boundaries with the flow. Assuming the flow through boundary  $B_2^e$  of Figure 7 is  $Q_{B_2}^e(t)$ , the flow attributed to node n from element e is:

$$Q_n^e(t) = \frac{1}{2} Q_{B_2}^e(t)$$

and the total flow through node n is:

$$Q_n(t) = \sum_e \frac{1}{2} Q_{B_2}^e(t) = \frac{1}{2} [Q_{B_2}^e(t) + Q_{B_2}^{e+1}(t)]$$

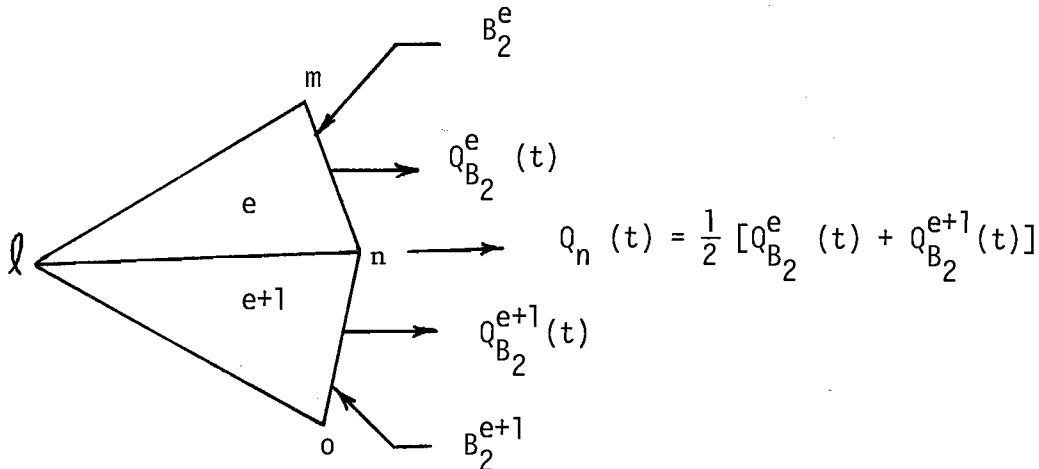


Figure 7: Schematic representation of flow through node n.

Let:

$$D_{nm} = \sum_e D_{nm}^e$$

and

$$AA_{nm} = \sum_e AA_{nm}^e$$

Then (24) becomes:

$$D_{nm} (\phi_m - \phi_{om}) + AA_{nm} * \phi_m - 1 * Q_n = 0 \quad (25)$$

In order to evaluate the convolutions in (25), let  $\Delta t$ , the time interval in the time integral, be small enough so that  $\phi$  and  $Q$  may be assumed to vary linearly with time:

$$1 * \phi_m = \int_{t-\Delta t}^t \phi_m(\tau) d\tau \approx \frac{\Delta t}{2} [\phi_m(t) + \phi_m(t - \Delta t)] \quad (26)$$

and

$$1 * Q_n \approx \frac{\Delta t}{2} [Q_n(t) + Q_n(t - \Delta t)] = \Delta t \bar{Q}_n \quad (27)$$

where

$$\bar{Q}_n = \frac{Q_n(t) + Q_n(t - \Delta t)}{2}$$

Substituting (26) and (27) into (25) one gets:

$$\begin{aligned} \frac{1}{2} [\phi_m(t) + \phi_m(t - \Delta t)] [D_{nm} + \frac{\Delta t}{2} AA_{nm}] &= \\ = \frac{\Delta t}{2} \bar{Q}_n + D_{nm} \phi_m(t - \Delta t) \end{aligned} \quad (28)$$

which by defining:

$$A_{nm} = D_{nm} + \frac{\Delta t}{2} AA_{nm}, \quad (29)$$

$$B_{nm} = \frac{\Delta t}{2} [Q_n] + D_{nm} \phi_m(t - \Delta t) \text{ and } X_m = \frac{1}{2} [\phi_m(t) + \phi_m(t - \Delta t)],$$

(27) becomes:

$$A_{nm} X_m = B_n \quad (30)$$

The set of equations represented by (30) are  $N$  simultaneous linear algebraic equations with  $N$  unknowns,  $X_m$ , where  $N$  represents the total number of nodal points in the system.

2. Boundary Conditions: Two types of boundary conditions are of particular interest in ground-water hydrology: i) constant flow boundary, ii) constant potential boundary. The treatment of both of these boundaries with equation (30) has been demonstrated by Javandel and Witherspoon (1969). In the solution of the models to be presented a constant flow boundary was considered most often. However, a few runs were made with a constant potential boundary.

3. Implementation of the Finite Element Method: Based on the formulation of the finite element method and assumptions presented in the previous sections, a computer program capable of handling any anisotropic and nonhomogeneous confined system, was written and tested for various hypothetical models with constant potential boundary condition and constant flow rate boundary condition. Then a field problem with constant flow boundary condition was solved with this program. It must be mentioned that most of the results presented in the following sections are not unique to this study (see Javandel and Witherspoon, 1968-a,b and 1969, and Neuman and Witherspoon, 1971). What has been accomplished in this incomplete study is the understanding of the finite element method, its advantages, and its disadvantages.

a - Properties of Matrices  $D$  and  $AA$ : The properties of  $D$  and  $AA$  play a very important role in the feasibility of the solution of equation (30) for large systems. In these matrices  $n, m = 1, \dots, N$ . Thus, were these matrices dense, a storage of  $2(N \times N)$  would have been required to store elements of  $AA$  and  $D$  (because  $\Delta t$  in (29) has to be very small in the early time steps and increase gradually, both  $D$  and  $AA$  must be accessible throughout the time steps). But, as was



mentioned earlier, the nonzero elements in  $AA_{nm}$  and  $D_{nm}$  are only due to the nodes which share an element with the node whose  $aa_{nm}$  and  $D_{nm}$  are being calculated, where lower case letters refer to elements of the corresponding matrices, i.e.,  $n$  (see Figure 8). Therefore, these matrices are sparse.

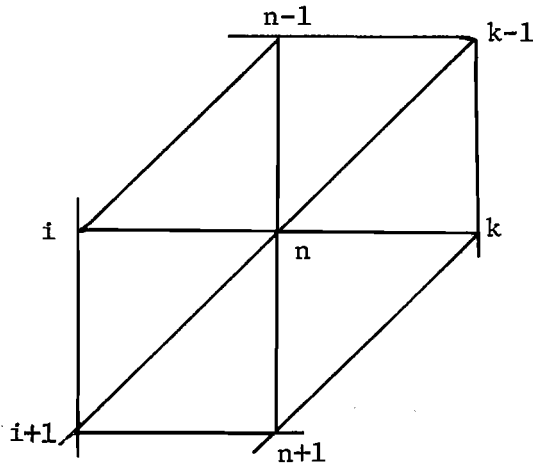


Figure 8: Schematic representation of the nodes which contribute to the elements  $aa_{nm}$  and  $d_{nm}$ , where  $m = i, i+1, n-1, n, n+1, k-1,$  and  $k$ .

Another property of  $AA_{nm}$  and  $D_{nm}$  is that they are symmetric, i.e.,  $d_{nm} = d_{mn}$  and  $aa_{nm} = aa_{mn}$ . This is because in equations (21) and (22) the terms and constants used for node  $n$  are the same as for node  $m$ .

These properties facilitate the computer storage and retrieval of  $aa_{nm}$  and  $d_{nm}$  greatly. The storage required for these matrices depends on the number of nodes immediately surrounding the node in question. In Figure 8, when computations are being performed for node  $n$ , storage must be provided at most for seven

coefficients in the matrix equations 24-a and b and, because of symmetry, a storage for four coefficients will be sufficient. The retrieval of these coefficients may be achieved by providing a pointer vector,  $w$ , which indicates the node number,  $m$ , whose contribution to the node  $n$  is being stored in location  $j$ ,  $j = 1, 2, 3, 4$ , of the four locations. This method of storage and retrieval is somewhat close to the George's (1971) Method 1 for the compact storage scheme for sparse matrices. The difference between the method used here and that of George is that in this study the pointer vector,  $w$ , contains the node whose contribution to  $n$  is calculated, while in George's method this pointer contains the differences between the node numbers whose contributions to  $n$  is being calculated and node  $n$ . Thus, following the numbering system of Figure 8, the pointer vector,  $w$ , in this study for node  $n$  contains:

$$w(n, 1) = n$$

$$w(n, 2) = n + 1$$

$$w(n, 3) = k - 1$$

$$w(n, 4) = k$$

Any time it is required to retrieve the contribution of, say,  $m$  to  $n$ , i.e.,  $aa_{nm}$  and  $d_{nm}$ , a search may be conducted in vector  $w(n, j)$ ,  $j = 1, \dots, 4$ , to find which member of  $w$  contains  $m$ . The element denoted by  $j$  will direct us to the locations of  $aa_{nm}$  and  $d_{nm}$ .

b - Mesh Construction: The mesh construction in this study started by dividing the system into elements with square or triangular cross-sections. Then each element with square cross-sections was temporarily subdivided into two elements with triangular cross-sections, such as shown in Figure 8 to calculate the coefficients required. This type of decomposition generates equations which

are identical with those derived by the well known finite difference method (Allen, 1955). Several other decompositions were tried, but this scheme was found to be easier to handle and required less storage than others. For example, had the decomposition of Figure 9 been used, the coefficients of node  $n$  would have required a storage of  $2 \times 5 = 10$  to store  $a_{nm}$  and  $d_{nm}$ , where  $m = n, o, p, q, \text{ and } i$ .

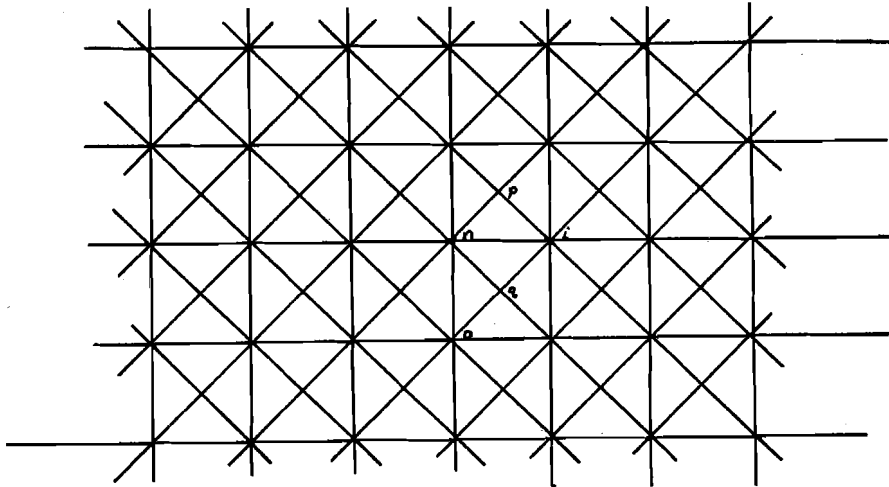


Figure 9: Schematic representation of irregular elements  
(After: Zienkiewicz and Cheung, 1970)

At the same time the storage requirement for node  $p$  is  $2 \times 3 = 6$ . The regular mesh system as defined by Zienkiewicz and Cheung (1970) and as presented in Figure 8 requires a storage of  $2 \times 4$  for every node in the system. Whereas the sums of the computer storage for all nodes are the same for the two decompositions, in practice the decomposition of Figure 9 requires more storage because of the irregularity in the meshes.

c - Change of Time Increment: In the formulation of the method, it was mentioned that because some of the time dependent variables such as  $\phi$  and  $Q$  are approximated over an increment of time by a linear function, the time increment,  $\Delta t$ , must be very small at the early time steps. However, because of economy,  $\Delta t$  must be changed as fast as possible. Therefore, in this study a specific point,  $p$ , in the system whose potential is expected to change rather fast with time was chosen and as soon as the relationship (31) was satisfied,  $\Delta t$  was multiplied by a constant greater than 1. In this study this constant was arbitrarily set equal to 1.5.

$$\frac{\phi_p(t) - \phi_p(t - \Delta t)}{\phi_p(t)} \leq E \quad (31)$$

During the early time steps  $E$  was set at 0.04 to 0.1 depending on the problem. But as time steps continued this constant was gradually increased to a maximum of 0.2. With this scheme it was possible to increase  $\Delta t$  from a fraction of a second to 10 to 15 hours in about 40 to 60 time steps.

d - Solution of the Linear Equations: This step of the computations was accomplished by the Gauss Elimination Method as presented by Crandall (1956), and Crout (1941). This technique seemed to be efficient for a small system of equations, say,  $N = 100$ . However, for a large system of equations, say,  $N = 1000$ , its deficiencies were evident. A system of 1000 equations took a prohibitive time of about 40 seconds per time step on the IBM 360/75. This excessive time encouraged a more efficient solution of large systems of linear equations. The results of this study are presented in section III of this report.

#### 4. Test of the Computer Program:

a - Homogeneous and Isotropic Systems: Two tests were made on the homogeneous and isotropic models. The results are plotted in Figures 10-a and 10-b

in the form of:

dimensionless potential =

$$\frac{2\pi K H}{Qg} \Delta\phi \quad \text{vs.}$$

dimensionless time =

$$\frac{K t}{S_s r^2}$$

where: K = permeability

H = thickness of the formation

Q = total rate of injection

g = constant of gravity

$\Delta\phi$  = change of potential with respect to the start of injection, i.e.,  $\phi_t - \phi_0$   
where  $\phi_t$  and  $\phi_0$  are the potentials at times t and 0 respectively

t = time

$S_s$  = specific storage

r = distance of the point from the center of the well

$r_e$  = outer radius of the system

The solid lines in these figures which are copied from Javandel and Witherspoon (1968-b) are the analytical solutions to the respective problems given by Muskat (1946) and Muller and Witherspoon (1965).

Figure 10-a is the solution of a finite radial system with an impermeable outer boundary and constant potential outer boundary, constant flow rate, and well of zero radius. The finite element solution matches that of the analytical solution. For a point with  $r = \frac{r_e}{8.9}$ , where  $r_e$  is the outer radius, the analytical solution was not readily available, and time was not spent to calculate this curve.

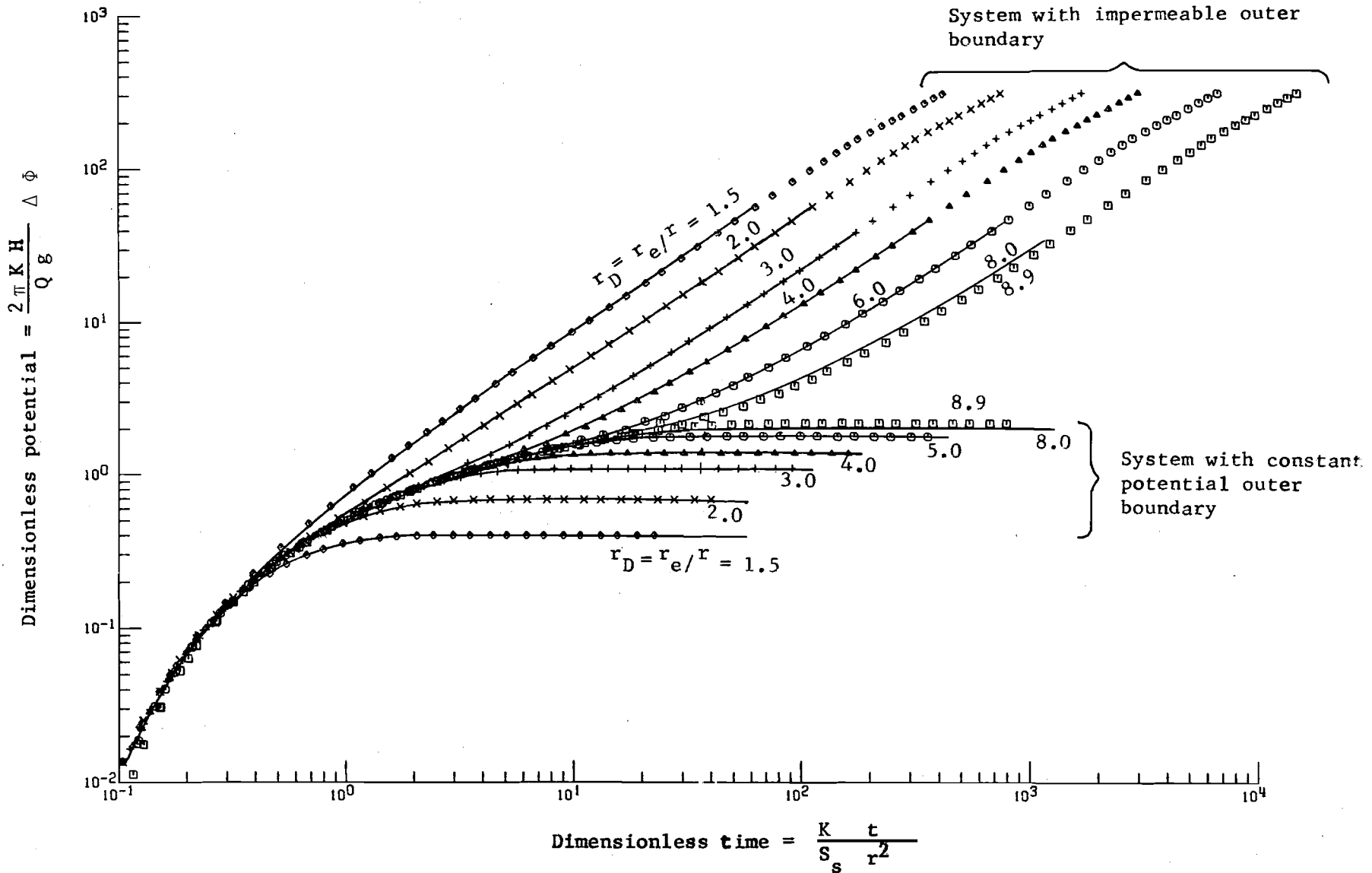


Figure 10-a: Results obtained by finite element method (symbols) as compared with analytical solutions (solid lines after Muskat, 1946) for impermeable and constant potential outer boundary system, with well of zero radius and constant rate of injection.

$$\text{Dimensionless potential} = \frac{2 \pi K H}{Q s} \Delta \phi$$

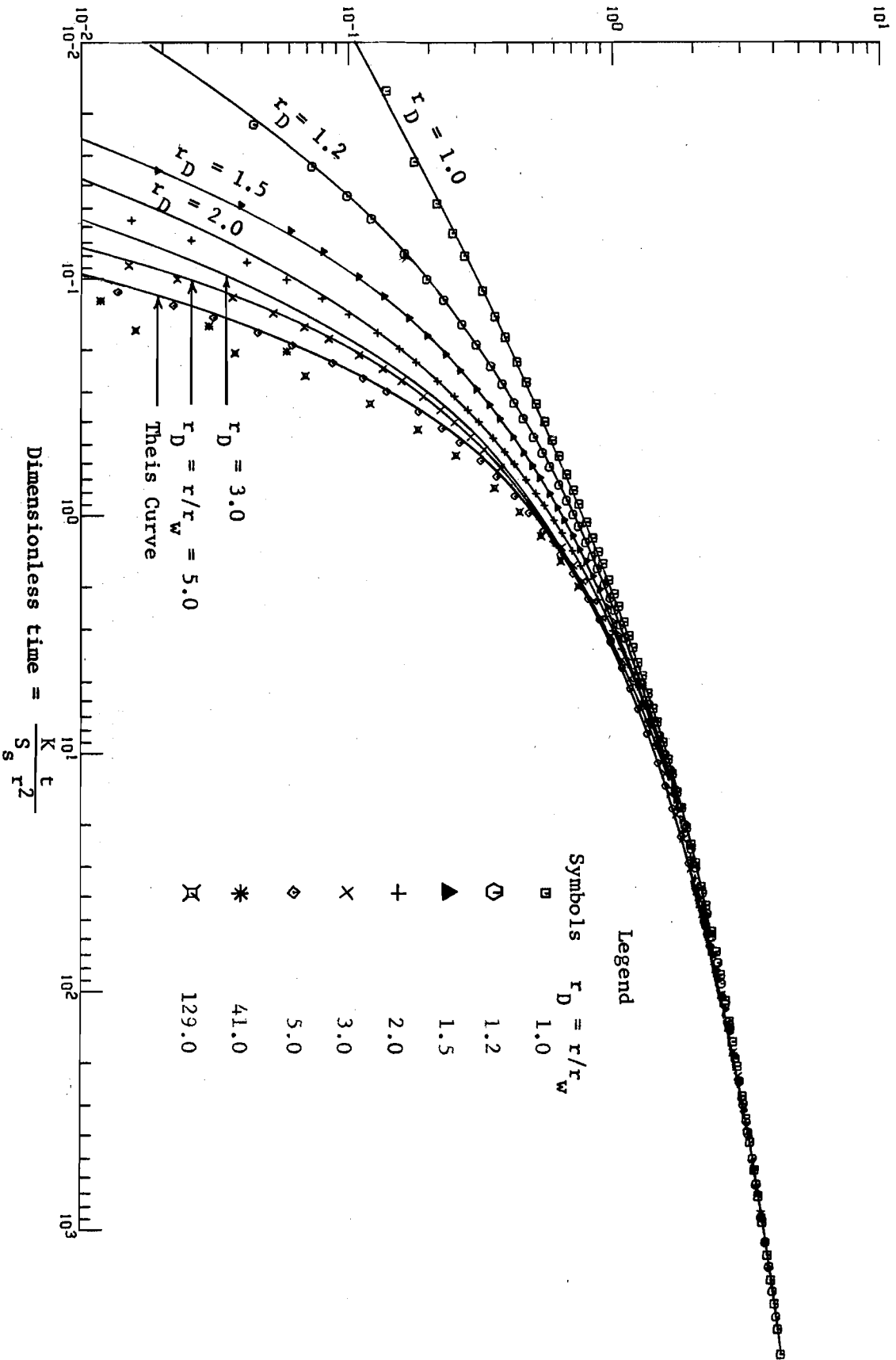


Figure 10-b: Results obtained by finite element method (symbols) as compared with analytical solutions (solid lines, after Muller and Witherspoon, 1965) for infinite radial system, with well of radius 0.25 ft. and constant rate of injection.

