Submitted to Water Resources Research

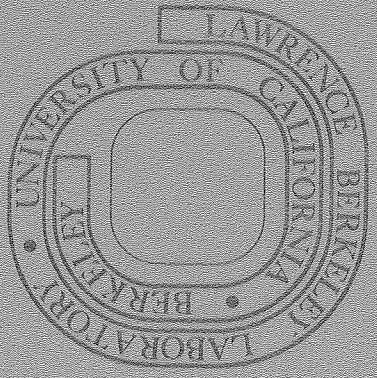
LBL-7005
Preprint

FINITE ELEMENT METHOD FOR SUBSURFACE HYDROLOGY USING A MIXED EXPLICIT-IMPLICIT SCHEME

T. N. Narasimhan, S. P. Neuman, and
P. A. Witherspoon

August 1977

Prepared for the U. S. Department of Energy
under Contract W-7405-ENG-48



FINITE ELEMENT METHOD FOR SUBSURFACE HYDROLOGY USING A MIXED EXPLICIT-IMPLICIT SCHEME

T. N. Narasimhan,[†] S. P. Neuman,[‡] P. A. Witherspoon[§]

August 1977

ABSTRACT

The mixed explicit-implicit Galerkin finite element method developed previously by the authors is shown to be ideally suited for a wide class of problems arising in subsurface hydrology. These problems include confined saturated flow, unconfined flow under free surface conditions subject to the Dupuit assumption, flow in aquifers which are partly confined and partly unconfined, axisymmetric flow to a well with storage, and flow in saturated-unsaturated soils. A single computer program, entitled FLUMP, can now handle all of these problems. The mixed explicit-implicit solution strategy employed in the program insures a high level of accuracy and computation efficiency in most cases. It eliminates many of the difficulties that groundwater hydrologists have been encountering in trying to simulate extensive aquifer systems by finite elements. Some of the outstanding features of this solution strategy include an automatic control of time step size, reclassification of nodes from explicit to implicit during execution, automatic adjustment of the implicit time weighting factor, and the treatment of boundary conditions and source terms as arbitrary functions of time of the state of the system. Five examples are presented to demonstrate the versatility and power of this new approach. A purely physical derivation of the finite element equations which does not rely on the Galerkin formalism is also included in one of the appendices.

[†]Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720.

[‡]Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona 85721.

[§]Lawrence Berkeley Laboratory and Department of Materials Sciences and Engineering, University of California, Berkeley, California 94720.

INTRODUCTION

A large class of problems in subsurface hydrology are governed by linear or quasilinear parabolic partial differential equations of the type

$$\nabla \cdot (\underline{\underline{K}} \nabla h) - q = C \frac{\partial h}{\partial t} \qquad (1)$$

where h is the dependent variable, $\underline{\underline{K}}$ is the second-rank symmetric positive-definite tensor, q is the source term, and C is a fluid mass capacity coefficient. For example, in the case of confined saturated flow, h represents the hydraulic head, $\underline{\underline{K}}$ the hydraulic conductivity (or transmissivity), and C the specific storage (or storage coefficient). In the case of flow in an unconfined aquifer, h represents the hydraulic head, $\underline{\underline{K}}$ the transmissivity (i.e., the product of hydraulic conductivity and saturated thickness), and C the specific yield. The resulting expression is known as the Boussinesq equation. If h is replaced by the volumetric water content, θ, $\underline{\underline{K}}$ by the diffusion coefficient, $D(\theta)$, and C is set equal to unity, then (1) becomes the well known diffusivity equation for unsaturated flow in a homogeneous soil. Richards' equation for saturated-unsaturated flow is obtained from (1) simply by writing $h = \psi + Z$ where ψ is pressure head and Z is elevation above an arbitrary datum; by allowing $\underline{\underline{K}}$ to vary in a prescribed fashion with θ, while at the same time treating θ as a prescribed function of ψ; and by setting $C = \partial\theta/\partial\psi$. In the case of flow in a confined saturated medium, the coefficients $\underline{\underline{K}}$ and C are prescribed at each point in the medium independently of h, and therefore (1) is generally linear in the dependent variable. On the other hand, Boussinesq, diffusion and Richards' equations are nonlinear because at least one of the coefficients, $\underline{\underline{K}}$ and C, varies with h.

When (1) is discretized by the Galerkin finite element method, the result is a system of first-order linear or quasilinear differential equations of the form [cf. Neuman, 1975]

$$\underline{\underline{A}}\,\underline{h} + \underline{\underline{D}}\,\dot{\underline{h}} = \underline{Q} \tag{2}$$

where $\underline{\underline{A}}$ is the conductance (or stiffness) matrix, $\underline{\underline{D}}$ is the capacity matrix, $\underline{h}$ is the dependent variable vector (e.g., hydraulic head, water content), and $\underline{Q}$ is a vector representing sources or sinks. The term $\dot{h}$ represents time derivatives of $\underline{h}$.

In most finite element schemes (e.g., Javandel and Witherspoon [1968], Pinder and Frind [1972]) the capacity matrix $\underline{\underline{D}}$ includes non-zero off-diagonal terms (i.e., it is nondiagonal), and there is growing evidence in the literature suggesting that this may lead to conceptual as well as numerical difficulties. For example, a recent analysis by Narasimhan [1976] indicates that a nondiagonal $\underline{\underline{D}}$ matrix may upset the maintenance of local mass balance in the vicinity of each node, although overall balance over the entire finite element grid may still be preserved. Fujii [1973] showed that when $\underline{\underline{D}}$ is nondiagonal, the time increment Δt in the numerical scheme must not be too large or too small if the maximum principle is to be preserved (the maximum principle states that, in the absence of internal sources q, the maximum value of h must occur at the initial time or at the boundary). When $\underline{\underline{D}}$ is diagonal, the permissible value of Δt is larger and there is no lower limit. This may perhaps explain why, as our experience indicates (2) sometimes yields physically unrealistic values of $\underline{h}$ when there is a sudden and drastic change in $\underline{Q}$ (e.g., the rate of pumpage from certain wells in an aquifer), and why

this can be remedied by diagonalizing the $\underline{\underline{D}}$ matrix as has been done by Wilson [1968], Emery and Carson [1977], Neuman [1973, 1975], Neuman et al. [1976], and others. In fact, Neuman [1973] found by experience that in solving the highly nonlinear problem of saturated-unsaturated ground water flow with the aid of finite elements, the rate of convergence can be improved dramatically by diagonalizing the $\underline{\underline{D}}$ matrix. Neuman used triangular elements, but the same happens when one uses isoparametric elements [Segol, 1976, personal communication]. Similar conclusions have also been reached by Mercer and Faust [1976] in connection with two-phase immiscible flow in porous media.

Neuman and Narasimhan [1977] described an alternative form for (2) with a diagonal capacity matrix which, they showed, has many advantages over the traditional approach. In this new approach, the differential equation (2) is discretized in time by finite differences which enables one to treat them either explicitly, or implicitly, or by an optimum combination of both schemes. The implicit equations are solved by a point iterative technique rather than by a direct method such as Gaussian elimination or Cholesky decomposition. This eliminates one of the major difficulties that groundwater hydrologists have faced in trying to utilize existing finite element computer programs to simulate flow in extensive aquifers. Neuman and Narasimhan [1977] derived local stability and convergence criteria for the explicit-implicit scheme, as well as convergence criteria for the proposed point iterative technique. Their theory assumes that the matrix $\underline{\underline{A}}$ is diagonally dominant, and it is shown that one can always construct a finite element mesh satisfying this condition.

Narasimhan et al. [1977] have incorporated these new ideas into a powerful and versatile computer program entitled FLUMP. They applied it to various linear problems and concluded that the mixed explicit approach is highly efficient for problems that might otherwise involve matrices with large bandwidths (as in the case of extensive aquifers); problems in which the boundary conditions or source terms vary often and rapidly with time as in the case of variable pumpage rates; and problems characterized by a significant spatial variability of element sizes and material properties. They showed that the mixed scheme is capable of yielding highly accurate results to such linear problems, and that it often requires a lesser amount of computer time than other explicit or iterative implicit methods. Neuman et al. [1976] applied FLUMP to various quasilinear problems involving unconfined flow under free surface and saturated-unsaturated conditions. Their results indicated that the mixed approach is especially well suited for quasilinear problems in which the coefficients, the source terms, and the boundary conditions vary with the dependent variable.

The purpose of this paper is to review briefly the theory of the mixed explicit-implicit iterative finite element method and to demonstrate its flexibility and reliability in dealing with a variety of problems arising in subsurface hydrology. The power of the new method is demonstrated by five examples including the build-up of a groundwater ridge, infiltration into an unsaturated soil, drainage of a saturated-unsaturated system, radial flow to a well with well-bore storage, and areal flow in an extensive stream-aquifer system. All of these examples were computed with the aid of a single computer program, FLUMP. A purely

physical derivation of the finite element equations not relying on the Galerkin formalism is also presented in Appendix B.

EXPLICIT - IMPLICIT FORMULATION

In order to discretize (1) in space we adopt a network of triangular elements for plane flow and a network of concentric rings of constant triangular cross-section for axisymmetric problems (Fig. 1). Each triangle is further subdivided along its medians into three subdomains of equal areas, each of which is associated with the nearest nodal point. The collection of all such subdomains associated with any given node n (see shaded area in Fig. 1) is referred to as exclusive subregion, R_n, of n.

In each individual element, e, the dependent variable is described approximately by

$$h(\underline{x},t) = \sum_n h_n(t)\, \xi_n^e(\underline{x}) \qquad (3)$$

where h_n are the values of h at the nodal coordinates, x_n, and $\xi_n^e(\underline{x})$ are linear shape functions of the space coordinates defined in Appendix A and satisfying $\xi_n^e(\underline{x}_m) = \delta_{nm}$, δ_{nm} being Kroenecker delta. With this we are ready to apply the Galerkin method to (1). However, since this method is applicable only at a given instant of time, the time derivative $\partial h/\partial t$ must be determined independently of the Galerkin orthogonalization process, as pointed out earlier by Neuman [1973]. Thus, instead of replacing h in the time derivative by (3) as is usually done in the conventional finite element approach, one is justified to define the nodal values of $\partial h/\partial t$ as averages over the exclusive subregions, R_n, associated with

each node. This leads to a system of first-order linear or quasilinear differential equations of the form given in (2), but with a diagonal capacity matrix. The individual terms of the matrices $\underline{\underline{A}}$ and $\underline{\underline{D}}$ are given by [Neuman, 1973]

$$A_{nm} = \sum_e \frac{\alpha}{4\Delta} \left[K_{xx} b_n b_m + K_{xy} (b_n c_m + b_m c_n) + K_{yy} c_n c_m\right] \tag{4}$$

$$D_{nm} = \delta_{nm} \sum_e \frac{\alpha\Delta}{3} C \tag{5}$$

where Δ is area of triangle, $\alpha = 1$ for plane flow and $\alpha = 2\pi\bar{r}$ for axisymmetric flow ($\bar{r}$ being average radius to the centroid of triangle), b and c are geometric coefficients defined in Appendix A, and the summation sign applies to all elements adjacent to nodal point n.

The finite element differential equations (2), with $\underline{\underline{A}}$ and $\underline{\underline{D}}$ being defined as in (4) and (5), can also be derived from purely physical considerations (that is, without using the Galerkin formalism). Such a derivation is given in Appendix B [see also Narasimhan, 1976]. It is clear from the physical derivations that $-A_{nm}$ represents the rate of fluid transfer into R_n, the exclusive subregion of node n, due to a unit difference in head between nodes m and n. On the other hand, D_{nn} is the fluid mass capacity [Narasimhan and Witherspoon, 1977] of R_n, denoting the ability of R_n to take fluid into storage due to a unit change in head. The fact that D_{nn} is diagonal matrix simply means that the amount of fluid in R_n is a function only of h_n.

If we replace the time derivatives in (2) by finite differences and introduce a weighting factor, θ, we obtain a system of simultaneous linear or quasilinear algebraic equations of the form

$$\underline{\underline{A}}\,[\theta \underline{h}^{k+1} + (1-\theta)\,\underline{h}^{k}] + \underline{\underline{D}}\,\frac{\underline{h}^{k+1} - \underline{h}^{k}}{\Delta t} = \underline{Q} \tag{6}$$

where $0 \leq \theta \leq 1$, Δt is the time increment, and k indicates the number of time steps. Defining a new term

$$\lambda_{nm} = \frac{A_{nm}\,\Delta t}{D_{nn}} \tag{7}$$

and recognizing from (A6) in Appendix A that

$$\lambda_{nn} = -\sum_{m \neq n} \lambda_{nm} \tag{8}$$

we can rewrite (6) as

$$h_n^{k+1} - h_n^k = \theta \sum_{m \neq n} \lambda_{nm} \left(h_n^{k+1} - h_m^{k+1}\right) + (1-\theta) \sum_{m \neq n} \lambda_{nm} \left(h_n^k - h_m^k\right) + \frac{Q_n\,\Delta t}{D_{nn}}$$

$$n = 1, 2, \ldots, N \tag{9}$$

where N is the total number of nodes and the summation is taken over all nodes m in the immediate neighborhood of N. When $\theta = 0$, all the values of h_n^{k+1} can be calculated explicitly from the system of equations (9), which now corresponds to a forward difference scheme in time. When $\theta = 1$, the result is a fully implicit backward difference scheme, whereas $\theta = 1/2$ corresponds to a time-centered or Crank-Nicolson scheme. It is important to recognize that conventional finite element schemes with a nondiagonal $\underline{\underline{D}}$ matrix are inherently implicit [Narasimhan and Witherspoon,

1976] in that h_n^{k+1} can never be computed explicitly from the time-discrete form of (2). Furthermore our formulation in (9) does not include any diagonal terms of the matrix $\underline{\underline{A}}$, which leads to a considerable saving in computer time and storage, especially in dealing with nonlinear problems where the matrix must be recomputed at each time step.

STABILITY AND CONVERGENCE

For the purpose of defining stability and convergence criteria, let us assume that $\lambda_{nm} \leq 0$ for all $m \neq n$. Since $\lambda_{nn} = -\sum_{m \neq n} \lambda_{nm}$, this implies that $\underline{\underline{\lambda}}$ is diagonally dominant, i.e.,

$$\lambda_{nn} = \sum_{m \neq n} |\lambda_{nm}| \quad . \tag{10}$$

Later in the text we will show that one can always construct a finite element mesh which satisfies this requirement at all nodes. Neuman and Narasimhan [1977] showed that when (10) is satisfied, the finite element scheme in (9) is unconditionally stable at node n for all values of θ which are not less than 0.5. When $\theta < 0.5$, stability is conditioned upon the criteria

$$\lambda_{nn} \leq \frac{1}{1 - 2\theta} \qquad \text{or} \qquad \Delta t \leq \frac{D_{nn}}{(1-2\theta)A_{nn}} \tag{11}$$

For example, the explicit version of (9) is stable at any node at which the ratio between capacity and conductance is large enough so that $\Delta t \leq D_{nn}/A_{nn}$ for any given Δt. Conversely, the explicit scheme can be made stable at all nodes by choosing a sufficiently small Δt.