Figure 10-b is the solution of an infinite radial system,  $r_e = \infty$ , with a well of finite radius and constant flow. The difficulty with respect to  $r_e = \infty$  was overcome by assuming a finite  $r_e$  which could be handled with the computer memory available, and temporal solution was continued until the time where deviations from the Theis Curve, such as in Figure 10-a, were noticed in the solution. Whereas the analytical solutions for points close to the well match those of the finite element method, there are noticeable amounts of variation between the results of the analytical method and the finite element method for points with  $r = 2.0 r_w$  or larger, where  $r_w =$  radius of the well. This may be attributed to the mesh size and the length of the time increment. The length of the time increment,  $\Delta t$ , may affect the accuracy because of the linearity assumed in the evaluation of the convolutions of equations (26) and (27). Apparently this is not the cause of the variations mentioned above, otherwise the results for nodes with  $r = 2.0 r_w$  would show such variations too. The mesh sizes, for the areas in the system, such as areas adjacent to the well, with high and non-linear variations of  $\phi$ , can affect the accuracy considerably. During the early time steps, the gradient of  $\phi$  in these areas is very high. Because  $\phi$  is approximated over an element by a linear function, i.e., equation (16), care must be exercised to assure that such an approximation is valid. One way that this may be achieved is to use small mesh sizes in these areas. An examination of the mesh sizes of the model used to obtain data in Figure 10-b revealed that the mesh size for nodes with  $r_D = r/r_w = 2.0$ , in Figure 10-b, where  $r_D$  is a dimensionless quantity, increases by a factor of 5 over that of the node with  $r_D = r/r_w = 1.0$  within a distance less than a foot.

b - Layered Systems: In this section two models of confined leaky systems are tested.



i) Two-Layered System With Constant Potential Boundary at the Well:

The first model tested was a two-layered confined, leaky axisymmetric system with a pond of radius  $r_w = 80.0$  feet located in the center of the model and other geometrical and physical properties given in Figure 11. Katz (1960, p. 86) has derived an expression for two-layered systems and Javandel and Witherspoon (1968a) have calculated and evaluated this expression for four points with  $r_D = r/r_w = 2.0, 3.0, 4.0,$  and  $5.0$  and presented them together with the finite element solutions. In Figure 11 these four curves are presented together with the finite element solutions obtained in this study for the same points and additional points with  $r_D = 1.25$  and  $1.01$ . The numerical solutions for the first four points match very closely those of the analytical solutions, and the point with  $r_D = 1.25$  seems to follow the trend of these curves. However, the solution for  $r_D = 1.01$  encounters oscillations as shown by the dashed lines. This oscillation may be associated with the mesh sizes on the vertical surface of the system adjacent to the pond. An attempt to resolve this point by using finer mesh sizes on the vertical surface did not change the results substantially. Other tests with a well of radius  $r_w = 0.25$  feet replacing the pond produced no oscillations and results matched the analytical solutions.

ii) Three-Layered System with Constant Injection Rate Boundary Condition:

The next model was a three-layered, confined system with constant injection rate at the well. Consider a system such as shown in Figure 12, with the origin at the top of the injection layer (layer 1) and the well bore.

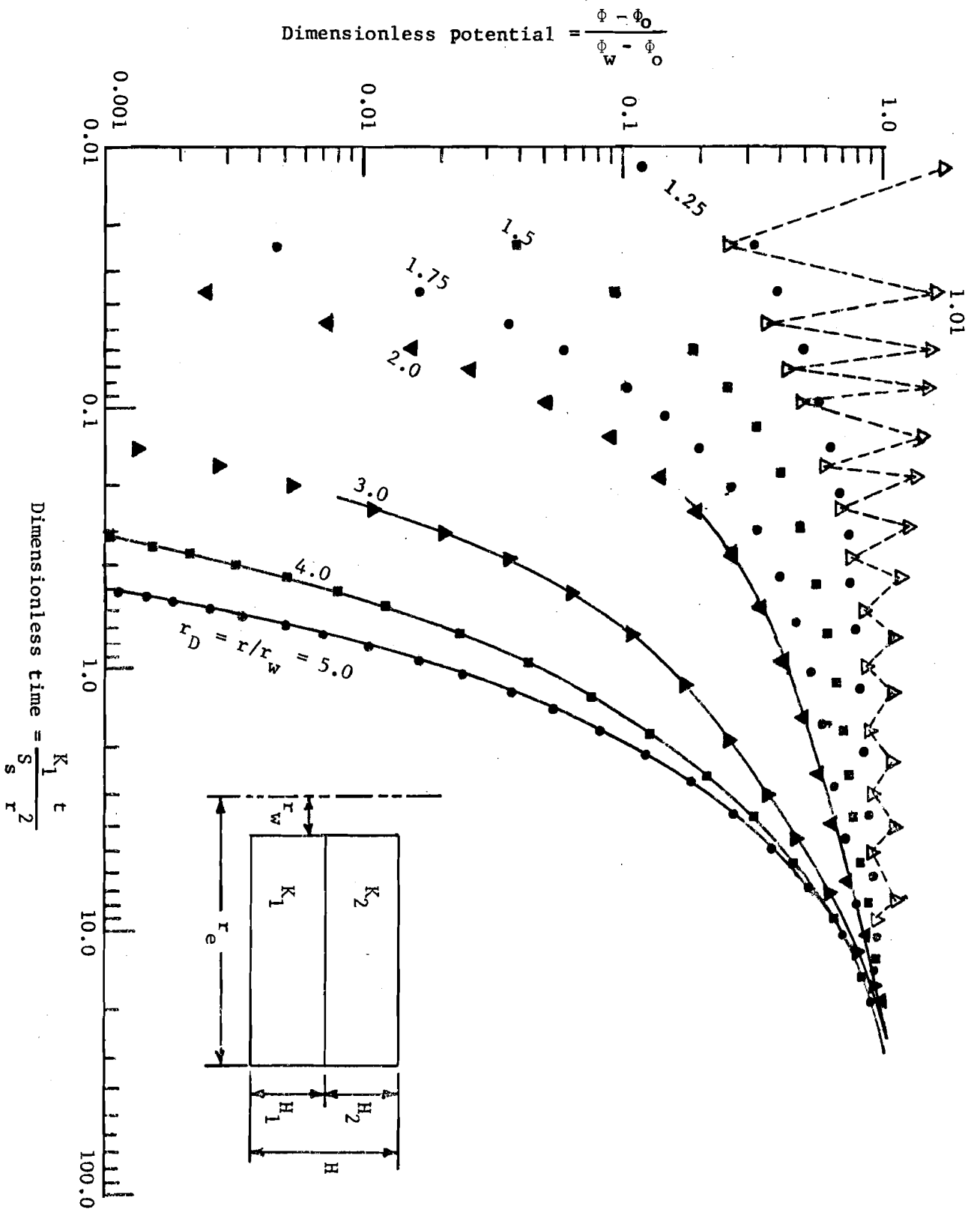


Figure 11: Results of the finite element solution for a two-layered system adjacent to a pond with constant head and radius  $r_w = 80.0$  ft. ( $r_w/H = 10.0$ ,  $r_e/r_w = 5.0$ ,  $H_2/H_1 = 1.0$ , and  $K_2/K_1 = 2.0$ ) (symbols) as compared with the analytical solutions (solid lines, after Katz, 1960).

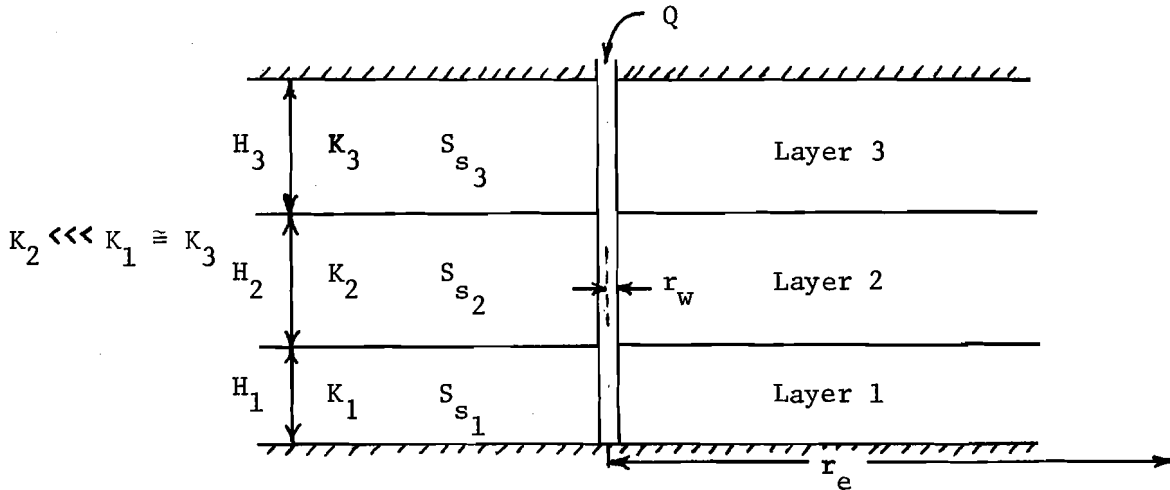


Figure 12: Schematic representation of a three-layered system  
(After Neuman and Witherspoon, 1969a)

Assume injection is taking place through a well of radius  $r_w$  into a homogeneous and isotropic layer which is infinite in horizontal extent, i.e., layer 1. This layer is bounded by an impermeable layer at the bottom and by layer 2 ( $K_2 \lll K_1$ ) on the top. Layer 3 with permeability and specific storage properties similar to layer 1 is located on top of the layer 2. The system, which is elastic, is initially saturated with a homogeneous fluid, and the density and viscosity of the injected fluid is the same as the native fluid.

Jacob (1946) presented a solution to such a problem by assuming that leakage is proportional to the potential change across layer 2. This in effect assumes that layer 2 has achieved a steady state, i.e., storage in this layer can be ignored and that the flow through layer 2 is vertical. Hantush (1960a) removed the implicit assumption about storage in layer 2 and assumed that change of potential in layer 3 remains zero; he gave a modified solution for nonsteady flow

in a leaky system in which the effect of the storage of layer 2 was considered. Neuman and Witherspoon (1969a,b) have dropped both of these assumptions, i.e., their model allows for the change of potential in layer 3 and does not ignore storage in layer 2. They evaluated their solution in terms of the dimensionless parameters:

$$r/B_{i2} = r \sqrt{\frac{K_2}{K_i H_i H_2}} \quad (32)$$

and

$$\beta_{i2} = \frac{r}{4H_i} \sqrt{\frac{K_2 S_{s2}}{K_i S_{si}}} \quad (33)$$

where  $i = 1$  or  $3 =$  indices given to layers 1 and 3

$K_i$  or  $S_{s_i}$  = permeability or specific storage of layer  $i$ , respectively

$r$  = radial distance of the point whose  $r/B_{i2}$  and  $\beta_{i2}$  are being calculated

$H_i$  = thickness of layer  $i$  (see Figure 12)

Equation (32) gives a measure of the ratio of the permeabilities of layer 2, the aquitard, and the other two layers. Equation (33) gives a measure of the same parameters and the effects of storage. Neuman and Witherspoon (1969a,b) have evaluated their solutions for several combinations of  $\beta_{i2}$  and  $r/B_{i2}$  and compared their analytical results with those of the finite element method.

To test the capabilities of the computer program written for this study, a system such as Figure 12 was set up with the following parameters:

$$r_e = 30,000.0 \text{ feet}$$

$$H = 40.0 \text{ feet}$$

$$H_2/H_1 = H_2/H_3 = 2$$

$$r_w = 0.$$

$$K_1/K_2 = K_3/K_2 = 5000$$

$$S_{s_2}/S_{s_1} = S_{s_2}/S_{s_3} = 8.0$$

$$S_{s_2} = 1.0 \times 10^{-6} \text{ ft}^{-1}$$

Total number of mesh points = 469

Total number of elements = 436

These are essentially the same parameters that Neuman and Witherspoon (1971) used to evaluate their finite element solutions. With these parameters, when  $\beta_{i2} = r/B_{i2} = 0.01$ , then  $r = 10.0$  feet; when  $\beta_{i2} = r/B_{i2} = 0.1$ , then  $r = 100$  feet; and when  $\beta_{i2} = r/B_{i2} = 1.0$ , then  $r = 1000$  feet. The injection period was divided into a) build-up period, and b) decay-period. In the build-up period injection continued at a desired rate of  $0.1 \text{ ft}^3/\text{sec}$  for 31.9 hours. At the end of this period the flow rate calculated at  $r = 0$ . from the potential distribution was  $.09999 \text{ ft}^3/\text{sec}$ . At this time injection was stopped and the calculation for 44.8 hours of the decay-period was started. Figures 13, 14, and 15 present dimensionless head increase  $h_D = \frac{4 \pi K_1 H_1}{Q} \Delta h$  vs. dimensionless time  $t_D = \frac{K_1}{S_{s_1}} \frac{t}{r^2}$  obtained by the finite element method (symbols) as compared with the analytical solutions (solid lines) of Neuman and Witherspoon (1969a) for points with  $r = 10$  feet (fig. 13),  $r = 100$  feet (fig. 14) and  $r = 1000$  feet (fig. 15) at elevations  $z = -5.0$  (layer 1, injection zone), 4.0, 10.0, 16.0 (layer 2, aquitard) and 25.0 feet (layer 3) for each  $r$ . These figures also represent the data obtained

from the decay period for the same points as the injection period. Because of the lack of time and funds no attempt was made to compare the decay data with the analytical solutions. In Figure 13 some discrepancies are observed between the analytical solutions and the finite element solutions. It is observed that the finite element solutions for short times deviate noticeably from the analytical solution. This may be due to three factors. One factor is that the analytical solutions for short times can be considered as approximations for the actual change of head in the system. Neuman and Witherspoon (1969a) defined the limit of the short injection times as:

$$t_{\ell} \leq \frac{1.6 \beta_{i2}^2}{(r/B_{i2})^4} \quad (34)$$

The other factors could be the mesh size used in the model and the length of the time increment during the early time steps. Neuman and Witherspoon (1969a) also observed that their numerical solution was on or below the analytical solution. This was specifically noticed for  $\beta_{i2} = r/B_{i2} = 0.01$  in layer 2 (aquitarde) where the numerical solution was found to be about 5% below the analytical solution. The same type of errors may be observed in Figures 14 and 15.

The data for the decay-period clearly show the effect of the elasticity of the layers. Since layer 2 (aquitarde) is 8 times more elastic than layers 1 or 3, it is noticed that  $\Delta h$  for nodes within this layer does not start to decrease immediately after the injection stops. Rather, as expected, the data show a small increase in  $\Delta h$  before it starts decreasing. This is more noticeable for points closer to the injection well (Figures 13 and 14) than for the nodes far away from the well (Figure 15). The time coordinate of the decay curves starts from  $t_D = 1.0$

$$\text{Dimensionless increase in head} = h_D = \frac{4 \pi K_1 H_1}{Q} \Delta h$$

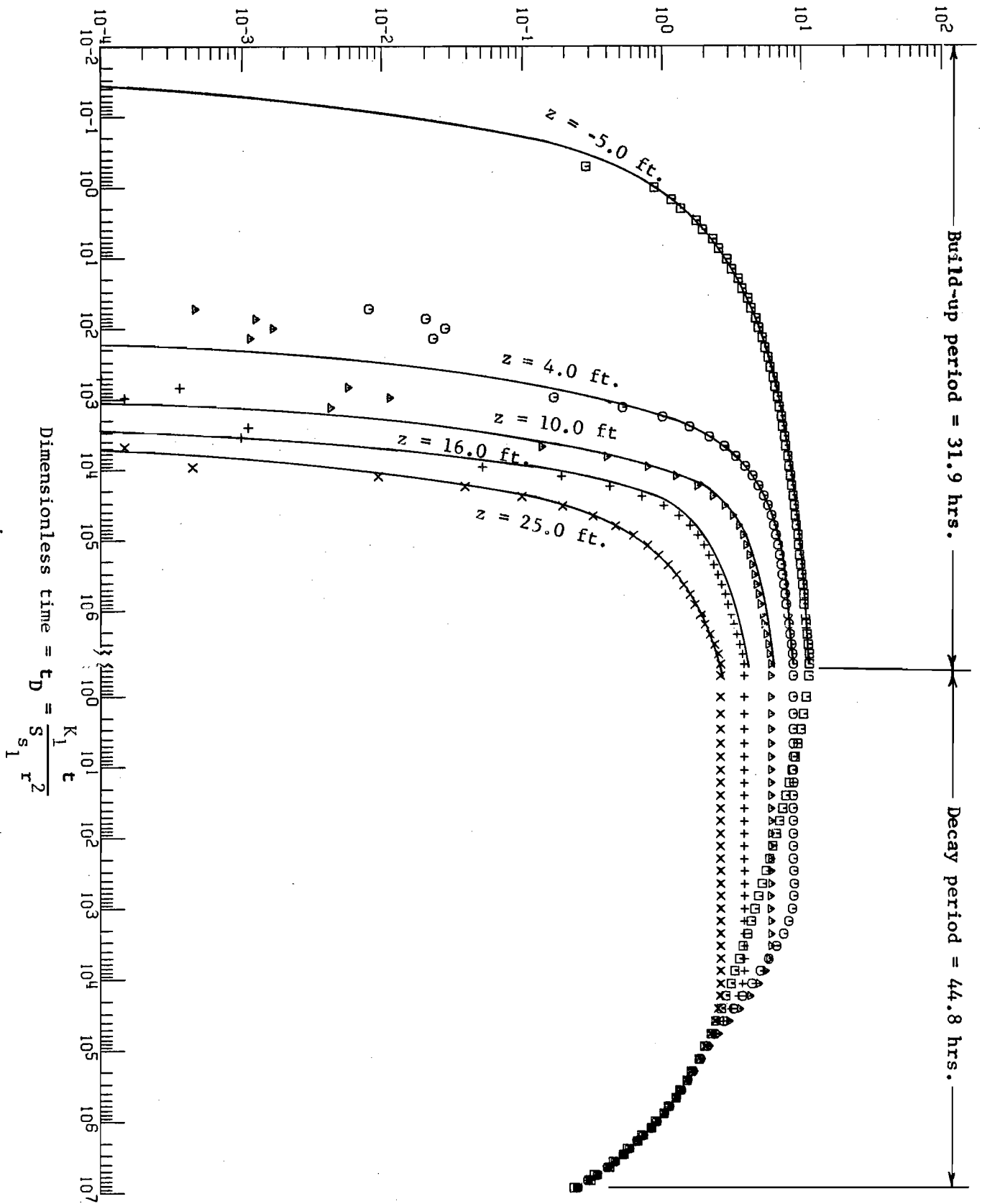


Figure 13: Results obtained by finite element method (symbols) as compared with analytical solutions (solid lines, after Neuman and Witherspoon, 1971) for a three-layered leaky system with  $\beta_{12} = \beta_{32} = 0.01$  and constant injection rate.

$$\text{Dimensionless increase in head} = h_D = \frac{4\pi K_1 H_1}{Q} \Delta h$$

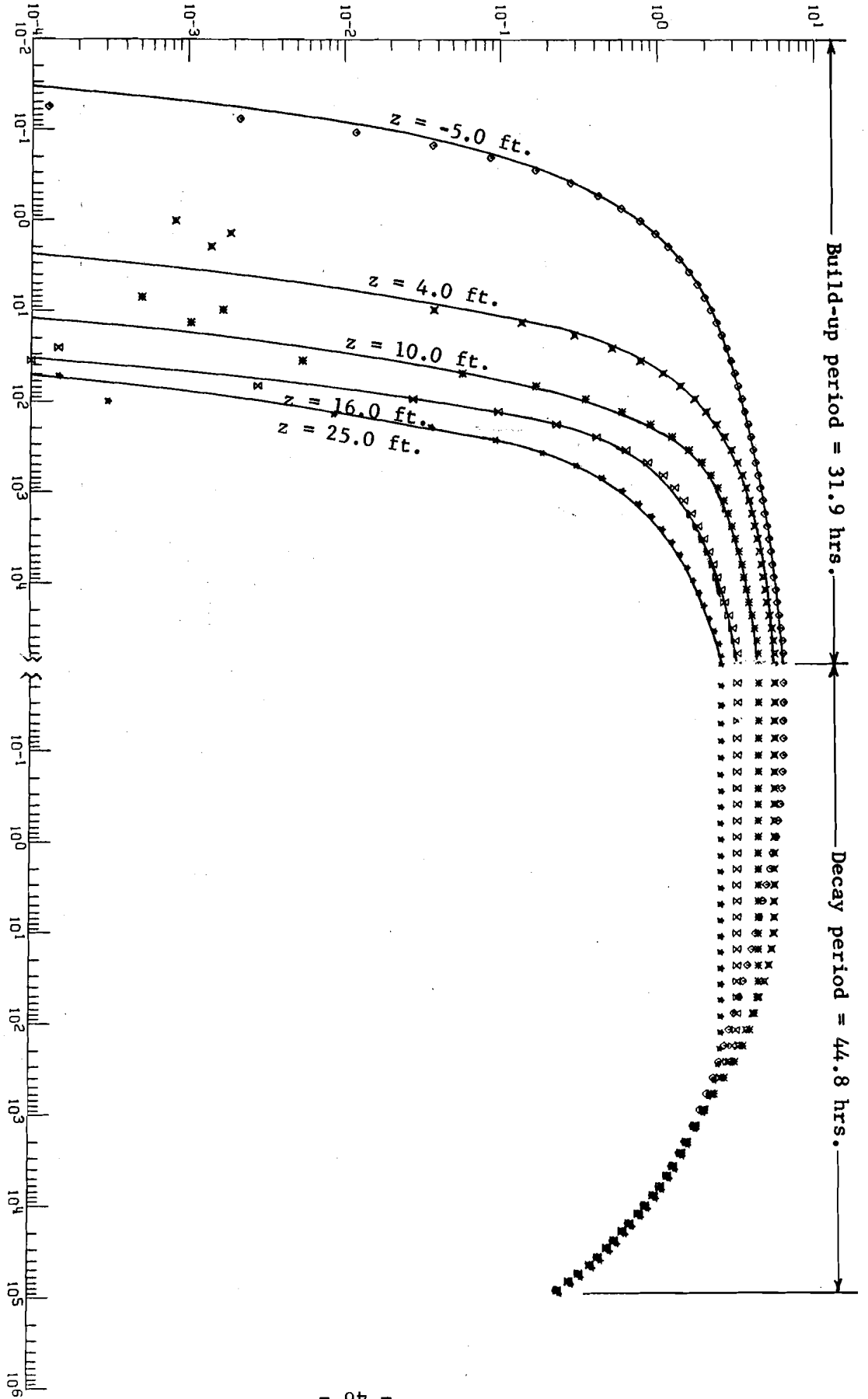


Figure 14: Results obtained by finite element method (symbols) as compared with analytical solutions (solid lines, after Neuman and Witherspoon, 1971) for a three-layered leaky system with  $\beta_{12} = \beta_{32} = 0.1$  and constant rate of injection.



$$\text{Dimensionless increase in head} = h_D = \frac{4 \pi K_1 H_2}{Q} \Delta h$$