The stability criterion $\Delta t \leq D_{nn}/A_{nn}$ for the explicit scheme has a

simple physical interpretation. According to Appendix B, $\Delta t\, A_{nn}$ is the amount of fluid entering in the exclusive subregion of node n, R_n, when heads at all adjacent nodes, h_m, exceed h_n by unity. On the other hand, D_{nn} is the capacity of the subregion to absorb fluid when h_n changes by one unit. Thus, the above stability criterion merely states that the amount of fluid entering R_n must not exceed the capacity of R_n to absorb fluid. A value of Δt in excess of what is prescribed by the stability criterion would imply that h_n must change by more than unity, which is contrary to the maximum principle (recall that in the explicit scheme, the values of h_m remain fixed during a time step). Note that the identification of the stability criterion based on physical considerations focuses attention on the invariant nature of stability.

We mention that (11) includes the term $(1-2\theta)$ in the denominator; Fujii [1973] derived a global criterion for the maximum principle associated with (9) which includes the term $(1-\theta)$ in the denominator. This implies that our local stability criterion is less restrictive than the global criterion of Fujii.

Neuman and Narasimhan [1977] have also investigated an important question: under what conditions does the numerical scheme in (9) converge to the exact solution of (1) as the mesh is made finer and finer in space and time? Many conventional finite element formulations are known to converge in the mean; however, they seldom guarantee convergence at each point, as is the case with many finite difference schemes. Neuman and Narasimhan [1977] were able to demonstrate that, in the linear case, the explicit-implicit scheme in (9) converges to the exact solution of (1) at each node at which stability and certain local

symmetry conditions are satisfied.

POINT ITERATIVE METHOD

The particular iterative scheme adopted in this work is known as the point acceleration method and was developed originally by Evans et al. [1954]. As will be seen below, it differs from the more familiar point successive over-relaxation technique [Gambolati, 1975]; the latter can be viewed as an extension of the Gauss-Seidel method, whereas the former is more closely related to the point Jacobi method. The acceleration method is readily amenable to an analysis of pointwise convergence and is therefore well suited for the mixed explicit-implicit scheme proposed in this work. This by no means implies that one would not be able to achieve an equally fast, perhaps even faster, rate of convergence with other iterative techniques, such as the point successive over-relaxation method.

The system of equations (9) can be rewritten as

$$\Delta h_n = \sum_{m \neq n} \lambda_{nn} \underbrace{[(h_n^k - h_m^k)}_{\text{Explicit part}} + \theta \underbrace{(\Delta h_n - \Delta h_m)}_{\text{Implicit part}}] + Q_n \, \Delta t / D_{nn}$$

$$n = 1, 2, \ldots N \qquad (12)$$

where $\Delta h_n = h_n^{k+1} - h_n^k$ and it is seen that the implicit part vanishes where $\theta = 0$. The acceleration method consists of introducing the following substitutions into (12):

$$\Delta h_n \text{ (left side)} \rightarrow \Delta h_n^{j+1}$$

$$\Delta h_n \text{ (right side)} \rightarrow (1 + g) \, \Delta h_n^{j+1} - g \, \Delta h_n^j$$

$$\Delta h_m \text{ (right side)} \rightarrow \Delta h_m^j$$

where j is number is iterations and g is acceleration factor. Solving for Δh_n^{j+1}, the acceleration algorithm takes the form

$$\Delta h_n^{j+1} = \frac{\sum_{m\neq n} \lambda_{nm}(h_n^k - h_m^k) - \theta \sum_{m\neq n} \lambda_{nm}(g\,\Delta h_n^j + \Delta h_m^j) + Q_n \Delta t/D_{nn}}{1 - \theta(1+g) \sum_{m\neq n} \lambda_{nm}} \tag{13}$$

The same algorithm can also be written in terms of the residuals ε as

$$\varepsilon_n^{j+1} = \frac{\theta \sum_{m\neq n} \lambda_{nm}\,(g\,\varepsilon_n^j + \varepsilon_m^j)}{1 - \theta(1+g) \sum_{m\neq n} \lambda_{nm}} \tag{14}$$

where $\varepsilon_n^{j+1} = \Delta h_n^{j+1} - \Delta h_n^j$.

The reader may easily recognize the fact that when g is set equal to zero, (13) and (14) reduce to the point Jacobi algorithm. As a matter of contrast, when the relaxation factor in the point successive over-relaxation algorithm is set equal to unity, it reduces to the Gauss-Seidel algorithm.

Neuman and Narasimhan (1977) showed that, for linear problems, the acceleration method converges at any node at which $\underline{\lambda}$ is locally diagonally dominant (we mentioned earlier that it is always possible to construct a mesh which satisfies this requirement), provided that $g \geq 0$. For optimum results, g should not exceed 1 and should usually be less than 0.5. Experience indicates that near-zero values of g may cause difficulties and the optimum value tends to be in the vicinity of 0.2.

The optimum value of g appears to be more stable (i.e., to have a narrower range of values) than the optimum value of the relaxation factor ω in the point successive over-relaxation method.

LOCAL DIAGONAL DOMINANCE

The purpose of this section is to show that one can always construct a finite element network of triangles so as to guarantee that $\underline{\underline{\lambda}}$ will be diagonally dominant. To do so for isotropic domains, we will follow an approach suggested earlier by Gambolati (1973). However, we will follow a simpler and more practical line of reasoning for anisotropic domains.

Consider a triangular element in a plane described by the coordinate system x, y as shown in Fig. 2. Then it is easy to show that

$$\begin{aligned} b_1 &= -a_1 \sin\gamma & c_1 &= a_1 \cos\gamma \\ b_2 &= a_2 \sin(\gamma-\beta_3) & c_2 &= -a_2 \cos(\gamma-\beta_3) \\ b_3 &= a_3 \sin(\gamma+\beta_2) & c_3 &= -a_3 \cos(\gamma+\beta_2) \end{aligned} \tag{15}$$

where b and c are geometric coefficients defined in Appendix A and all other terms are defined in Fig. 2. If $\underline{\underline{A}}^e$ is the contribution of this triangle to the global matrix $\underline{\underline{A}}$ then Eqs. (4) and (15) imply that, in the isotropic domain,

$$\underline{\underline{A}}^e = \frac{K}{4\Delta}\begin{bmatrix} a_1^2 & -a_1a_2\cos\beta_3 & -a_1a_3\cos\beta_2 \\ -a_1a_2\cos\beta_3 & a_2^2 & -a_2a_3\cos\beta_1 \\ -a_1a_3\cos\beta_2 & -a_2a_3\cos\beta_1 & a_3^2 \end{bmatrix} \tag{16}$$

where Δ is area of triangle and K is the scalar equivalent of $\underline{\underline{K}}$. From Appendix A we know that

$$A^e_{nn} = -\sum_{m \neq n} A^e_{nm} \tag{17}$$

indicating that $\underline{\underline{A}}^e$ is diagonally dominant if and only if all the off-diagonal terms in (16) are negative. This means that none of the angles β may exceed 90°, and therefore a sufficient (though not always necessary) condition for the global matrix $\underline{\underline{A}}$ network consists entirely of right and/or acute angled triangles. The same, of course, holds for $\underline{\underline{\lambda}}$.

Next, consider an anisotropic domain with principal hydraulic conductivities K_1 and K_2 oriented parallel to the x and y coordinates, respectively. Then, according to (4),

$$A^e_{nm} = \frac{1}{4\Delta} (K_1 b_n b_m + K_2 c_n c_m) \tag{18}$$

we can now define a new set of coordinates $x' = x/(K_1/K_2)^{1/2}$ and $y' = y$, so that in the transformed domain of x' and y', Eq. (1) will take the form

$$(K_1 K_2)^{1/2} \nabla^2 h = \left(\frac{K_1}{K_2}\right)^{1/2} \left(C \frac{\partial h}{\partial t} + q\right) \tag{19}$$

In other words, the original anisotropic domain in the x, y plane has been transformed into an equivalent isotropic domain with conductivity $(K_1 K_2)^{1/2}$ in the x', y' plane merely by expanding or contracting one of the coordinate axes. If b' and c' are the equivalents of b and c

in the transformed domain, then it is easy to verify that Eq. (18) can also be written as

$$A^e_{nm} = \frac{(K_1 K_2)^{1/2}}{4\Delta'} (b'_n\, b'_m + c'_n\, c'_m) \tag{20}$$

Let a, β, γ in Fig. 2 transform into a', β', γ' in the x' y' plane. Then it is immediately obvious from (15), (16), and (19) that

$$\underline{\underline{A}}^e = \frac{(K_1 K_2)^{1/2}}{4\Delta'} \begin{bmatrix} (a'_1)^2 & -a'_1 a'_2 \cos\beta'_3 & -a'_1 a'_3 \cos\beta'_2 \\ -a'_1 a'_2 \cos\beta'_3 & (a'_2)^2 & -a'_2 a'_3 \cos\beta'_1 \\ -a'_1 a'_3 \cos\beta'_2 & -a'_2 a'_3 \cos\beta'_1 & (a'_3)^2 \end{bmatrix} \tag{21}$$

Thus, following the same line of reasoning as before, it is evident that the global matrix $\underline{\underline{A}}$ is locally diagonally dominant whenever the local network in the transformed domain consists entirely of right and/or acute angled triangles. Again, so is $\underline{\underline{\lambda}}$.

It is easy to show that the same holds true when the principal conductivities K_1 and K_2 are not parallel to the x and y coordinates. For this purpose, it is sufficient to recognize that the solutions of (1) and (9) at any given point in space are independent of the choice of coordinates. Thus $\underline{\underline{\lambda}}$ must remain invariant under a rotation of coordinates and therefore, if it is diagonally dominant in a set of coordinates which is parallel to K_1 and K_2, it must also remain diagonally dominant in another set of coordinates oriented at an angle to the first one. One can therefore transform any anisotropic domain into an equivalent isotropic domain merely by expanding or contracting it parallel to K_1 by the amount $(K_1/K_2)^{1/2}$. If all the triangles in the transformed domain are constructed without obtuse angles, $\underline{\underline{\lambda}}$ will be

diagonally dominant.

In a composite material consisting of several segments with different degrees and orientations of anisotropy, each segment must be transformed separately parallel to its own principal direction of conductivity. Here, in addition to ensuring that all the triangles in the transformed domain are free of obtuse angles, one must also make sure that corresponding nodal points at both sides of a material interface will coincide with each other when the meshes are transformed back into the original plane. This is usually not a very difficult task, as has been demonstrated by an example in Narasimhan et al. [1977].

MIXED EXPLICIT IMPLICIT SOLUTION STRATEGY

The local nature of the stability criteria for (9), together with the use of a point iterative technique, suggest the interesting possibility of solving the finite element equations explicitly at some nodes and implicitly at some other nodes during a single time step. If Δt satisfies the stability condition (11) for some node n, then (9) can be solved explicitly for Δh_n at that node. At nodes which do not satisfy the stability criterion, Δh_n is determined iteratively by using the algorithms in (13) and (14). A final correction is then made to the Δh_n values calculated explicitly, when required, to conserve mass. We refer to this approach as a mixed explicit-implicit solution strategy.

The mixed strategy is very useful in dealing with meshes characterized by a significant spatial variability of element sizes and material properties. For example, if the region of interest consists of two materials having different conductivities and capacities, it

may be possible to solve explicitly in one material and implicitly in the other. The mixed approach is also useful when there is a sudden temporal change in boundary conditions. In this case it is often desirable to use small Δt values for a short period of time until the system reaches a certain level of equilibrium, otherwise there may be a loss of accuracy. The attractive possibility of using an explicit solution procedure during this period may lead to significant savings in computer time.

The idea of combining explicit and implicit calculations in a single time step was previously used in conjunction with an integrated finite difference scheme by Edwards [1972]. The procedure has been incorporated by Edwards into a powerful computer program, called TRUMP, which can handle multidimensional steady state and transient temperature distributions in complex, non-uniform, and isotropic systems, taking into account conduction, convection, and radiation. The program has also been applied by Edwards [1969] to darcian fluid flow in porous media. The conduction aspects of TRUMP are based on a set of algebraic equations which have the same general form as (9). This made it possible for us to develop a computer program which combines the advantages of the finite element method (such as the ability to treat anisotropic regions with complex geometry) with the remarkable logic and facilities of TRUMP. The new program is called FLUMP, as a mnemonic for Finite element and trUMP.

In addition to the mixed explicit-implicit solution strategy, the user of FLUMP has the option of using a fully explicit forward difference scheme ($\theta = 0$), a time-centered Crank Nicolson scheme

($\theta = 0.5$), or a fully implicit backward difference scheme ($\theta = 1.0$), throughout any part of the solution process. However, in practice these options are seldom used because FLUMP has the facility to adjust the weight factor θ automatically during execution in a manner that ensures a high level of accuracy at each time step. This, as well as other special features of FLUMP, are described briefly below.

ITERATIVE APPROACH

The iterative approach is based on (13) and (14). During a given time step, Δt, the algorithm is applied only to a selected number of nodes (called implicit nodes) which cannot be treated explicitly without endangering stability. For the first iteration ($j = 0$), the initial guesses used are $\Delta h_n^o = \dot{h}_n \Delta t$ and $\Delta h_m^o = \dot{h}_m \Delta t$, where $\dot{h}_n$ and $\dot{h}_m$ are estimated time derivatives (a method for obtaining these derivatives is described later in the text). The values of Δh_n^1 are then calculated using (13), and the first set of residuals is determined from the relation $\varepsilon_n^1 = \Delta h_n^1 - \dot{h}_n \Delta t$. The next set of residuals, ε_n^2, is calculated with (14) and Δh_n^2 is found from the relation $\Delta h_n^{j+1} = \Delta h_n^j + \varepsilon^{j+1}$. This procedure is continued until convergence is achieved or until 80 iterations have been completed. In this latter case, the calculations are repeated with half the original value of Δt. This may continue until Δt reaches a minimum specified value, in which case execution terminates with a diagnostic message.

Two convergence criteria must be satisfied simultaneously before terminating the iterations. The first criterion is

$$\max_n |\varepsilon_n^j| \leq 10^{-4} \Delta h_{des} \tag{22}$$

where Δh_{des} is the desired maximum change in h at any node during a time step (as will be seen later, Δt is adjusted during execution to maintain the maximum change in h near the value of Δh_{des}, and less than $2\Delta h_{des}$). The second criterion is based on the net correction to fluid content

$$\Delta H^j_{net} = \left| \sum_n D_{nn} \varepsilon^j_n \right|$$

and the fluid capacity

$$D_{net} = \sum_n D_{nn}$$

of all implicit nodes (the summation is taken over all implicit nodes). The iterative procedure is stopped when (22) is satisfied together with (23) below,

$$\Delta H^j_{net} < 10^{-5} D_{net} \Delta h_{des} \qquad (23)$$

i.e., the net error in fluid content is less than 10^{-5} of the amount of fluid required to change h at each implicit node by the amount Δh_{des}. Experiments by Edwards [1972] on a large number of sample problems using TRUMP indicate that the net cumulative error in the average value of h tends to be no more than 0.01 Δh_{des} after several hundred time steps; the cumulative error at individual nodes does not usually exceed 0.1 Δh_{des}, and is much less if some values of h are fixed at the boundary of the system.