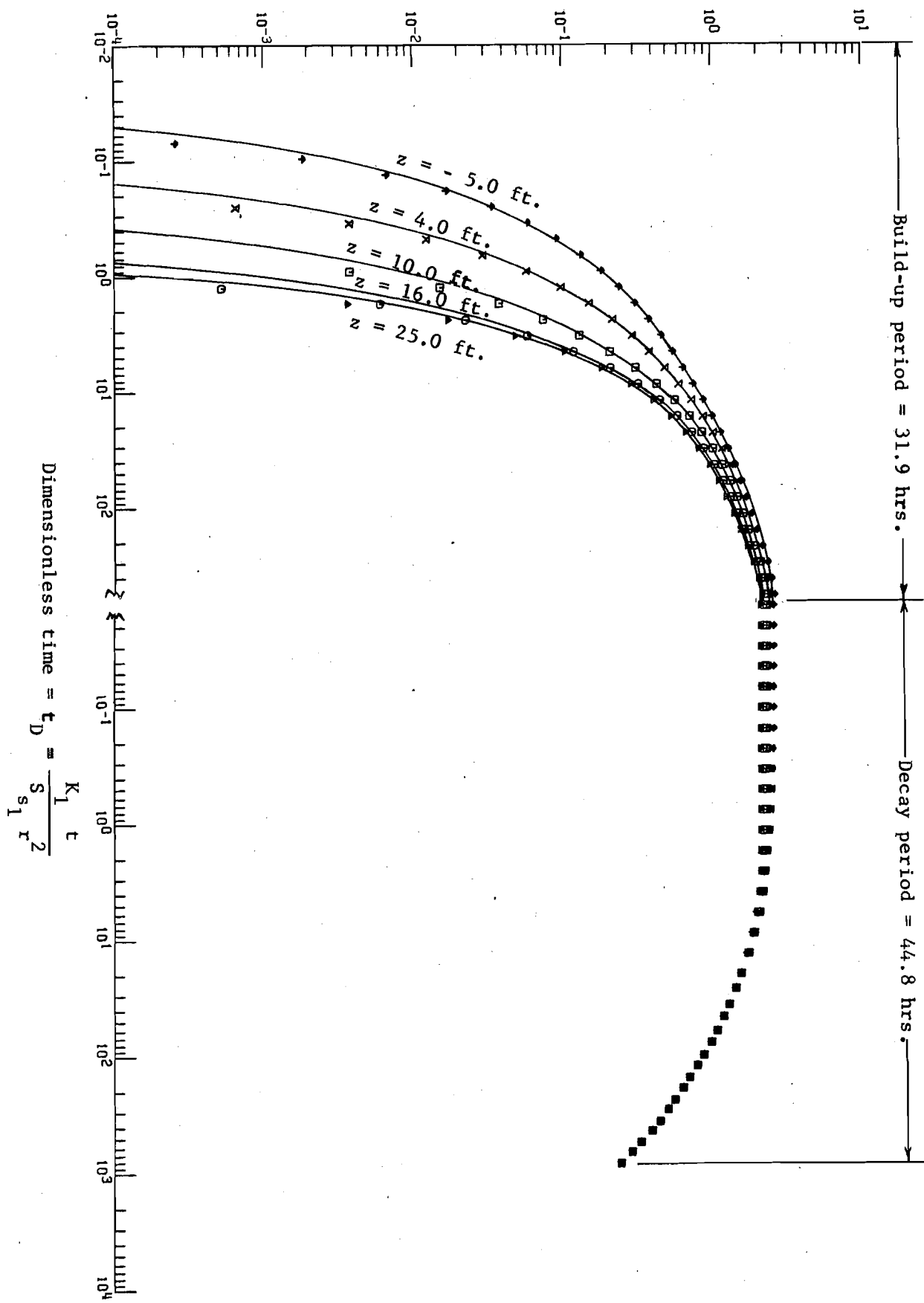


Figure 15: Results obtained by finite element method (symbols) as compared with analytical solutions (solid lines, after Neuman and Witherspoon, 1971) for a three-layered leaky system with  $\beta_{12} = \beta_{32} = 1.0$  and constant injection rate.

$$\text{Dimensionless time} = t_D = \frac{K_1 t}{S_1 r^2}$$

or 0.01 rather than continuation of the  $t_D$  for the build-up curves. The reason is that when data of the decay-period was plotted with  $t_D$  as the continuation of the  $t_D$  of the build-up period, no trend was visible because of the closeness of data points. Figure 16 shows such a plot for  $r = 10.0$  feet. Figure 17 represents the approximate contours of the equal  $\Delta h$  for this model at the end of the build-up period. After 44.8 hours of decay-period these contour lines looked as presented in Figure 18.

c - Application of the Finite Element Method to a Field Problem: Up to this point we have described briefly the theory of the finite element method, and evaluated its accuracy and some of the problems involved in its implementation. In this section the finite element method is applied to the Mt. Simon sandstone currently used for injection of liquid wastes by Jones and Laughlin Corporation at Hennepin, Illinois.

i) Geology: Deep permeable geologic formations appear to have considerable potential for waste disposal. The deep formations contain highly saline brines, which in some areas approaches 10-times the salinity of sea water. The isolation from surface environments which have resulted in the formation of the brines is of particular interest in deep well disposal, as this is exactly the environment desired for injection. Bergstrom (1968) outlined the possible disposal reservoirs (Figure 19), and suggested regions where there were individual disposal zones and associated confining formations which may serve to isolate the formation fluids and injected wastes.

Existing hydrogeologic and structural-geologic data are available so that the mathematical model developed in this research will represent a realistic picture of potential deep well disposal. Of primary concern is the continued isolation of the disposal zone from man's environment. The pressure build-up and

$$\text{Dimensionless increase in head} = h_D = \frac{4 \pi K_1 H_1}{Q} \Delta h$$

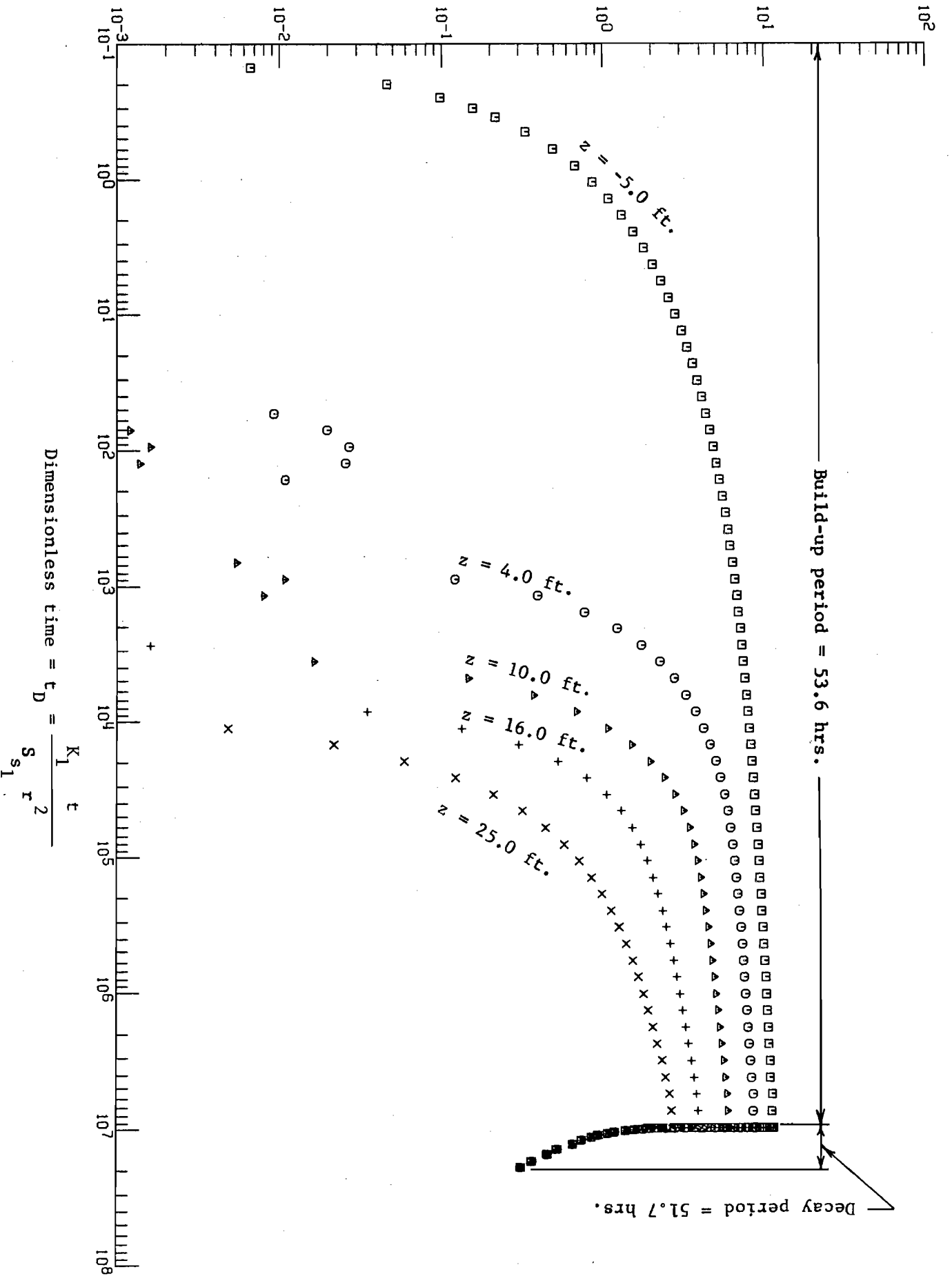


Figure 16: Results obtained by finite element method for a three-layered leaky system with  $\beta_{12} = \beta_{32} = 0.01$ , constant injection rate, and decay period timed continuously after build-up period.

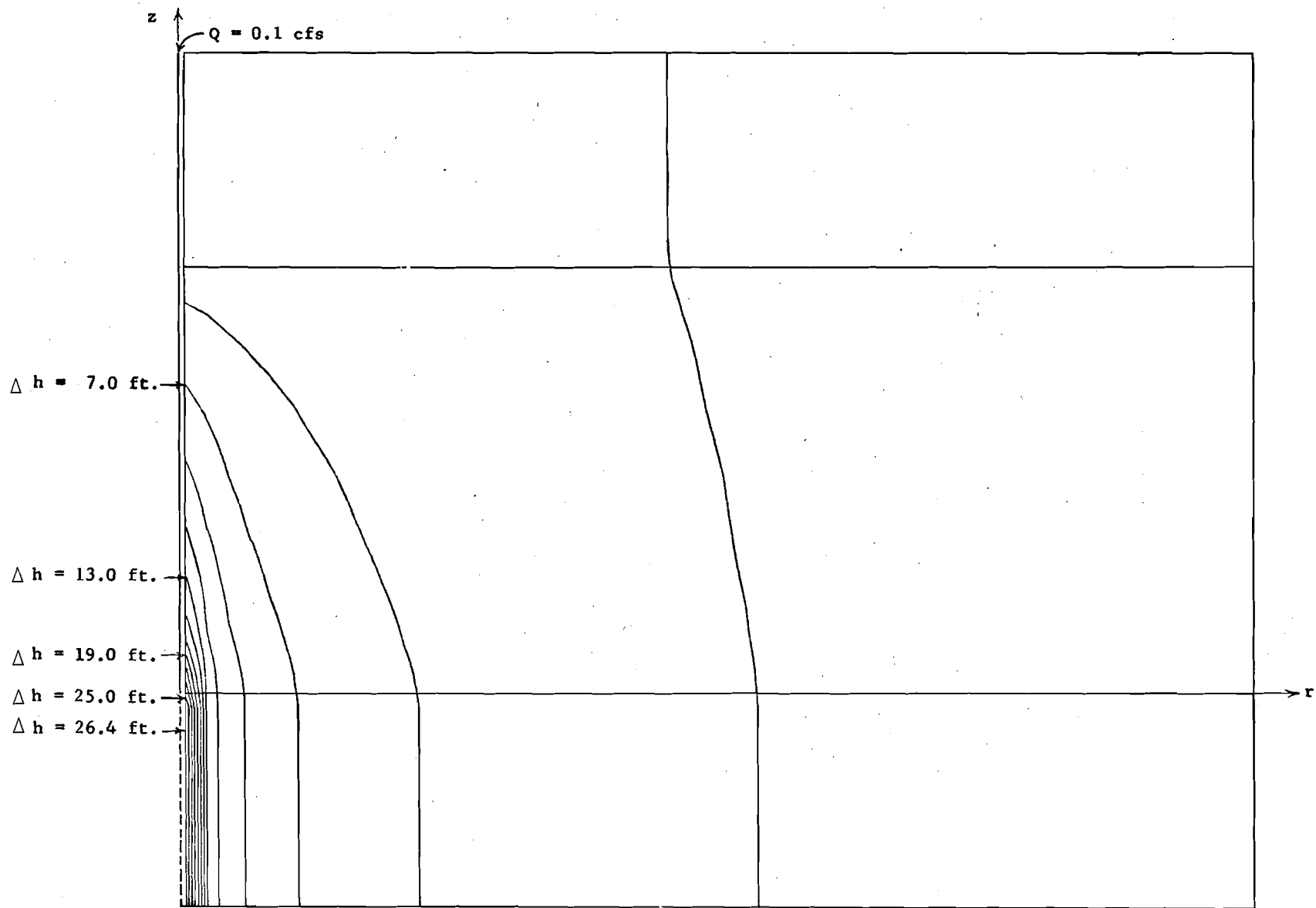


Figure 17: Approximate contours of equal  $\Delta h$  for the three-layered leaky system at the end of build-up period.

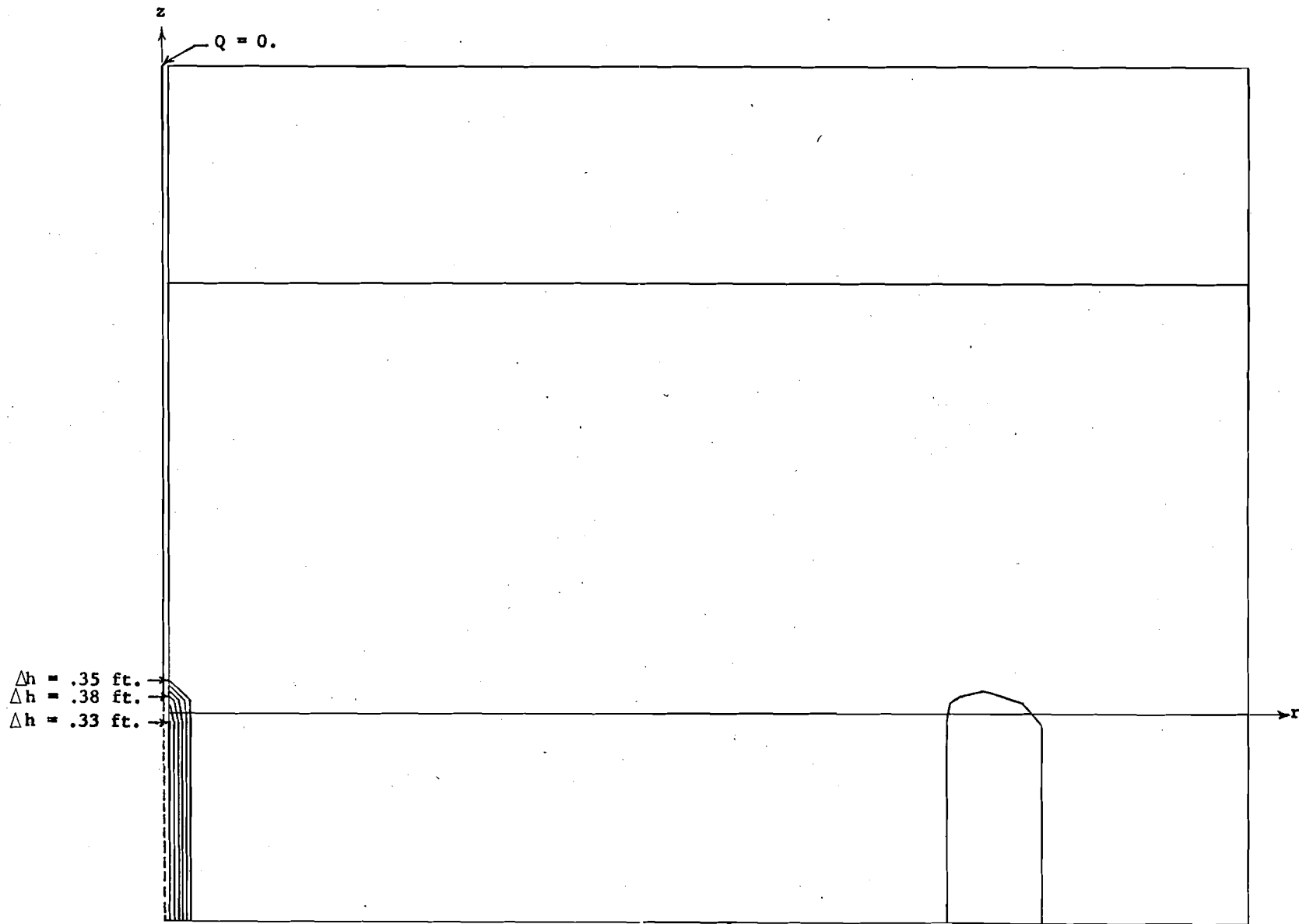


Figure 18: Approximate contours of equal  $\Delta h$  for the three-layered leaky system at the end of decay period.

ultimate fate of wastes in the disposal formations is dependent on the geologic and hydrogeologic parameters, and the accuracy of the results is dependent upon these.

The existing data on the physical properties of the Mt. Simon sandstone are for the upper zones, generally 400 feet or less, of the formation. Existing sources of data are deep water wells in the northern part of the state, underground gas storage tests, deep structural tests in oil producing regions, and existing disposal wells and tests. Maps of the geologic structure of the Illinois Basin have been published (Bond, 1972) and are adequate for use in defining the geometry of the Mt. Simon sandstone.

The Mt. Simon sandstone is the disposal zone of most interest because of its isolation from surface environments and its great thickness and because its native waters are brines in most of Illinois. (In this report, Mt. Simon refers to the hydrologic unit consisting of the Mt. Simon Formation and the contiguous sandstones in the lower Eau Claire Formation; the latter sandstone may be as thick as 400 feet.) The Mt. Simon varies in thickness from less than 1000 feet to over 2500 feet, and is generally greater than 1500 feet thick in most of the region where it would be considered as a liquid waste injection zone (Figure 20). However, the sandstone has been shown to be absent in the top of the Mt. Simon; the Mt. Simon varies from less than 1000 feet below sea level to over 12,000 feet (Figure 21).

The Mt. Simon is dominantly a fine- to coarse-grained sandstone with some beds of shale and siltstone. The formation is moderately compact and friable in the north and northwest, where it also yields potable water. Deep in the basin and on its eastern margin, the sandstone is well cemented and quartzitic. In these regions, the formation has a relatively low porosity and permeability.

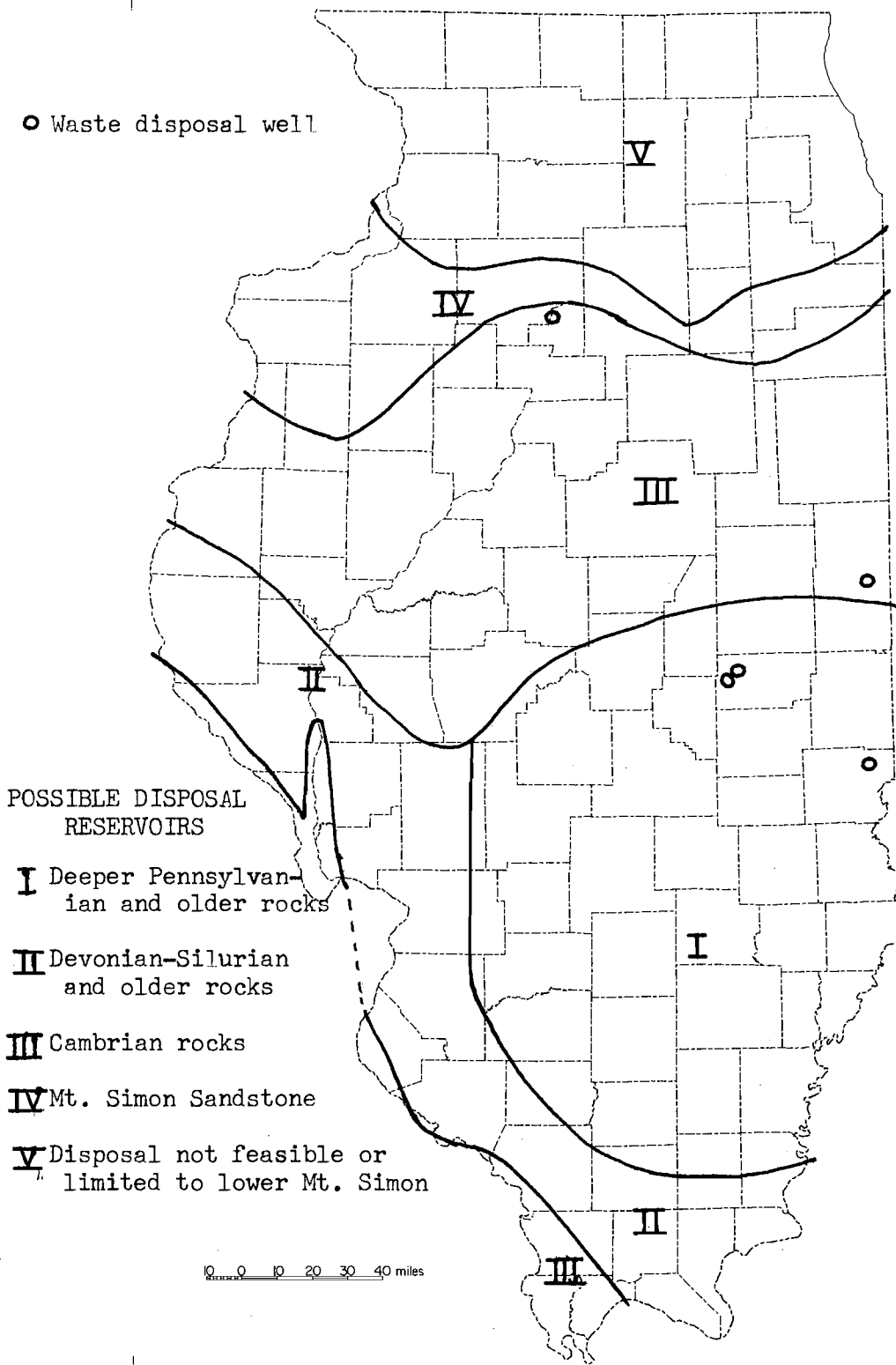


Figure 19- Possible disposal reservoirs for liquid wastes. The southern boundaries of areas V and IV are the approximate southern limits of use of the Mt. Simon and Ironton-Galesville Sandstones, respectively, as sources of potable water.