After having completed the iterative procedure for all implicit nodes, one must now correct the values of h at all explicit nodes connected to implicit nodes, according to (see (12))

$$\Delta h_n^{\text{explicit corrected}} = \Delta h_n^{\text{explicit}} + \theta \sum_{m \neq n} \lambda_{nm}(\Delta h_m - \Delta h_n) \qquad (24)$$

where the summation is taken only over implicit nodes. This correction is necessary for a correct material balance. Since internal fluxes between adjacent subregions are calculated simultaneously with the h values, these fluxes must also be corrected in a similar manner.

AUTOMATIC CONTROL OF TIME STEP SIZE

The size of Δt is controlled automatically by several factors such as the lower and upper limits specified by the user (Δt_{low} and Δt_{high}, respectively), the desired maximum change in h at any node during a time step (Δh_{des}), the smallest time step allowed at any explicit node by the stability criterion (Δt_{stab}), the average number of iterations required for convergence, and the desired interval between printed outputs.

The first time step is always 10^{-12} and is used primarily for checking the input data, establishing time derivatives, and determining Δt_{stab} (the latter is recalculated whenever the conductance or capacitance matrices change). The maximum allowed time step, Δt_{max}, is then set equal to 2/3 of Δt_{stab} or Δt_{high}, whichever is smallest (the use of 2/3 of Δt_{stab} instead of Δt_{stab} greatly increases the accuracy in coarse meshes). The minimum allowed time step, Δt_{min}, is set equal to Δt_{low} or 10^{-10}, whichever is greater. If Δt_{low} is equal to or greater than Δt_{max}, the value of Δt_{min} is reduced to slightly less than Δt_{max} so as to prevent the input value from causing instabilities. The default value for Δt_{low} is taken to be $\Delta t_{max}/100$.

During the subsequent calculations the size of Δt is gradually adjusted to obtain a maximum change in h close to Δh_{des} and not exceeding $2\Delta h_{des}$, to maintain the maximum change in any tabulated material property in nonlinear problems near 1% or less and not exceeding 2%, and to prevent the number of iterations from averaging more than 40, the maximum allowed being 80. The technique for doing this has been designed by Edwards [1972] to cause a rapid decrease in Δt when the above limits are exceeded, with a more gradual increase in Δt when changes are relatively slow. For this purpose, let δ be either the largest percentage change property, or 1/40 of the number of iterations required for convergence, whichever is larger. We then calculate the ratio

$$R = \Delta h_{des} / \max \left(\max_n |\Delta h_n|, \delta \, \Delta h_{des} \right) \tag{25}$$

and if $R \leq 0.5$ and $\Delta t \geq 1.01 \, \Delta t_{min}$, the entire computation for the recent time step is repeated with a modified value of Δt. If all the nodes in the mesh are set to be implicit, R is reduced by a factor of 100 for the first time step to start the calculation out smoothly. If $R \leq 1.0$, the new time step is calculated according to

$$\Delta t_{new} = R^2 \, \Delta t_{old} \tag{26}$$

whereas if $R > 1$, the formula is

$$\Delta t_{new} = 0.5 \, (1 + R) \, \Delta t_{old} \quad . \tag{27}$$

In both cases the adjustments are subject to the constraints $\Delta t_{min} \leq \Delta t_{new} \leq \Delta t_{max}$ and $0.5 \, \Delta t_{old} \leq \Delta t_{new} \leq 2.0 \, \Delta t_{old}$.

An additional adjustment in the size of Δt may be required in order to meet a desired interval between printed outputs.

RECLASSIFICATION OF NODES FROM EXPLICIT TO IMPLICIT

If the recent time step was equal to Δt_{max} and less than Δt_{high}, the stability limits of all explicit nodes are tested. All explicit nodes with stability limits equal to or less than 1.8 Δt_{max} are then reclassified as implicit nodes. Since $\Delta t_{max} = (2/3)\ \Delta t_{stab}$, the reclassification affects all explicit nodes having stability limits from 1.0 to 1.2 times Δt_{stab}. This range was chosen empirically by Edwards [1972] in an effort to minimize the required computation time for a large group of test problems using TRUMP.

The rate at which the nodes are reclassified from explicit to implicit depends on the input parameter Δh_{des}; the larger this parameter, the faster the increase in the size of Δt, and therefore the stability limits of most nodes are reached earlier.

ESTIMATION OF TIME DERIVATIVES

The initial guess of h for the iterative procedure requires a preliminary estimate of the time derivatives, $\dot{h}$. In nonlinear problems, the time derivatives are also used to estimate the average values of h to be used in evaluating h-dependent parameters. Rather than saving h_n values from several preceding time steps, which could be used to calculate more accurate time derivatives, a simpler method is used which requires less memory space and machine time and is sufficiently accurate for most problems.

The time derivatives for any time step $t_{k+1} - t_k$ are estimated

from the maximum rates of change in h occurring during the two preceding time steps, $\Delta t_k = t_k - t_{k-1}$ and $\Delta t_{k-1} = t_{k-1} - t_{k-2}$. For this purpose let us define the two ratios

$$R_k = \max_n \left| \frac{h_n^k - h_n^{k-1}}{\Delta t_k} \right| \Big/ \max_n \left| \frac{h_n^{k-1} - h_n^{k-2}}{\Delta t_{k-1}} \right|$$

$$R_t = \frac{\Delta t_k + \Delta t_{k+1}}{\Delta t_{k-1} + \Delta t_k} \tag{29}$$

where $0.5 \leq R_t \leq 2.0$ because, as will be recalled, Δt is not allowed to vary from one time step to another by more than a factor of 2. If $R_k < 1.0$, the maximum rate of change in h is decreasing with time, and the estimate is based on the assumption that all h values are approaching equilibrium exponentially according to the formula $h(t) = h(o)e^{-\alpha t}$. Since $e^{-\alpha} = [h(t)/h(o)]^{1/t}$, it follows that

$$R_{k+1}^{est} = R_k^{\ R_t} \tag{30}$$

where $R_{k+1}^{est} \leq 1$ is the estimated value of R_{k+1}. If $R_k > 1$, the maximum rate of change in h is increasing, and the estimate is based on the assumption that all h values vary quadratically according to the formula $h(t) = h(o) + \dot{h}(o)\, t + \alpha t^2$. Since $\dot{h}(t)/\dot{h}(o) - 1 = 2\alpha t/\dot{h}(o)$, it follows that

$$R_{k+1}^{est} = 1 + (1 - R_k^{\ -1})\, R_t \tag{31}$$

where $1 \leq R_{k+1}^{est} \leq 3$ due to the limits imposed on R_t. Equation (31) gives a more conservative estimate of the maximum rate of change in h than (30) does. The estimated time derivative at each node is calculated as the product of the actual derivative during the previous time step

and R_{k+1}^{est},

$$\dot{h}_n = R_{k+1}^{est} \frac{h_n^k - h_n^{k-1}}{\Delta t_k} \quad . \tag{32}$$

Numerical experiments with TRUMP led Edwards [1972] to conclude that it is advisable to keep $R_{k+1}^{est} = 1.0$ during the first two time steps (a) at the beginning of each problem, (b) after repeating a time step with a modified Δt, and (c) after a node has been reclassified from explicit to implicit. Edwards further concluded that the time derivatives should be set equal to zero or a very small number during the initial time step ($\Delta t = 10^{-12}$) as well as when they change sign. It was also found that more accurate results can be obtained for implicit nodes having stability limits smaller than Δt_k by estimating their time derivatives during the first two time steps according to

$$\dot{h}_n = (h_n^k - h_n^{k-1})/\Delta t_k \left(1 - \sum_{n \neq m} \lambda_{nm}\right) \tag{33}$$

where the values of λ_{nm} correspond to the time step just completed, Δt_k.

Estimation of Implicit Weight Factor

In most implicit procedures, it is customary to employ either a time-centered scheme with $\theta = 0.5$, or a backward difference scheme with $\theta = 1.0$. In FLUMP, θ is allowed to be zero for explicit nodes or to vary between 0.57 and 1.0 for implicit nodes. Experience indicates that small oscillations caused by rapid changes in boundary conditions or variable parameters tend to persist when θ is close to 0.5. The lower limit of 0.57 was chosen empirically by Edwards [1972] to eliminate persistent oscillations and to optimize the stability and accuracy of a large

number of test problems using TRUMP.

The average value of h at any node during a time step is calculated in the program as $\bar{h}_n = h_n^k + \theta\,(h_n^{k+1} - h_n^k)$. Let us assume that h approaches equilibrium exponentially. Then for small time steps and for time steps during which the slope of h remains nearly constant, the correct average value is obtained with $\theta = 0.5$. On the other hand, for large time steps near equilibrium, the correct average value is obtained with $\theta = 1.0$. Thus, θ should be in the vicinity of 0.57 during the period when rapid changes in h take place and should gradually shift toward 1.0 as equilibrium is approached, otherwise there may be a loss of accuracy. One way to accomplish this is by using the empirical formula

$$\theta = \max[0.57, \max(1.0, R_{k+1}^{est})/(1.0 + R_{k+1}^{est})] \quad . \tag{34}$$

Experiments conducted by Edwards [1972] on a large number of problems using TRUMP have shown that approach to equilibrium is usually too rapid when a forward difference or time-centered scheme is used, and much more accurate results can be obtained with a variable θ. His experiments also showed that θ should be set equal to 1.0 during the initial time step ($\Delta t = 10^{-12}$) as well as during any time step following a rejected time step. This enables nodes with small stability limits to reach equilibrium with their neighbors when there is a rapid change in a boundary condition or a variable parameter without overshoot which may lead to damped oscillations.

As mentioned earlier, the computer program also provides an option to fix the value of θ at 0.5, or 1.0 for the entire period of computation, corresponding to explicit forward difference, time-centered, or backward

difference schemes, respectively. However, this tends to reduce accuracy and increase computer time and is therefore not advisable. The purpose of including these options is to allow the calculational results and machine time to be compared with other methods using a fixed value of θ.

ADDITIONAL FEATURES OF FLUMP

The printed output of FLUMP provides information on the nodal values of h, Δh, and estimated value of $\dot{h}$ at discrete time intervals specified by the user. Additional information includes the amount of fluid contained in the exclusive subdomain of each node, change in the amount of fluid in each subdomain during Δt as well as from the start of the problem, total fluid content in the system, flux across the boundary of the system, and the net flux into or out of the exclusive subdomain of each node. This makes it possible to maintain a continuous check on material balance in the subdomain of each node as well as in the system as a whole.

The program also includes a built-in safety feature to warn the user about nodes at which the matrix $\underline{\underline{\lambda}}$ is not diagonally dominant. If the degree of deviation from local diagonal dominance is significant, there is a risk that the solution may be locally unstable (if the node is explicit) and inaccurate, and that convergence will be relatively slow. The problem can always be remedied by locally redesigning the finite element mesh according to the guidelines given earlier in the text. Since the numbering of nodes and elements is completely arbitrary (as opposed to direct methods such as Gaussian elimination or Cholesky decomposition in which numbering has an effect on the band width), local modifications

of the mesh can be easily introduced merely by changing a few cards in the data deck.

APPLICATION TO SUBSURFACE HYDROLOGY PROBLEMS

The nature of our solution process makes it ideally suited for problems encountered under confined, unconfined, and saturated-unsaturated subsurface flow conditions. In the program, such problems are treated simply by specifying selected values of nodal conductivities, capacities, and sources or sinks as tabulated functions either of time or of the dependent variable, h. The boundary conditions can also be controlled in a similar manner by tabulating them as functions of time or h. For saturated-unsaturated flow problems one has the option of tabulating the variables not as a function of total hydraulic head, h, but as a function of pressure head, $\psi = h - z$, where z is elevation. These features of FLUMP are demonstrated below by five examples involving various degrees of complexity and nonlinearity.

BUILDUP OF A GROUNDWATER RIDGE UNDER FREE SURFACE CONDITIONS

Our first example concerns the growth of a groundwater ridge under free surface conditions and the influence of uniform recharge from an infinite strip. This problem has been studied experimentally by Marino [1967] with the aid of a vertical viscous analogue (Hele-Shaw) model, and has been treated analytically by Hantush [1967]. Figure 3 is a schematic diagram of the hydraulic conditions at the bottom and on both sides. It represents a fictitious porous medium having a hydraulic conductivity $K = 0.42$ cm/sec and a specific yield $S_y = 1.00$. Initially, the water table in the model was at an elevation of 11.25 cm.

At time t = 0, a uniform recharge rate of 0.056 cm/sec was inducted over a segment of the model extending from x = 0 to x = 23.8 cm, and this situation was maintained for a prolonged period. This caused the water table to rise at the rate indicated by the solid curves in Fig. 3.

If one assumes that flow underneath the water table is essentially horizontal and that the recharging water reaches the water table instantaneously at t = 0, the flow in the model is governed by the one-dimensional Boussinesq equation

$$\frac{\partial}{\partial x}\left(T(h)\,\frac{\partial h}{\partial x}\right) - q = S_y\,\frac{\partial h}{\partial h} \quad . \tag{35}$$

Here h is the height of the water table above the bottom of the model, q is the specific rate of recharge, and T(h) = Kh is the transmissivity. An analytical solution for a linearized version of this problem has been published by Hantush [1967] and some of his results are shown by triangles in Fig. 3.

To solve (35) with FLUMP we used a finite element mesh of quadrilaterals as shown in Fig. 4. In the program each quadrilateral is automatically divided into two triangles as indicated by the dashed line in Fig. 4. The numerical results for t = 60, 300, and 540 are shown by circles in Fig. 3, and it is seen that they compare favorably with Hantush's analytical solution. The theoretical results also show a good fit with Marino's experimental data at early time. At later values of time, the fit is good at distances exceeding 50 cm from the crest of the ridge but is somewhat less satisfactory near the crest. This may be due to the development of significant vertical hydraulic gradients under the recharging area; such gradients are not taken into account in the Boussinesq

equation.

INFILTRATION INTO AN UNSATURATED SOIL COLUMN

The second example deals with the early absorption stage of infiltration into a vertical unsaturated soil column. Initially the volumetric water content in the soil is $\theta = 0.2376$. At time $t = 0$, the top of the column is brought to a state of saturation at $\theta_{sat} = 0.4950$, and this situation is maintained indefinitely. If one is interested only in the early absorption stage of the resulting infiltration process when gravity is not important (or if the soil column is horizontal), then the flow in the soil is governed by the diffusion equation

$$\frac{\partial}{\partial z}\left(D(\theta)\,\frac{\partial \theta}{\partial z}\right) = \frac{\partial \theta}{\partial t} \tag{36}$$

Here z is depth below the soil surface and $D(\theta)$ is hydraulic diffusivity.

The highly nonlinear relationship between D and θ is given in Table 1. Figure 5 shows the finite element mesh employed in solving this problem with FLUMP. The resulting water content profile at $t = 10^6$ sec is shown in Fig. 6. The numerical results are seen to be in excellent agreement with an analytical solution derived for this problem by Philip [1969], and with numerical results obtained by Van der Ploeg and Benecke [1974] with the aid of CSMP.

DRAINAGE FROM A SATURATED UNSATURATED SYSTEM

The third example is devoted to two-dimensional flow in a vertical plane where part of the porous medium is saturated and part is unsaturated. Vauclin [1975] and Vauclin et al. [1975] described a sophisticated

Table 1. D - θ relationship for adsorption problem.

θ	D, cm^2/sec
0.2376	0
0.2440	1.59×10^{-4}
0.2569	2.13×10^{-4}
0.2698	2.82×10^{-4}
0.2826	3.29×10^{-4}
0.2955	3.73×10^{-4}
0.3084	4.51×10^{-4}
0.3213	6.31×10^{-4}
0.3341	7.91×10^{-4}
0.3470	7.96×10^{-4}
0.3599	7.20×10^{-4}
0.3727	6.69×10^{-4}
0.3856	6.66×10^{-4}
0.3985	6.80×10^{-4}
0.4113	7.43×10^{-4}
0.4242	8.55×10^{-4}
0.4371	9.99×10^{-4}
0.4500	1.31×10^{-3}
0.4628	1.98×10^{-3}
0.4757	3.34×10^{-3}
0.4886	1.046×10^{-2}

and highly accurate drainage experiment in a sand box shown schematically in Fig. 7. Initially, the water table in the box was at an elevation of 143 cm. At time t = 0, the water level in one of the reservoirs attached to the sand box was lowered to an elevation of 80 cm, and this situation was maintained indefinitely. The resulting hydraulic gradient caused the water table to drop gradually at the rate indicated by the solid curves in Fig. 7. During the drainage process the sand remained saturated below the water table; the region above the water table had varying degrees of saturation.

Under this condition, flow in the model can be described by the Richard's equation

$$\frac{\partial}{\partial x}\left(K(\theta)\ \frac{\partial \psi}{\partial x}\right) + \frac{\partial}{\partial z}\left(K(\theta)\ \frac{\partial \psi}{\partial z}\right) = C(\theta)\ \frac{\partial \psi}{\partial t}\ . \tag{37}$$

Here x and z are horizontal and vertical coordinates, respectively; ψ is pressure head defined as $\psi = h - z$ where h is total hydraulic head; K is hydraulic conductivity; and C is specific moisture capacity defined as $C = d\theta/d\psi$. The highly nonlinear relationships between ψ,θ,K, and C are given in Table 2.

Figure 9 shows the finite element network used in simulating the experiment with FLUMP. As seen in Fig. 7, the numerical results for the position of the water table are in excellent agreement with the corresponding experimental data. Figure 8 compares the cumulative depletion from storage as obtained by direct measurement, as computed with FLUMP, and as obtained by an alternating direction implicit finite difference approach by Vauclin et al. [1975]; here again, the numerical results are in fairly good agreement with each other as well as with the experimental data.

Table 2. Relationship between ψ, θ, K, and C for drainage problem.

ψ cm	θ	c, 1/cm	K, cm/sec
-100	0.018		
-90		7×10^{-4}	
-80	0.032		1.94×10^{-7}
-75		1.3×10^{-3}	
-70	0.045		
-65		2.0×10^{-3}	
-60	0.065		6.94×10^{-6}
-55		3.1×10^{-3}	
-50	0.096		1.94×10^{-5}
-45		4.6×10^{-3}	
-40	0.142		6.94×10^{-5}
-30		5.9×10^{-3}	8.33×10^{-4}
-25			1.73×10^{-3}
-20	0.261		3.72×10^{-3}
-17.5		4.2×10^{-3}	
-15	0.282		
-12.5		2.5×10^{-3}	
-10	0.294		1.02×10^{-2}
-7.5		1.0×10^{-3}	
-5	0.299		
-4			1.12×10^{-2}
-2.5		1.6×10^{-4}	
0	0.300	1.0×10^{-5}	1.12×10^{-2}

CONFINED AXISYMMETRIC FLOW TO A WELL WITH WELL-BORE STORAGE

In groundwater hydrology, wells are often treated as line sources having negligible radii. However, when the aquifer transmissivity is small, the water stored in a well may have a considerable effect on discharge and on the hydraulic head in the aquifer [Papadopulos and Cooper, 1967; Narasimhan, 1968; Kipp, 1973; Neuman, 1975]. The problem of simulating well-bore storage under saturated-unsaturated conditions by finite elements was previously discussed by Neuman [1975] and Neuman et al. [1974]. In the present paper, we consider an alternative approach based on the explicit-implicit formulation in (9).

Consider a partially penetrating well together with a superimposed finite element network as shown in Fig. 10. The well has a radius r_w and discharges at a rate Q_w which may vary with time in a prescribed fashion. If the aquifer is confined, the head h_w in the well is the same at all the nodal points n = 1, 2, ..., 6, neglecting the storage capacity of the porous medium near the well in comparison to that of the well itself, we obtain

$$h_w^{k+1} - h_w^k = \frac{\Delta t}{D_w}\left|\theta \sum_n \sum_{m\neq n} A_{mn}\left(h_w^{k+1} - h_m^{k+1}\right) + (1-\theta)\sum_n \sum_{n\neq m} A_{nm}\left(h_w^k - h_m^k\right) + Q_w\right| \tag{38}$$

where D_w, the capacity of the well, is equal to πr_w^2; the first summation is taken over all nodes n = 1,2, ..., 6 along the well-bore. The aquifer has a transmissivity of $9.29 \times 10^{-3} m^2/sec$, and a storage coefficient of 10^{-3}. The well is located at the center of an impermeable circular boundary having a radius of 400 m. It discharges at a constant rate of $1.42 \times 10^{-3} m^3/sec$. To solve this problem with the mixed explicit-implicit scheme, the flow region was discretized into 130 nodes and 100 elements.

The radial spacing of the nodes increased logarithmically from the center.

Figure 11 shows a comparison between the numerical results and two analytical solutions, one developed by Papadopulos and Cooper [1967] for drawdown in a large diameter well, and the other developed by Muskat [1966] for an impermeable circular boundary. Their solution is included for reference purposes. It is seen that the numerical results compare favorably with the analytical solution of Papadopulos and Cooper. The agreement with Muskat's solution is slightly less favorable, possibly because the analytical solution disregards well storage.

Recently, Jargon [1976] pointed out that well-bore storage may also affect the drawdowns at observation wells located at a distance from the pumping well. Figure 12 shows the variation of drawdown with time at a distance of r = 12.5 m from the center of the pumping well.

EXTENSIVE STREAM-AQUIFER SYSTEM

As mentioned earlier, our new approach eliminates many of the difficulties previously encountered in trying to simulate extensive aquifer systems by finite elements. To illustrate the capabilities of FLUMP in treating such systems, we have considered a hypothetical basin as shown in Fig. 13. The basin is 20 km in length, 10 km in width, and it consists of an unconfined aquifer in contact with a stream and its tributary. The aquifer material consists of four different sedimentary units having distinct permeabilities. The thickness of the sediments varies in a step-like fashion from 20 m near the margins to 100 m below the streams. The transmissivity is assumed to vary linearly with the

saturated thickness, as shown in Fig. 14. Water levels in the streams are taken to be constant with time, decreasing monotonically at a slow rate in the downstream direction. The hydraulic contact between the aquifer and the streams is such that water can flow either from the streams into the aquifer, or vice versa, depending on the difference in head between the aquifer and the stream, as shown in Fig. 15. The aquifer is tapped by 21 wells which are indicated by circles in Fig. 13.

The system was discretized into 806 elements with 770 nodal points as illustrated in Fig. 16. The first phase of the simulation consisted of determining the steady state distribution of hydraulic heads in the aquifer prior to pumpage, due to constant head along the streams and along segments of the aquifer boundary. The resulting piezometric surface is depicted in Fig. 17. It can be seen that under steady state conditions, water flows from the stream into the aquifer east of the 106 m contour, whereas the reverse is true in the downstream portion of the basin west of this contour.

The second phase was to investigate what happens if all the 21 wells are turned on simultaneously, each pumping at a rate of 5451 m^3/day (1000 gpm), while head along the streams and the boundaries remains the same as before. Figure 18 shows the distribution of hydraulic heads in the aquifer after one year of pumping. It is seen that steep cones of drawdown have developed around all major pumping centers. In the left central part of the basin, the mutual interference of wells has caused a regional decline in water levels over a significant portion of the basin. Moreover, the pumping wells have captured much of the flow which was originally directed into the streams, so that in the

downstream part of the basin in the streams are now recharging the aquifer.

The problem treated in this example is nonlinear due to the dependence of transmissivities and source terms along the streams on the hydraulic head in the aquifer. For this reason, we found it necessary to restrict the size of each time step to approximately one day. This time step was less than the stability limit of all but 32 nodes in the mesh, so that most of the nodes were treated explicitly during the entire calculation. Each time step required an average of 1.45 seconds of computer time on the CDC 7600. We also ran the same problem with the Crank-Nicolson scheme in which all the nodes are treated implicitly. Here the results were similar to those of the mixed explicit-implicit scheme. However, each time step required on the average 2.89 seconds, i.e., twice as much computer time as the mixed scheme.

APPENDIX A

Consider the matrix $\underline{\underline{A}}$ as defined in (4) for plane flow (i.e., $\alpha = 1$). A typical term contributed by a single triangle such as that shown in Fig. 2 has the form

$$A^e_{nm} = \frac{1}{4\Delta}\left[K_{xx}b_nb_m + K_{xy}(b_nc_m + b_mc_n) + K_{yy}c_nc_m\right] \tag{A1}$$

where

$$\begin{aligned} b_1 &= y_2 - y_3 & c_1 &= x_3 - x_2 \\ b_2 &= y_3 - y_1 & c_2 &= x - x_3 \\ b_3 &= y_1 - y_2 & c_3 &= x_2 - x_1 \end{aligned} \tag{A2}$$

If K_1 and K_2 are the two principal conductivities, and σ is the angle between K_1 and x coordinate, then it can be shown that

$$\begin{aligned} K_{xx} &= K_1\cos^2\sigma + K_2\sin^2\sigma \\ K_{yy} &= K_1\sin^2\sigma + K_2\cos^2\sigma \\ K_{xy} &= (K_1 - K_2)\sin\sigma\cos\sigma \end{aligned} \tag{A3}$$

Substituting (A3) into (A1) and rearranging, we obtain

$$\begin{aligned} A^e_{nm} = \frac{1}{4\Delta}\big[&K_1(b_n\cos\sigma + c_n\sin\sigma)(b_m\cos\sigma + c_m\sin\sigma) \\ &+ K_2(b_n\sin\sigma - c_n\cos\sigma)(b_m\sin\sigma - c_m\cos\sigma)\big] \end{aligned} \tag{A4}$$

indicating that $A^e_{nn} \geq 0$, i.e., the diagonal terms of A^e and $\underline{\underline{A}}$ are always nonnegative.

Furthermore, recognizing that $b_1 + b_2 + b_3 = 0$ and $c_1 + c_2 + c_3 = 0$, we find that (A4) can also be written for $n = m = 1$ as

$$\begin{aligned} A^e_{11} = &-\frac{K_1}{4S}(b_1\cos\sigma + c_1\sin\sigma)\left[(b_2 + b_3)\cos\sigma + (c_2 + c_3)\sin\sigma\right] \\ &-\frac{K_2}{4S}(b_1\sin\sigma - c_1\cos\sigma)\left[(b_2 + b_3)\cos\sigma - (c_2 + c_3)\sin\sigma\right] \\ &= -A^e_{12} - A^e_{13} \end{aligned} \tag{A5}$$

Thus, in general we have

$$A^e_{nn} = -\sum_{m\neq n} A^e_{nm} \qquad \text{and} \qquad A_{nn} = \sum_{m\neq n} A_{nm} \tag{A6}$$

where the summation is taken over all nodes other than n.

The linear shape functions $\xi^e_n(\underline{x})$ in Eq. (3) are given by

$$\xi^e_n = \frac{1}{2}(a_n + b_n x + c_n y) \tag{A7}$$

where b_n and c_n are as in (A2) and

$$\begin{aligned} a_1 &= x_2y_3 - x_3y_2 \\ a_2 &= x_3y_1 - x_1y_3 \\ a_3 &= x_1y_2 - x_2y_1 \end{aligned} \tag{A8}$$

Note that

$$\frac{\partial \xi^e_n}{\partial x} = \frac{b_n}{2\Delta} \; ; \qquad \frac{\partial \xi^e_n}{\partial y} = \frac{c_n}{2\Delta} \tag{A9}$$

APPENDIX B

To derive the finite element equations (2) on the basis of purely physical considerations (i.e., without the Galerkin formalism), let us confine our attention to the exclusive subregion, R_n, associated with a given node n (Fig. 19). According to (3), h inside the particular triangle (n12) is given by the linear interpolation formula

$$h(\underline{x},t) = h_n(t)\ \xi_n^e(\underline{x}) + h_1(t)\ \xi_1^e(\underline{x}) + h_2(t)\ \xi_2^e(\underline{x}) \qquad (B1)$$

Since the functions $\xi^e(\underline{x})$ are linear in the coordinates, $\underline{X}$, (B1) implies that the gradient of h, ∇h, is constant inside the triangle. Furthermore, $\underline{\underline{K}}$ is assumed to be constant in each element, and therefore the flux $\underline{q} = - \underline{\underline{K}}\ \nabla h$ must be uniform inside the triangle. Thus, the sum of the flow rates across the line segments $\overline{AG}$ and $\overline{BG}$ must be equal to the flow rate across $\overline{AB}$. In other words, the flow rate q_{n12} into R_n, contributed by the triangle (n12), is given by

$$q_{n12} = \underline{\underline{K}}\ \nabla h \cdot \underline{n}_{12}\ \overline{AB} \qquad (B2)$$

where $\underline{n}_{12}$ is a unit normal pointing out of the triangle as shown in Fig. 20.

Now $\xi_n^e\ (\underline{x})$ is by definition unity at node n and zero at nodes 1 and 2. Therefore, the derivative of $\xi_n^e\ (\underline{x})$ parallel to the vector $\underline{n}_{12}$ is - 1/L where L is the distance between node n and the line segment joining nodes 1 and 2. The gradient of $\xi_n^e\ (\underline{x})$ is therefore given by

$$\nabla \xi_n^e = - \frac{\underline{n}_{12}}{L} \qquad (B3)$$

Solving (B3) for $\underline{n}_{12}$ and substituting into (B2) we obtain

$$q_{n12} = - \underline{\underline{K}} \nabla h \cdot \nabla \xi_n^e \overline{AB} \, L \tag{B4}$$

However, $\overline{AB}$ L is equal to the area of the triangle, Δ, and from (B1) we see that $\nabla h = h_n \nabla \xi_n^e + h_1 \nabla \xi_1^e + h_2 \nabla \xi_2^e$. Therefore (B4) can be rewritten as

$$q_{n12} = - \Delta \underline{\underline{K}} \left(h_n \nabla \xi_n^e + h_1 \nabla \xi_1^e + h_2 \nabla \xi_2^e \right) \cdot \nabla \xi_n^e \tag{B5}$$

Similar expressions can be developed for q_{n23}, q_{n34}, q_{n45}, and q_{n51}. The sum of all of these terms gives the net flow rate into R_n, q_n, according to

$$q_n = - \sum_m A_{nm} h_m \tag{B6}$$

where $A_{nm} = \sum_e \Delta \underline{\underline{K}} \nabla \xi_n^e \cdot \nabla \xi_m^e$ (B7)

and m takes on the values n, 1, 2, ..., 5.