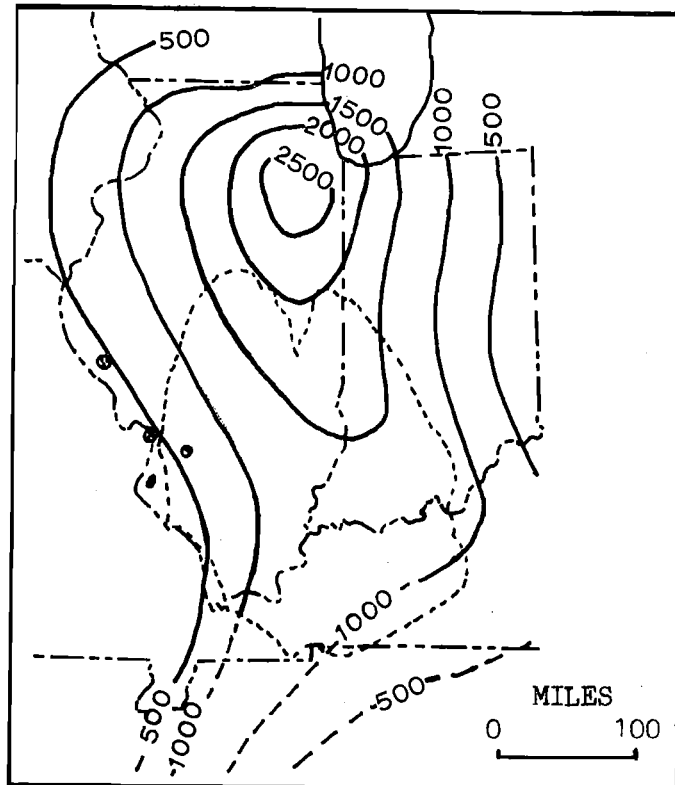
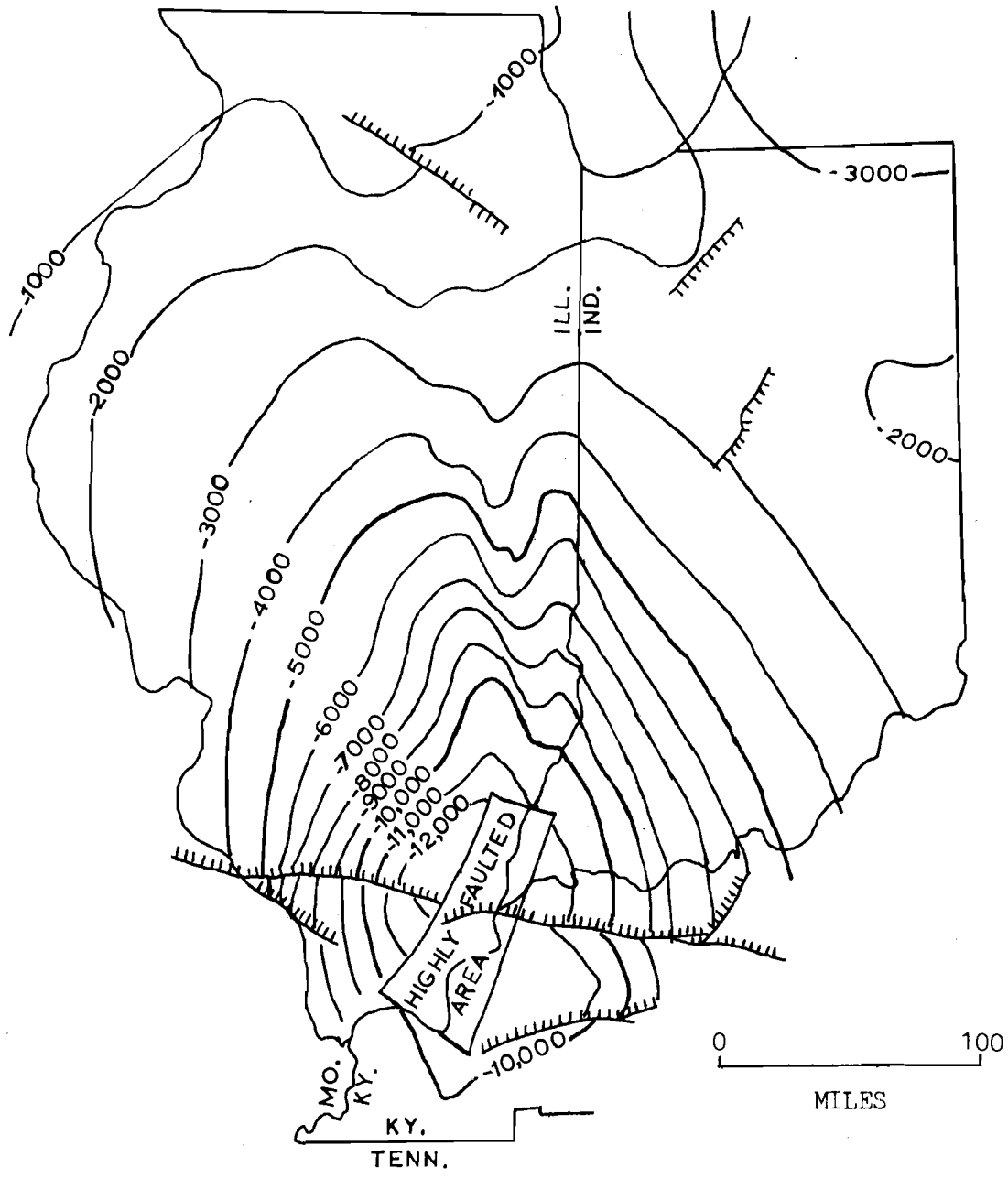


Figure 20: Thickness of the Mt. Simon sandstone, contour interval 500 feet. Dashed line represents the deep part of the Illinois basin. (after Bell et al., 1964).





—3000— Contour, interval 1000 feet, datum mean sea level

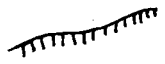
 Fault, downthrown side indicated

Figure 21: Structure on top of Mt. Simon Formation. (After Bond, 1972)

The geology in the Hennepin region was described by Cady (1919). McComas (1968) related the geology of the area to planning and environmental problems. The following brief discussion is taken from these publications.

The Jones & Laughlin disposal well is situated in a structurally low trough about 15 miles west of the crest of the La Salle anticlinal belt (Figure 22). A brief description and lithic log of the well is shown in Figure 23. At Hennepin the Mt. Simon aquifer is about 1800 feet thick. The well is open through the entire aquifer and through most of the overlying Eau Claire. Cores were taken over seven different intervals in the aquifer, totaling about 275 feet of core. Measurements of horizontal permeability and porosity were made by a commercial laboratory on the core. Measured permeabilities varied from about 5 to 500 millidarcys; porosity measurements varied between 8 and 22 percent. Measured porosity correlated well with geophysical logs (sonic and porosity logs) and the data were easily extrapolated for the entire section. Permeabilities were compared to electric and gamma logs and estimated for the whole aquifer.

No porosities or vertical permeabilities were measured on cores of the Eau Claire Formation, which serves as the confining layer. However, data from underground gas storage structures in northern Illinois suggest vertical permeabilities vary from 0.003 to 0.000006 millidarcys and porosity from 7 to 15 percent. Values were again assigned to specific layers on the basis of the geophysical logs.

Water resource studies (Hoover and Schicht, 1967, and Walton and Csallany, 1962) in the La Salle region, 15 miles northeast, show that the Iron-ton-Galesville sandstone has an average horizontal permeability of 500 millidarcys and 18 percent porosity. These values were assumed to be representative of the formation at Hennepin. These data are summarized in Figure 24.

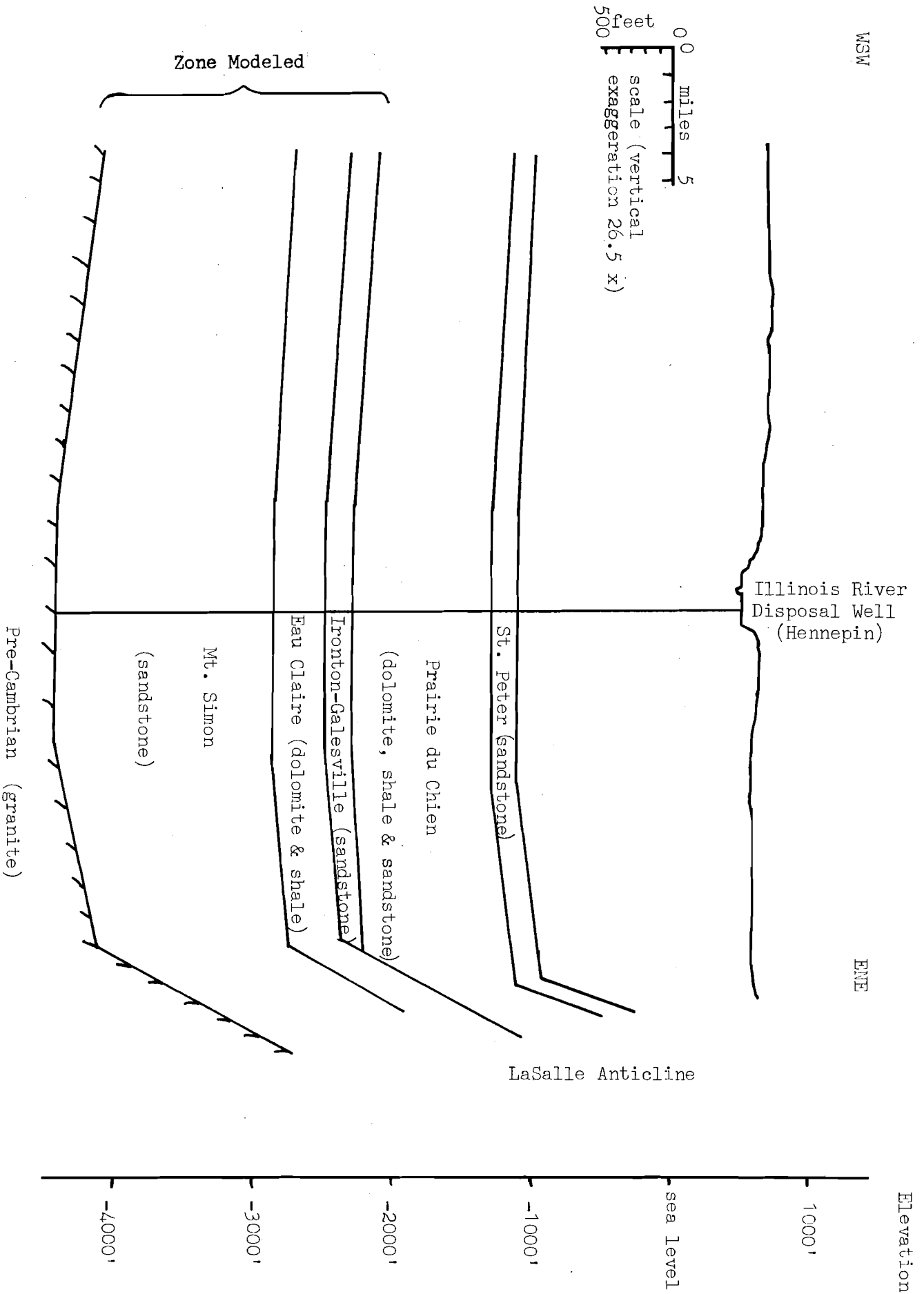
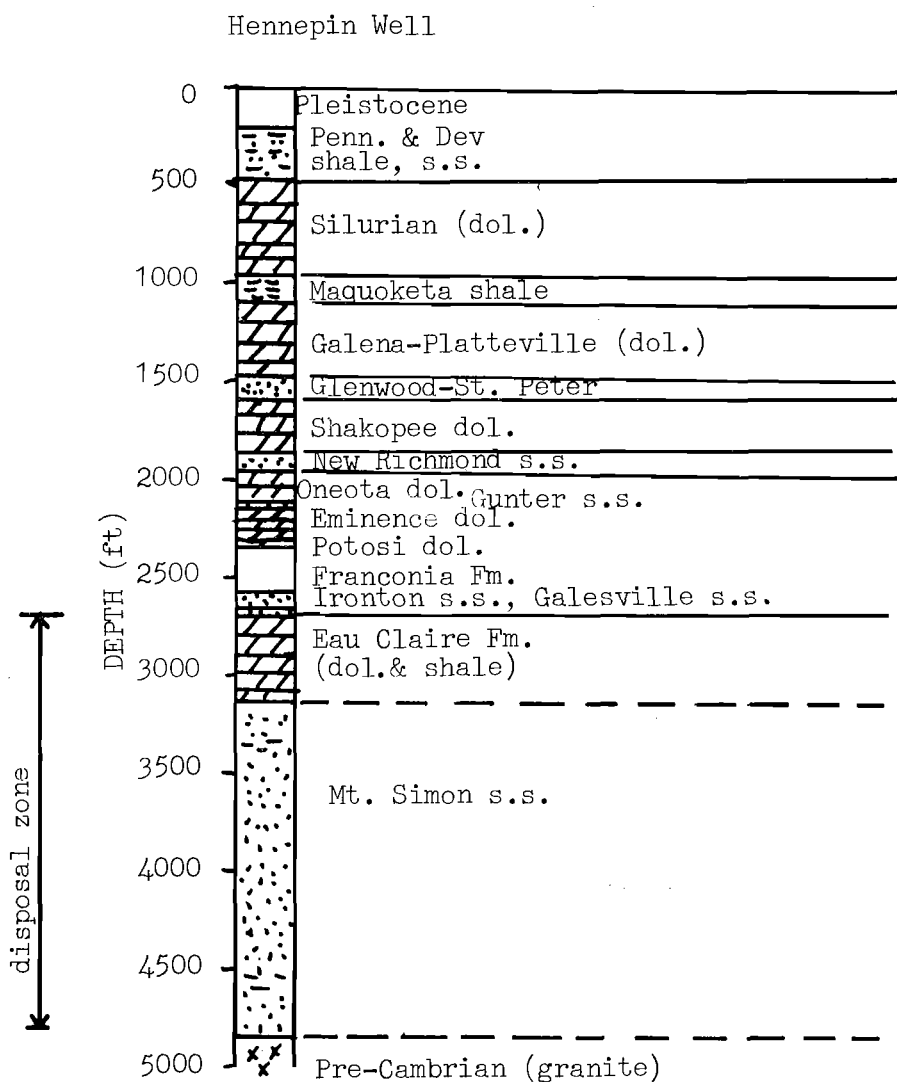


Figure 22. Geologic cross section (west-southwest: east-northeast) through the Hennepin region showing the section modeled in this study (Modified from Cady, 1919, and McComas, 1958).



Hennepin deep disposal well  
SW, SW, 3-32N-2W  
T.D. 4877'

Figure 23. Geologic section at the Jones & Laughlin waste disposal well, Hennepin, Illinois.

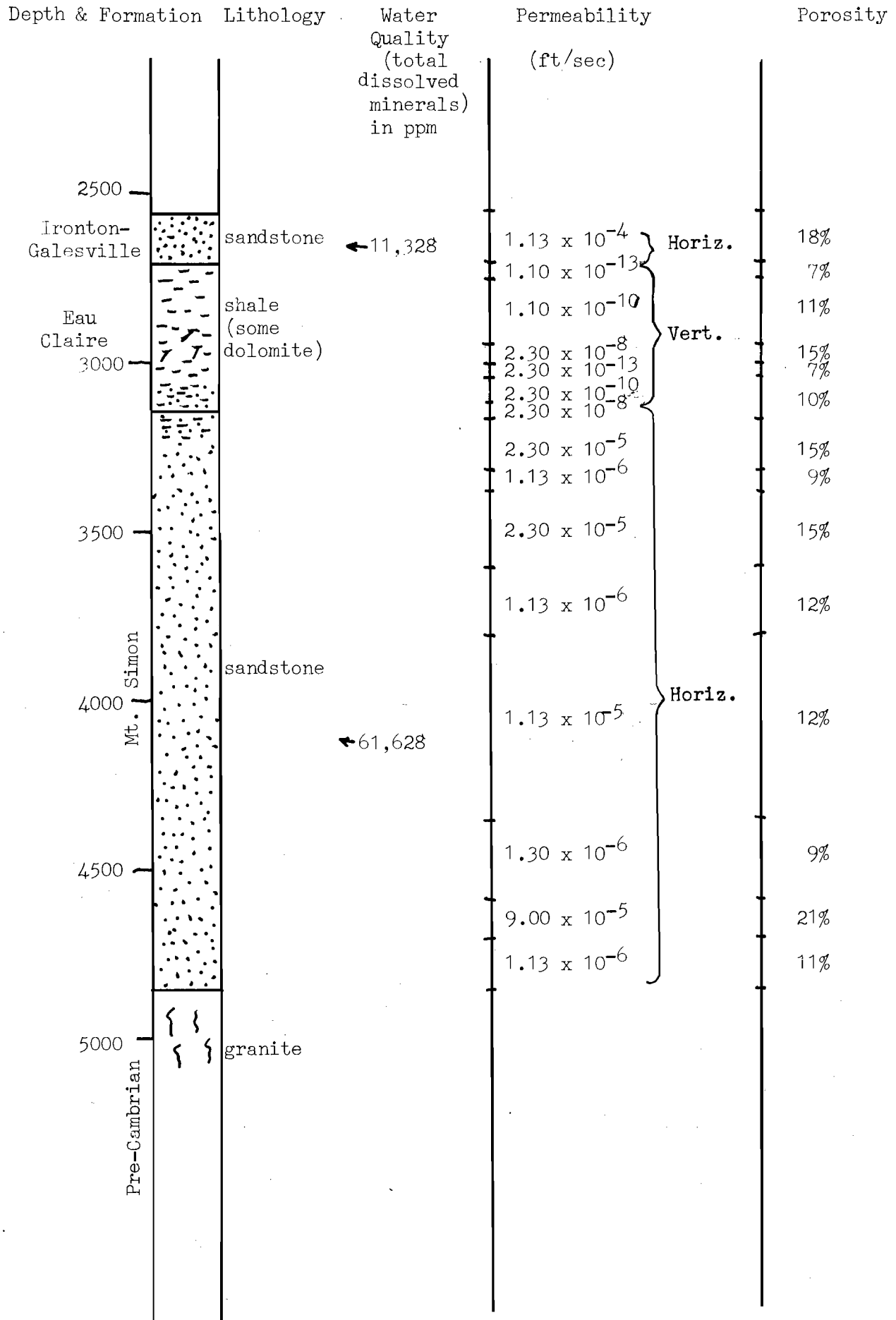


Figure 24. Details of the modeled zone in the Jones & Laughlin well at Hennepin, Illinois.

The Jones & Laughlin disposal well was cased into the top of the Eau Claire Formation. Although the injection tubing extends another 1500 feet down into the Mt. Simon, fluid is free to migrate up the open hole to the casing point. It is not necessary to protect the water quality in the Ironton-Galesville as the total dissolved minerals are in excess of the 10,000 ppm limit of the Illinois Environmental Protection Agency for useable waters. Nevertheless, it is probably undesirable to have large volumes of waste enter the overlying formation, since it is used as a fresh-water source only 15 miles away.

ii) Solution by Finite Element Method: Based on the geologic studies of the previous section, a model containing 864 nodes and 840 elements representing a 2291-foot geologic section was set up. This idealized model consisted of a 152-foot thick layer of Ironton-Galesville underlain by Eau Claire (410 feet thick) and Mt. Simon (1729 feet thick), and it extended 10,500 feet radially. On the basis of the geological studies, this system was divided into 15 layers, each layer possessing anisotropic but homogeneous properties (Figure 25). The vertical permeabilities were set equal to 0.01 of the horizontal permeabilities.

Assuming that the density and the viscosity of the injected fluid are the same as those of the native fluid, the hydraulic properties of each layer were calculated. Table 3 summarizes these properties and shows the extent of each layer to which these properties apply. It is fully realized that the assumption of unique viscosity and density is not absolutely valid. There are vertical variations of density and perhaps viscosity in the native fluids, and the properties of the injected fluid may vary from these fluids. Due to injection, there also exist density currents which change these properties with time. Because of the limitations of this study, these considerations, which require substantial studies, could not be incorporated into the model.

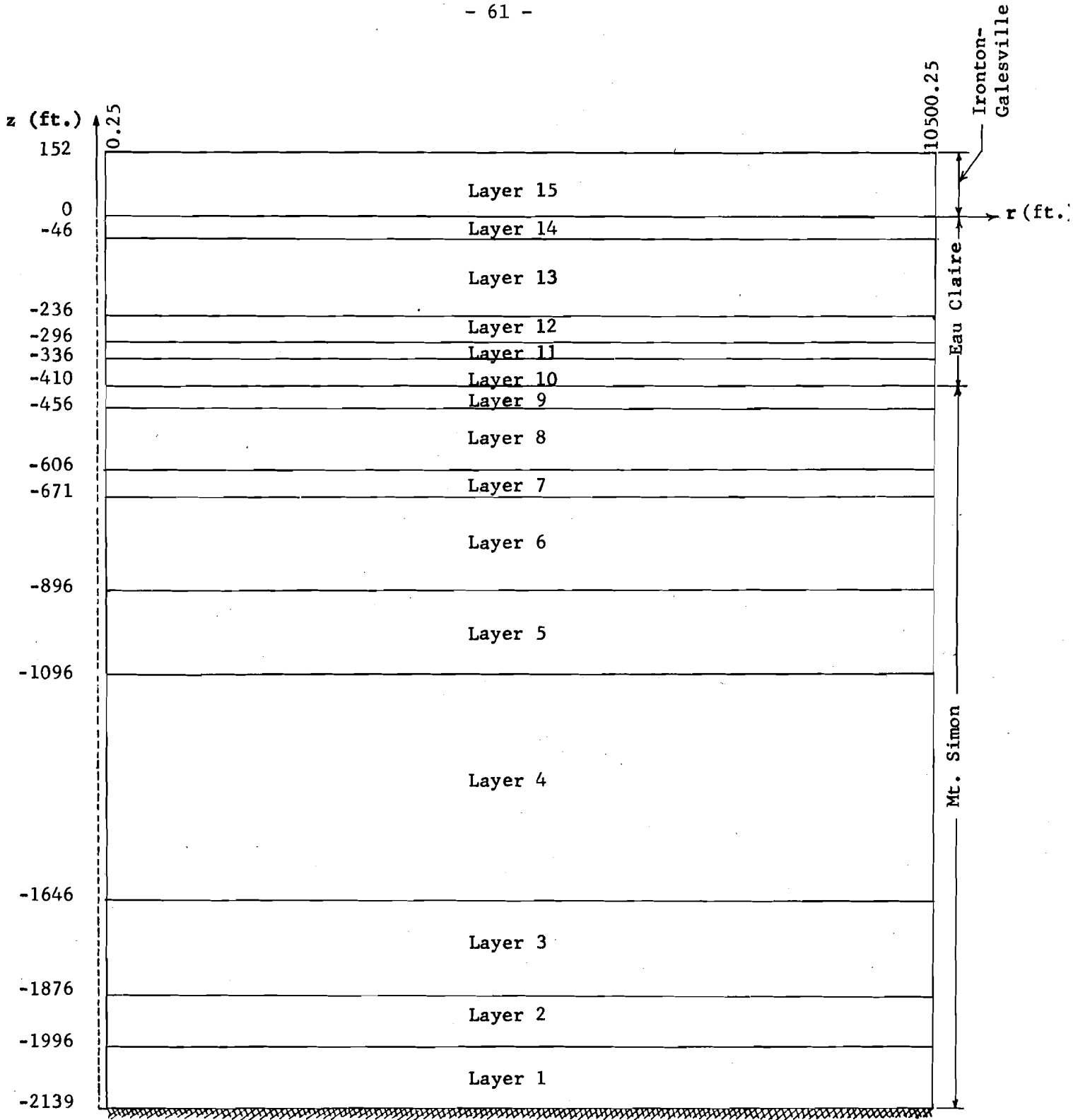


Figure 25: Schematic representation of the idealized layered system used to analyze the Jones and Laughlin Corporation's injection well at Hennepin, Ill.