Substituting (A9) into (B7) leads immediately to the final expression for A_{nm} as given in (4) for $\alpha = 1$. According to (A6), q_n can also be rewritten as

$$q_n = - \sum_{m \neq n} a_{nm} (h_m - h_n) \tag{B8}$$

This indicates that $- A_{nm}$ represents the new flow rate into R_n due to unit difference in head between nodes m and n. The same result was derived by Narasimhan (1975) by evaluating the Galerkin spatial integral over the triangular element e.

The fluid capacitance of the segment NAGB of R_n is equal to $C\Delta/3$ where C is the capacitance of a unit area. Therefore, the total capacitance

of R_n, D_{nn}, is given by

$$D_{nn} = \sum_e \frac{\Delta}{3} C \tag{B9}$$

which corresponds to the diagonal terms of (5) when $\alpha = 1$. Therefore, a complete expression of inarterial balance for R_n can be written as

$$D_{nn} \frac{dh_n}{dt} = q_n + Q_n \tag{B10}$$

where Q_n is the net rate at which fluid is generated inside R_n by internal sources when (B6) is substituted into (B10), we finally obtain

$$\sum_m A_{nm} h_m + D_{nn} \frac{dh_n}{dt} = Q_n \tag{B11}$$

which is the nth equation in the set of finite element equations 2. A similar derivation applies to the axisymmetric case where $\alpha \neq 1$.

ACKNOWLEDGEMENT

This work was partly supported by the U. S. Department of Energy.

REFERENCES

Edwards, A. L., "Trump Computer Program: Calculation of Transient Laminar Fluid Flow in Porous Media, "Rept. UCRL-50664, Lawrence Livermore Laboratory, California, 1969.

Edwards, A. L., "Trump: A Computer Program for Transient and Steady State Temperature Distributions in Multidimensional Systems," Rept. UCRL-14754, Rev. 3, Lawrence Livermore Laboratory, California, 1972.

Emery, A. F. and W. W. Carson, "An Evaluation of the Use of the Finite-Element Method in the Computation of Temperature," Jour. Heat Transfer, Trans. ASME, 93, pp. 136-145, 1971.

Evans, G. W., R. J. Brousseau, and R. Keirstead, "Instability Considerations for Various Difference Equations Derived from the Diffusion Equation," Rept. UCRL-4476, Lawrence Livermore Laboratory, California, 1954.

Fujii, H., "Some Remarks on Finite Element Analysis of Time-Dependent Field Problems," Theory and Practice in Finite Element Structural Analysis, University of Tokyo Press, pp. 91-106, 1973.

Gambolati, G., "Diagonally Dominant Matrices for the Finite Element Method in Hydrology," Int. Jour. Num. Meth. Engng. 6, pp. 587-591, 1973.

Gambolati, G., "Use of the Over-Relaxation Technique in the Simulation of Large Groundwater Basins by the Finite Element Method," Int. Jour. Num. Meth. Engng. 9, pp. 219-233, 1975.

Hantush, M. S., "Growth and Decay of Groundwater Mounds in Response to Uniform Percolation," Water Resources Res. 3, no. 1, pp. 227-234, 1967.

Jargon, J. R., "Effect of Well-bore Storage and Well-bore Damage at the Active Well on Interference Test Analysis," Jour. Pet. Tech., XXVIII, pp. 851-858, 1976.

Javandel, I. and P. A. Witherspoon, "Application of the Finite Element Method to Transient Flow in Porous Media," Soc. Petr. Eng. Jour. 8, p. 241, 1960.

Kipp, K. L., Jr., "Unsteady Flow to a Partially Penetrating Finite Radius Well in an Unconfined Aquifer," Water Resources Res. 9, no. 2, p. 448, 1973.

Marino, M. A., "Hele-Shaw Model Study of the Growth and Decay of Groundwater Ridges," Jour. Geophys. Res. 72, no. 4, pp. 1195-1205, 1967.

Mercer, J. W. and C. R. Faust, "The Application of Finite-Element Techniques to Immiscible Flow in Porous Media,"Proc. First Intl. Conf. on Finite Elements in Water Resources, Princeton, Penrech Press, 1.21-1.58, 1977.

Muskat, M., Flow of Homogeneous Fluids Through Porous Media, J. W. Edwards Inc., Ann Arbor, Michigan, 1966.

Narasimhan, T. N., "Pumping Tests on Open Wells in the Palar Alluvium near Madras City, India -- An Application of the Papadopulos-Cooper Method," Bull. Int. Assn. Sci. Hydr. 13, 91-105, 1968.

Narasimhan, T. N., "A Unified Numerical Model for Saturated-Unsaturated Groundwater Flow," Ph.D. Dissertation, University of California, Berkeley, 1975.

Narasimhan, T. N., "A Perspective on Numerical Analysis of the Diffusion Equation," Preprint No. LBL-4482, Lawrence Berkeley Laboratory, Berkeley, California, 1976.

Narasimhan, T. N. and P. A. Witherspoon, "An Integrated Finite Difference Method for Analyzing Fluid Flow in Porous Media," Water Resources Res. 12, 57-64, 1976.

Narasimhan, T. N. and P. A. Witherspoon, "Numerical Model for Saturated-Unsaturated Flow in Deformable Porous Media, "Water Resources Res. 13, 657-664, 1977.

Narasimhan, T. N., S. P. Neuman and A. L. Edwards, "Mixed Explicit-Implicit Iterative Finite Element Scheme for Diffusion-Type Problems: II. Solution Strategy and Examples," Int. J. Num. Meth. Engng. 11, pp. 325-344, 1977.

Neuman, S. P., R. A. Feddes, and E. Bresler, "Finite Element Simulation of Flow in Saturated-Unsaturated Soils Considering Water Uptake by Plants," 3rd Annual Report, Project No. ALO-SWC-77, Hydrodynamics and Hydraulic Eng. Laboratory, Technion, Haifa, 1974.

Neuman, S. P., R. A. Feddes, and E. Bresler, "Finite Element Analysis of Two Dimensional Flow in Soils Considering Water Uptake by Roots. I. Theory, "Soil Sci. Soc. Am. Proc. 39, no. 2, pp. 224-230, 1975.

Neuman, S. P., "Galerkin Method of Analyzing Nonsteady Flow in Saturated-Unsaturated Porous Media," Finite Elements in Fluids 1, Ch. 10, Gallagher, R. H., J. T. Oden, C. Taylor, O. C. Zienkiewicz, ed., Wiley, London, 1975.

Neuman, S. P., "Saturated-Unsaturated Seepage by Finite Elements," Jour. Hydraulic Div., ASCE 99, pp. 2233-2250, 1975.

Neuman, S. P., T. N. Narasimhan, and P. A. Witherspoon, "Application of Mixed Explicit-Implicit Finite Element Method to Nonlinear Diffusion Type Problems," Proc. Intern. Conf. on Finite Elements in Water Resources, Princeton, 1976 (in press).

Neuman, S. P., and T. N. Narasimhan, "Mixed Explicit-Implicit Iterative Finite Element Scheme for Diffusion-Type Problems: I. Theory, "Int. J. Num. Method. Engng. 11, pp. 309-323, 1977.

Papadopulos, I. S. and H. H. Cooper, Jr., "Drawdown in a Well of Large Diameter," Water Resources Res. 3,pp. 241-244, 1967.

Philip, J. R., "Theory of Infiltration," Advances in Hydroscience, Ven Te Chow, ed., 5, p. 216, 1969.

Pinder, G. F. and E. O. Frind, "Application of Galerkin Procedure to Aquifer Analysis," Water Resources Res. 8, p. 108, 1972.

Van der Ploeg, R. R., and P. Benecke, "Unsteady, Unsaturated, N-Dimensional Moisture Flow in Soil: A Computer Simulation Program," Soil Sci. Soc. Am. Proc. 38, no. 6, pp. 881-885, 1974.

Vauclin, M., "Etude Expérimentale et numerique du drainage de nappes à surface libre: Influence de la zone non saturee," Doctoral Thesis, University of Grenoble, France, 1975.

Vauclin, M. G. Vachaud, and J. Khanji, "Two Dimensional Numerical Analysis of Transient Water Transfer in Saturated-Unsaturated Soils," Computer Simulation of Water Resources Systems, G. C. Vansteenkiste, ed., p. 103, North-Holland Publishing Company, Amsterdam, and American Elsevier, New York, 1975.

Wilson, E. L., "The Determination of Temperatures within Mass Concrete Structures," SESM Rept. No. 68-17, Dept. of Civil Engineering, University of California, Berkeley, 1968.

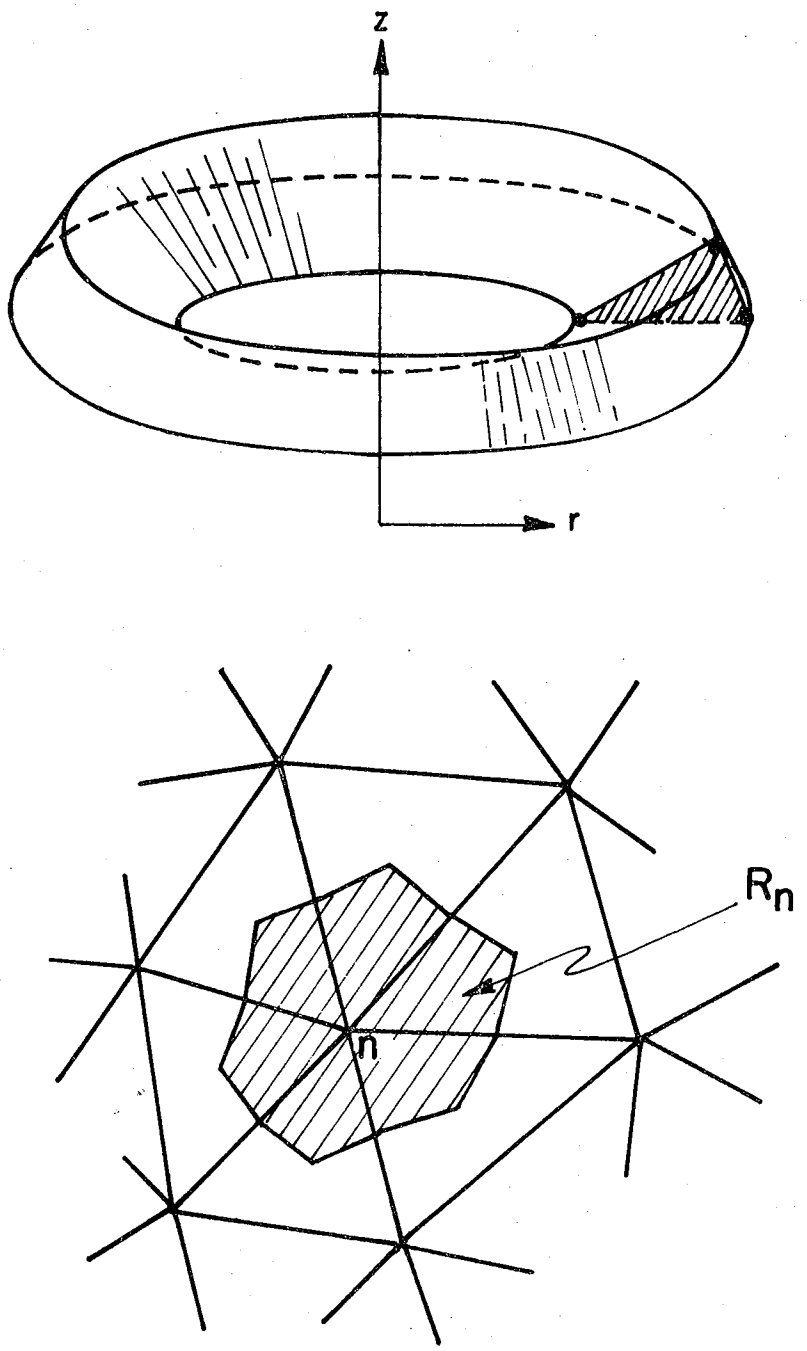


Fig. 1. Exclusive subdomain R_n associated with node n.

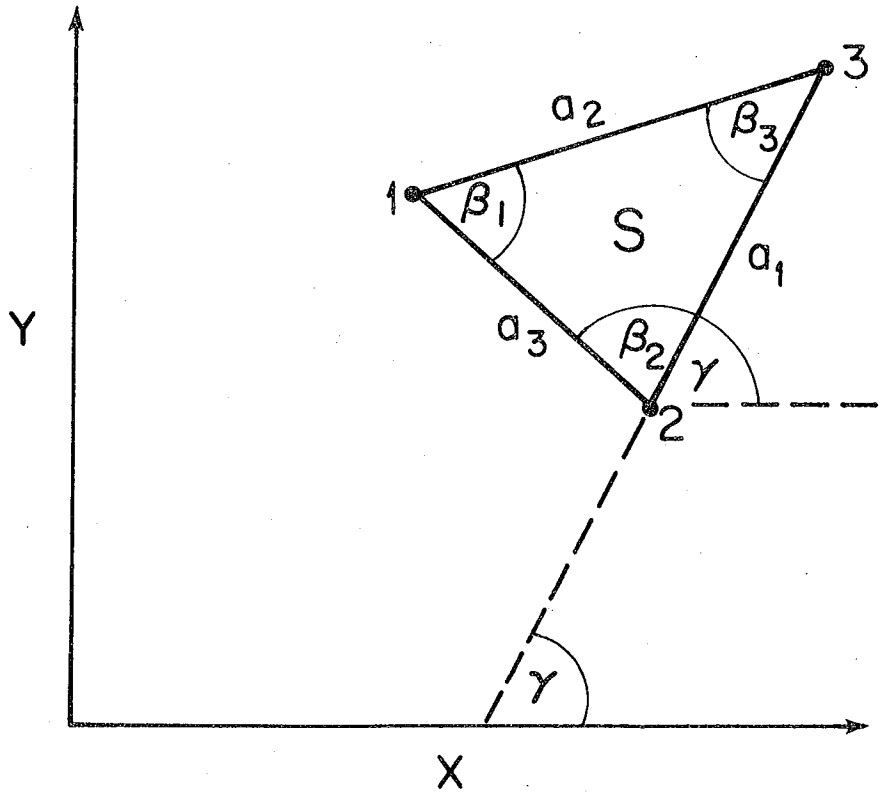


Fig. 2. Triangular element in a fixed coordinate system x,y.