	First Principal permeability $K'_x$ (ft/sec)	Second Principal permeability $K'_y$ (ft/sec)	Specific Storage $S_s$ (ft <sup>-1</sup> )	Angle between the Principal Axes and Horizontal and Vertical Axes (Degrees)	r-coordinate of the side of layer adjacent to well (ft)	r-coordinate of the side of layer away from well (ft)	z-coordinate of the bottom of layer (ft)	z-coordinate of the top of layer (ft)
1	0.11300E-07	0.11300E-09	0.57000E-06	0.0	0.25	10500.25	-2139.00	-1996.00
2	0.90000E-06	0.90000E-08	0.10900E-05	0.0	0.25	10500.25	-1996.00	-1876.00
3	0.11300E-07	0.11300E-09	0.47000E-06	0.0	0.25	10500.25	-1876.00	-1646.00
4	0.11300E-06	0.11300E-08	0.62000E-06	0.0	0.25	10500.25	-1646.00	-1096.00
5	0.11300E-07	0.11300E-09	0.62000E-06	0.0	0.25	10500.25	-1096.00	-896.00
6	0.23000E-06	0.23000E-08	0.78000E-06	0.0	0.25	10500.25	-896.00	-671.00
7	0.11300E-07	0.11300E-09	0.47000E-06	0.0	0.25	10500.25	-671.00	-606.00
8	0.23000E-06	0.23000E-08	0.78000E-06	0.0	0.25	10500.25	-606.00	-456.00
9	0.23000E-09	0.23000E-11	0.52000E-06	0.0	0.25	10500.25	-456.00	-410.00
10	0.23000E-09	0.23000E-11	0.52000E-06	0.0	0.25	10500.25	-410.00	-336.00
11	0.23000E-12	0.23000E-14	0.36000E-06	0.0	0.25	10500.25	-336.00	-296.00
12	0.23000E-07	0.23000E-09	0.78000E-06	0.0	0.25	10500.25	-296.00	-236.00
13	0.11000E-09	0.11000E-11	0.57000E-06	0.0	0.25	10500.25	-236.00	-46.00
14	0.11000E-12	0.11000E-14	0.36000E-06	0.0	0.25	10500.25	-46.00	0.0
15	0.11300E-05	0.11300E-07	0.97000E-06	0.0	0.25	10500.25	0.0	152.00

Table 3: Physical and geometrical properties of the model of Jones and Laughlin Corporation's injection well.



The injection in the Jones & Laughlin well at Hennepin usually takes place at the rates of about 130-180 gpm for periods of 10-20 hours through an open hole 6 inches in diameter which starts at the top of the Eau Claire Formation and extends to the bottom of the Mt. Simon sandstone. A total of 51 nodal points was set up on the vertical face of the geological section adjacent to the well. The number of nodes was reduced to 30 at a distance of 3000 feet from the well. Figure 25 represents the layers and their coordinates with respect to a z-axis passing through the center of the well and an r-axis passing through the interface between the Ironton-Galesville and the Eau Claire Formations.

The model was first tested as a totally homogeneous and isotropic system with wells penetrating the entire section to evaluate the mesh sizes. Figure 26 presents the data obtained from this test (symbols) as compared to the Theis's Curve (solid line). From calculations using the data in Table 3, Figures 27, 28, and 29 show the temporal change of pressure of the anisotropic, homogeneous model for nodal points at the interfaces of Ironton-Galesville and Eau Claire, and Eau Claire and Mt. Simon, and points in the middle of Mt. Simon, respectively, at distances specified on the figures. The build-up period lasted for 130.6 hours and was followed by 103.4 hours of decay period. To make the data points in the decay period distinguishable in Figure 27, the logarithmic time scale for this period was re-started as in Figures 13-15. Figure 27 shows the pressure increase at the interface of the Ironton-Galesville and Eau Claire and surface of open hole after 130.6 hours of continuous injection at the rate of 200 gpm is about 448.5 psi. A decay period of 103.4 hours reduces this increase of pressure to 5.2 psi. These data, shown in Figure 27, are more significant than the rest of the data presented in Figures 28 and 29.

The protection of the Eau Claire sandstone, the confining layer, from possible damage due to hydraulic fracturing must be assured since the fracturing

$$\text{Dimensionless potential} = \frac{2\pi K H}{Q s} \Delta\phi$$

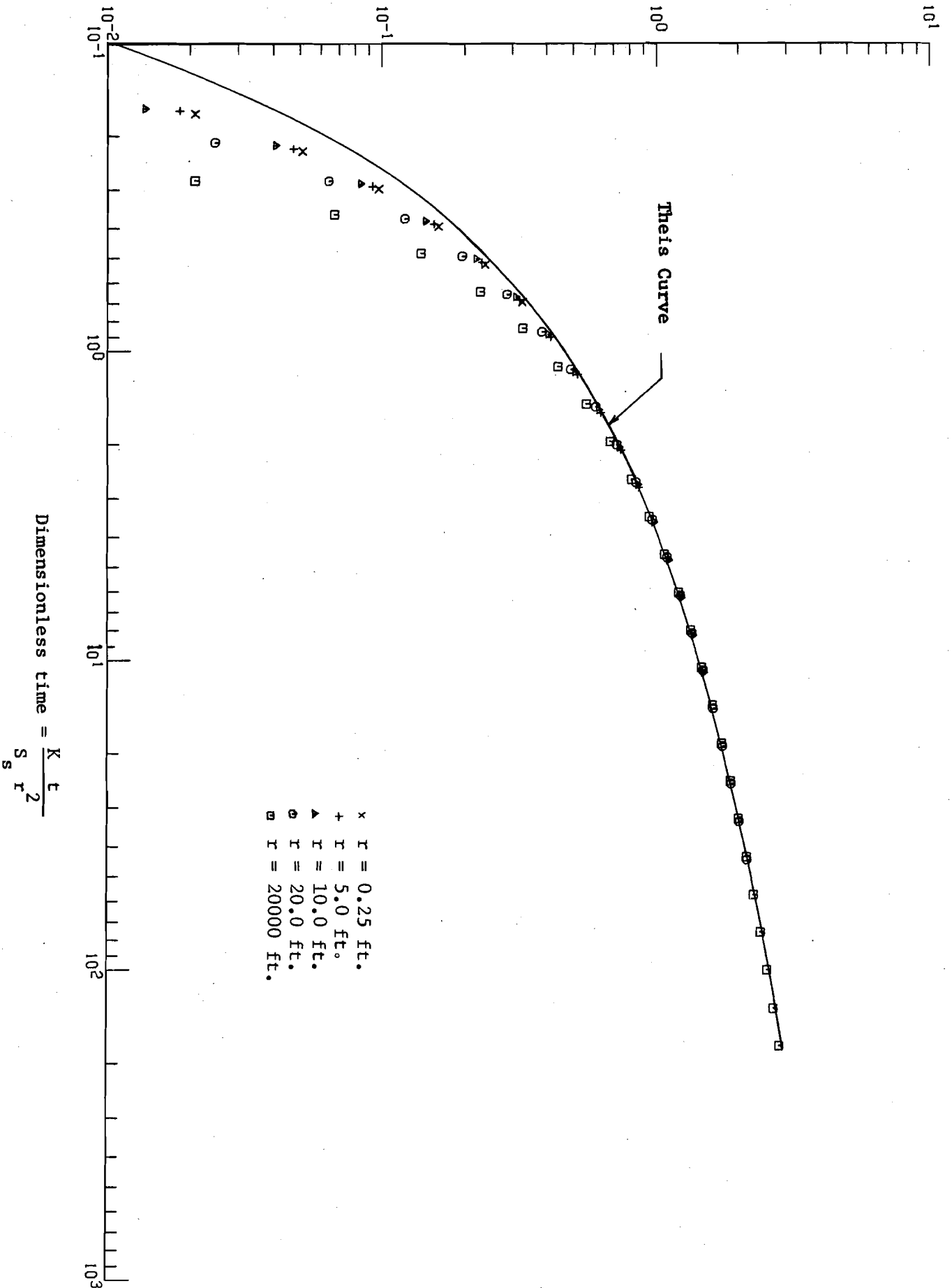


Figure 26: Results of finite element method (symbols) for the grid spacing used in the Jones and Laughlin Corporation's injection well with homogeneous and isotropic properties, as compared with analytical solution (Theis, 1935).

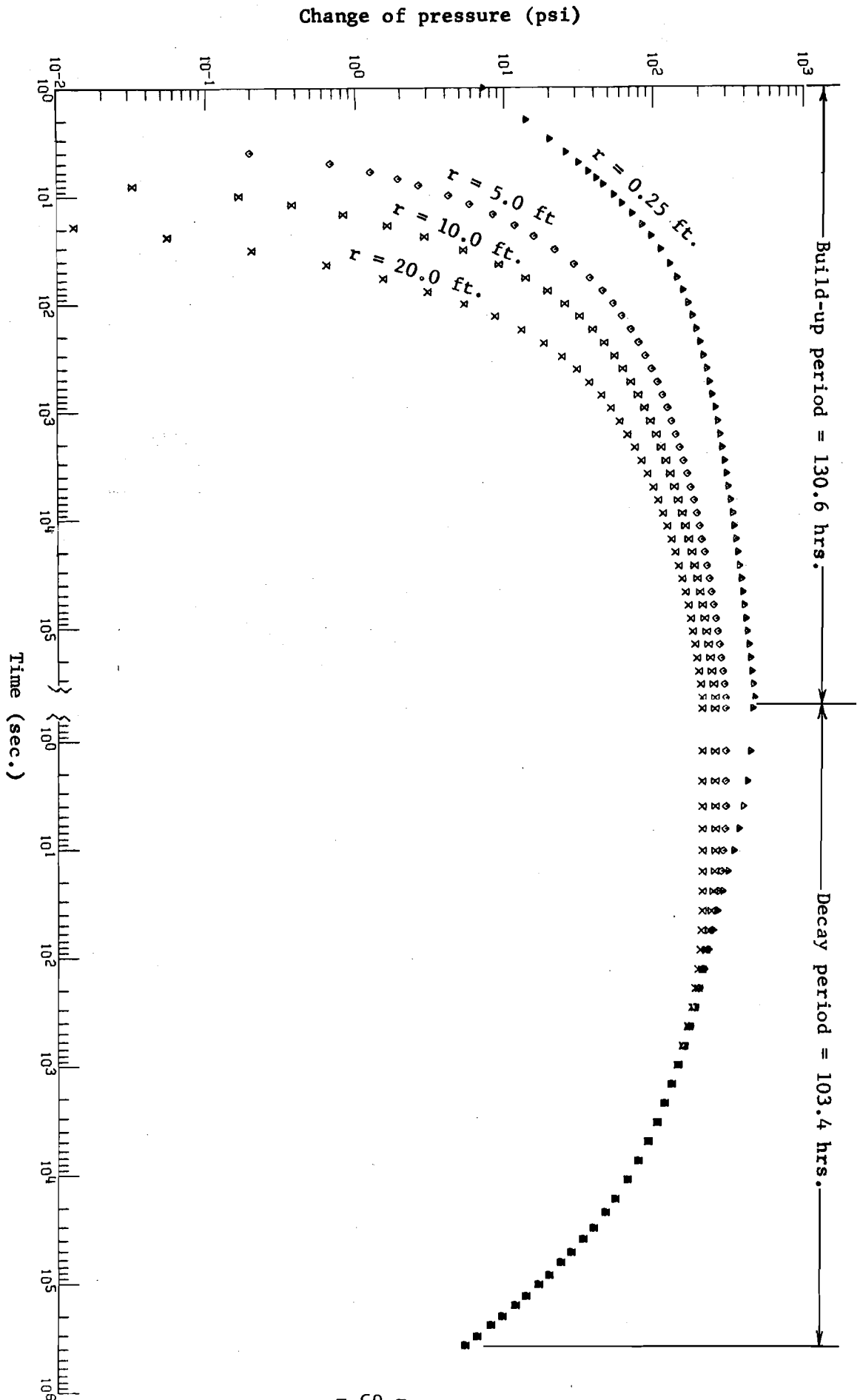


Figure 27: Change of pressure vs. time calculated at the interface between Ironton-Galesville and Eau Claire for different distances from the center of model.

Change of pressure (psi)

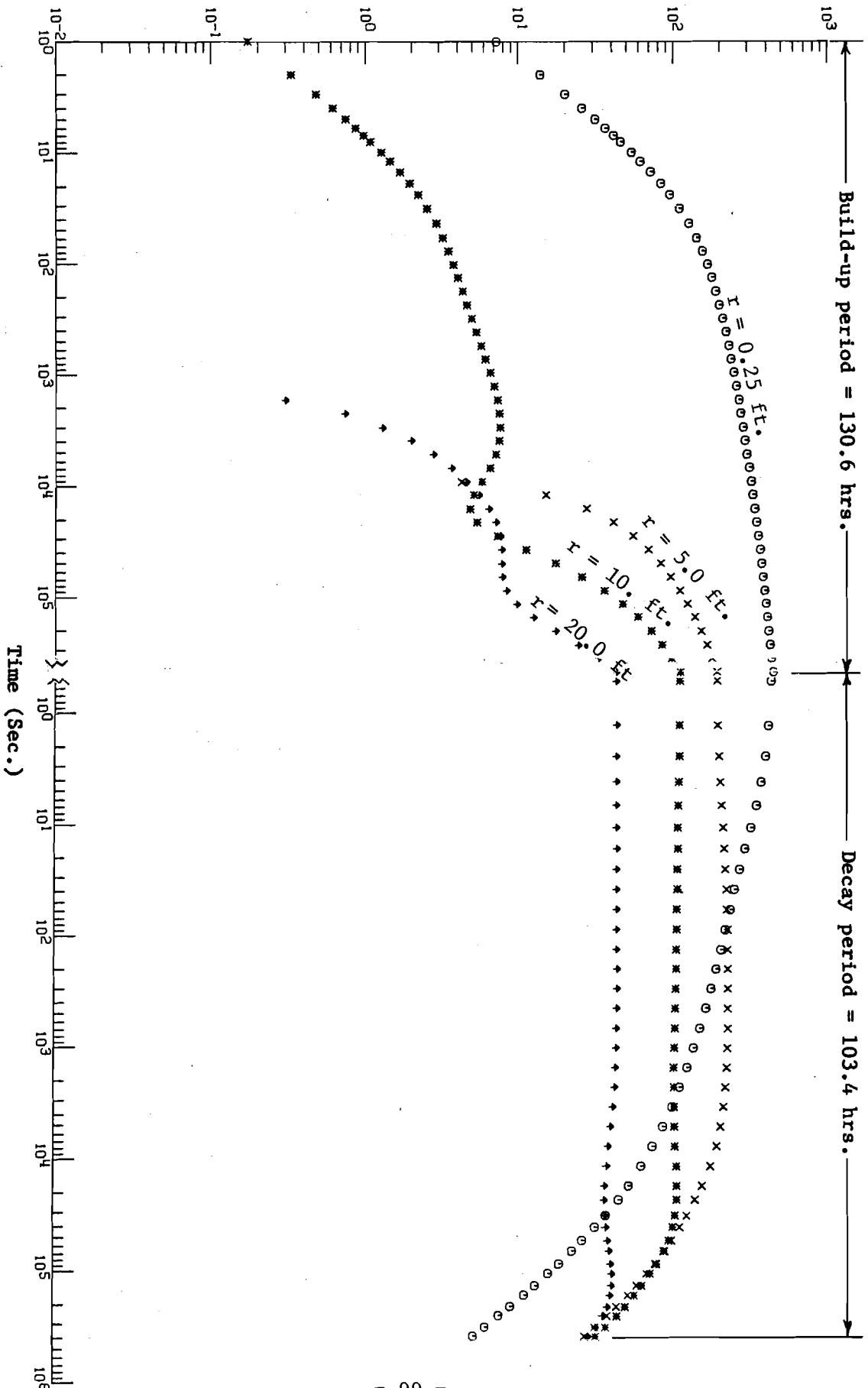


Figure 28: Change of pressure vs. time calculated at the interface between Eau Claire and Mt. Simon for different distances from the center of model.

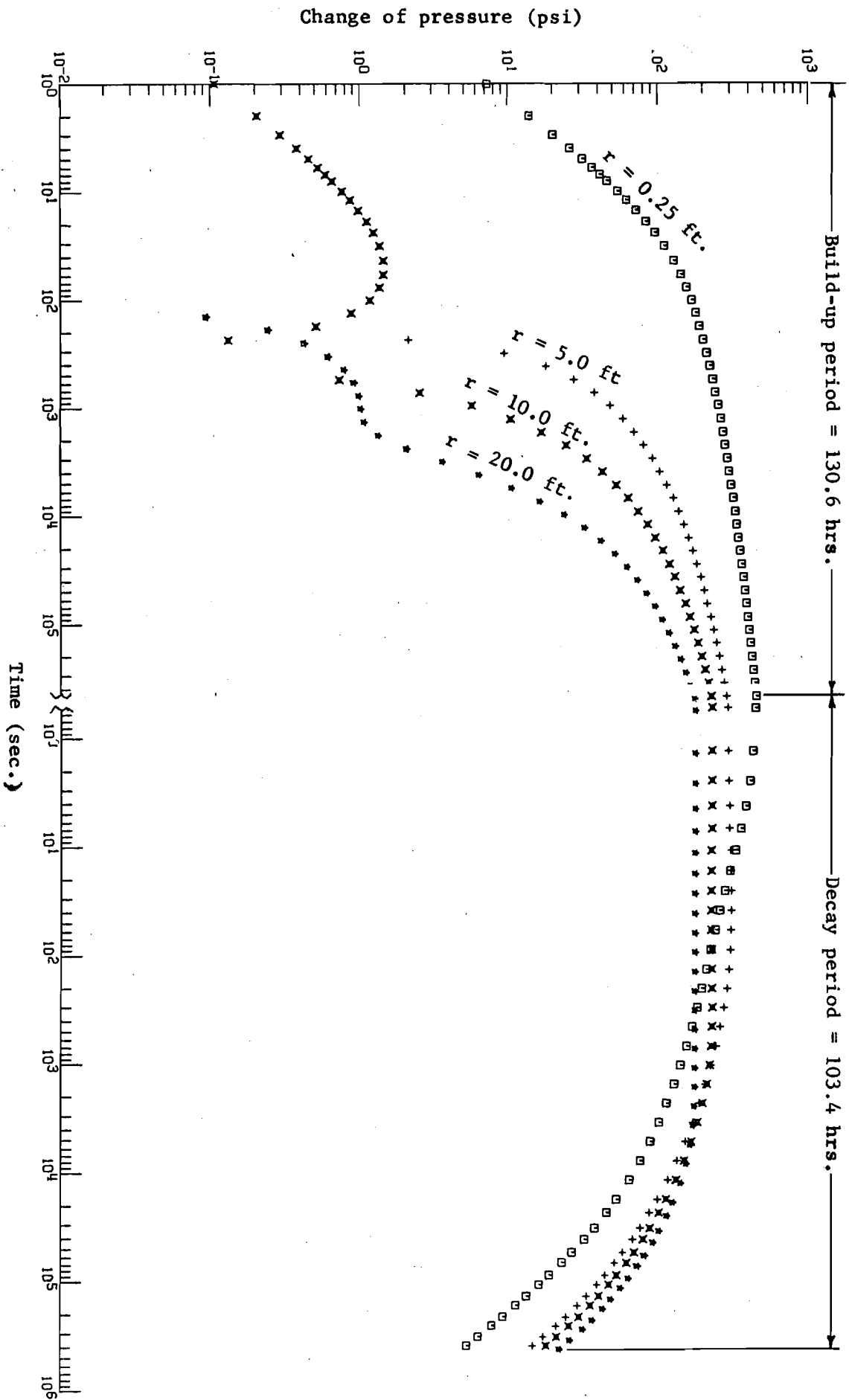


Figure 29 : Change of pressure vs. time calculated at the middle of Mt. Simon for different distances from the center of model.

could result in a large quantity of waste seeping into the Ironton-Galesville aquifer. If the undisturbed fluid pressure which existed within the rocks before injection started had been known, then using the data in Figure 27 we would have been able to predict the minimum injection pressure which would have produced hydraulic fracturing. The length of the continuous injection time required to cause such a fracture could have been estimated using the procedure summarized by Hubbert and Willis (1957). However, since these data are not available, only a gross approximation is possible.

Bond (1972) has tabulated the densities and the elevations of fresh-water head in the Ironton-Galesville for some wells in Illinois and Indiana. From the two data points given by Bond in the Troy Grove gas storage field in La Salle County and the Lake Bloomington gas storage field in Livingston County (approximately 20 and 30 miles, respectively, from Hennepin), the height of the column of fresh water above the interface of the Ironton-Galesville due to undisturbed pore pressure,  $p$ , may be approximated to be 2700 feet, which is equivalent to about 1170 psi. Figure 27 indicates that, subject to the assumptions made in this study, the increase in pressure,  $\Delta P$ , at the surface of the well at the interface between Ironton-Galesville and Eau Claire is about 448 psi at the end of the build-up period. Thus, the total pressure:

$$P = p + \Delta P = 1170 + 448 = 1618 \text{ psi}$$

which when divided by depth  $z = 2704$  feet is  $\frac{P}{z} = 0.60$  psi/ft.

On the basis of the following assumptions 1) incipient normal faulting, which may be justified for northern Illinois, 2) an overburden pressure per unit depth of 1.0 psi/ft., and 3) an original pore fluid pressure per unit depth of 0.46 psi/ft., Hubbert and Willis (1957) calculated the minimum value of the injection pressure which may cause fracture for Gulf Coast rocks to be about

0.64 psi/ft. Realizing that the length of the injection period and the rate of injection are both much greater than practiced at the well (130.6 hours with an injection rate of 200 gpm as compared to a maximum of 20.0 hours with an injection rate of 130 to 190 gpm normally practiced), and that the more indurated rocks of the Illinois Basin have greater strength than the rock of the Gulf Coast, we may conclude that the possibility of damage due to hydraulic fracturing is small.

The results in Figures 8 and 9 show irregularities such as those in Figures 13, 14, and 15 during early stages of injection. Again, it may be speculated that these discrepancies are due to a combination of coefficients of permeability and storage, time increment, and mesh sizes. It must be mentioned that the nodes used to obtain the data of Figures 8 and 9 are located in layers with much lower coefficient of permeability than that of Figure 27. The effect of low coefficient of permeability may also be observed in the decay curves. After 103.4 hours of decay all the changes of pressure of all the nodes in Figure 7 are about the same, while that of Figures 8 and 9 still show noticeable differences between themselves and the node at the surface of the well, i.e.,  $r = 0.25$  ft. Careful examination of the decay curves of Figures 8 and 9 shows that the change of pressure keeps increasing for a while after the build-up period ends for nodes close to the well ( $r = 5.0$  feet). The coefficients of storage for the elements surrounding these nodes are not unusually higher than that of the rest of the elements. The coefficients of permeabilities of the elements surrounding these nodes and the high gradient of pressure are apparently the cause of this increase in  $\Delta P$  after the build-up period ends.

Figure 30 represents the contour of equal  $\Delta P$  within a radial distance of 40.25 feet from the center of the well at the end of the build-up period. The high pressure gradients in the area around the bottom of casing immediately focusses our attention on these areas. If there ever occurs a failure due to

hydraulic fracturing, most probably it will be in this area. Figure 31 shows the approximate contours of the equal  $\Delta P$  at the end of decay period. The high pressure gradient at the bottom of the casing has disappeared and possible damage to upper layers is minimized by the end of the time period used in the calculation for the decay period. Because of the high permeability of the Ironton-Galesville, pressure in this formation is dissipated faster than the rest of the system, and there seems to be little danger of fracturing overlying formations.



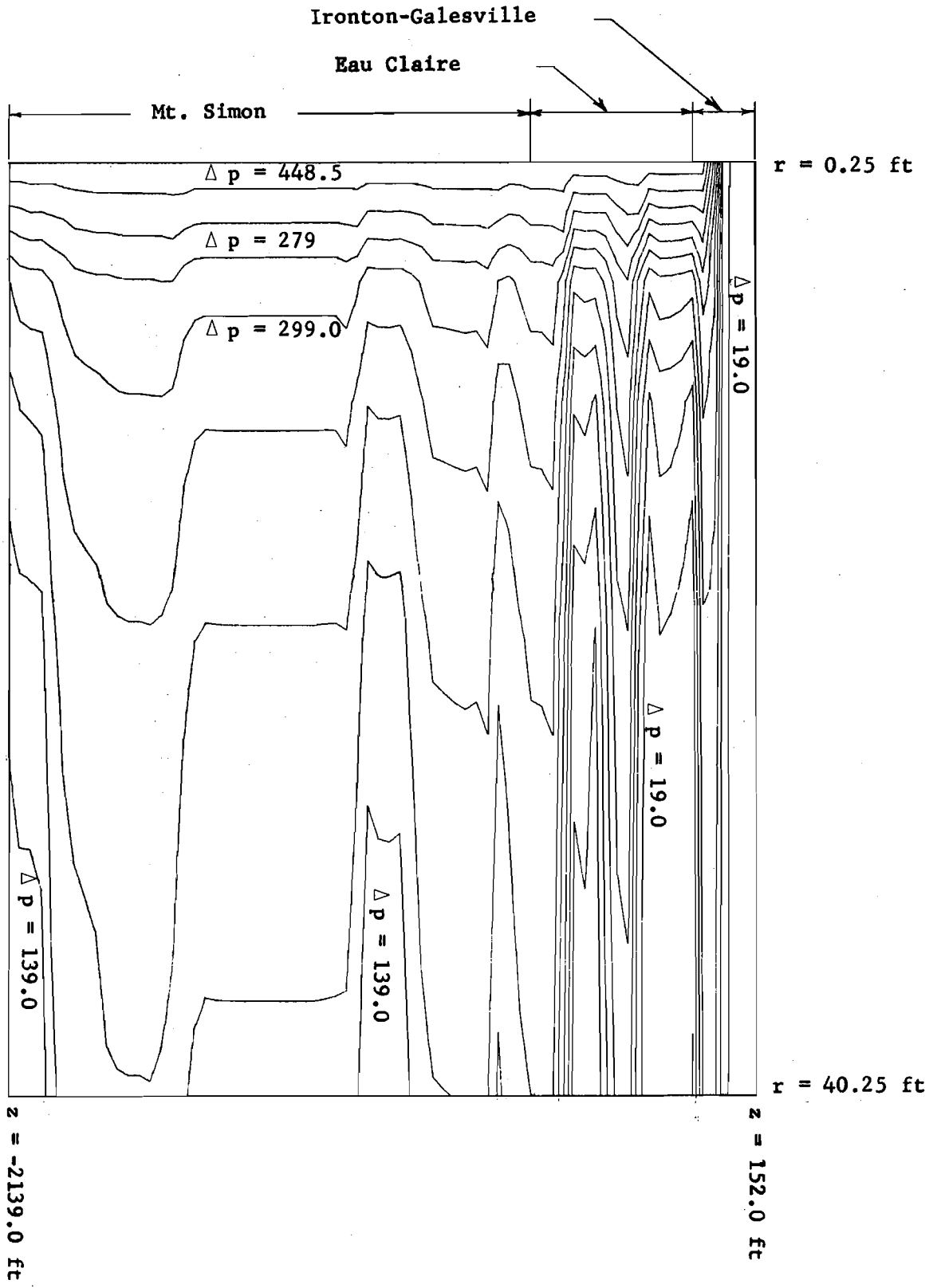


Figure 30 : Approximate contours of equal  $\Delta p$  (psi) for the model of Jones and Laughlin Corporation's injection well at the end of the build-up period.

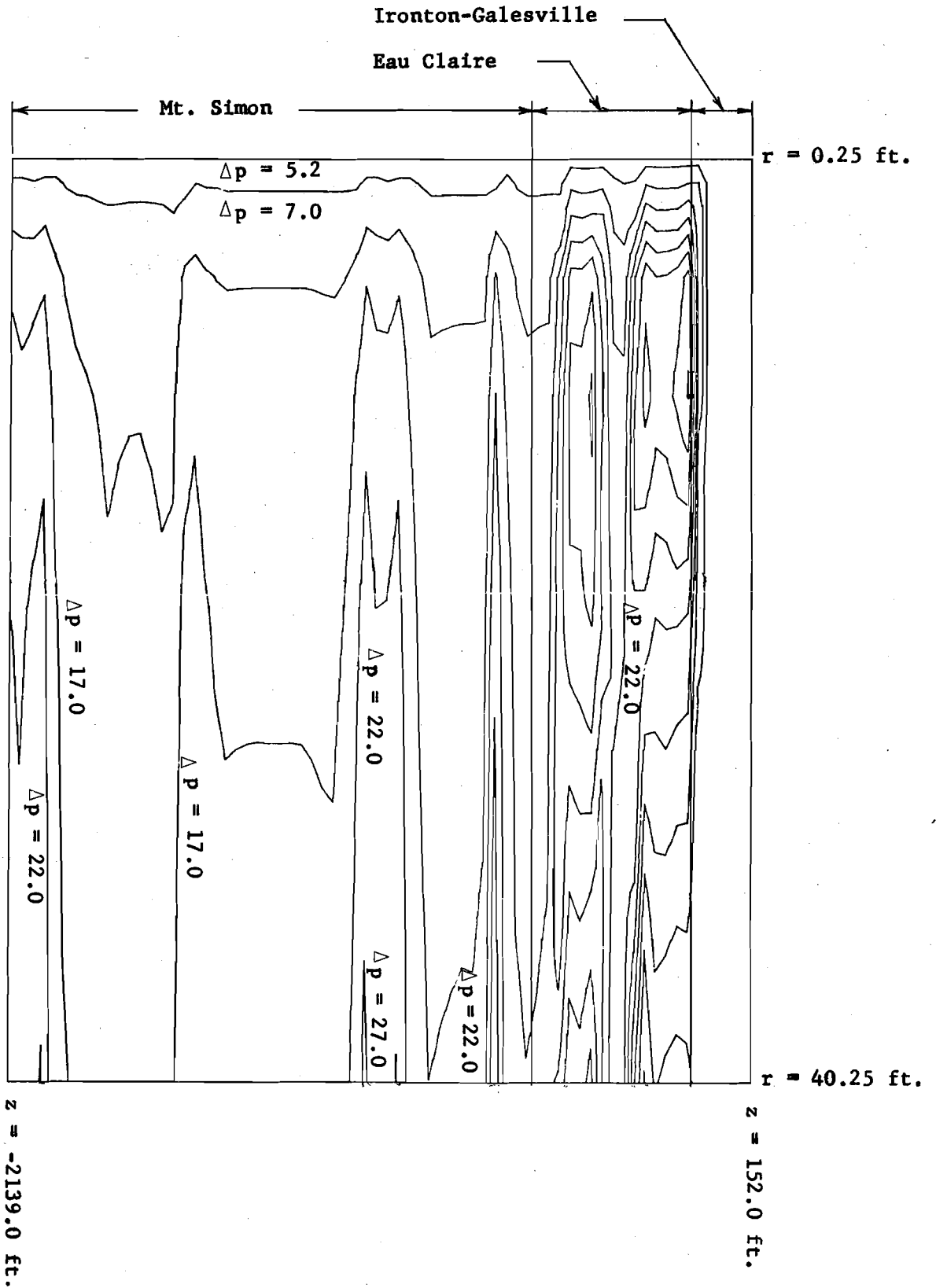


Figure 31: Approximate contours of equal  $\Delta p$  (psi) for the model of Jones and Laughlin Corporation's injection well at the end of decay period.

### III. AN ADAPTIVE - CHEBYSHEV - FACTORIZATION METHOD FOR THE SOLUTION OF LINEAR EQUATIONS ARISING FROM THE NUMERICAL SOLUTION OF FLOW EQUATION

#### A. Introduction

The solution of steady state, transient or eigenvalue problems by the finite element method leads ultimately to a system of linear equations here called a finite element system. We shall describe a factorization method to factor "approximately" the matrix of the system into sparse upper and lower triangular matrices. This factorization may be used with an iteration due to Diamond (1971) for the efficient solution of the system. In this paper, Diamond's algorithm is called an adaptive Chebyshev factorization (ACF) algorithm.

It is customary to solve finite element systems by a direct method such as Gaussian elimination or the equivalent Cholesky LU-decomposition. Iterative methods have not been popular for solving finite element systems for a variety of reasons (George, 1971): (a) finite element systems are of lower order than finite difference systems when each yields the same accuracy; (b) iterative refinement may be used with direct methods to gain estimates of the condition of the system; and (c) it is common for more than one right hand side to be processed.

These three reasons are important but sound somewhat like arguments that may once have been made in favor of rail travel over airlines. As the magnitude of the problem increases, these reasons fail to justify the blind application of direct methods. Consider for example, the flow equations that arise in ground-water problems. These are time dependent problems that may be solved by approximating the space derivatives by

the finite element method to obtain a system of ordinary differential equations. The solution of this system is then computed by discretizing the time derivative. If the backward difference method is used, each time step requires the solution of a finite element system. Typical studies require 1000 time steps. With time dependent coefficients, no economy is possible from solving the same system repeatedly by a direct method. Currently, only two-dimensional studies are considered feasible. Some technique more efficient than conventional direct methods is essential to make three-dimensional problems practical.

Another reason why direct methods are used exclusively at the present is that the iterative method with which direct methods are being compared is successive-over-relaxation (SOR), which is often disappointing in practice due to slow convergence and the need for a "good" relaxation parameter. One may expect SOR to converge somewhat slowly for finite element systems since this is characteristic of SOR for finite difference systems. It is natural to try to apply Alternating-Direction-Implicit (ADI) to these systems since they are faster for difficult finite difference systems. However, the increased complexity of finite element systems makes matrix commutativity difficult to study and eigenvalue estimates laborious to obtain, both essential to the success of ADI. Nevertheless, it should be noted that Douglas and Dupont (1971) have described an ADI procedure for finite element systems.

Other arguments are also made in favor of Gaussian elimination over iterative methods (George, 1971). One is that storage is becoming cheap and plentiful. For years, many people have favored iterative methods

for solving finite difference equations primarily because they require less storage. Slow convergence was tolerated to gain this advantage. Because of hardware improvements such as bulk core storage (Fuchel and Heller, 1969) and features such as virtual memory (Daley and Dennis, 1969), the importance of conserving storage is fading away. Yet this cannot mean the unlimited use of Gaussian elimination. As storage problems become minor, run time for Gaussian elimination remains a formidable barrier to the solution of large systems that low cost memory cannot affect. Run time can be reduced only by more efficient methods than Gaussian elimination, such as the direct methods discussed by Buzbee, Golub and Neilson (1970) or else rapidly convergent iterative methods.

A final argument in favor of direct methods is that data management is simpler for elimination schemes than for iterative methods when the triangular mesh is arbitrary. The result of an arbitrary mesh is a sparse matrix with irregular structure. Techniques for efficient storage of sparse matrices adapt to elimination schemes more easily than to iterative methods. Nevertheless, fast iterative methods would compel reconsideration of this reasoning. An efficient iterative method for regular meshes would justify study of the complex problem of applying iterative methods to the solution of systems resulting from arbitrary meshes.

## B. Factorization Methods

Let  $Au = s$  be the system of equations obtained by a finite element approximation, where  $A$  is a positive definite and, therefore, symmetric square matrix. If  $B$  is a matrix for which the Cholesky LU-factors of  $A + B$  are sparse then the sequence defined by

$$(A+B)u^{(i+1)} = (A+B)u^{(i)} - \tau_i (Au^{(i)} - s), \quad (35)$$

for  $\tau_i$  a parameter, is efficient to compute. Therefore, if the sequence converges to  $u$  and converges rapidly, the result is a practical method to approximate the solution. Methods to solve  $Au = s$  that combine an iteration such as (35) with an algorithm to construct  $A + B$  with sparse LU factors are called factorization methods. For example, successive-over-relaxation may be written in form of (35) with elements of  $B$  explicitly equal to constant multiples of elements of  $A$ . Matrix  $A + B$  is called an approximate factorization (of  $A$ ) and (35) is called a factorization procedure. If  $A + B$  is positive definite a sequence of parameters  $\tau_0, \dots, \tau_n$  may always be chosen to obtain convergence. Moreover, an optimum sequence of parameters,  $\tau_0, \dots, \tau_n$  may be computed to minimize the error after  $n$  steps. Optimum parameters result from a routine use of Chebyshev minimax theory. Two requirements are essential to apply the theory. One is that  $A + B$  be positive definite. This implies that the eigenvalues of  $(A+B)^{-1}A$  are positive, since this matrix is similar to a positive definite matrix. Unless these eigenvalues are positive there is no convenient, effective method to apply Chebyshev theory to obtain decent

parameters. If they are positive it is trivial to compute optimum parameters using Chebyshev theory from the largest and smallest eigenvalues of  $(A+B)^{-1}A$ .

The second requirement is the practical one that the largest and smallest eigenvalues of  $(A+B)^{-1}A$  be available to compute with. It is a common difficulty with iterative methods for solving linear systems that a numerical value of some parameter is required in advance of running the iteration. For example, an accurate value of the optimum relaxation parameter is essential in order that SOR run efficiently. Also, precise eigenvalue estimates are critically important in order to obtain optimum or even reasonably effective ADI iteration parameters for solving finite difference equations. In each case, a priori estimates valid for a class of matrices are often satisfactory to use. The difficulty with estimating iteration parameters is more acute for factorization procedures because the techniques that work for ADI or SOR are no longer effective. The regularity of the matrices that appear in SOR or ADI makes it possible to obtain good a priori estimates in most cases, whereas the complicated properties of the elements of B in factorization procedures, hinder effective estimation of extreme eigenvalues of  $(A+B)^{-1}A$ , so that a priori estimates of Chebyshev parameters to use in factorization procedures are generally unsatisfactory. In cases when a priori computation of SOR or ADI parameters is not possible, methods have been devised for the direct computation of the necessary extreme eigenvalues that, although a nuisance, are much less expensive than direct numerical approximation of the

eigenvalues of  $(A+B)^{-1}A$  for factorization procedures, which could cost as much as running the iteration. The amount of computational work required to obtain accurate bounds of the eigenvalues of  $(A+B)^{-1}$  is so great that factorization methods are not efficient unless the labor of executing the iteration yields these bounds as a by-product. Efficiency requires that the work of executing the iteration be combined with the work of estimating the eigenvalues. The adaptive-Chebyshev-factorization algorithm of Diamond (1971) achieves this. It estimates eigenvalue bounds and optimal parameters dynamically by using the fact that without optimal Chebyshev parameters, the iterates may be combined to approximate an extreme eigenvector.

The idea of estimating parameters dynamically is used in some SOR programs (Forsyth and Wasow, 1960) to obtain an increasing sequence of estimates of the SOR parameter  $w$ . Therefore, the idea is not new although the details are completely different. What does seem to be new is its potential for practical applications. Experimental comparisons of Diamond's algorithm to other iterative methods have been made (Diamond, 1971). For each test, Diamond's algorithm achieved roughly two orders of magnitude greater accuracy for the same computational work.

One of the most appealing features of Diamond's algorithm is that no preliminary eigenvalue estimates are necessary. Only trivial input parameters are required, and the user is freed of the task of careful numerical preparations, essential to the effective use of SOR and ADI. This feature should be stressed as well as the promise of great practical utility.



### C. Adaptive Algorithms

Diamond's algorithm is an example of an adaptive or dynamic algorithm. The characteristic property of adaptive numerical algorithms is that the sequence of executed steps is determined dynamically while the algorithm is running. The logical control feature of the machine is exploited in a more fundamental way than by traditional numerical algorithms, which may be called static algorithms. In adaptive algorithms, logical control statements (If ... statements in the FORTRAN version of the algorithm) direct the choice of one chunk of computation over another, depending on the outcome of certain trials, whereas in static algorithms, the use of these statements is superficial, for example to halt an iteration or to flag errors. In a traditional algorithm, if the path of computation is to be altered by some test, the test can usually be formalized as a mathematical function, that is, the path the computations take may be predicted by a formula. For this reason, static algorithms may be analyzed more easily than adaptive algorithms to determine, for example, the rate of convergence. In adaptive algorithms the quantities necessary for analytic study are not available until certain trials are performed and the outcome of these trials is too complex to formalize mathematically.

Recently developed algorithms to solve initial value problems in ordinary differential equations are an example of adaptive algorithms in common use. These algorithms select at each step the next stepsize to be used, the order of the method and the method, depending on estimates of certain parameters based on numerical results from the preceding step.

D. Approximate Factorizations.

The ACF algorithm of Diamond requires only that  $A+B$  be positive definite, but for execution to be efficient the  $LL^*$  factors of the Cholesky decomposition must be sparse and convergence must be rapid. It is trivial to choose  $B$  to make the  $LL^*$  factors sparse. For, let  $L$  be any sparse lower triangular matrix and let  $B = LL^* - A$ . Also,  $A+B = LL^*$  is positive definite. The difficult requirement is to achieve rapid convergence.

To describe the algebraic criteria for rapid convergence, let  $e_n = u - u_n$  be the error in approximating the solution of  $Au = s$  by (35). The error satisfies, as can be seen by writing out the expression for  $e_n$ , then for  $e_{n-1}$ , etc.,

$$e_n = P_n [(A + B)^{-1}A] e_0$$

$$P_n(x) = \prod_{j=0}^{n-1} (1 - \tau_j x).$$

Let  $\lambda_1 \geq \dots \geq \lambda_n$  be the eigenvalues of  $(A + B)^{-1}A$ . If  $\lambda_1$  and  $\lambda_n$  happen to be known, Chebyshev theory yields an optimum set of  $\tau$ -parameters for which  $e_n$  is smallest as measured by the  $A$ -norm, defined to be

$$\|e_n\|_A = \|Ae_n\|_2$$

If  $\tau_0, \dots, \tau_{n-1}$  are optimum, then (Dupont, et al., 1968)

$$\|e_n\|_A \leq \|e_0\|_A / T_n\left(\frac{p+1}{p-1}\right),$$

where  $T_n(x) = x^n + \dots$  is the Chebyshev polynomial of degree  $n$  and

$$p = \lambda_1 / \lambda_n \geq 1$$

is the  $p$ -condition number of  $(A + B)^{-1}A$ . Let

$$y_0 = \frac{p+1}{p-1} .$$

For  $|y_0| > 1$ , we have (Forsyth and Wasow, 1960)

$$T_n(y_0)^{-1} = 2 / [(y_0 + (y_0^2 + 1)^{1/2}) + (y_0 - (y_0^2 - 1)^{1/2})]^n .$$

This quantity grows small as  $p$  grows close to 1. Therefore, to improve the rate of convergence of a Chebyshev factorization procedure means simply to select  $B$  so as to improve the  $p$  - condition of  $(A + B)^{-1}A$ .

### E. The Finite-Difference Case.

In this section we shall describe various factorization methods for the solution of the linear systems that arise from finite-difference approximations. It is convenient to call these systems finite-difference systems. Systems that arise from finite element approximations will be called finite-element systems. Factorization methods for the solution of finite element or finite difference systems will be called finite element or finite difference factorization methods respectively. Other terms that will be used are finite difference or finite element approximate factorization and finite difference or finite element factorization procedure.

Our objective is to devise an efficient finite element factorization method by drawing on the techniques due to Buleev (1960) and Stone (1968) that have yielded efficient finite difference factorization methods. In particular, we shall use an idea of Stone originally proposed by Buleev to obtain a finite approximate factorization to construct  $A + B$  (where  $A$  is a finite element system), for which the  $p$ -condition number of  $(A + B)^{-1}A$  is small. Only the idea is presented here. No study has yet been made of the mathematical properties of Stone's finite difference approximate factorization or of its extension to the finite element approximate factorization that we present. The idea and the reasons why it works are intuitive. No estimates of the  $p$ -condition number will be given, only experimental evidence that is small.

We shall call  $B$  the auxiliary matrix. Stone's idea is that the components of  $Bv$  should be small relative to the components of the finite

difference system  $Av$  when the components of  $v$  are the values of polynomial  $V(x,y)$  taken at the gridpoints of the system. In fact, Stone's auxiliary matrix has the property that  $Bv = 0$  when  $V$  is a first degree polynomial in  $x$  and  $y$ . When  $Bv$  is small, it is plausible that one step of

$$(A + B) u^{(n+1)} = (A + B) u^{(n)} - (Au^{(n)} - s)$$

is nearly the same as solving

$$Au^{(n+1)} = s.$$

Stone's auxiliary matrix is nonsymmetric, and there is no simple way to apply Chebyshev theory. We shall construct a finite element approximate factorization for which the auxiliary matrix is symmetric and  $A + B$  is positive definite. If  $B$  is also small relative to  $A$ , and  $A + B$  is approximately  $A$ , it is plausible that  $(A + B)^{-1}A$  is approximately the identity matrix with  $p$ -condition number close to unity. If so, Chebyshev iteration converges rapidly.

The technique of Buleev and Stone to make  $Bv$  small relative to  $Av$  is to construct  $B$  such that the components of  $v$  cancel (next section) to form the components of  $Bv$ . Their technique yields an auxiliary matrix for which the components of  $h^2 Bv$  approximate a second derivative of  $V(x,y)$ , where  $h$  is the grid spacing. Buleev's construction approximates a second directional derivative whereas Stone's approximates the mixed  $x$ - $y$  derivative,

$$\partial^2 V(x,y) / \partial x \partial y.$$

Other finite difference factorization algorithms have been proposed. Dupont, et al. (1968) studied an algorithm for which the components of  $h^2 Bv$  approximate the first order directional derivative in the direction  $-\pi/4$ . None of these auxiliary matrices directly satisfies the condition

that the components are of the same order of magnitude. However, there is a property of the mixed derivative that makes  $\langle Bu, u \rangle$  smaller than  $\langle Au, u \rangle$  when  $B$  is Stone's auxiliary matrix. This is explained below.