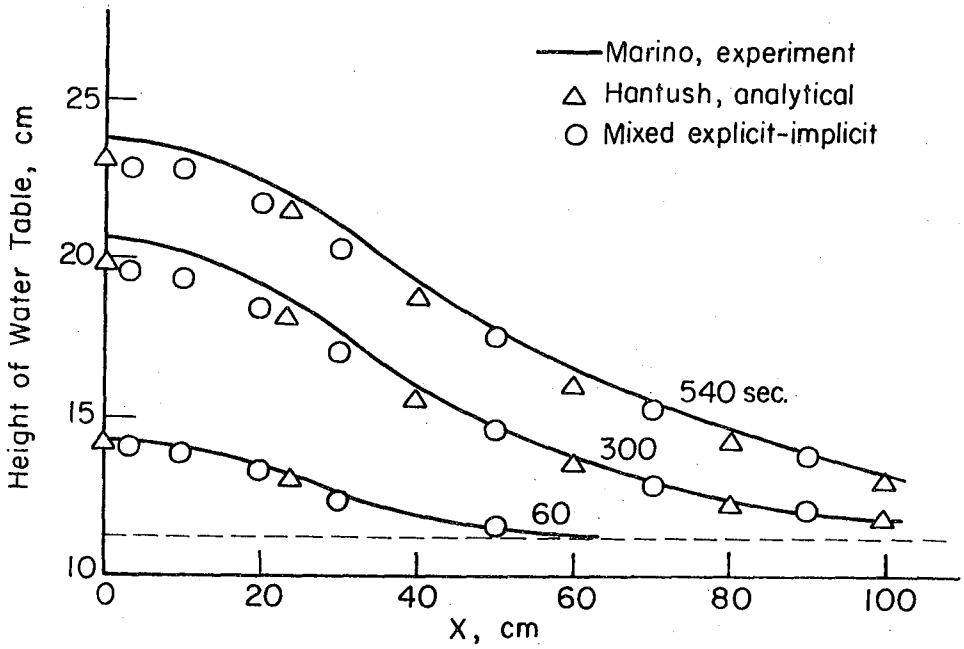


Fig. 3. Growth of groundwater ridge in Marino's (1967) experiment and theoretical results.

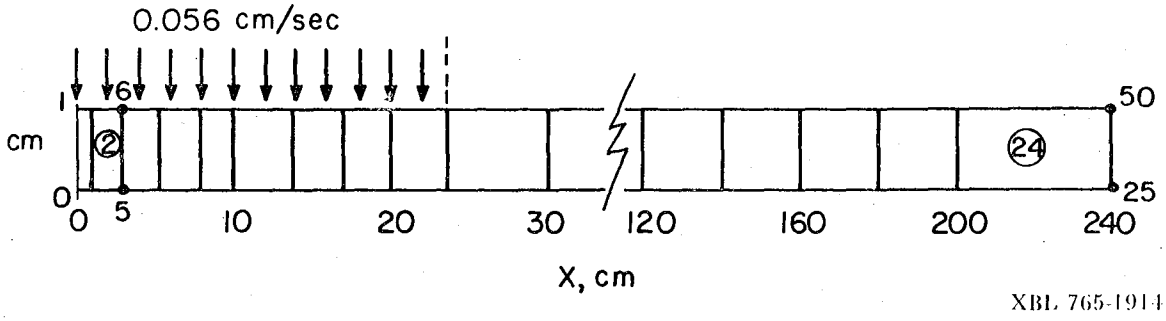
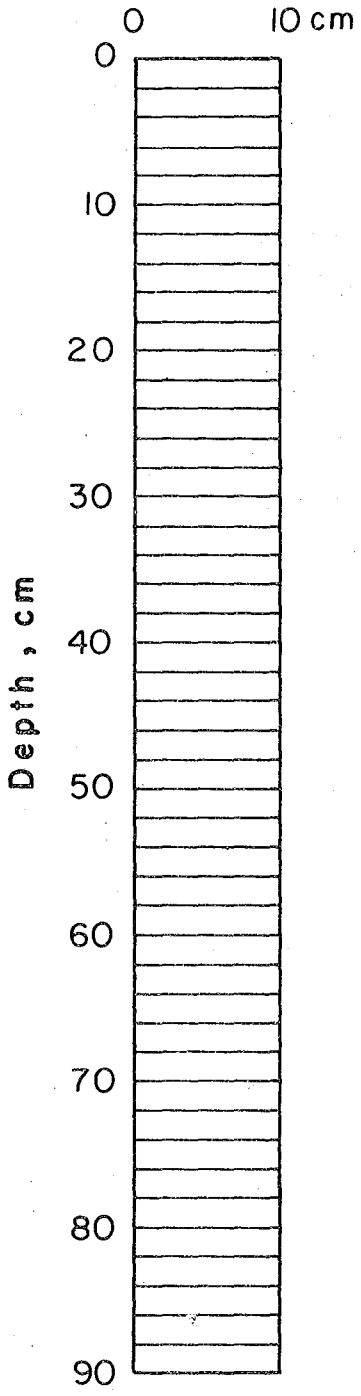


Fig. 4. Finite element mesh used to simulate Marino's (1967) experiment.



XBL 765-1913

Figure 5. Finite element mesh for adsorption problem.

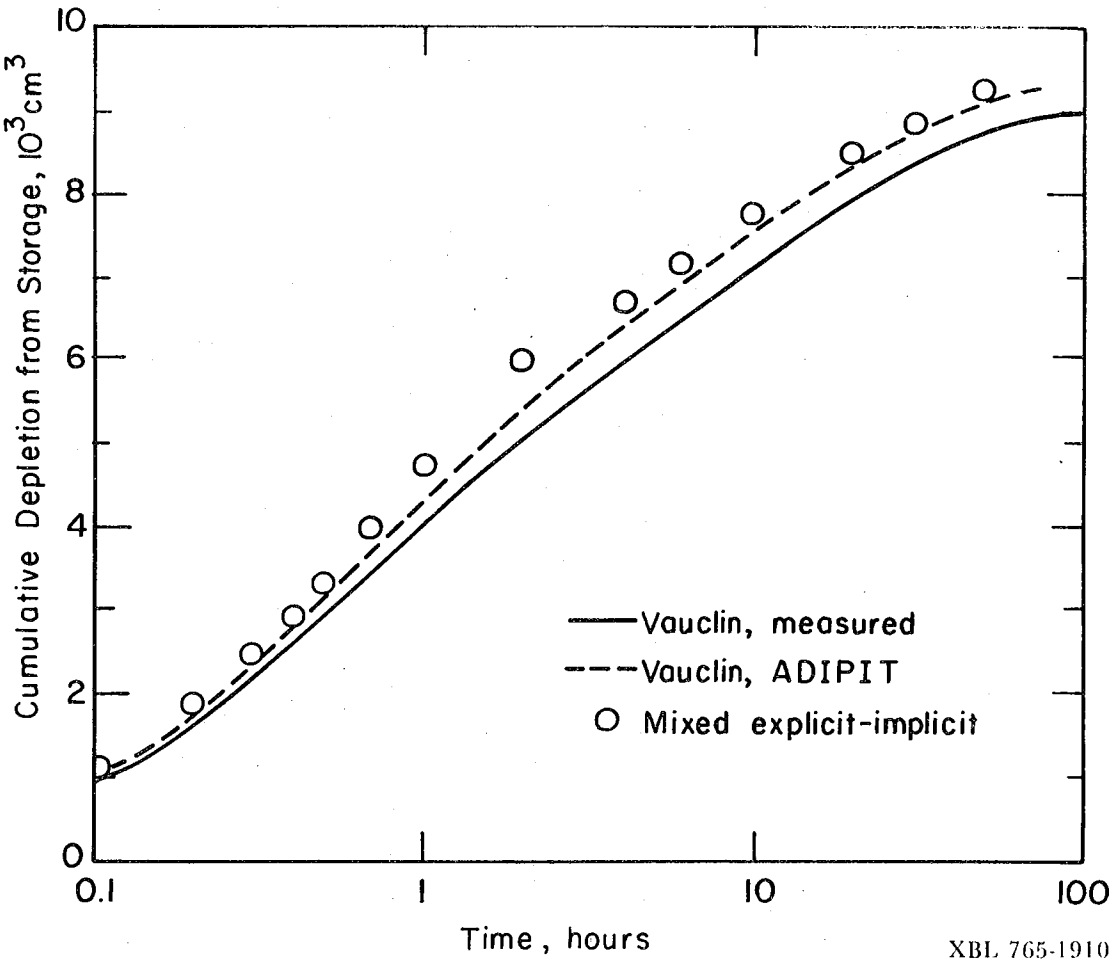
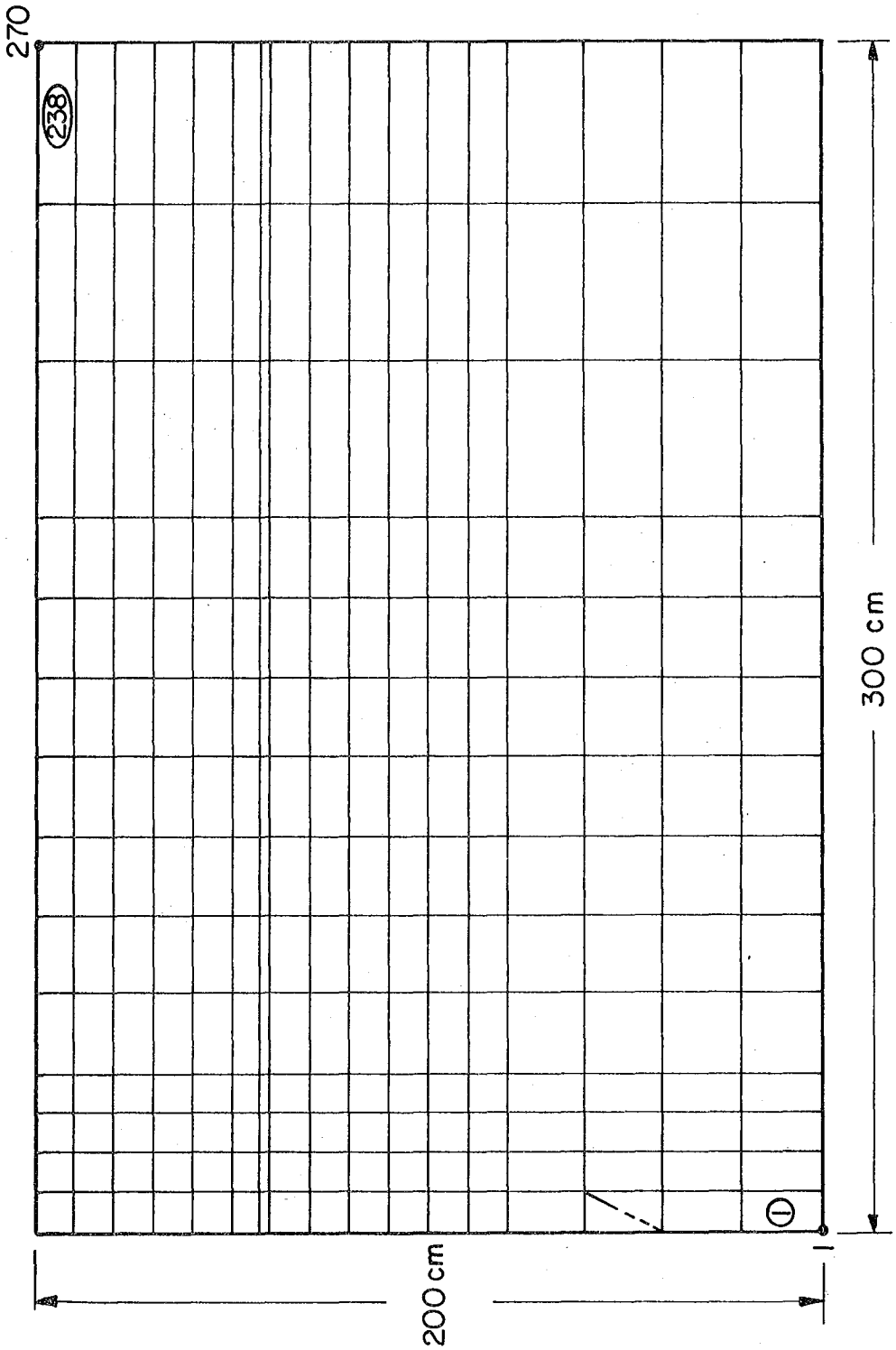
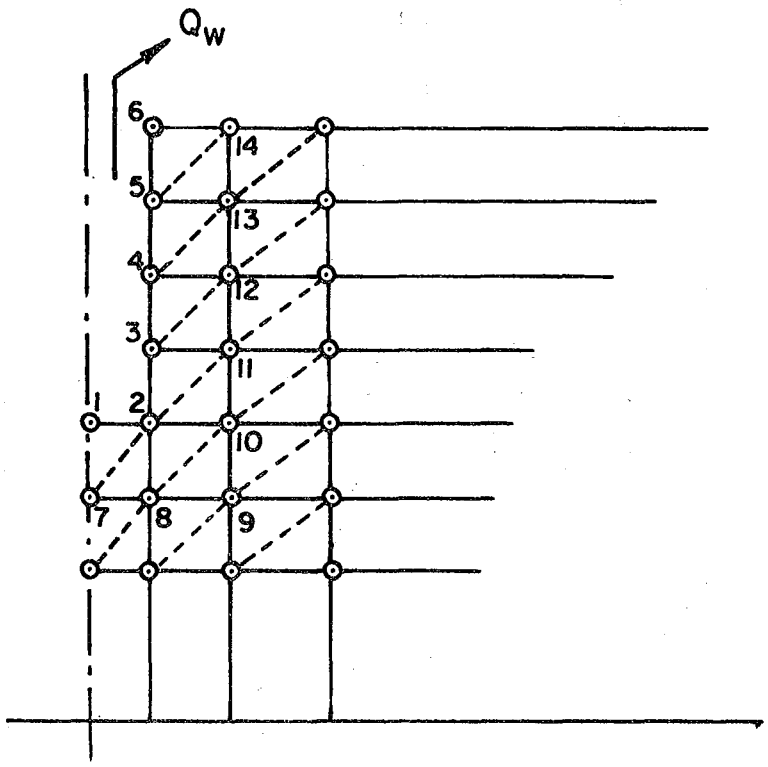


Fig. 8. Cumulative depletion from storage in drainage experiment of Vauclin (1975) and theoretical results.



XBL 765-1909

Figure 9. Finite element mesh used to simulate Vauclin's experiment (1975).



XBL 779-6032

Fig. 10. Finite element mesh around a well of finite diameter.