Suppose the region of definition of the boundary value problem is the square

$$R = \{(x,y) = -\pi \leq x \leq \pi, y \leq \pi\}.$$

Let  $f$  and  $g$  be any twice continuously differentiable functions. Suppose  $f(x,y)$  is zero on the boundary of  $R$  and the normal derivative of  $g$  is zero on the boundary of  $R$ . The Fourier series for  $f$  and  $g$  are

$$f(x,y) = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} a_{jk} \sin jx \sin ky$$

and

$$g(x,y) = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} b_{jk} \cos jx \cos ky$$

The Fourier series for the mixed derivatives are

$$\partial^2 f(x,y) / \partial x \partial y = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} jk a_{jk} \cos j\pi x \cos k\pi y$$

and

$$\partial^2 g(x,y) / \partial x \partial y = \sum_{k=0}^{\infty} \sum_{j=1}^{\infty} jk b_{jk} \sin j\pi x \sin k\pi y$$

It follows that

$$\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(x,y) \partial^2 f(x,y) / \partial x \partial y \, dx dy = 0$$

and

$$\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} g(x,y) \partial^2 g(x,y) / \partial x \partial y \, dx dy = 0.$$

If the system of difference equations

$$Au = s$$

results from the finite difference approximation of either Dirichlet's problem or Neumann's problem defined on  $R$ , then the components of  $u$  may be assumed to be the values at the gridpoints of a twice continuously differentiable function that vanishes on the boundary of  $R$  or whose normal derivative vanishes on the boundary respectively. For  $A + B$  constructed by Stone's algorithm, we have

$$|\langle Bv, v \rangle h^2| \doteq \left| \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} v(x,y) \partial^2 v(x,y) / \partial x \partial y \, dx dy \right| = 0.$$

The components of  $Bv$  are the values of  $v$ , and  $h$  is the grid size. The precise way in which the left side approximates the right requires further study. If it is accepted intuitively, it is plausible that  $|\langle Bv, v \rangle|$  is small in comparison to  $\langle Av, v \rangle$ .

Stone's factorization method, that is, his factorization algorithm together with the iterative procedure he recommends, is well known by the name Stone has given it, Strongly Implicit Procedure or SIP. The idea to make the components of  $Bu$  small and the approximation of  $B$  to the  $x - y$  derivative does not alone give an effective method for solving finite difference systems. The impressive success of the SIP is due also to Stone's iteration. Stone's iteration, is:

$$(A + B_{\alpha_n}) u^{(n+1)} = (A + B_{\alpha_n}) u^{(n)} - t_n (Au^{(n)} - s),$$

where  $\alpha_n$  is a parameter determined by a formula of Stone's. This parameter is thought to enhance cancellation among the components of  $v$  thereby diminishing the magnitude of the components of  $Bv$ . No more precise reason is known why Stone's parameters are effective. Analysis is difficult because the complicated dependence of  $B$  on  $\alpha$  obscures the relation between  $A + B_{\alpha_n}$  and  $A + B_{\alpha_{n+1}}$ .

To increase the rate of convergence, Stone recommends reversing the order of the gridpoints at alternate steps of the iteration. Again, only intuitive reasons are known why this technique works.

Stone's techniques are instinctive, ingenious, and strikingly effective in solving difficult finite-difference systems. Tests show that for the same amount of computational work, the error achieved by SIP is smaller by the factor  $10^{-4}$  than the error achieved by alternating-direction-implicit iterations (ADI) in solving finite difference systems for which the elements are random quantities corresponding to a differential equation with discontinuous coefficients.

SIP is attractive to users not only because it converges rapidly but also because it is convenient. Convergence parameters may be computed from a formula which Stone gives that requires only simple input parameters available from the problem. This formula does not require difficult eigenvalue estimates such as are necessary for ADI or SOR.

Convenience to the user is a feature common to both SIP and the ACF algorithm of Diamond. However, there are two distinctions



between these methods. First, ACF rests on mathematical theory, whereas the techniques of Stone in SIP appear to be oracular. To explain the second distinction, the Stone approximate factorization may be modified to yield a finite difference approximate factorization for which  $A + B$  is positive definite. This positive definite approximate factorization may be used with ACF. The result was referred to in section above. In comparison tests, ACF achieves roughly two orders of magnitude greater accuracy than SIP in solving the same finite difference system.

F. Approximate Factorization of Finite Element Matrices.

In this section we shall attempt to determine an auxiliary matrix  $B$  for a regular finite element matrix  $A$  in such a way that  $B$  is a discretization of the operator  $\partial^2/\partial x \partial y$ . Elements of the matrices will correspond to grid points of the mesh, so we begin with a precise description of the mesh system.

Let  $h = 1/(1+n)$ , and let  $D_h = \{(jh, kh) : 1 \leq j, k \leq n\}$  be the grid system over the unit square  $D$ , where

$$D = [0,1] \times [0,1].$$

Denote grid point  $(jh, kh)$  by  $(j, k)$ . Order the grid points in a left-to-right, down-to-up fashion:  $(j_1, k_1)$  precedes  $(j_2, k_2)$  if  $k_1 < k_2$  or if  $k_1 = k_2$  and  $j_1 < j_2$ . Let  $u$  be an  $n^2$ -vector whose components  $u_{jk}$ ,  $(j, k)$  a gridpoint, are ordered the same way as the gridpoints. Matrix  $A$  is defined by

$$\begin{aligned} (Au)_{jk} = & S_{j-1, k-1} u_{j-1, k-1} + R_{j, k-1} u_{j, k-1} + \\ & + Q_{j+1, k-1} u_{j+1, k-1} + P_{j-1, k} u_{j-1, k} + O_{jk} u_{jk} + \\ & + P_{jk} u_{j+1, k} + Q_{jk} u_{j-1, k+1} + R_{jk} u_{j, k+1} + S_{jk} u_{j+1, k+1}. \end{aligned}$$

In matrix form, A looks like

$$A = \begin{bmatrix} 0 & P & \cdot & \cdot & \cdot & Q & R & S \\ P & 0 & P & & & & & \\ \cdot & & & & & & & \\ \cdot & & & & & & & \\ \cdot & & & & & & & \\ Q & & & & & & & \\ R & & & & & & & \\ S & & & & & & & \end{bmatrix}$$

where all elements not show are zero. Diagonals Q starts at column n-1 and row 1, and column 1 and row n-1, R at column n and row 1, and column 1 and row n S at column n+1 and row 1, and column 1 and row n+1, and P at column 2 and row 1, and column 1 row 2. Certain elements in the P, Q, R and S diagonals are zero. More precisely,

$$R_{nk} = P_{nk} + S_{nk} = Q_{nk} = 0$$

for  $1 \leq k \leq n-1$ . Matrix A arises from the Galerkin method applied to the differential equation

$$Lu = (a_1(x,y) u_x(x,y))_x + (a_2(x,y) u_y(x,y))_y = f$$

defined in D. It also arises from an application of the nine point difference operator applied to

$$Lu = f.$$

The objective will be to determine B such that A + B has sparse LU factors and such that

$$(BU)_{jk} = \frac{\partial^2 u}{\partial x \partial y} (x_j, y_k) h^2.$$

Recall that Gaussian elimination is an algorithm for solving a system of linear equations by first reducing the system to upper triangular form, called the forward reduction, and then solving the resulting upper triangular system for the unknowns, called the backward substitution. If  $A$  is the matrix of the system, Gaussian elimination is equivalent to factoring  $A$  as

$$A = LU$$

where  $L$  is lower triangular and  $U$  is upper triangular, then solving

$$Au = s$$

by solving

$$Lw = s$$

and

$$Ux = w.$$

The forward reduction is equivalent to computing the LU factors and  $w$ .

The backward substitution is equivalent to solving

$$Ux = w.$$

We begin the determination of  $B$  by knowing the general form of the nonzero elements of  $A + B$  will take.

$$A+B = \begin{bmatrix} 0_{11} & P_{11} & 0 \dots 0 & & R_{11} & & S_{11} \\ P_{11} & 0_{21} & P_{21} & 0 \dots 0 & Q_{21}+G_{21} & R_{21}-G_{21} & S_{21} \\ 0 & P_{21} & 0_{31} & P_{31} & & Q_{31}+G_{31} & R_{31}-G_{31} & S_{31} \\ \vdots & & & & & & & \\ 0 & & & & & & & \\ R_{11} & Q_{21}+G_{21} & 0 \dots & & & & & \\ S_{11} & & & & & & & \end{bmatrix}$$

To display elements of rows  $n+i$ ,  $0 \leq i$ , requires a figure too large to fit on the page. Row  $(k-1)n+j$ , corresponding to gridpoint  $(j,k)$  is:

$$\begin{aligned} & \dots S_{j-1,k-1} \quad R_{jk-1}^{-G_{j,k-1}} \quad Q_{j+1,k-1}^{+G_{jH,k-1}} \quad \dots \\ & \dots P_{j-1,k}^{-G_{j,k-1}} \quad 0_{jk}^{+2G_{j,k-1}} \quad P_{jk}^{-G_{jH,k-1}} \\ & \dots Q_{jk}^{+G_{jk}} \quad R_{jk}^{-G_{jk}} \quad S_{jk} \quad \dots \end{aligned}$$

The formulas for computing B are determined by the requirement that Gaussian elimination be efficient, or equivalently that the LU factors be sparse. The first step of Gaussian elimination is to eliminate elements  $P_{11}$ ,  $R_{11}$  and  $S_{11}$  from column 1. This yields

$$(A+B)^{(1)} = \begin{array}{c} \text{col} \\ \text{n+1} \\ \left[ \begin{array}{cccccccc} 0_{11} & P_{11} & 0 & \dots & 0 & 0 & R_{11} & S_{11} \\ 0 & 0_{21} - \frac{P_{11}^2}{0_{11}} & P_{21} & 0 & \dots & 0 & Q_{21} + G_{21} - \frac{P_{11}R_{11}}{0_{11}} & R_{21} - G_{21} - \frac{P_{11}S_{11}}{0_{11}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & P_{21} & 0_{21} & P_{31} & 0 & \dots & \dots & \dots \\ \text{row n+1} & 0 & 0_{21} + G_{21} - \frac{P_{11}R_{11}}{0_{11}} & 0 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ R_{21} - G_{21} - \frac{P_{11}S_{11}}{0_{11}} & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{array} \right] \end{array}$$

At the next step, an appropriate multiple of the second row is added to rows 3, n+1 and n+2 to eliminate elements  $P_{21}$ ,

$$g = Q_{21} + G_{21} - P_{11}R_{11}/0_{11},$$

and

$$R_{21} - G_{21} - P_{11}S_{11}/0_{11}.$$

Note that  $g$  also is the second element in column n+1. The third element of column n+1 is zero before the second row is added to the third, but when a multiple of the second is added to the third, the same multiple of  $g$  is added to zero in position n+1 of the third row, thereby replacing a zero element by a nonzero element. At the next step when a multiple of the third row is added to the fourth, a fourth nonzero element is created in column n+1. Ultimately column n+1 is filled with nonzero elements from row 1 to the diagonal at row n+1. The same phenomenon fills the remaining columns with a total of n+1 nonzero elements each. This causes the

resulting reduced upper triangular system to contain  $n+1$  nonzero elements per row thereby making the backward substitution very inefficient. Another computational difficulty is that elimination of  $g$  from row  $n+1$  yields a nonzero quantity as the third element of row  $n+1$ . Each succeeding step yields a nonzero quantity in positions four, ... of row  $n+1$ . Moreover, the same phenomenon affects rows  $n+2$ ,  $n+3$ , ... . Ultimately this means that each step of Gaussian elimination requires the elimination of  $n+1$  elements, thereby making the forward reduction very inefficient.

Now suppose

$$G_{21} = P_{11} R_{11} / O_{11} - Q_{21}$$

Then  $g = 0$  and the addition of a multiple of the second row to the third, and the third to the fourth, etc. would not create a column of nonzero elements. Moreover, it would not be necessary to eliminate element  $g$  from row  $n+1$ . The difficulties with the remaining steps of Gaussian elimination that cause the forward reduction and the backward substitution to be inefficient could also be avoided by appropriate values of  $G_{31}$ , ... . The value of  $G_{31}$ , for example, is determined by

$$-S_{11} P_{11} / O_{11} - P_{21} (R_{21} - G_{21}) / O_{21} + Q_{31} + G_{31} = 0.$$

A general recursive formula may be written to define  $G_{jk}$  which is obtained in the same way as the formulas for  $G_{21}$  and  $G_{31}$ . This formula is omitted here because it has no practical value. Instead of computing  $G_{21}$ ,  $G_{31}$ , ... directly, thereby executing the forward reduction part of Gaussian elimination, we shall compute the LU factors of  $A+B$ .







Row  $n+2$  follows.

	col 1	2	3	
row $n+2$	$a_{22}$	$a_{22}e_{11}+b_{22}$	$b_{22}e_{21} \dots$	
	col $n+1$		col $n+2$	col $n+3$
row $n+2$	$a_{22}f_{11}+c_{22}$	$a_{22}g_{11}+b_{22}f_{21}+c_{22}e_{12}+d_{22}$		$b_{22}g_{21}+d_{22}e_{22}$
	col $2n+1$	col $2n+2$	col $2n+3$	
row $n+2$	$c_{22}f_{12}$	$c_{22}g_{12}+d_{22}f_{22}$	$d_{22}g_{22}$	

In general, row (j,k) is:

$$\begin{array}{ccccccc}
 (j-1,k-1) & (j,k-1) & (j+1,k-1) & \text{col}(j-1,k) & & \text{col}(j,k) & \\
 a_{jk} & a_{jk}^e a_{j-1,k-1} + b_{jk} & b_{jk}^e a_{j,k-1} \cdots & a_{jk}^f a_{j-1,k} + c_{jk} & & a_{jk}^g a_{j-1,k} + b_{jk}^f a_{j,k-1} + c_{jk}^e a_{j-1,k} + d_{jk} & \\
 & & & & & & \\
 & & \text{col}(j+1,k) & (j-1,k+1) & \text{col}(j,k+1) & (j+1,k+1) & \\
 & & b_{jk}^g a_{j,k-1} + d_{jk}^e a_{jk} \cdots & c_{jk}^f a_{j-1,k} & c_{jk}^g a_{j-1,k} + d_{jk}^f a_{jk} & d_{jk}^g a_{jk} & 
 \end{array}$$

Let  $N = n^2$ . Because row  $N-n$  of  $U$  possesses fewer nonzero elements than the preceding rows, the formation of the product of  $L$  and  $U$  is slightly altered. The product is formed as shown below.

$$\begin{array}{l}
 \text{row } N-n-1 \\
 N-n \\
 N-n+1
 \end{array}
 \left( \begin{array}{cccc}
 a_{n-1,n-1} & b_{n-1,n-1} & \cdots & c_{n-1,n-1} \\
 & a_{n,n-1} & & b_{n,n-1} \cdots c_{n,n-1} & d_{n,n-1} \\
 & & a_{1n} & b_{1n} & & c_{1n} & d_{1n}
 \end{array} \right)
 \begin{array}{l}
 \text{row } (n-1,n-1) \\
 \text{row } (n,n-1) \\
 \text{row } (1,n)
 \end{array}$$

L

$$\begin{array}{l}
 \text{row } N-n-1
 \end{array}
 \left( \begin{array}{cccc}
 1 & e_{n-1,n-1} & \cdots & f_{n-1,n-1} & g_{n-1,n-1} \\
 & 1 & & e_{n,n-1} & \cdots & f_{n,n-1} \\
 & & & \text{col}(N-n) & & \text{col } N
 \end{array} \right)$$

U

Rows N-n-1 and N-n of the product are shown as follows:

$$\begin{array}{r}
 \text{row } N-n-1 \quad \dots a_{n-1,n-1} \quad a_{n-1,n-1} \quad e_{n-2,n-2} + b_{n-1,n-1} \quad \dots \quad \underbrace{a_{n-1,n-1} f_{n-2,n-2} + c_{n,n-1} \dots}_{\text{col } N-n-2} \\
 \text{row } N-n-1 \quad \underbrace{a_{n,n-1} g_{n-2,n-2} + b_{n,n-1} f_{n-1,n-2} + c_{n,n-1} f_{n-2,n-1} + d_{n-1,n-1} \dots}_{\text{col } N-n-1} \\
 \text{row } N-n-1 \quad \underbrace{\dots c_{n-1,n-1} f_{n-2,n-1}}_{\text{col } N-2} \quad \underbrace{c_{n-1,n-1} g_{n-2,n-1} + d_{n,n-1} f_{n-1,n-1}}_{\text{col } N-1} \quad \underbrace{d_{n-1,n-1} g_{n,n-1}}_{\text{col } N}
 \end{array}$$

Only columns N-n-1 and N-n of row N-n are shown:

$$\underbrace{c_{n,n-1} f_{n-1,n-1} \quad c_{n,n-1} g_{n-1,n-1} + d_{n,n-1} f_{n,n-1}}_{\text{col } N}$$

Note that the element in column N of row N-n+1 is  $c_{1n} f_{n,n-1} = 0$  since  $c_{1n} = 0$ . (The fact that  $c_{1n} = 0$  follows from the construction of B in §6 where the elements of B may be verified to satisfy  $G_{1k} = 0$ . Since  $c_{1k} = Q_{1k} + G_{1k}$  and  $Q_{1k} = 0$ , we have  $c_{1k} = 0$ .)

The algorithm for computing the LU decomposition of  $A+B$  directly from the elements of  $A$  can now be given.

$$\begin{array}{l}
 \text{row 1} \\
 d_{11} = O_{11} \\
 d_{11}e_{11} = P_{11} \\
 d_{11}f_{11} = R_{11} \\
 d_{11}g_{11} = S_{11} \\
 \\
 \text{row 2} \\
 c_{21} = P_{11} \\
 c_{21}e_{11} + d_{21} = O_{21} \\
 d_{21}e_{21} = P_{21}
 \end{array}$$

The next element in row 2 of LU is  $c_{21}f_{11}$ . This quantity is determined by the equations above. Since the corresponding element of  $A+B$  is  $Q_{21} + G_{21}$ , we have

$$c_{21}f_{11} = Q_{21} + G_{21}. \quad (36)$$

The next element in row 2 of LU is  $c_{21}g_{11} + d_{21}f_{21}$ . Element  $f_{21}$  is undetermined. Since

$$c_{21}g_{11} + d_{21}f_{21} = R_{21} - G_{21}$$

we have from (36)

$$c_{21}g_{11} + d_{21}f_{21} = R_{21} + Q_{21} - c_{21}f_{11},$$

which may be solved for  $f_{21}$ . Finally,

$$d_{21}g_{21} = s_{21}.$$

The equations above may be combined in one set:

$$c_{21} = P_{11}$$

$$c_{21}e_{11} + d_{21} = O_{21}$$

$$d_{21}e_{21} = P_{21}$$

$$c_{21}g_{11} + d_{21}f_{21} = R_{21} + Q_{21} - c_{21}f_{11}$$

$$d_{21}g_{21} = s_{21}.$$

The remaining formulas up through row  $n$  are no different. We next consider row  $n+1$ . We have

$$c_{12} = R_{11}.$$

A question now arises because the next element of row  $n+1$  of LU is determined as is also the corresponding element of  $A+B$ . Therefore, we should have

$$b_{12}e_{11} = Q_{21} + G_{21}.$$

But  $G_{21}$  has been computed to be  $G_{21} = c_{21}f_{11} - Q_{21}$ . Therefore,

$$Q_{21} + G_{21} = c_{21}f_{11},$$

that is,

$$b_{12}e_{11} = Q_{21} + G_{21} \tag{37}$$

is valid if LU is symmetric.

Lemma.

$$b_{12} e_{11} = Q_{21} + G_{21} \cdot$$

Proof: The formulas given above, i.e.,

$$d_{11} = O_{11}$$

$$d_{11} e_{11} = P_{11}$$

$$d_{11} f_{11} = R_{11}$$

$$d_{11} g_{11} = S_{11}$$

and for  $j \leq n$ ,

$$c_{j1} = P_{j-1,1}$$

$$c_{j1} e_{j-1,1} + d_{j1} = O_{j1}$$

$$d_{j1} e_{j1} = P_{j1}$$

$$c_{j1} g_{j-1,1} + d_{j1} f_{j1} = R_{j1} + Q_{j1} - c_{j1} f_{j-1,1}$$

$$d_{j1} g_{j1} = S_{j1}$$

and

$$c_{12} = R_{11}$$

determine the first  $n$  rows of  $L$  and  $U$  together with  $l$  element of  $L$ . These have the property that when  $LU$  is formed the elements of the first  $n$  rows of  $LU$  are equal to the corresponding elements of  $A+B$ . Now, let  $L'U'$  be the "LU" - decomposition of  $A+B$ , with  $A+B$  constructed directly according to the method described in section C. Certainly, the elements in the first  $n$  rows of  $L'U'$  together with the first element of row  $n+1$  are equal to the



corresponding elements of  $A+B$ . Therefore  $LU = L'U'$ , i.e.,  $LU$  and  $L'U'$  are  $LU$ -decompositions of the same matrix. Since the  $LU$ -decomposition of any matrix is unique, and since  $L'U'$  is symmetric (because  $L'U' = A+B$ ) it follows that  $LU$  is symmetric and therefore (37) is valid.

In general the formulas obtained by equating rows  $(j,k)$  of  $LU$  and  $A+B$  are

$$a_{jk} = S_{j-1,k-1}$$

$$\begin{aligned} a_{jk}e_{j-1,k-1} + b_{jk} &= R_{j,k-1}^{-G_{j,k-1}} \\ &= R_{j,k-1}^{-(Q_{j,k-1} - c_{j,k-1}f_{j-1,k-1})} \end{aligned}$$

$$\begin{aligned} a_{jk}f_{j-1,k} + c_{jk} &= P_{j-1,k}^{-G_{j,k-1}} \\ &= P_{j-1,k}^{-(Q_{j,k-1} - c_{j,k-1}f_{j-1,k-1})} \end{aligned}$$

$$\begin{aligned} a_{jk}g_{j-1,k} + b_{jk}f_{jk-1} + c_{jk}e_{j-1,k} + d_{jk} &= O_{jk} + 2G_{j,k-1} \\ &= O_{jk} + 2(Q_{j,k-1} - c_{j,k-1}f_{j-1,k-1}). \end{aligned}$$

$$\begin{aligned} b_{jk}g_{jk-1} + d_{jk}e_{jk} &= P_{jk}^{-G_{j+1,k-1}} \\ &= P_{jk}^{-(Q_{j+1,k-1} - c_{j+1,k-1}f_{j,k-1})} \end{aligned}$$

$$\begin{aligned} c_{jk}g_{j-1,k} + d_{jk}f_{jk} &= R_{jk}^{-G_{jk}} \\ &= R_{jk} + Q_{jk} - C_{jk}f_{j-1,k} \end{aligned}$$

$$d_{jk}g_{jk} = S_{jk}.$$

## H. Experimental Results.

We have described an approximate factorization method in the preceding sections for a regular finite element system, and it may be used in the Diamond ACF algorithm to solve these systems. Experimental results are satisfactory, but because regular systems do not occur in modeling fluid flow in porous media, we have omitted these results. Instead, we shall summarize experimental results for an approximate factorization devised for an irregular finite element system. It is more complicated than the one presented in the preceding sections, but the ideas embodied in it are the same. The details are included in the thesis of Y.J. Kim of the Department of Computer Science of the University of Illinois, U-C, to be completed August, 1973 (Kim, 1973).

The flow problem to which Kim's factorization is applied results from the injection of liquid waste into a porous aquifer. The liquid enters the aquifer from a well bore and dissipates under pressure. The square in Figure 32 represents a cross section of the aquifer, assuming axial symmetry, with the well bore at the left. The half of the square nearer the bore is triangulated with a fine mesh; the other half is triangulated with a course mesh. When an elliptic differential equation is discretized by the finite element method with the triangulation in Figure 32, the result is an irregular matrix as shown in Figure 33 with nonzero elements marked by "x". The ACF algorithm was tested with Kim's factorization applied to the solution of the systems shown in Figure 33 with the number of unknowns equal to 570, 1600, 5130, 8732, and 10660. See Figure 34.

We shall compare the numerical labor required to solve the system by the ACF Kim algorithm and by Gaussian elimination when the number of

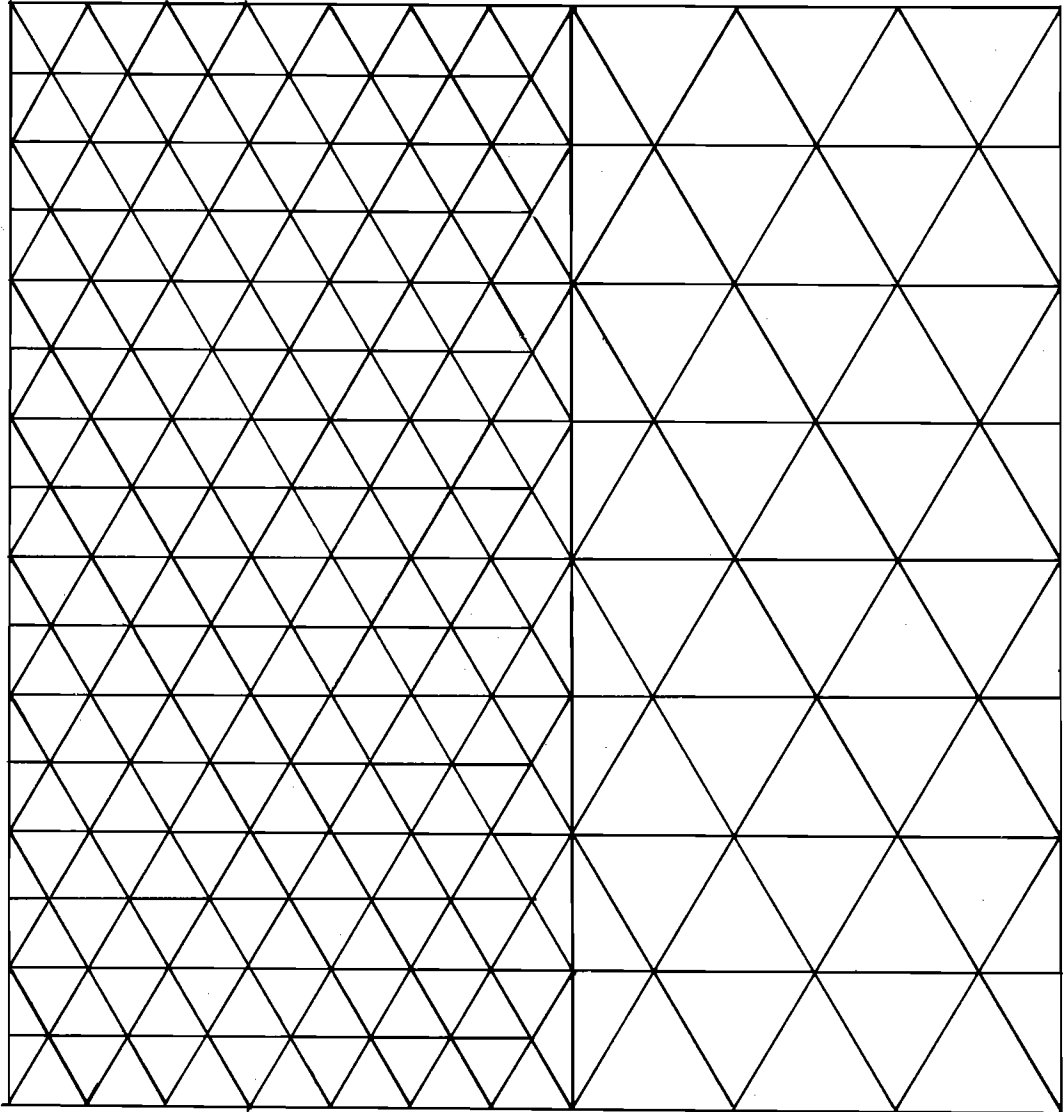


Figure 32: Triangularization of domain

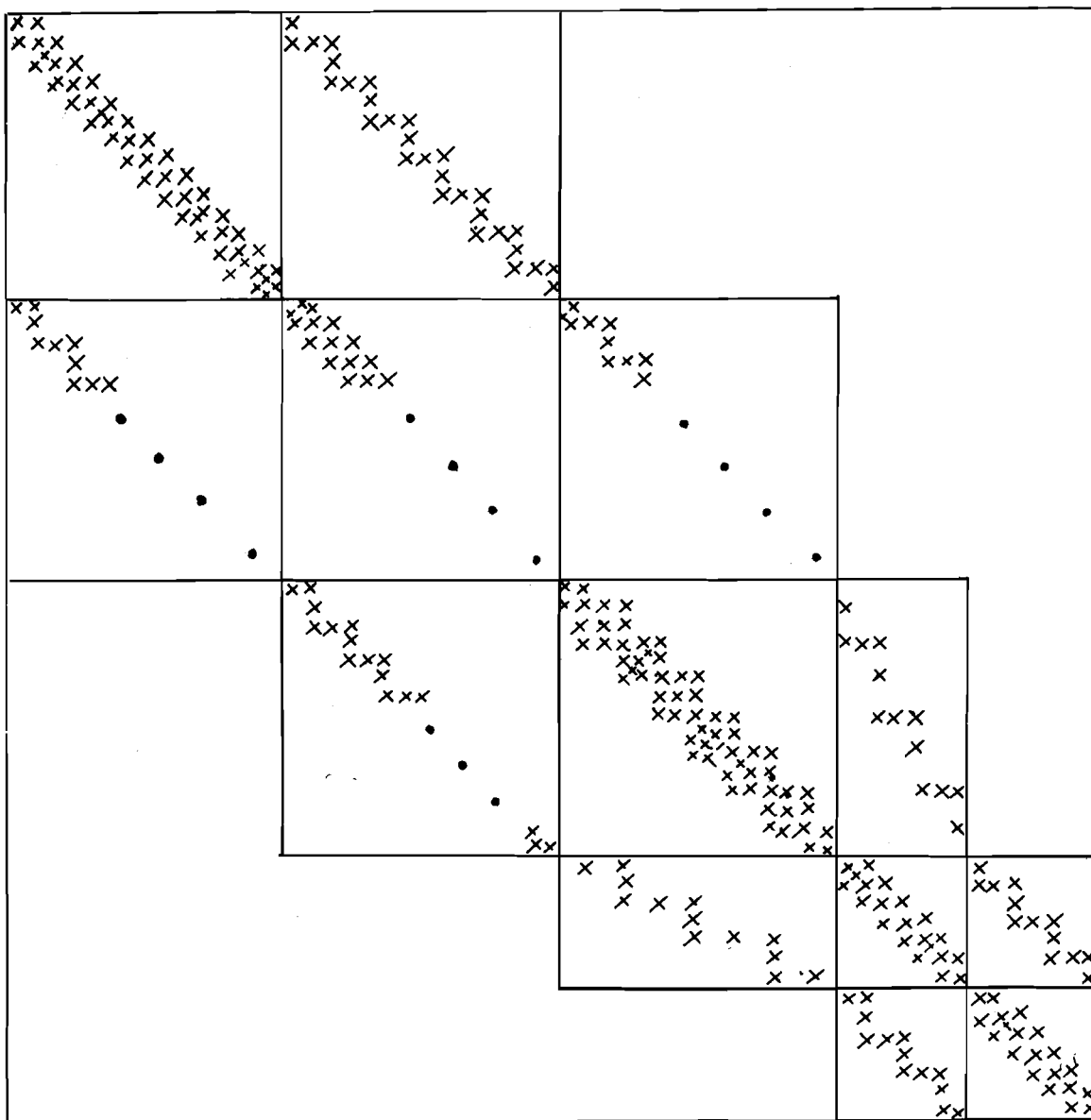


Figure 33: Non-zero elements of sparse matrix

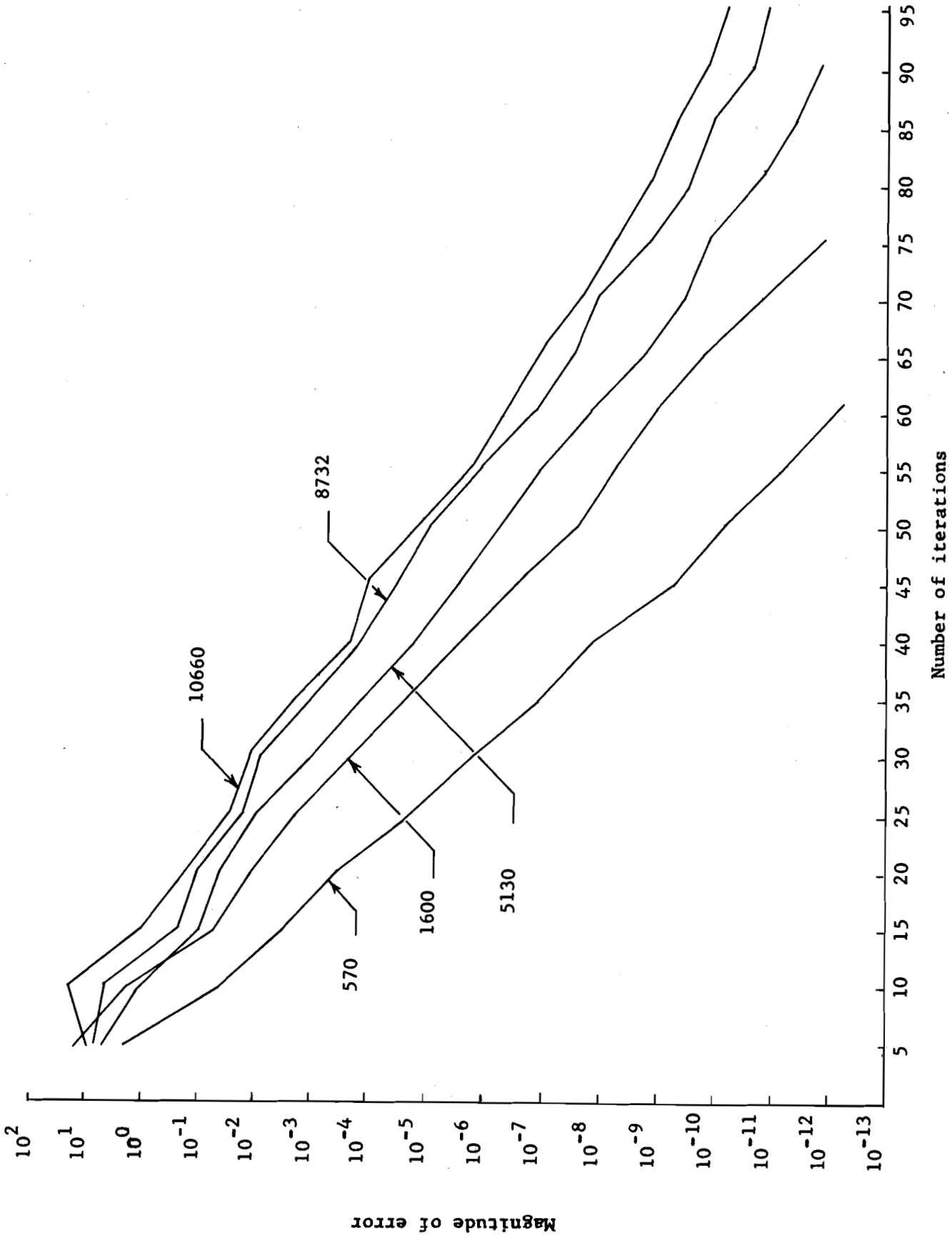


Figure 34 : Magnitude of error vs. number of iterations for 570, 1600, 5130, 8732, and 10660 unknowns.

unknowns is  $N = 10660$ . We shall use  $N \approx 10^4$ .

We shall only count multiplications and count a division as a multiplication. The number of additions is ignored because it is proportional to the number of multiplications. The number of multiplications to execute each step of the iteration is approximately  $19N$ . If  $k$  is the number of iterations, the total number of multiplications is roughly  $15N + k19N$ . From Figure 34 the number of iterations required to attain accuracy equal to  $10^{-6}$  is 60. Although Gaussian elimination in theory is exact, accuracy in practice is determined by the accuracy of the machine. We shall take it to be  $10^{-6}$  for Gaussian elimination also. The total number of multiplications is therefore

$$15N + k19N = (15+60 \cdot 19)N \approx 1150N$$

The number of multiplications required by an efficient coding of Gaussian elimination is approximately

$$N^2 + 4NN^{1/2} + N$$

For  $N = 10^4$ , the number of multiplications for the ACF-Kim method is approximately

$$10^3 \cdot 10^4 = 10^7.$$

For  $N = 10^4$ , the number of multiplications for Gaussian elimination is

$$10^8 + 4 \cdot 10^6 + 10^4.$$

The ACF-Kim method for this particular test is roughly ten times as efficient as Gaussian elimination.

Another advantage to the use of Gaussian elimination is a reduction in memory requirements. For the iterative method the number of memory spaces is proportional to  $N$  whereas for Gaussian elimination, memory is proportional to  $N^2$ . For the University of Illinois computer system, Gaussian elimination code for flow problems is limited to 2000 unknowns, whereas tests up to 10660 unknowns, as shown in Figure 34, have been made with ACF-Kim code.

The work is not complete. Further experimental tests of the method must be made to discover its limitations. At the present time, the algorithm is too complex for a casual user to adapt it to a different problem. Finally, additional programming is necessary in order to make certain minor modifications in the code as it is now written, so it will be compatible with the subroutines developed for section II-4 of this report.

## IV. REFERENCES

- Allen, D. N. de G., 1955, Relaxation Methods, McGraw-Hill, p. 199.
- Bell, A. H., Elwood Atherton, T. C. Buschbach, and D. H. Swann, 1964, Deep oil possibilities in the Illinois Basin, Ill. State Geol. Survey Circ. 368, 38 p.
- Bergstrom, R. E., 1968, Feasibility of subsurface disposal of industrial wastes in Illinois, Ill. State Geol. Survey Circ. 426, 17 p.
- Bond, D. C., 1972, Hydrodynamics in deep aquifers of the Illinois Basin, Ill. State Geol. Survey Circ. 470, Urbana, IL, 72 p.
- Buleev, N. I., 1960, A numerical method for the solution of two-dimensional and three-dimensional equations of diffusion, Mat. Sb., 51, no. 2, pp. 227-238.
- Buzbee, B. L., G. H. Golub, and C. W. Nielson, 1970, On direct methods for solving Poisson's equation, Siam Journal on Numerical Analysis, 7, pp. 627-656.
- Cady, G. H., 1919, Geology and mineral resources of the Hennepin and La Salle Quadrangles, Ill. State Geol. Survey Bull. No. 37, 136 p.
- Crandall, S. H., 1956, Engineering Analysis, McGraw-Hill, p. 27.
- Crout, P. D., 1941, A short method for evaluating determinants and solving systems of linear equations with real and complex coefficients, Trans. AIEE, 60, pp. 1235-1240.
- Daley, R. C., and Jack B. Dennis, 1969, Virtual memory processes and sharing in multics, Comm. Assoc. Comput. Mach., 11, pp. 306-312.
- Diamond, M. A., 1971, An economical algorithm for the solution of elliptic difference equations independent of user-supplied parameters, Rep. 492, Dept. of Comp. Sci., University of Illinois, Urbana.



- Douglas, Jim, Jr., and Todd Dupont, 1971, Alternating direction Galerkin methods on rectangles, in Numerical Solution of Partial Differential Equations, edited by Bert Hubbard, Academic Press, New York.
- Dupont, Todd, R. P. Kendall, and H. H. Rachford, Jr., 1968, An approximate factorization procedure for solving self-adjoint elliptic equations, Siam Journal on Numerical Analysis, 5, pp. 559-573.
- Forsyth, G. E., and W. R. Wasow, 1960, Finite difference methods for partial differential equations, John Wiley, New York.
- Fuchel, Kurt, and Sidney Heller, 1969, Considerations in the design of a multiple computer system with an extended core storage, Comm. Assoc. Comput. Mach. 11, pp. 334-340.
- Garder, A. O., D. W. Peaceman, and A. L. Pozzi, Jr., 1964, Numerical calculation of multidimensional miscible displacement by the method of characteristics: Society of Petroleum Engineers Journal, vol. 4, no. 1, March 1964, pp 26-36.
- George, J. A., 1971, Computer implementation of the finite element method, STAN-CS-71-208, Computer Science Dept., School of Humanities and Sciences, Stanford University, 222 p.
- Gurtin, M. E., 1964, Variational principles for linear initial-value problems, Qtly. Appl. Math, v. 22, p. 252.
- Hantush, M. S., 1960, Modification of the theory of leaky aquifers, Jour. Geophysical Res. 65, 3713.
- Hoopes, J. A., and D. R. F. Harleman, 1965, Waste water recharge and dispersion in porous media, Dept. of C. E., School of Eng., M. I. T., Hydrodynamics Lab. Report no. 75, 166 p.
- Hoover, L. R., and R. J. Schicht, 1967, Development in deep sandstone aquifers along the Illinois River in La Salle County, Ill. State Water Survey Rept. of Investigation no. 59, 23 p.

- Hubbert, M. K., and W. W. Rubey, 1959, Mechanics of fluid-filled porous solids and its application to overthrust faulting, Bull., GSA 70.
- Hubbert, M. K., and D. G. Willis, 1957, Mechanics of hydraulic fracturing, Trans. Soc. of Pet. Engin. of AIME, v. 210, pp. 153-168.
- Jacob, C. E., 1946, Radial flow in a leaky artesian aquifer, Trans. Amer. Geophys. Union, 27, 198.
- Javandel, Iraj, and P. A. Witherspoon, 1968a, Application of the finite element method to transient flow in porous media. Society of Petroleum Eng. Jour., vol. 8, no. 3, pp. 241-252.
- Javandel, Iraj, and P. A. Witherspoon, 1968b, Analysis of transient fluid flow in multi-layered systems, Water Resources Center Contribution No. 124, University of Calif., Berkeley, 119 p.
- Javandel, Iraj, and P. A. Witherspoon, 1969, A method for analyzing transient fluid flow in multi-layered aquifers, Water Resources Research, vol. 5, no. 4, pp. 856-869.
- Katz, M. L., 1960, Fluid flow and heat transfer in stratified systems, Ph.D. thesis, University of Michigan, 347 p.
- Kim, Y. J., 1973, Adaptive-Chebyshev-factorization methods for the solution of linear systems arising from the finite element method, Ph.D. thesis, to be issued as a Department of Computer Science Report, University of Illinois at Urbana-Champaign.
- McComas, M. R., 1968, Geology related to land use planning in the Hennepin Region, Ill. State Geol. Survey Circ. 422, 24 p.
- Muller, T. D., and P. A. Witherspoon, 1965, Pressure interference effects within reservoirs and aquifers, Trans. AIME 234, 471

- Muskat, M., 1946, The flow of homogeneous fluids through porous media, J. W. Edwards, Ann Arbor, pp. 621-675.
- Nalluswami, Marappagounder, 1971, Numerical simulation of general hydrodynamic dispersion in porous medium, Ph.D. dissertation, Colorado State University, Fort Collins, Colo., 138 p.
- Neuman, S. P., and P. A. Witherspoon, 1969a, Theory of flow in a confined two aquifer system, Water Resources Research, vol. 5, no. 4, pp. 803-816.
- Neuman, S. P., and P. A. Witherspoon, 1969b, Applicability of current theories of flow in leaky aquifers, Water Resources Research, vol. 5, no. 4, pp. 817-829.
- Neuman, S. P., and P. A. Witherspoon, 1971, Appendix A of sea-water intrusion: Aquitards in the coastal ground-water basin of Oxnard Plain, Ventura County, Bull. no. 63-4 State of California, The Resources Agency, Dept. of Water Resources, pp. 159-359.
- Reddell, L. D., and D. K. Sunada, 1970, Numerical simulation of dispersion in groundwater aquifers, Hydrology Paper no. 41, Colorado State University, Fort Collins, Colo., June 1970, 79 p.
- Slichter, C. S., 1905, Field measurements of the rate of movement of underground waters: Water Supply and Irrigation Paper no. 140, U.S.G.S., U. S. Government Printing Office, Washington, D. C., 122 p.
- Spiegel, M. R., 1961, Theory and problems of statistics, Schaum Publishing Co., New York, 359 p.
- Stone, H. L., 1968, Iterative solution of implicit approximations of multi-dimensional partial differential equations, Siam Journal on Numerical Analysis 5, pp. 530-558.

- Theis, C. V., 1935, The relationship between the lowering of piezometric surface and the rate and duration of discharge using ground-water storage, Trans. AGU, vol. 16, Part II, pp. 519-524.
- U. S. Environmental Agency, Office of Water Programs, 1972, Subsurface water pollution, A selective annotated bibliography, Part I: Subsurface Waste Injection.
- Van Everdingen, R. O., and R. A. Freeze, 1971, Subsurface disposal of waste in Canada, Tech. Bull. no. 49, Inland Water Branch, Dept. of Environment, Ottawa, Canada, 63 p.
- Van Poolen, H. K., and D. B. Hoover, 1970, Waste disposal and earthquakes at the Rocky Mountain Arsenal, Derby, Colorado, Jour. Pet. Tech., pp. 983-993.
- Walton, W. C., and Sandor Csallany, 1962, Yields of deep sandstone wells in Illinois, Ill. State Water Survey Rept. of Investigation 43, 47 p.
- Warner, D. L., 1965, Deep-well injection of liquid waste: A review of existing knowledge and an evaluation of research needs, U. S. Dept. of Health, Education, and Welfare, Div. of Water Supply and Pollution Control, Robert Taft Sanitary Engineering Center, Pub. No. 999-WP-21, 55 p.
- Wilson, E. L., and Robert E. Nickell, 1966, Application of the finite element method to heat conduction analysis, Nuclear Engineering and Design 4, pp. 276-286, North-Holland Publishing Co., Amsterdam.
- Zienkiewicz, O. C., and Y. K. Cheung, 1970, The finite element method in structural and continuum mechanics, McGraw-Hill, London.