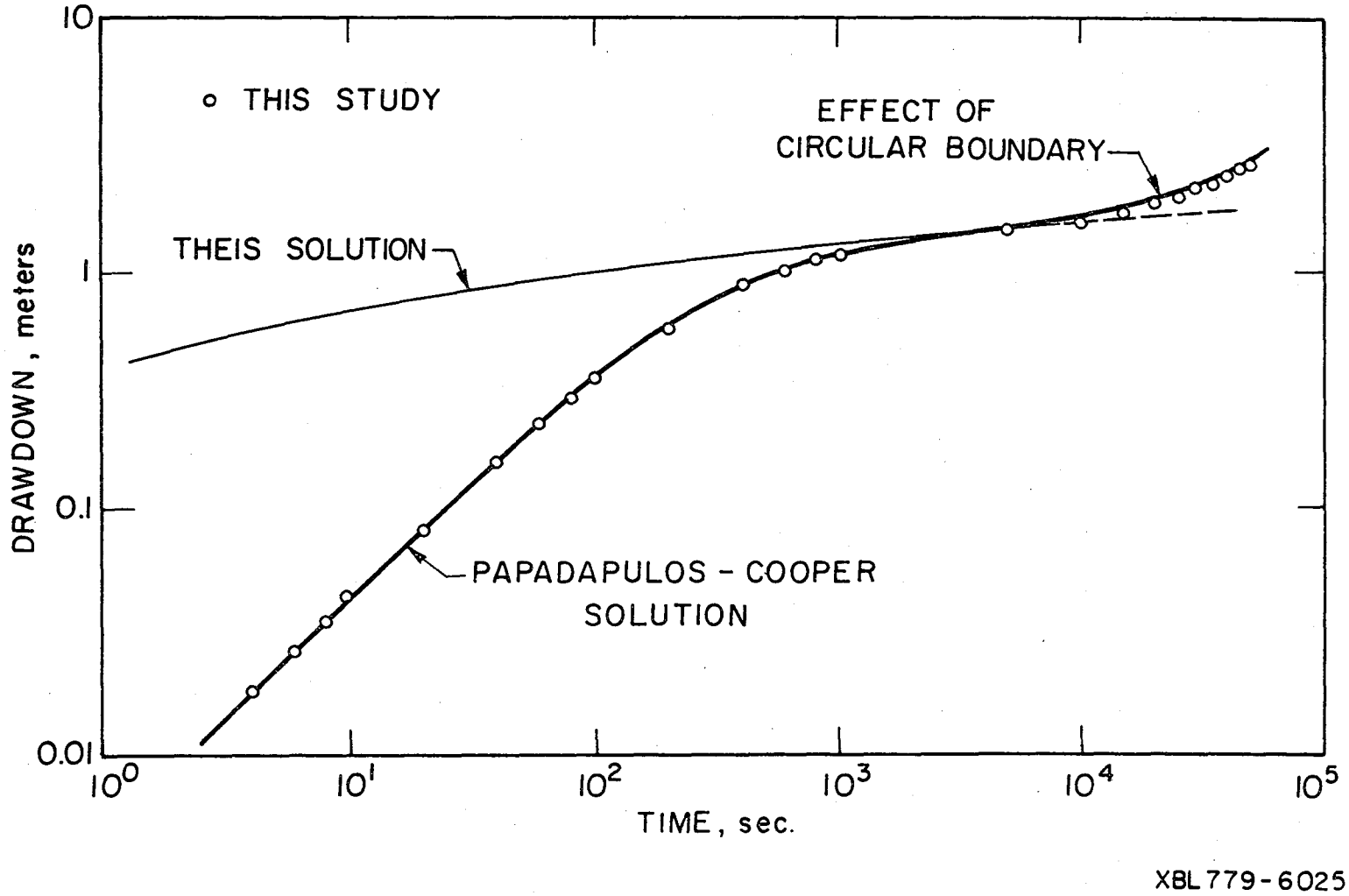


Fig. 11. Drawdown in a finite diameter well: comparison of analytical and numerical results.

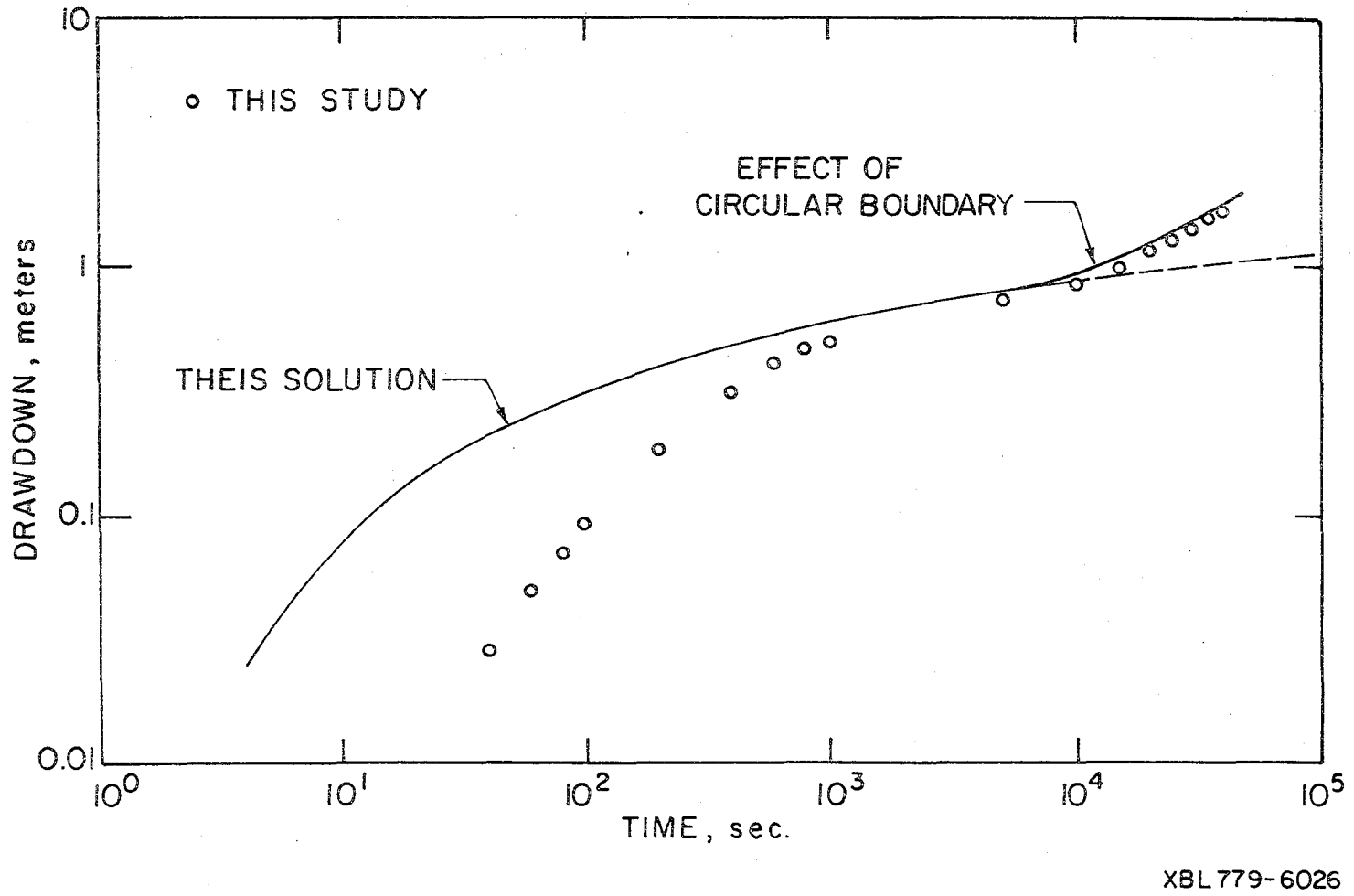


Fig. 12. Effect of well-bore storage at an observation point 12.5 m away from the well.

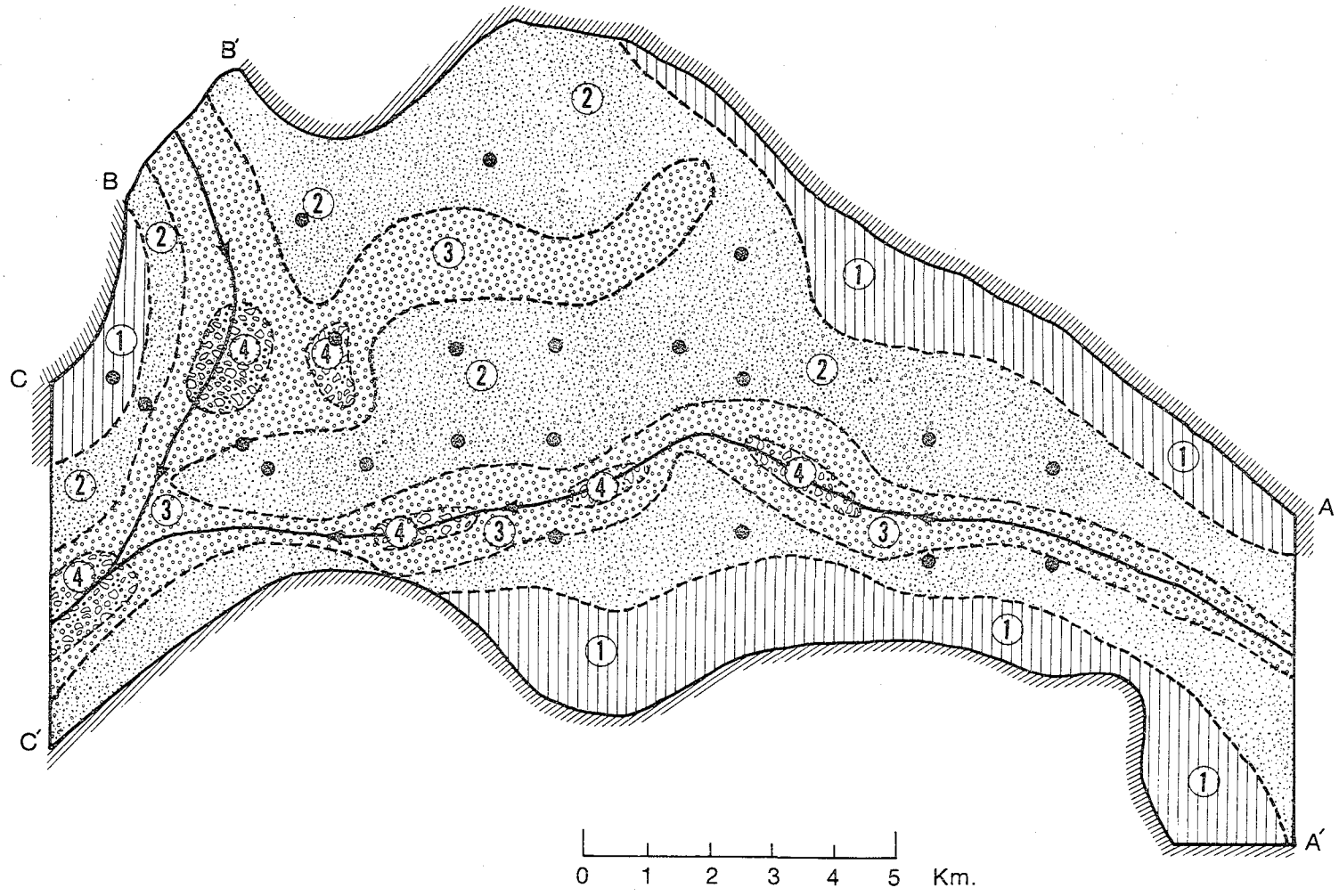


Fig. 13. Hypothetical groundwater basin: distribution of materials.

HYPOTHETICAL GROUNDWATER BASIN

Dependance of Transmissibility on Hydraulic Conductivity and Hydraulic Head.

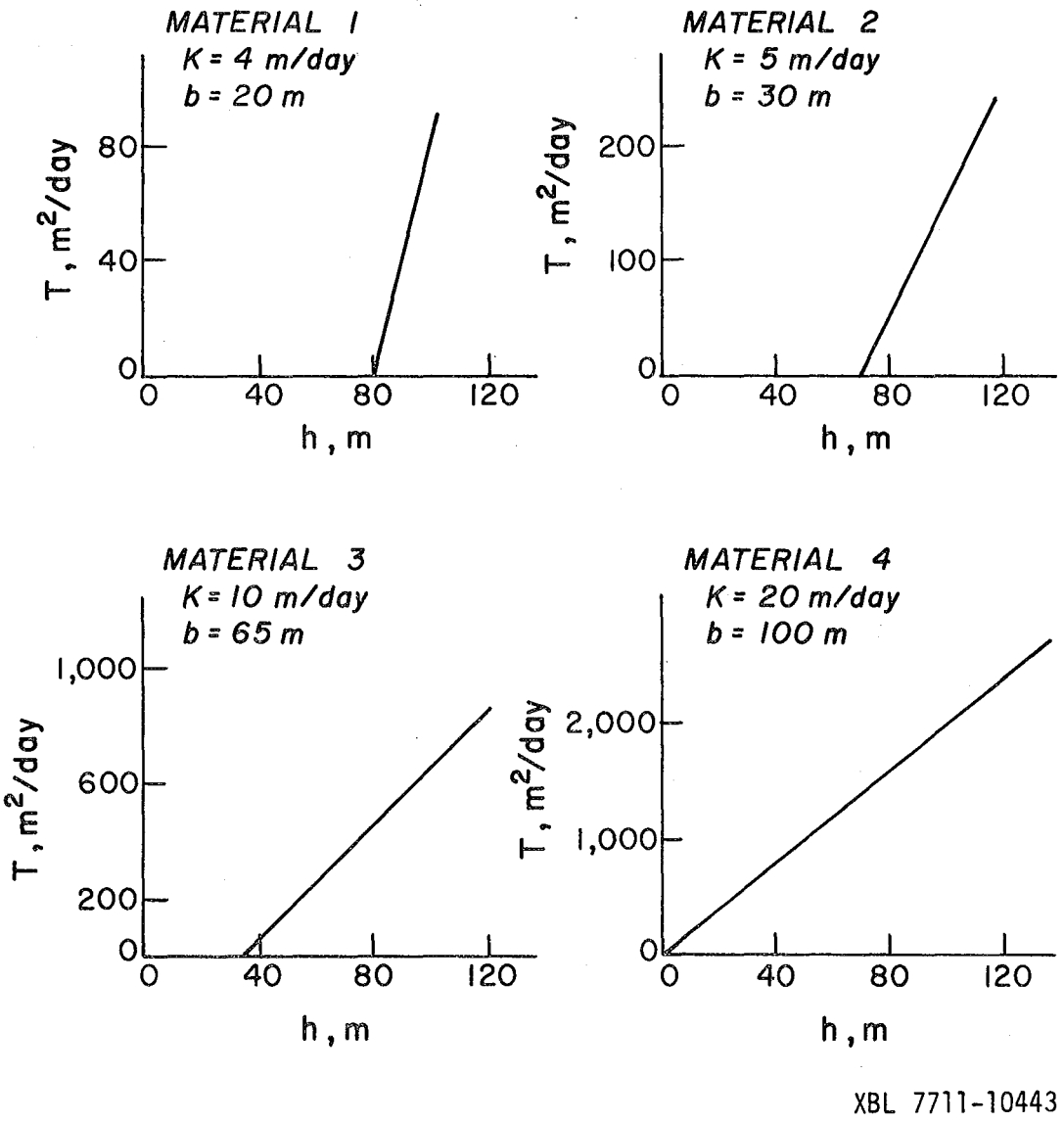


Fig. 14. Hypothetical groundwater basin: dependence of transmissibility on hydraulic head.

HYPOTHETICAL GROUNDWATER BASIN

Relation between River Stage, Water Table, and Infiltration from River.

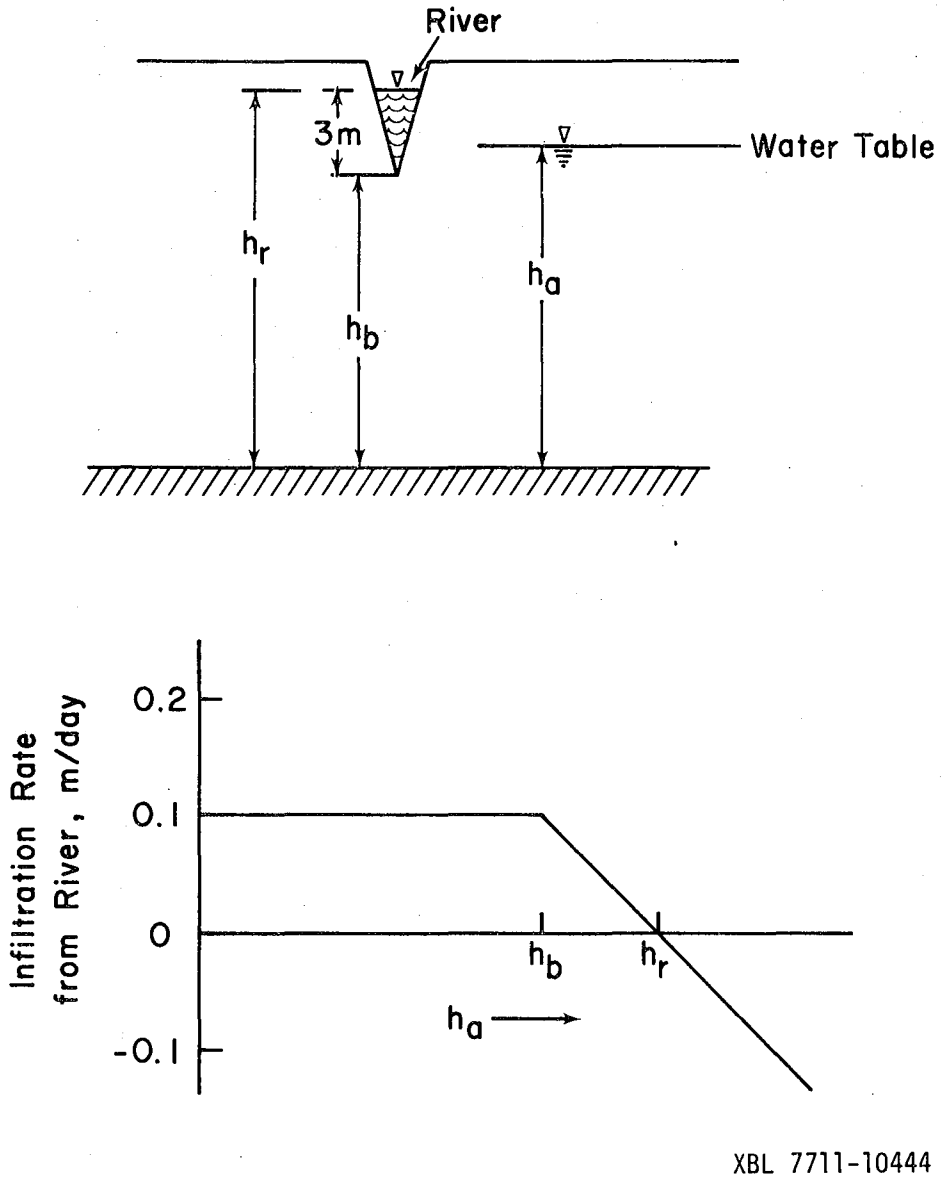
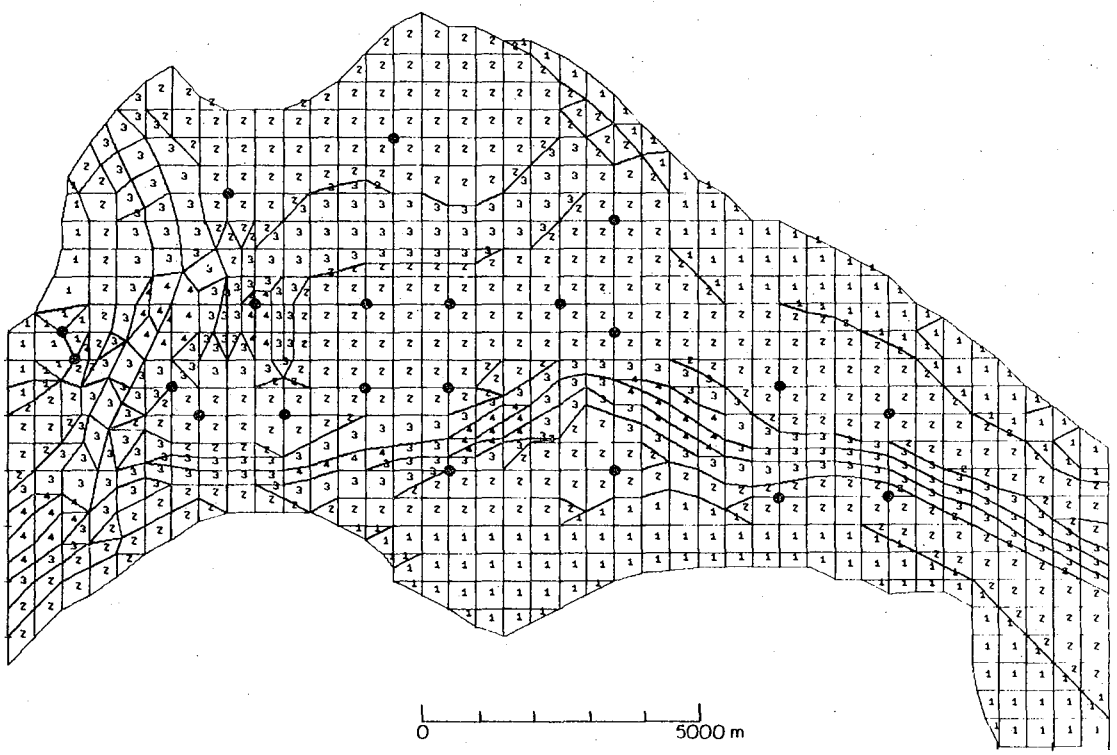


Fig. 15. Hypothetical groundwater basin: relation between river stage, water table and infiltration.



XBL 779-6028

Fig. 16. Hypothetical groundwater basin: finite element mesh used.

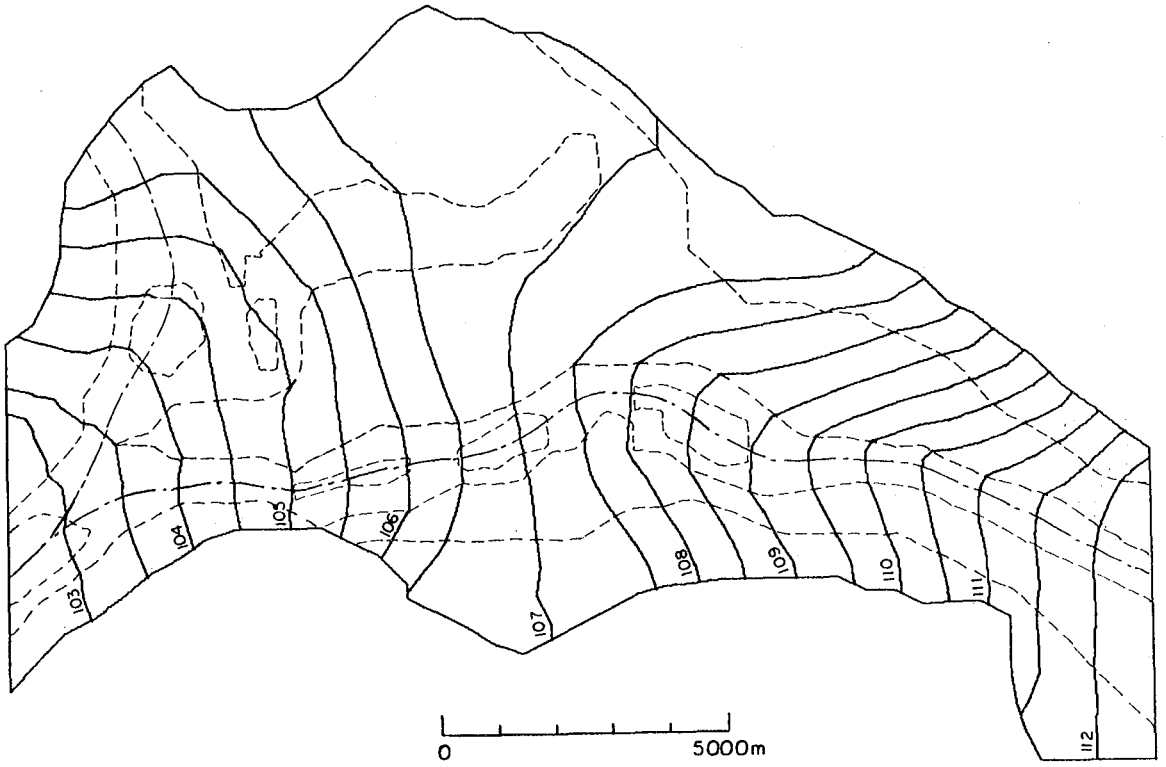


Fig. 17. Hypothetical groundwater basin: steady state distribution of hydraulic head.

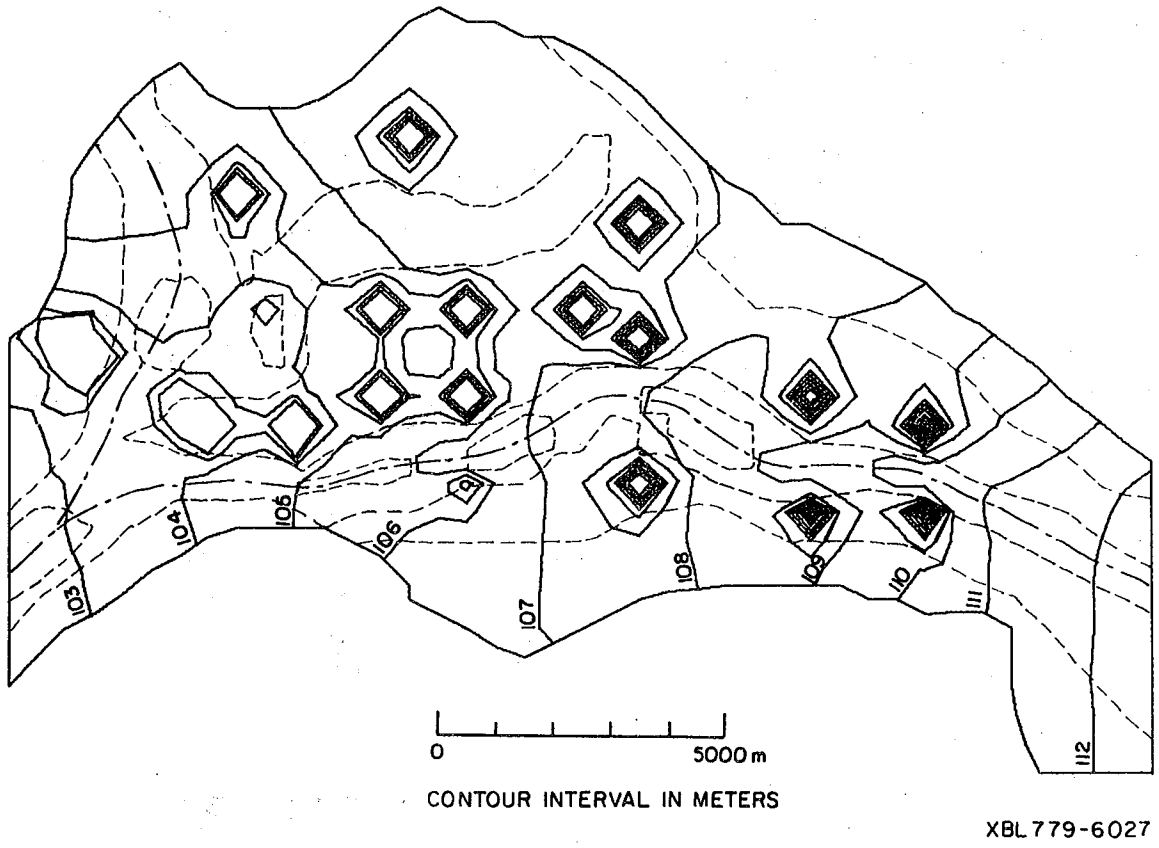
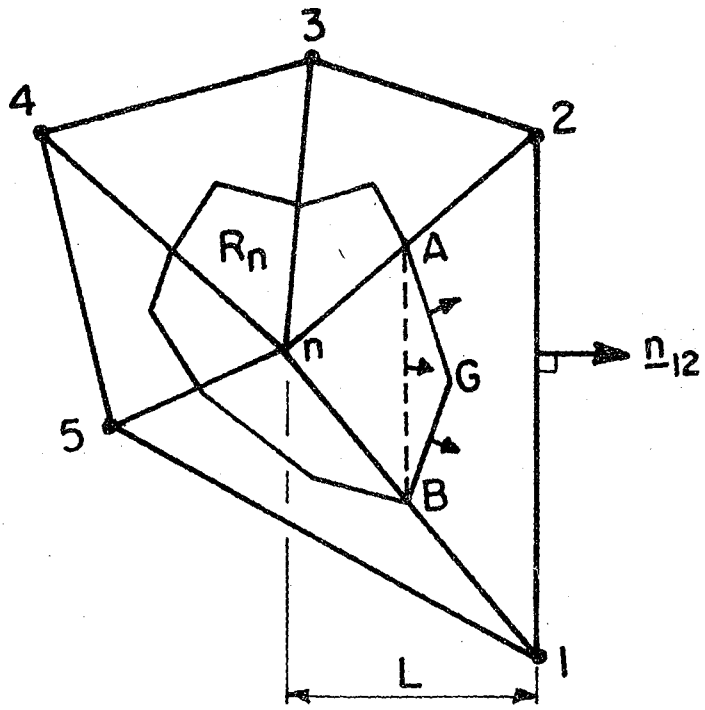


Fig. 18. Hypothetical groundwater basin: distribution of hydraulic head after one year with 21 wells pumping.



XBL779-6030

Fig. 19. Exclusive subregion R_n associated with node n.

This report was done with support from the United States Energy Research and Development Administration. Any conclusions or opinions expressed in this report represent solely those of the author(s) and not necessarily those of The Regents of the University of California, the Lawrence Berkeley Laboratory or the United States Energy Research and Development Administration.